IMCIC'12

and the events organized in its context

icEME 2012
The 3rd International Conference on Engineering and Meta-Engineering

Spring DeMset 2012
The 2nd International Conference on Design and Modeling in Science, Education, and Technology

IDRE&C 2012
International Symposium on Interdisciplinary Research, Education, and Communication

PROCEEDINGS

Post-Conference Edition

Edited by:

Nagib Callaos
Belkis Sánchez
Michael J. Savoie
Mohammad Siddique
Andrés Tremante
C. Dale Zinn

IIIS

Organized by
International Institute of Informatics and Systemics
Member of the International Federation for Systems Research (IFSR)

COPYRIGHT

ISBN- 978-1-936338-57-3

IMCIC'12

PROGRAM COMMITTEE

Chairs: Nagib Callaos (Venezuela)
Jorge Baralt (Venezuela)
Hsing-Wei Chu (USA)
Michael J. Savoie (USA)

Abdel-Qader, Ikhlas	Western Michigan University	USA
Acharya, Sushil	Robert Morris University	USA
Adamopoulou, Evgenia	National Technical University of Athens	Greece
Afandizadeh Z., Shahriar	Iran University of Science and Technology	Iran
Aguirre-Muñoz, Zenaida	Texas Tech University	USA
Akashi, Takuya	Yamaguchi University	Japan
Aksoy, M. S.	King Saud University	Saudi Arabia
Alam, Hassan	BCL Technologies Inc.	USA
Alameh, Kamal	Edith Cowan University	Australia
Alhayyan, Khalid N.	University of South Florida	USA
Alshara, Osama	Higher Colleges of Technology	UAE
Alvarado Moore, Karla	University of Central Florida	USA
Amaral, Luis	University of Minho	Portugal
Ammann, Walter J.	SLF	Switzerland
Anantharaj, Valentine	Mississippi State University	USA
Andina, Diego	Polytechnic University of Madrid	Spain
Anwar, Faisal	Columbia University	USA
Aoki, Toru	Shizuoka University	Japan
Arikan, Emin	Montclair State University	USA
Aruga, Masahiro	Teikyo Heisei University	Japan
Assaf, Mansour	University of Trinidad and Tobago	Trinidad and Tobago
Astakhov, Vadim	University of California San Diego	USA
Aukstakalnis, Nerijus	Kaunas University of Technology	Lithuania
Auvinen, Anssi	University of Tampere	Finland
Balasubramanian, K.	European University of Lefke	Turkey
Bangert, Patrick	Algorithmica Technologies	USA
Barros-Justo, José Luis	University of Vigo	Spain
Baruah, Debendra C.	Tezpur University	India
Belcher, E. Christina	Trinity Western University	Canada
Benbouziane, Mohamed	University of Tlemcen	Algeria
Benedicenti, Luigi	University of Regina	Canada
Bennett, Leslie	University of Louisville	USA
Bermeo, José	University of the Andes	Colombia
Bernardino, Jorge	Coimbra Institute of Engineering	Portugal
Berndt, Donald J.	University of South Florida	USA
Bezuglov, Anton	University of South Carolina	USA
Bhat, Talapady N.	National Institute of Standards and Technology	USA
Bidarra, José	Universidade Aberta	Portugal
Blair, Madelyn	Pelerei, Inc.	USA
Blanchard, Richard	Loughborough University	UK
Bolboaca, Sorana Daniela	Iuliu Hațieganu University of Medicine and Pharmacy	Romania
Bönke, Dietmar	Reutlingen University	Germany

Borchers, Carsten H. J.	University of Hanover	Germany
Botto, Todd	Quinnipiac University	USA
Bradbury, Kyle J.	Tufts University	USA
Bradl, Peter	University of Würzburg	Germany
Branski, L. K.	University of Texas Medical Branch	USA
Bueno, Newton Paulo	Federal University of Viçosa	Brazil
Burke, David	Robert Morris University	USA
Burnett, Andrea	University of the West Indies	Barbados
Burton, Gideon O.	Brigham Young University	USA
Butrous, Nasir	Australian Catholic University	Australia
Calenzo, Patrick	IM2NP	France
Cao, Hong	Southern Medical University	China
Cárdenas, Henry E.	Louisiana Tech University	USA
Cardoso, Eduardo	ITESM	Mexico
Carrasquero, José Vicente	Simon Bolivar University	Venezuela
Carvalho, Marco	Institute for Human and Machine Cognition	USA
Castaño, Bonifacio	University of Alcala	Spain
Cázares-Rangel, Víctor	ITESM	Mexico
Cerny, Václav	University of West Bohemia in Pilsen	Czech Republic
Chae, Hui Soo	Columbia University	USA
Chang, Wen-Kui	Tunghai University	Taiwan
Chen, Huei-Huang	Tatung University	Taiwan
Chen, Lisa Y.	I-Shou University	Taiwan
Chen, Mei-Yung	National Taiwan Normal University	Taiwan
Chen, Shih-Chih	National Taiwan University	Taiwan
Chen, Yuhua	University of Houston	USA
Cherinka, R.	The MITRE Corporation	USA
Cherry, Barbara	Indiana University	USA
Chiou, Yin-Wah	Nanhua University	Taiwan
Cho, Vincent	Hong Kong Polytechnic University	Hong Kong
Choi, Yu-Lee	Chungbuk National University	South Korea
Choo, Jinboo	Korea Electric Power Research Institute	South Korea
Chopra, Deepak	Netaji Subhas Institute of Technology	India
Chou, Andrew	Kainan University	Taiwan
Chou, Hsueh-Cheng	National Taiwan Normal University	Taiwan
Chung, Jen-Yao	IBM Research	USA
Clarke, Tim	University of Wales Institute Cardiff	UK
Coffman, Michael G.	Souther Illinois University Carbondale	USA
Contreras, Sebastián	University of the Andes	Colombia
Cripe, Billy	Oracle Corporation	USA
Curran, Kevin	University of Ulster	UK
Das, Asesh	Pennsylvania College of Technology	USA
De Volder, Dennis	Western Illinois University	USA
Debnath, Rameswar	Neuroscience Research Institute	Japan
Demestichas, K.	National Technical University of Athens	Greece
Desa, Shakinaz	Sultan Idris University of Education	Malaysia
Dhall, Sudarshan	University of Oklahoma	USA
Diallo, Saikou Y.	Old Dominion University	USA
Dierneder, Stefan	Linz Center of Mechatronics GmbH	Austria
Djukom, C. D.	University of Texas Medical Branch at Galveston	USA
Dolan, Dan	South Dakota School of Mines & Technology	USA

Doran, Michael V.	University of South Alabama	USA
Dosi, Vasiliki	University of Ioannina	Greece
Duman, Hakan	University of Essex	UK
Dunning, Jeremy	Indiana University	USA
Dziong, Zbigniew	University of Québec	Canada
Edrington, Chris	Florida State University	USA
Edwards, Stephen H.	Virginia Tech	USA
El-Halafawy, Farag Z.	Menoufiya University	Egypt
Elmahboub, W. M.	Hampton University	USA
El-Mezyani, Touria	Florida State University	USA
Emdad, F.	University of Texas Medical Branch	USA
Erkollar, Alptekin	University of Applied Sciences	Austria
Eshragh, Sepideh	University of Delaware	USA
Eunni, Madhuri	University of Kansas	USA
Fabbri, G.	Sapienza University of Rome	Portugal
Faghri, Ardeshir	University of Delaware	USA
Fang, Chun-Hsiung	National Kaohsiung University of Applied Sciences	Taiwan
Fisher, Wendy	The Open University	UK
Flores, Juan	University of Michoacan	Mexico
Florescu, Gabriela C.	ICI	Romania
Fogelholm, Mikael	Academy of Finland Communications	Finland
Fougeres, Alain-Jerome	University of Technology of Belfort-Montbéliard	France
Fox, Kelly	Texas Tech University	USA
Fu, Shih-Lung	National Taiwan University	Taiwan
Fu, Xiuju	Institute of High Performance Computing	Singapore
Fu, Yonggang	Shanghai Jiao Tong University	China
Fuhrer, Patrik	University of Fribourg	Switzerland
Fujikawa, Takemi	University of Western Sydney	Australia
Fukumi, Minoru	University of Tokushima	Japan
Fumizawa, Motoo	Shonan Institute of Technology	Japan
Fúster-Sabater, Amparo	Spanish Council for Scientific Research	Spain
Gagnon, Francois	School of High Technology	Canada
Ganapathi, Nanthini	Research Triangle Institute	USA
Ganchev, Ivan	University of Limerick	Ireland
García-Atanacio, César	National Nuclear Research Institute	Mexico
Gebre, Biruk A.	Stevens Institute of Technology	USA
Georgescu, Vasile	University of Craiova	Romania
Gesekus, Tim	Deutsche Flugsicherung	Germany
Geyer-Schulz, Andreas	University Karlsruhe	Germany
Goel, Gaurav	Netaji Subhas Institute of Technology	India
Gordon, Richard	University of KwaZulu Natal	South Africa
Goriachkin, Oleg	Volga State University of Telecommunications and Informatics	Russian Federation
Goulding, Tom	Daniel Webster College	USA
Gregory, Mark A.	Royal Melbourne Institute of Technology University	Australia
Guevara López, Miguel	University of Porto	Portugal
Guiasu, Radu Cornel	York University	Canada
Guiasu, Silviu	York University	Canada
Gulez, Kayhan	Yildiz Technical University	Turkey
Guo, Dagang	National University of Singapore	Singapore
Guo, Liwen	Stevens Institute of Technology	USA
Gupta, Neha	Maharaja Agrasen Institute of Technology	India

Gustavsson, Rune	Blekinge Institute of Technology	Sweden
Haba, Cristian-Győző	Gheorghe Asachi Technical University of Iaşi	Romania
Hammond, Bruce R.	Saint Leo University	USA
Hangai, Seiichiro	Tokyo University of Science	Japan
Hartono, Rachmat	BCL Technologies Inc.	USA
Hashimoto, Shigehiro	Osaka Institute of Technology	Japan
Hasnaoui, Salem	National Engineering School of Tunis	Tunisia
Heiserich, Gerd	University of Hannover	Germany
Hemmelman, Brian	South Dakota School of Mines & Technology	USA
Hendel, Russell Jay	Towson University	USA
Henninger, Michael	PH Weingarten	Germany
Herget, Josef	University of Applied Sciences of Eastern Switzerland	Switzerland
Herr, Stephan	Deutsche Flugsicherung	Germany
Hetzer, Dirk	T-Systems Nova Deutsche Telekom	Germany
Hochin, Teruhisa	Osaka Prefecture University	Japan
Hofmeister, Helge	BASF it Services	Belgium
Hoh, Carl	University California San Diego	USA
Horne, Jeremy	Maricopa Community College	USA
Hovakimyan, Anna	Yerevan State University	Armenia
Hsu, Pai-Hui	National Taiwan University	Taiwan
Hsu, Shu-Mei	Tatung University	Taiwan
Hu, Kai-Wei	Tunghai University	Taiwan
Huang, Pin Chia	I-Shou University Kaohsiung	Taiwan
Huang, Ruhua	Wuhan University	China
Huang, Sheng He	University of Southern California	USA
Iembo, Rosanna	University of Calabria	Italy
Ishikawa, Hiroshi	NUIS	Japan
Ito, Akinori	Tohoku University	Japan
Iyyunni, Chakradhar	Larsen & Toubro	USA
Jäntschi, Lorentz	Academic Direct Organization	Romania
Jenq, John	Monclair State University	USA
Jiménez R., Lourdes	University of Alcala	Spain
Johnson, Mark	Benet Laboratories	USA
Jones, Paul	University of Cincinnati	USA
Jonson, Mark	University of New Mexico	USA
Jung, Kyung Im	Samsung Electronics, R. O. Korea	South Korea
Kadoch, Michel	School of High Technology	Canada
Kao, Chih-Yang	Yuan Ze University	Taiwan
Karamat, Parwaiz	The Open Polytechnic of New Zealand	New Zealand
Kasapoglu, Ercin	Hacettepe University	Turkey
Kaszubiak, J.	University Magdeburg	Germany
Katsikas, Sokratis	University of Piraeus	Greece
Katsuyama, Tomoo	Numazu National College of Technology	Japan
Khatri, Anil	Bowie State University	USA
Kim, Jeongdae	Information and Communications University	South Korea
Kim, Kyungwoo	University of Florida	USA
Kim, Seok-Man	Chungbuk National University	South Korea
Kim, Seong-Gon	University of Florida	USA
Kim, Yeo Jin	Samsung Electronics, R. O. Korea	South Korea
Kim, Yong Hak	Korea Electric Power Research Institute	South Korea
Kim, Yong-Gi	Gyeongsang National University	South Korea

Kira, Dennis S.	Concordia University	Canada
Kitagawa, Kazuhiro	World Wide Web Consortium Keio	Japan
Klapp, Jaime	National Nuclear Research Institute	Mexico
Kobal, Damjan	University of Ljubljana	Slovenia
Kouki, A.	School of High Technology	Canada
Krakowska, Monika	Jagiellonian University	Poland
Kreisler, Alain	University of Paris-Sud	France
Kromrey, Jeffrey D.	University of South Florida	USA
Kronreif, Gernot	Austrian Research Centers	Austria
Kulkarni, Arun	University of Texas at Tyler	USA
Kumar, Aman	BCL Technologies Inc.	USA
Kung, C. M.	I-Shou University	Taiwan
Kurita, Takio	Hiroshima University	Japan
Kutter, Anna K.	University of Education Weingarten	Germany
Kwok, Kenneth	National University of Singapore	Singapore
Lahlouhi, Ammar	University of Biskra	Algeria
Laksmivarahan, S.	University of Oklahoma	USA
Latawiec, Krzysztof J.	Opole University of Technology	Poland
Latchman, Haniph A.	University of Florida	USA
Lau, Newman	Hong Kong Polytechnic University	Hong Kong
Lee, Hwajung	Radford University	USA
Lee, Jae-Suk	Gwangju Institute of Science and Technology	South Korea
Lee, Nam Ho	Korea Electric Power Research Institute	South Korea
Lee, Suk Bong	Samsung Electronics, R. O. Korea	South Korea
Lee, Yih-Jiun	Chien Kuo Technology University	Taiwan
Letellier, T.	Victor Segalen Bordeaux 2 University	France
Leu, Fang-Yie	Tunghai University	Taiwan
Li, Xiaoping	National University of Singapore	Singapore
Lin, Feng-Tyan	National Taiwan University	Taiwan
Lin, Hong	University of Houston-Downtown	USA
Lin, Huan Yu	National Chiao Tung University	Taiwan
Lin, Shu-Chiung	Tatung University	Taiwan
Linares, Oscar	Francisco Jose of Caldas District University	Colombia
Lind, Nancy	Illinois State University	USA
Litvin, Vladimir	California Institute of Technology	USA
Liu, Jun	University of Ulster	UK
Liu, Kuo-Shean	Tatung University	Taiwan
Liu, Shih-Chi	Tatung University	Taiwan
Livne, Nava L.	University of Utah	USA
Livne, Oren E.	University of Utah	USA
Long, Changjiang	Huazhong University	China
López Zafra, Juan M.	Complutense University of Madrid	Spain
Lou, Shi Jer	NPUST	Taiwan
Loutfi, Mohamed	University of Sunderland	UK
Love C., Gloria	Dillard University	USA
Lowe, John	University of Bath	UK
Lowry, Pam	Lawrence Technological University	USA
Lyell, Margaret	Intelligent Automation, Inc.	USA
Lyuu, Yuh-Dauh	National Taiwan University	Taiwan
Ma, Yongqing	Victoria University of Wellington	New Zealand
Machotka, Jan	University of South Australia	Australia

Maggiore, Giulio	Telecom Italia	Italy
Mahendran, Francis	Motorola Software Group Singapore	Singapore
Mahgoub, Ahmed G.	Alexandria University	Egypt
Mak, Peng Un	University of Macau	Macau
Makaruk, Aleksander	Vienna University of Technology	Austria
Manley, Denis	Dublin Institute of Technology	Ireland
Mansikkamäki, Pauliina	Tampere University of Technology	Finland
Marino, Mark	Erie County Community College	USA
Marques Cardoso, A. J.	University of Coimbra	Portugal
Martínez, Pablo	Laboratory for Advanced Brain Signal Processing	Japan
Mascari, Jean-François	National Research Council Italy	Italy
Masoumi, Nasser	University of Tehran	Iran
Mathews, Brian	University of Bedfordshire	UK
Matsumoto, Kazunori	KDDI R&D Laboratories Inc.	Japan
Matsuno, Akira	Teikyo University	Japan
Mbobi, Aime Mokhoo	Centennial College	Canada
Mehrabian, Ali	University of Central Florida	USA
Michaelis, B.	University Magdeburg	Germany
Migliarese, Piero	University of Calabria	Italy
Miller, R.	The MITRE Corporation	USA
Minnaert, Eric	South Dakota School of Mines & Technology	USA
Mitchell, Charles	Grambling State University	USA
Mitrovic, Frane	University of Split	Croatia
Moosavian, Ali A.	K.N. Toosi University of Technology	Iran
Moreau, Alban	University of Brest	France
Morii, Hisashi	Shizuoka University	Japan
Mostafaeipour, Ali	Yazd University	Iran
Muknahallipatna, Suresh	University of Wyoming	USA
Munitic, Ante	University of Split	Croatia
Muñoz, Humberto	Southern University	USA
Mura, Gianluca	Polytechnic University of Milan	Italy
Muraleedharan, Rajani	Syracuse University	USA
Naddeo, Alessandro	University of Salerno	Italy
Nadworny, Margaret	Global Software Group, Motorola	USA
Nagalakshmi, Vadlamani	Gandhi Institute of Technology and Management	India
Nagaoka, Tomoyuki	Hoshino Resort Ltd.	Japan
Nahmens, Isabelina	University of Central Florida	USA
Nair, V. S. Sukumaran	Southern Methodist University	USA
Nakashima, Takuya	Shizuoka University	Japan
Natriello, Gary	Columbia University	USA
Nave, Felecia M.	A&M University	USA
Nawawi, S. W.	University Technology of Malaysia	Malaysia
Nazmy, Taymoor M.	Ain Shams University	Egypt
Nedic, Zorica	University of South Australia	Australia
Nelwamondo, Fulufhelo	University of the Witwatersrand	South Africa
Neo, Yoichiro	Shizuoka University	Japan
Noh, Bong Nam	Chonnam National University	South Korea
Nousala, Susu	University of Melbourne	Australia
Novikov, Oleg	Tomko Inc.	Russian Federation
Nyan, M. N.	National University of Singapore	Singapore
Obrebski, Jan B.	Warsaw Polytechnic	Poland

O'Brien, Ann	Loughborough University	UK
Ocelka, Tomas	Institute of Public Health Ostrava	Czech Republic
Oh, Yun Sang	Samsung Electronics, R. O. Korea	South Korea
Olla, Phillip	Madonna University	USA
Ong, Soh-Khim	National University of Singapore	Singapore
Ong, Vincent Koon	University of Bedfordshire	UK
Osadciw, Lisa Ann	Syracuse University	USA
Ostergaard, Soren Duus	IBM Europe	Denmark
Overmeyer, Ludger	University of Hannover	Germany
Oya, Hidetoshi	Shonan Institute of Technology	Japan
Ozdemir, Ahmet S.	Marmara University	Turkey
Palesi, Maurizio	University of Catania	Italy
Paré, Dwayne E.	University of Toronto Scarborough	Canada
Passos, Nelson	Midwestern State University	USA
Patel, Nishit	Stevens Institute of Technology	USA
Pécatte, Jean-Marie	University Paul Sabatier-Toulouse 3	France
Pedrycz, Witold	University of Alberta	Canada
Peng, Jian	University of Saskatchewan	Canada
Pfeifer, Michael	Technical University of Dortmund	Germany
Phillips, C. Dianne	NorthWest Arkansas Community College	USA
Phillips, James	Louisiana Tech University	USA
Pierce, Ryan M.	Arkansas State University	USA
Pieters, Cornelis P.	University for Humanistics	Netherlands
Pitzer, Erik	University of Applied Sciences Hagenberg	Austria
Platt, Glenn	CSIRO	Australia
Pochiraju, Kishore	Stevens Institute of Technology	USA
Podaru, Vasile	Military Technical Academy	Romania
Polanski, Andrzej	Silesian University of Technology	Poland
Poon, Gilbert	University of Waterloo	Canada
Postolache, Octavian	Institute of Telecommunications	Portugal
Potente, Till	RWTH Aachen University	Germany
Praus, Petr	Charles University in Prague	Czech Republic
Pulcher, Karen L.	University of Central Missouri	USA
Pun, Daniel	Central Queensland University	Australia
Putnam, Janice	University of Central Missouri	USA
Rachev, Boris	Technical University of Varna	USA
Rahman, Anuar Abdul	Pusat Tenaga Malaysia	Malaysia
Rahman, Fuad	BCL Technologies Inc.	USA
Repa, Václav	University of Economics in Prague	Czech Republic
Revetria, Roberto	University of Genoa	Italy
Reyes G., Leonardo	Carnegie Mellon University	USA
Reyes-Méndez, Jorge J.	Metropolitan Autonomous University	Mexico
Riaz M., Mohammad	University College of Engineering and Technology	Pakistan
Rigaud, Bernard	Jean-François Champollion University	France
Rodríguez, María D.	University of Alcala	Spain
Rossignol, R.	Victor Segalen Bordeaux 2 University	France
Saadane, Rachid	Institut Eurecom	France
Sahara, Tomohiro	Osaka Institute of Technology	Japan
Sala, Nicoletta	University of Italian Switzerland Largo Bernasconi	Switzerland
Salazar, Dora	Texas Tech University	USA
Saleh, Magda M.	Alexandria University	Egypt

Šaruckij, Mark	Cracow University of Economics	Poland
Sato, Tomoaki	Hirosaki University	Japan
Savva, Andreas	University of Nicosia	Cyprus
Sax, Eric	Mercedes-Benz Technology Gmbh	Germany
Schaeffer, Donna M.	Marymount University	USA
Schlette, Christian	RWTH Aachen University	Germany
Schrader, P. G.	University of Nevada	USA
Schulz, Lennart	Institute of Transport and Automation Technology	Germany
Schumacher, Jens	University of Applied Sciences Vorarlberg	Austria
Sert, Yasemin	University of South Florida	USA
Seyed-Abbassi, Behrooz	University of North Florida	USA
Shao, Shiyun	National University of Singapore	Singapore
Shaw, Jill	The Open University	UK
Shayeghi, Hossien	Iran University of Science	Iran
Shim, Eung Bo	Korea Electric Power Research Institute	South Korea
Shin, Jungpil	University of Aizu	Japan
Shiraishi, Yoshiaki	Nagoya Institute of Technology	Japan
Shum, Kwok	Stanford University	USA
Sim, Sang Gyoo	Samsung Electronics, R. O. Korea	South Korea
Singh, Madhusudan	Netaji Subhas Institute of Technology	India
Sinkunas, Stasys	Kaunas University of Technology	Lithuania
Smith III, Charles V.	University of South Alabama	USA
Soeiro, Alfredo	University of Porto	Portugal
Song, Hong Jun	University of Sydney	Australia
Srisawangrat, Songsiri	University of Regina	Canada
Srivastava, Sanjeev	Florida State University	USA
Srivastava, Smriti	Netaji Subhas Institute of Technology	India
Stasytyte, Viktorija	Vilnius Gediminas Technical University	Lithuania
Su, J. L.	Quanzhou Normal University	China
Su, Wei	National University of Defense Technology	USA
Su, Wen-Ray	NCDR	Taiwan
Sulc, Bohumil	Czech Technical University in Prague	Czech Republic
Sun, Baolin	Wuhan University	China
Suomi, Reima	Turku School of Economics	Finland
Szygenda, Stephen A.	Southern Methodist University	USA
Takemi, Fujikawa	University of Western Sydney	Australia
Tam, Wing K.	Swinburne University of Technology	Australia
Tarnikova, Yuliya	BCL Technologies Inc.	USA
Tay, Francis E. H.	National University of Singapore	Singapore
Taylor, Stephen	Sussex University	UK
Tchier, Fairouz	King Saud University	USA
Teng, Chia-Chi	Brigham Young University	USA
Teshigawara, Yoshimi	Soka University	Japan
Thede, Anke	University of Karlsruhe	Germany
Theologou, Michael	National Technical University of Athens	Greece
Thornton, Mitchell A.	Southern Methodist University	USA
Till, Robert	John Jay College	USA
Traum, Maria	Johannes Kepler University	Austria
Treven, Sonja	University of Maribor	Slovenia
Tsai, Li-Hung	Tatung University	Taiwan
Tseng, Ko Ying	Tatung University	Taiwan

Turnitsa, Charles D.	Old Dominion University	USA
Ucan, Osman N.	Istanbul University	Turkey
Väänänen, Kalervo	University of Turku	Finland
Vanka, Sita	Kakatiya University	India
Vasinek, Vladimir	Technical University of Ostrava	Czech Republic
Venkataraman, S.	AIAMED	India
Verlinde, Patrick	Royal Military Academy	Belgium
Verma, Pramode	University of Oklahoma	USA
Viloria, Orlando H.	Simon Bolivar University	Venezuela
Visram, Zabin	University of Warwick	UK
Vodovozov, Valery	St. Petersburg State Electrotechnical University	Russian Federation
Volkov-Husovic, T.	University of Belgrade	Yugoslavia
Voss, Andreas	Dortmund University of Technology	Germany
Vrána, Stanislav	Czech Technical University in Prague	Czech Republic
Wang, Chuan-Ju	National Taiwan University	Taiwan
Wang, Lei	University of Houston	USA
Wang, Ling	University of Oklahoma	USA
Wells, Harvey	King's College London	UK
Wilson, Ralph	Florida State University	USA
Woodhead, Steve	University of Greenwich	UK
Woodthorpe, John	The Open University	UK
Wu, Chun Yin	Tatung University	Taiwan
Wu, Shang-Yu	NCDR	Taiwan
Wu, Sheng-You	National Changhua University of Education	Taiwan
Xu, L.	National University of Singapore	Singapore
Yang, Hung Jen	National Kaohsiung Normal University	Taiwan
Yang, Sung Un	Syracuse University	USA
Yao, Wenbing	Brunel University	UK
Yap, K. L.	National University of Singapore	Singapore
Yasser, Muhammad	Chiba University	Japan
Yazawa, Toru	Tokyo Metropolitan University	Japan
Yilmaz, Levent	Auburn University	USA
Yoshida, Eri	Toyohashi University of Technology	Japan
Young, Jieh-Shian	National Changhua University of Education	Taiwan
Yu, Chong Ho	University of California at Berkeley	USA
Yu, Ke	National University of Singapore	Singapore
Yu, Xin	University of Bath	UK
Yuen, Fei Lung	University of Hong Kong	Hong Kong
Zadeh, Jeff	Virginia State University	USA
Zaretsky, Esther	Givat Washington College of Education	Israel
Zelinka, Tomas	Czech Technical University in Prague	Czech Republic
Zeller, Andrew J.	Norisbank GmbH	Germany
Zhang, Jinfang	University of Waterloo	Canada
Zhang, Qingyu	Arkansas State University	USA
Zhang, Xiaozheng Jane	California Polytechnic State University	USA
Zheng, Huiyong	University of Michigan	USA
Zhonghua, Fang	Shanghai Institute of Technical Physics	USA
Zhu, Hui	Soochow University	China
Zobaa, Ahmed	Cairo University	Egypt
Zolotová, Iveta	University of Technology in Kosice	Slovakia

ADDITIONAL REVIEWERS

Abe, Akinori	Communication Science Laboratories	Japan
Abuzir, Yousef	Al-Quds Open University	Palestine
Aguilar, José	University of the Andes	Venezuela
Ajit Kumar, Gopee	University of Technology, Mauritius	Mauritius
Allali, Mohamed	Chapman University	USA
Balas, Valentina	Aurel Vlaicu University of Arad	Romania
Baldazzi, Giuseppe	University of Bologna	Italy
Bartolini, Sandro	University of Siena	Italy
Berke, József	Dennis Gabor Applied University	Hungary
Bevinakoppa, Savitri	Melbourne Institute of Technology	Australia
Bjelkemyr, Marcus	Royal Institute of Technology	Sweden
Bocu, Razvan	University College Cork	Ireland
Borza, Sorin	Lucian Blaga University of Sibiu	Romania
Buglione, Luigi	Engineering IT	Italy
Caldararu, Florin	ECOSEN Ltd.	Romania
Campos, Ana	Institute for the Development of New Technologies	Portugal
Cao, Yiwei	Aachen University	Germany
Castellini, Horacio	National University of General Sarmiento	Argentina
Cázares Rangel, Víctor M.	Autonomous University of Nuevo Leon	Mexico
Ceglowski, Andrzej	Monash University	Australia
Chandra, Vigyan	Eastern Kentucky University	USA
Chen, Chih-Chun	University College London	UK
Chen, Zengqiang	Nankai University	China
Ching, Wai-Ki	University of Hong Kong	Hong Kong
Chiu, Hung-Wen	Taipei Medical University	Taiwan
Chung, Wen-Yaw	Chung Yuan Christian University	Taiwan
Comyn-Wattiau, Isabelle	Essec Business School	France
Corcuera, Pedro	University of Cantabria	Spain
Dajani, Khalil	Albany State University	USA
De la Puente, Fernando	University of Las Palmas de Gran Canaria	Spain
Desai, Sachi	U.S. Army Research Laboratory	USA
Diaconescu, Stefan	Softwin SRL	Romania
Díaz-Otero, Francisco J.	University of Vigo	Spain
Dimililer, Kamil	Near East University	Cyprus
Dorn, Marcio	Federal University of Rio Grande do Sul	Brazil
Dutta, Sourav	IEEE Bombay	India
Dvornik, Josko	University of Split	Croatia
El-Gendy, Hazem	Misr International University	Egypt
Faulín, Javier	Public University of Navarre	Spain
Fikret Ercan, M.	Singapore Polytechnic	Singapore
García Marco, Francisco J.	University of Zaragoza	Spain

Genge, Bela	Petru Maior University	Romania
Gervino, Gianpiero	National Institute of Nuclear Physics	Italy
Giampapa, Joseph	Carnegie Mellon University	USA
Gong, Jianhua	Chinese Academy of Sciences	China
Hanakawa, Noriko	Hannan University	Japan
Hashimoto, Shigehiro	Osaka Institute of Technology	Japan
Hespel, Christiane	National Institute of Applied Sciences of Rennes	France
Hlynka, Myron	University of Windsor	Canada
Ho, Liangwei	Collex Communication Corporation	Taiwan
Hong, Daniel W.	KT Network Technology Lab.	South Korea
Hong, Daniel Won-Kyu	Korea Telecom, Network Technology Laboratory	South Korea
Hou, Chun-Ju	Southern Taiwan University of Technology	Taiwan
Hsu, Chun-Fei	Chung Hua University	Taiwan
Huang, Hanming	Guangxi Normal University	China
Hunek, Wojciech	Opole University of Technology	Poland
Idowu, Peter Adebayo	Obafemi Awolowo University	Nigeria
Ingber, Lester	Lester Ingber Research	USA
Ishikawa, Hiroshi	NUIS	Japan
Ishimura, Naoyuki	Hitotsubashi University	Japan
Janota, Ales	University of Zilina	Slovakia
Johnson, Mark	The Aerospace Corporation	USA
Kacarska, Marija	Methodius University in Skopje	Macedonia
Kaminska, Aniela	Poznan University of Technology	Poland
Karadimitriou, Kosmas	Louisiana State University	USA
Kim, Dohoon	Kyung Hee University	South Korea
Krechel, Dirk	University of Applied Science Wiesbaden	Germany
Kugiumtzis, Dimitris	Aristotle University of Thessaloniki	Greece
Lambiotte, Renaud	Imperial College London	UK
Lee, Chang Won	Hanyang University	South Korea
Lehner, Karlheinz	Ruhr-University of Bochum	Germany
Liang, Chih-Chin	National Central University	Taiwan
Lipikorn, Rajalida	Chulalongkorn University	Thailand
Lipinski, Piotr	Technical University of Lodz	Poland
Liu, Ruey-Tarng	National Chung Hsing University	Taiwan
Lorenz, Pascal	University of Haute Alsace	France
Maluckov, Aleksandra	University of Nis	Yugoslavia
McConnell, George	SELEX ELSAG	UK
Memon, Qurban	United Arab Emirates University	UAE
Mihaita, Niculae	Bucharest University of Economics	Romania
Milani, Mattia	Polytechnic University of Turin	Italy
Moratal, David	Polytechnic University of Valencia	Spain
Moreira Vidal, Raul	University of Porto	Portugal
Mozar, Stefan	Central Queensland University	Australia
Nagar, Atulya	Liverpool Hope University	UK
Nahm, In Hyun	Sun Moon University	South Korea
Narasimhan, Lakshmi	University of Newcastle	Australia
Nordquist, Peter	Southern Oregon University	USA

Olson, Patrick C.	National University	USA
Ophir, Dan	University Center Ariel in Samaria	Israel
Ortiz Sosa, Lourdes Maritza	Andrés Bello Catholic University	Venezuela
Ostrowski, David	Ford Motor Company	USA
Pal, Bijay Baran	University of Kalyani	India
Pavone, Mario	University of Catania	Italy
Poltronieri, Palmiro	National Research Council	Italy
Provaznik, Ivo	Brno University of Technology	Czech Republic
Pruna, Simion	Telemedicine Center	Romania
Qureshi, Waqar Shahid	Air University Islamabad	Pakistan
Rabaea, Adrian	North University of Baia Mare	Romania
Raibulet, Claudia	University of Milano-Bicocca	Italy
Rajan, Amala V. S.	Higher Colleges of Technology	UAE
Ramdass, Kemlall Ramsaroop	University of Johannesburg	South Africa
Resta, Marina	University of Genoa	Italy
Romero, Pedro	Rafael Belloso Chacin University	Venezuela
Rot, Artur	Wroclaw University of Economics	Poland
Ruan, Tongjun	New Mexico Institute of Mining and Technology	USA
S. Medeiros, Joao Paulo	Federal University of Rio Grande do Norte	Brazil
Salous, Sana	Durham University	UK
Satybaldiyeva, Ryskhan	International Information Technologies University	Kazakhstan
Segall, Richard S.	Arkansas State University	USA
Semboloni, Ferdinando	University of Florence	Italy
Shi, Lei	University of Maryland Catonsville	USA
Shing, Chen-Chi	Radford University	USA
Siemieniuch, Carys	Loughborough University	UK
Sizikov, Valery	State University of Informational Technologies	Russian Federation
Small, Michael	Hong Kong Polytechnic University	Hong Kong
Sokolov, Sergey	Keldysh Institute of Applied Mathematics	Russian Federation
Song, Bo	University of Illinois at Chicago	USA
Stancu, Stelian	University of Economics	Romania
Stanko, Zerajic	The University of Nis	Serbia
Tahami, Ehsan	Azad University of Mashhad	Iran
Toorani, Mohsen	University of Bergen	Norway
Turchiano, Biagio	Polytechnic of Bari	Italy
Vasilache, Simona	University of Tsukuba	Japan
Verber, Domen	University of Maribor	Solomon Islands
Watanabe, Eiichi	Fujita Health University	Japan
Wu, Tung-Xiung (Sean)	Shih Hsin University	Taiwan
Xochicale Rojas, Hugo A.	Technology University of Izúcar de Matamoros	Mexico
Yasseri, Taha	Budapest University of Technology and Economics	Hungary
Yin, Jianwen	Dell Computer Corporation	USA
Zemliak, Alexander	Autonomous University of Puebla	Mexico
Zhang, Yongfei	Beihang University	China
Zhu, Quanyan	University of Illinois at Urbana-Champaign	USA
Zyubin, Vladimir	Institute of Automation and Electrometry	Russian Federation

ADDITIONAL REVIEWERS FOR THE NON-BLIND REVIEWING

Arthmar, Rogerio	Federal University of Espirito Santo	Brazil
Astuti, Pudji	Bandung Institute of Technology	Indonesia
Attalla, Mohamed	McMaster University	Canada
Azeem, M. F.	Near Mangalore University	India
Baklanov, Alexander	Danish Meteorological Institute	Denmark
Berber, Aykut	Istanbul University	Turkey
Buttarazzi, Berta	University of Rome Tor Vergata	Italy
Cerveira Fronza, Fernanda	USP	Brazil
Chai, Xinsheng	South China University of Technology	China
Chakhmouradian, Anton	University of Manitoba	Canada
Chebanov, Sergey V.	Baltic State Technical University Voenmeh	Russian Federation
Chen, Chao	US Environmental Protection Agency	USA
Chen, Zengqiang	Nankai University	China
Choo, Raymond	University of South Australia	Australia
Cotsaftis, Michel	Ecole Centrale d`Electronique	France
Dajani, Khalil	Albany State University	USA
Dellenback, Steven	Southwest Research Institute	USA
Delorenzo, Gary	California University of Pennsylvania	USA
Díaz, Manuel	Autonomous University of Aguascalientes	Mexico
Dobrucky, Branislav	University of Zilina	Slovakia
Dreschhoff, Gisela	University of Kansas	USA
El Nasan, Adnan	American University in Dubai	UAE
Ellabidy, Magdy	Wentworth Institute of Technology	USA
Faro, Jose H.	Insper Institute of Education and Research	Brazil
Ferrari, Gianluigi	Università degli Studi di Parma	Italy
Fouda, Yasser	King Faisal University	Saudi Arabia
Giampapa, Joseph	Carnegie Mellon University	USA
González, Roberto	Autonomous University of Aguascalientes	Mexico
Guzmán, Ismenia	Pontificial Catholic University of Valparaíso	Chile
Hao, Yuxin	Beijing Information Science and Technology University	China
Hernández Corporales, Luís	Carlos III University of Madrid	Spain
Hershock, Peter D.	University of Hawai	USA
Hiraishi, Hironori	Akita National College of Technology	Japan
Hoppichler, Josef	Federal Institute for Less-Favoured and Mountainous Areas	Austria
Hosseinzadeh, Mehdi	Science and Research Branch of Azad University	Iran
Jelic-mrcelic, Gorana	Faculty of Maritime Studies	Croatia
Kassar, Abdul Nasser	Lebanese American University	Lebanon
Khodachenko, Maxim	Austrian Academy of Sciences	Austria
Knipp, Tammy	Florida Atlantic University	USA
Kresse, Adelheid	Medical University Graz	Austria
Lai, Chih-Chin	National University of Kaohsiung	Taiwan
Lambiotte, Renaud	Imperial College London	UK
Lamm, Ryan	Southwest Research Institute	USA

Longstaff, Patricia	Syracuse University	USA
Love, David W.	New Mexico Institute of Mining and Technology	USA
Majzoub, Sohaib	American University in Dubai	UAE
Melde, Gretchyn	The MITRE Corporation	USA
Miller, Bob	The MITRE Corporation	USA
Mizoguchi, Fumio	Tokyo University of Science	Japan
Morais, Carlos	University of Sao Paulo	Brazil
Moreno, Wilfrido	University of South Florida	USA
Nishiyama, Hiroyuki	Tokyo University of Science	Japan
Nuseibeh, Hasan	University of South Florida	USA
Oborkhale, Lawrence	American University of Nigeria	Nigeria
Olivos, Elena	University of the Frontier	Chile
Ophir, Dan	University Center in Ariel	Israel
Pandey, Abhishek	Indian Institute of Information Technology Allahabad	India
Patel, Dhanesh	The Maharaja Sayaji Rao University	India
Pekdemir, Isil	Istanbul University	Turkey
Pencek, Tom	University of South Florida	USA
Prawesh, Shankar	University of South Florida	USA
Prinetto, Paolo	Politecnico di Torino	Italy
Puthooran, Emjee	Indian Institute of Technology Roorkee	India
Rafiq, Qasim	Aligarh Muslim University	India
Rahardjo, Budi	Bandung Institute of Technology	Indonesia
Rahmani, Amir Masoud	Science and Research Branch of Azad University	Iran
Ramdass, Kemlall Ramsaroop	University of Johannesburg	South Africa
Reck, Ronald	RRecktek LLC	USA
Reel, Parminder Singh	Thapar University, India	India
Ristov, Panco	Faculty of Maritime Studies	Croatia
Sayed Ahmed, Khaled	King Faisal University	Saudi Arabia
Scott, Imhoff	Raytheon Intelligence and Information Systems	USA
Shaer, Bassam	University of West Florida	USA
Shishlenin, Maxim	Sobolev Institute of Mathematics SB RAS	Russian Federation
Shorgin, Sergey	Russian Academy of Sciences	Russian Federation
Smolin, Denis	American University of Nigeria	Nigeria
Soshnikov, Dmitry	Microsoft Russia	Russian Federation
Tatashev, Alexander	MTUCI	Russian Federation
Thiran, Jean-phillippe	EPFL	Switzerland
Tsay, Li-shiang	North Carolina A and T State University	USA
Tyrsin, Alexander	Yerevan State University	Russian Federation
Vargas, Javier	Instituto Tecnológico de Costa Rica	Costa Rica
Vilchez, Enrique	Universidad Nacional	Costa Rica
Vílchez, Enrique	National University of Costa Rica	Costa Rica
Wolfengagen, Vyacheslav	Institute for Contemporary Education JurInfoR-MSU	Russian Federation
Wu, Chih-Hung	National University of Kaohsiung	Taiwan
Yang, Xiaodong	Shenyang Institute of Aeronautical Engineering	China
Yarmohamadian, M. H.	Isfahan University of Medical Sciences	Iran
Yoshida, Nelson	University of Sao Paulo	Brazil
Zayed, Tarek	Concordia University	Canada
Zhang, Cynthia	University of North Carolina at Pembroke	USA
Zhao, Hong	Xiamen University	China
Zissis, Dimitrios	University of the Aegean	Greece

IMCIC'12

HONORARY CHAIR
William Lesso

PROGRAM COMMITTEE CHAIRS
C. Dale Zinn
José Ferrer

GENERAL CHAIRS
Nagib Callaos
Jorge Baralt
Hsing-Wei Chu
Michael J. Savoie

ORGANIZING COMMITTEE CHAIR
Belkis Sánchez

CONFERENCE PROGRAM MANAGER
Leonisol Callaos

HARDCOPY PROCEEDINGS PRODUCTION CHAIR
María Sánchez

TECHNICAL CONSULTANT ON COMPUTING SYSTEM / CD PROCEEDINGS PRODUCTION CHAIR
Juan Manuel Pineda

SUBMISSION QUALITY CONTROL SUPPORT
Leonardo Contreras

META REVIEWERS SUPPORT
Maria Sánchez
Dalia Sánchez

SYSTEM DEVELOPMENT, MAINTENANCE AND DEPLOYMENT
Dalia Sánchez
Keyla Guédez
Nidimar Diaz
Bebzabeth García

OPERATIONAL ASSISTANTS
Marcela Briceño
Cindi Padilla

HELP DESK
Louis Barnes
Sean Barnes
José Aponte

The 2nd International Conference on Design and Modeling in Science, Education, and Technology: DeMset 2012
in the context of
The 3rd International Multi-Conference on Complexity, Informatics and Cybernetics: IMCIC 2012

GENERAL CHAIR
Nagib Callaos
Hsing-Wei Chu

ORGANIZING COMMITTEE CHAIR
Andrés Tremante
Belkis Sánchez

PROGRAM COMMITTEE

Chairs:
Mohammad Siddique (USA)
C. Dale Zinn (USA)

Abdulwahed, Mahmoud	Loughborough University	UK
Acharya, Sushil	Robert Morris University	USA
Ahn, Kwang Il	Korea Atomic Energy Research Institute	South Korea
Alameh, Kamal	Edith Cowan University	Australia
Alhayyan, Khalid N.	University of South Florida	USA
Ali, Shahid	National University of Science and Technology	Pakistan
Al-Omari, Hussein	Applied Science University	Jordan
Alsayegh, Osamah	Kuwait Institute for Scientific Research	Kuwait
Ambrose, Jude Angelo	Northumbria University	UK
Annakkage, Udaya D.	University of Manitoba	Canada
Arimitsu, Yoshihiro	Meisei University	Japan
Aristarkhov, Vasily	Nizhniy Novgorod State Technical University	Russian Federation
Arndt, Angela E.	University of Cincinnati	USA
Aslam, Dean M.	Michigan State University	USA
Bang, Duck Je	Chungbuk National University	South Korea
Bell, Colin A.	Brunel University	UK
Bertel, Lykke	Aalborg University	Denmark
Blanchard, Richard	Loughborough University	UK
Bo, Hao	Shenyang Ligong University	China
Butler, Cary	US Army Corps of Engineers	USA
Call, Anson B.	Iowa State University	USA
Campbell, David J.	Siemens Medical Systems	USA
Carnahan, Heather	University of Toronto	Canada
Castillo A., Alejandro	CINVESTAV	Mexico
Chalupa, Milan	University of Defence in Brno	Czech Republic
Chang, Julian	Ching Yun University	Taiwan
Chang, Jyun-Wei	National Chiao Tung University	Taiwan
Chang, Kuo-Hwa	Chung Yuan Christian University	Taiwan
Chen, Chin-Ti	Institute of Chemistry Academia Sinica	Taiwan
Chen, Huifen	Chung Yuan Christian University	Taiwan
Chen, Mei-Yung	National Taiwan Normal University	Taiwan
Chen, Zong	Fairleigh Dickinson University	USA
Cheng, Yi-Chang	National Chiao Tung University	Taiwan

Chiu, Ting-Lan	University of Minnesota	USA
Cho, Kyoung-Rok	Chungbuk National University	South Korea
Cho, Young Duk	Blink Studios	South Korea
Choi, Yu-Lee	Chungbuk National University	South Korea
Chou, Yu-Hur	Tung Nan University	Taiwan
Chowdhury, Masud H.	University of Illinois at Chicago	USA
Chudý, Peter	Brno University of Technology	Czech Republic
Chung, Won Jee	Changwon National University	South Korea
Ciastellardi, Matteo	Polytechnic of Milan	Italy
Costa, F. F.	University of Bahia	Brazil
Cowan, Mark	US Army Corps of Engineers	USA
Cruciani, Andrea	University of Toronto	Canada
Curós-Vilà, Pilar	University of Girona	Spain
Das, Asesh	Pennsylvania College of Technology	USA
Davis, Karen C.	University of Cincinnati	USA
Dawoud, D. S.	University of KwaZulu Natal	South Africa
Dawoud, Peter D.	University of KwaZulu Nata	South Africa
D'Cruz, Carmo	Florida Institute of Technology	USA
De Castro L., Antonio C.	Federal University of Bahia	Brazil
De Kerckhove, Derrick	University of Toronto	Canada
Dierneder, Stefan	Linz Center of Mechatronics	Austria
Domínguez, Miguel A.	University of Vigo	Spain
Dubrowski, Adam	University of Toronto	Canada
Eilouti, Buthayna H.	Jordan University of Science and Technology	Jordan
Flores, Juan	University of Michoacan	Mexico
Fontana, M.	Federal University of Bahia	Brazil
Fujita, Naoyuki	Japan Aerospace Exploration Agency	Japan
Glovnea, Romeo P.	Brunel University	UK
Greca, Ardian	Georgia Southern University	USA
Guimarães, Diego	Federal University of Bahia	Brazil
Gunn, Rod	University of Glamorgan	UK
Guo, Weidong	Purdue University	China
Harvey, Adrian	University of Toronto	Canada
Hashimoto, Shigehiro	Kogakuin University	Japan
Hasnaoui, Salem	National Engineering School of Tunis	Tunisia
Hendel, Russell Jay	Towson University	USA
Herrnstadt, Steven	Iowa State University	USA
Hirvonen, Juhani	Technical Research Centre of Finland	Finland
Hirz, Mario	Graz University of Technology	Austria
Honzik, Petr	Brno University of Technology	Czech Republic
Hook, Derek J.	University of Minnesota	USA
Horvath, Gabor	University of Glamorgan	UK
Howell, P. R.	Pennsylvania State University	USA
Hsu, Yung-Chi	National Chiao Tung University	Taiwan
Huang, C. J.	Ching Yun University	Taiwan
Huang, Hsuan-Kuan	Industrial Technology Research Institute	Taiwan
Ishikawa, Koichiro	Japan Advanced Institute of Science and Technology	Japan
Ismail, Napsiah	Putra Malaysia University	Malaysia
Iyyunni, Chakradhar	Institute of Project Management, Larsen & Toubro	USA
Jackson, Gregory A.	University of Manitoba	Canada
Jamuar, Sudhanshu S.	Putra Malaysia University	Malaysia

Jun, Moon-Seog	Soongsil University	South Korea
Jung, Dong Won	Lee International	South Korea
Kao, Diana	University of Windsor	Canada
Katila, Sanda	Kent State University	USA
Kermanshahi, Shahab	Sharif University of Technology	Iran
Khaled, Pervez	University of Illinois at Chicago	USA
Kim, Chul	University of New South Wales	Australia
Kim, Dae-Jung	Soongsil University	South Korea
Kim, Ho Chong	Shinsung Solar Energy	South Korea
Kim, Hyung Nam	Virginia Tech	USA
Kim, Jae Min	Changwon National Univesity	South Korea
Kim, Jeong-Jai	Soongsil University	South Korea
Kim, Mijin	Duksung Womens University	South Korea
Kim, Seok-Man	Chungbuk National University	South Korea
Koo, Kil Mo	Korea Atomic Energy Research Institute	South Korea
Krylov, Vladimir	Intel Corporation	Russian Federation
Kucera, Pavel	Brno University of Technology	Czech Republic
Kulkarni, Arun	University of Texas at Tyler	USA
Kuragano, Tetsuzo	Meisei University	Japan
Kurita, Takio	AIST	Japan
Lau, Newman	Hong Kong Polytechnic University	Hong Kong
Lee, Choon Man	Changwon National University	South Korea
Lee, Eun-Hee	Chungbuk National University	South Korea
Lee, Eunoak	Duksung Women's University	South Korea
Lee, Je-Hoon	Chungbuk National University	South Korea
Lee, Jun-Ho	Samsung	South Korea
Lee, Seung-Min	Soongsil University	South Korea
Lee, Won Chang	Changwon National University	South Korea
Lee, Young-Sang	Samsung	South Korea
Lesaja, Goran	Georgia Southern University	USA
Leung, Lin	City University of New York	USA
Li, Xiaoping	National University of Singapore	Singapore
Lin, Hong	University of Houston-Downtown	USA
Lin, Huan Yu	National Chiao Tung University	Taiwan
Lin, J.	Ching Yun University	Taiwan
Lin, Sheng-Fuu	Nation Chiao Tung University	Taiwan
Liu, Yuyu	University Tsinghua	China
Locklin, R. H.	Penn State University	USA
López-M., Cuauhtemoc	University of Guadalajara	Mexico
Machado, I. D. R.	Federal University of Bahia	Brazil
Maheshwari, Bharat	University of Windsor	Canada
Mares, Cristinel	Brunel University	UK
Mariño, Perfecto	University of Vigo	Spain
Masunov, Artëm E.	University of Central Florida	USA
Mccann, Roy A.	University of Arkansas	USA
Merino, Miguel	Community Health Centre at Merino	Spain
Mikhaylov, Ivan A.	University of Central Florida	USA
Miranda de A., Cristina	University of the Basque Country	Spain
Mochizuki, Shuichi	Osaka Institute of Technology	Japan
Mohan, K. Krishna	Indian Institute of Technology Bombay	India
Moin, Lubna	Pakistan Navy Engineering College	Pakistan

Mokhov, Serguei A.	Concordia University	Canada
Morita, Yusuke	Osaka Institute of Technology	Japan
Moss, Gloria	University of Glamorgan	UK
Muñoz, Humberto	Southern University and A&M College	USA
Nagar, Atulya K.	Liverpool Hope University	UK
Nagy, Zoltan K.	Loughborough University	UK
Nam, Chang Soo	University of Arkansas	USA
Ndagije, Charles	National University of Rwanda	Rwanda
Ng, M. C. F.	Hong Kong Polytechnic University	Hong Kong
Ng, Min Shen	Putra Malaysia University	Malaysia
Nooshabadi, Saeid	Gwangju Institute of Science and Technology	South Korea
Oanta, Emil	Constanta Maritime University	Romania
Obrebski, Jan B.	Warsaw University of Technology	Poland
Oh, Kil Nam	Gwangju University	South Korea
Oh, Paul	Drexel University	USA
Okada, Masahide	Osaka Institute of Technology	Japan
Ono, Kohei	Osaka Institute of Technology	Japan
Ostergaard, Soren Duus	IBM Software Group	Denmark
Otero, Santiago	University of Vigo	Spain
Paiano, Roberto	University of Salento	Italy
Park, Uchang	Duksung Womens University	South Korea
Passos, Nelson	Midwestern State University	USA
Petkov, Emil	Northumbria University	UK
Pisupati, Sarma V.	Penn State University	USA
Platanitis, George	University of Ontario Institute of Technology	Canada
Podaru, Vasile	Military Technical Academy	Romania
Pop-Iliev, Remon	University of Ontario Institute of Technology	Canada
Pulimeno, Enrico	University of Salento	Italy
Ramírez C., Guillermo H.	Soka University	Japan
Raud, Zoja	Tallinn University of Technology	Estonia
Rebielak, Janusz	Wroclaw University of Technology	Poland
Riedling, Eveline	Technical University of Vienna	Austria
Robins, David	Kent State University	USA
Rojas-Moreno, Arturo	National University of Engineering	Peru
Romero, Pedro	Rafael Belloso Chacin University	Venezuela
Rossbacher, Patrick	Graz University of Technology	Austria
Rutherfoord, James K.	Chattahoochee Technical College	USA
Rutherfoord, Rebecca H.	Southern Polytechnic State University	USA
Rzucidlo, Pawel	Rzeszow University of Technology	Poland
Saad, Inès	University of Picardie Jules Verne	France
Sangiorgi, U. B.	Federal University of Bahia	Brazil
Sariyildiz, Sevil	Delft University of Technology	Netherlands
Sateesh Reddy, J.	Birla Institute of Technology and Science	India
Sato, Yoshishige	Tsuruoka National College of Technology	Japan
Savva, Andreas	University of Nicosia	Cyprus
Scharfe, Henrik	Aalborg University	Denmark
Seetharaman, R.	National Academy of Excellence	India
Shafahi, Yousef	Sharif University of Technology	Iran
Shaikh, Muzaffar	Florida Institute of Technology	USA
Sharieh, Ahmad	University of Jordan	Jordan
Shyu, Hsin-Yih Cindy	Tamkang University	Taiwan

Sidek, Roslina	University Putra Malaysia	Malaysia
Singh, Harwinder	Guru Nanak Dev Engineering College	India
Sissom, James	Southern Illinois University	USA
Smith-Jackson, Tonya	Virginia Polytechnic Institute and State University	USA
Snidvongs, Suravut	Thai Natural Products	Thailand
Snow, Richard K.	Embry-Riddle Aeronautical University	USA
Song, Yong Mann	Korea Atomic Energy Research Institute	South Korea
Srinivasan, S.	Indian Institute of Technology Madras	India
Srivastava, Smriti	Netaji Subhas Institute of Technology	India
Steiner, Bernhard	Johannes Kepler University	Austria
Su, J. L.	Quanzhou Normal University	China
Sulaiman, S.	Putra Malaysia University	Malaysia
Tang, S. H.	Putra Malaysia University	Malaysia
Teshigawara, Yoshimi	Soka University	Japan
Tomovic, Mileta M.	Purdue University	USA
Torres Román, D.	CINVESTAV	Mexico
Tseng, Shian-Shyong	National Chiao-Tung University	China
Uddin, Vali	National University of Science and Technology	Pakistan
Vander Biest, Alexis	Free University of Brussels	Belgium
Vasili, M. R.	Putra Malaysia University	Malaysia
Ventä, Olli	Technical Research Centre of Finland	Finland
Verma, A. K.	Indian Institute of Technology Bombay	India
Vodovozov, Valery	St. Petersburg State Electrotechnical University	Russian Federation
Vrána, Stanislav	Czech Technical University in Prague	Czech Republic
Wagiran, Rahman	Putra Malaysia University	Malaysia
Wang, Chuan-Ju	National Taiwan University	Taiwan
Wang, S. W.	Ching Yun University	Taiwan
Warwick, Jon	London South Bank University	UK
Whiteley, Rick	Calabash Educational Software	Canada
Wile, Gregory	Case Western Reserve University	USA
Wilson, Ralph	Florida State University	USA
Wong, Ben	Hong Kong Polytechnic University	Hong Kong
Xia, W. J.	Hong Kong Polytechnic University	Hong Kong
Yamaguchi, Akira	Meisei University	Japan
Yang, Joon Eon	Korea Atomic Energy Research Institute	South Korea
Yeh, Syh-Shiuh	National Taipei University of Technology	Taiwan
Yoon, Sang Hwan	Changwon University	South Korea
Yoon, Young Min	Shinsung Holdings	South Korea
Zaretsky, Esther	Givat Washington Academic College of Education	Israel
Zhang, Yuru	Beijing University of Aeronautics and Astronautics	China
Zheng, Huiyong	University of Michigan	USA

ADDITIONAL REVIEWERS

Alberto, Catalina	Universidad Nacional de Córdoba	Argentina
Allahverdi, Ali	Kuwait University	Kuwait
Arcos Martinez, Fabio O.	Corporacion Escuela Pedagógiva Experimental	Colombia
Bersuker, Isaak	University of Texas	USA

Bruns, F. Willhelm	University of Bremen	Germany
Castillo A., Alejandro	CINVESTAV	Mexico
Dobrucky, Branislav	University of Zilina	Slovakia
Eberhardt, John S.	DecisionQ Corporation	USA
Elsayed, Mostafa	McMaster University	Canada
Emmanoel, Carlo	Universidade Federal do Rio de Janeiro	Brazil
Ferracioli, Laercio	Universidade Federal do Espirito Santo	Brazil
Fombona, Javier	Universidad de Oviedo	Spain
Fraser, Benjamin	The College of Charleston	USA
Gao, Wei	Central University of Finance and Economics	China
Goyal, Megh R.	University of Puerto	Puerto Rico
Guo, Dong Mei	Central University of Finance and Economics	China
Iyyunni, Chakradhar	Institute of Project Management, Larsen & Toubro	USA
J Kailath, Binsu	IIITDandM Kancheepuram	India
Kawai, Kazuhisa	Toyohashi University of Technology	Japan
Kuang, Yu	Stanford University School of Medicine	USA
Martinez R., Carmen A.	Universidad Francisco Jose de Caldas	Colombia
Matsutomi, Tatsuo	Kinki University	Japan
Molina Andrade, Adela	Universidad Distrital	Colombia
Moon, Dosoo	University of Illinois at Urbana Champaign	USA
Ohtani, Takashi	Kinki University	Japan
Pa, Janakiraman	IIT Madras	India
Piaia, Miquela	Instituto Federal Farroupilha	Brazil
Rajappa, Gnanamoorthy	IIITDandM Kancheepuram	India
Reyes, Jaime	Universidad Distrital	Colombia
Rubin, Gustavo	Benemerita Universidad Autonoma de Puebla	Mexico
Russo, Rhett	New Jersey Institute of Technology	USA
Sado, Kazuhiro	Gunma University	Japan
Sandu, Ion	Gheorghe Asachi Technical University of Iasi	Romania
Sarper, Huseyin	Colorado State University-Pueblo	USA
Smith, Marilyn	Georgia Institute of Technology	USA
Srivastava, Akshay	Georgia Institute of Technology	USA

ADDITIONAL REVIEWERS FOR THE NON-BLIND REVIEWING

Aboul-Ela, Magdy	French University in Egypt	Egypt
Allia, Khedidja	USTHB	Algeria
Anish, Anthony	University of Alabama at Birmingham	USA
Bangert, Patrick	Algorithmica Technologies Corporation	USA
Dimililer, Kamil	Institute of Electrical and Electronics Engineers	Cyprus
Erkollar, Alptekin	University of Applied Sciences ETCOP	Austria
Finch, Aikyna	Stayer University	USA
Georgieva-T., Tsvetanka	University of Veliko Turnovo St Cyril and St. Methodius	Bulgaria
Giampapa, Joseph	Carnegie Mellon University	USA
Huang, En-Hsin	Alcatel-Lucent Technologies	USA
Hussein Ali A., Samaher	Babylon University	Iraq
Iribarne, Luís	University of Almeria	Spain

Lin, Chieh-Yu	Chang Jung Christian University	Taiwan
Minea, Alina Adriana	Gheorghe Asachi Technical University	Romania
Mullins, Michael	Aalborg University	Denmark
Ramos, Ana Luísa	University of Aveiro	Portugal
Schuermann, Volker	Bochum University of Applied Sciences	Germany
Sembera, Jan	Technical University of Liberec	Czech Republic
Sheng, Ye	University of California Berkeley	USA
Stanchev, Peter	Kettering University	USA
Strefezza, Miguel	Simon Bolivar University	Venezuela
Sulema, Yevgeniya	NTUU KPI	Ukraine
Vijayalakshmi, S.	Vellore Institute of Technology University	India
Vizureanu, Petrica	Gheorghe Asachi Technical University	Romania

International Symposium on Interdisciplinary Research, Education, and Communication: IDREC 2012
in the context of
The 3rd International Multi-Conference on Complexity, Informatics and Cybernetics: IMCIC 2012

GENERAL CHAIR
Nagib Callaos
Hsing-Wei Chu

ORGANIZING COMMITTEE CHAIR
Belkis Sánchez

PROGRAM COMMITTEE **Chairs:** C. Dale Zinn (USA)
José Ferrer (Venezuela)

Abd El-Aziem, Mostafa	Arab Academy for Science and Technology	Egypt
Abdel-Qader, Ikhlas	Western Michigan University	USA
Abe, Jair Minoro	Paulista University	Brazil
Acedo de B., María de L.	Simon Bolivar University	Venezuela
Acharya, Sushil	Robert Morris University	USA
Adamopoulou, Evgenia	National Technical University of Athens	Greece
Adascalitei, Adrian	Technical University	Romania
Afandizadeh Z., Shahriar	Iran University of Science and Technology	Iran
Aguilar C., José	University of the Andes	Venezuela
Aguirre-Muñoz, Zenaida	Texas Tech University	USA
Akashi, Takuya	Yamaguchi University	Japan
Aksoy, M. S.	King Saud University	Saudi Arabia
Al Hamouz, Sadeq	Amman Arab University	Jordan
Al Obaidy, Mohaned	Gulf College	Oman
Alam, Hassan	BCL Technologies Inc.	USA
Alameh, Kamal	Edith Cowan University	Australia
Alhayyan, Khalid N.	University of South Florida	USA
Alshara, Osama	Higher Colleges of Technology	UAE
Alvarado Moore, Karla	University of Central Florida	USA
Amaral, Luis	University of Minho	Portugal
Ambikairajah, E.	The University of New South Wales	Australia
Ammann, Walter J.	SLF	Switzerland
Anantharaj, Valentine	Mississippi State University	USA
Andina, Diego	Polytechnic University of Madrid	Spain
Anunciacão, Pedro	Polytechnic Institute of Setubal	Portugal
Anwar, Faisal	Columbia University	USA
Aoki, Toru	Shizuoka University	Japan
Arikan, Emin	Montclair State University	USA
Armendáriz, J. E.	Public University of Navarra	Spain
Arney, Chris	United States Military Academy	USA
Aruga, Masahiro	Tokai University	Japan
Ascoli, Aurelio	University of Milan	Italy
Assaf, Mansour	University of Trinidad and Tobago	Trinidad and Tobago

Astakhov, Vadim	University of California San Diego	USA
Aukstakalnis, Nerijus	Kaunas University of Technology	Lithuania
Auvinen, Anssi	University of Tampere	Finland
Babu, B. V.	Birla Institute of Technology and Science	India
Balasubramanian, K.	European University of Lefke	Turkey
Bangert, Patrick	Algorithmica Technologies	USA
Baruah, Debendra Chandra	Tezpur University	India
Belcher, E. Christina	Trinity Western University	Canada
Benbouziane, Mohamed	University of Tlemcen	Algeria
Benedicenti, Luigi	University of Regina	Canada
Bennett, Leslie	University of Louisville	USA
Bermeo, José	University of the Andes	Colombia
Bernardino, Jorge	Coimbra Institute of Engineering	Portugal
Berndt, Donald J.	University of South Florida	USA
Bezuglov, Anton	University of South Carolina	USA
Bhat, Talapady N.	National Institute of Standards and Technology	USA
Bidarra, José	University of Aberta	Portugal
Blair, Madelyn	Pelerei, Inc.	USA
Blanchard, Richard	Loughborough University	UK
Bogliolo, Alessandro	University of Urbino	Italy
Bohl, Oliver	University of Kassel	Germany
Bolboaca, Sorana Daniela	Iuliu Hațieganu University of Medicine and Pharmacy	Romania
Bönke, Dietmar	Reutlingen University	Germany
Borchers, Carsten H. J.	University of Hanover	Germany
Botto, Todd	Quinnipiac University	USA
Bradbury, Kyle J.	Tufts University	USA
Bradl, Peter	University of Würzburg	Germany
Branski, L. K.	University of Texas Medical Branch	USA
Bubnov, Alexej	Academy of Sciences and the Czech Republic	Czech Republic
Bueno, Newton Paulo	Federal University of Viçosa	Brazil
Burke, David	Robert Morris University	USA
Burnett, Andrea	University of the West Indies	Barbados
Burton, Gideon O.	Brigham Young University	USA
Butrous, Nasir	Australian Catholic University	Australia
Calenzo, Patrick	Im2np	France
Camara, Boubacar	UNESCO	Nigeria
Cao, Hong	Southern Medical University	China
Cárdenas, Henry E.	Louisiana Tech University	USA
Cardoso, Eduardo	ITESM	Mexico
Carrasquero, José Vicente	Simon Bolivar University	Venezuela
Castaño, Bonifacio	University of Alcala	Spain
Castaño-Moraga, Carlos	ULPGC	Spain
Cázares-Rangel, Víctor M.	ITESM	Mexico
Cedano, Karla	National Autonomous University of Mexico	Mexico
Cerny, Václav	University of West Bohemia	Czech Republic
Chae, Hui Soo	Columbia University	USA
Chang, Fengming M.	Asia University	Taiwan
Chang, Maiga	Chung Yuan Christian University	Taiwan
Chang, Wen-Kui	Tunghai University	Taiwan
Chao, Hui-Mei	Kao Yuan University	Taiwan
Chen, Huei-Huang	Tatung University	Taiwan

Chen, Lisa Y.	I-Shou University	Taiwan
Chen, Shih-Chih	National Taiwan University	Taiwan
Chen, Yuhua	University of Houston	USA
Cheng, Kuo-Sheng	National Cheng-Kung University	Taiwan
Cheng, Xiaochun	University of Reading	UK
Cherinka, R.	The MITRE Corporation	USA
Cherry, Barbara	Indiana University	USA
Chiou, Yin-Wah	Nanhua University	Taiwan
Chiu, Shao-I	Taipei College of Maritime Technology	Taiwan
Cho, Vincent	Hong Kong Polytechnic University	Hong Kong
Choi, Yu-Lee	Chungbuk National University	South Korea
Choo, Jinboo	Korea Electric Power Research Institute	South Korea
Chopra, Deepak	Netaji Subhas Institute of Technology	India
Chou, Andrew	Kainan University	Taiwan
Chou, Hsueh-Cheng	National Taiwan Normal University	Taiwan
Chung, Jen-Yao	IBM Research	USA
Clarke, Tim	University of Wales Institute Cardiff	UK
Coffman, Michael G.	Souther Illinois University Carbondale	USA
Contreras, Sebastián	University of the Andes	Colombia
Cripe, Billy	Oracle Corporation	USA
Curran, Kevin	University of Ulster	UK
Das, Asesh	Pennsylvania College of Technology	USA
David, Amos	Université Nancy II	France
De Volder, Dennis	Western Illinois University	USA
Debnath, Rameswar	Neuroscience Research Institute	Japan
Del Río, Antonio	National Autonomous University of Mexico	Mexico
Demestichas, K.	National Technical University of Athens	Greece
Desa, Shakinaz	Sultan Idris University of Education	Malaysia
Dhall, Sudarshan	University of Oklahoma	USA
Diallo, Saikou Y.	Old Dominion University	USA
Dierneder, Stefan	Linz Center of Mechatronics GmbH	Austria
Djukom, C. D.	University of Texas Medical Branch at Galveston	USA
Dolan, Dan	South Dakota School of Mines & Technology	USA
Doran, Michael V.	University of South Alabama	USA
Dosi, Vasiliki	University of Ioannina	Greece
Duman, Hakan	University of Essex	UK
Dunning, Jeremy	Indiana University	USA
Dziong, Zbigniew	University of Québec	Canada
Edwards, Stephen H.	Virginia Tech	USA
El-Halafawy, Farag Z.	Menoufiya University	Egypt
Elias, Rimon	German University in Cairo	Egypt
Elmahboub, W. M.	Hampton University	USA
El-Mezyani, Touria	Florida State University	USA
Emdad, F.	University of Texas Medical Branch	USA
Epps, J.	The University of New South Wales	Australia
Erkollar, Alptekin	University of Applied Sciences	Austria
Eshragh, Sepideh	University of Delaware	USA
Eunni, Madhuri	University of Kansas	USA
Fabbri, G.	Sapienza University of Rome	Portugal
Faghri, Ardeshir	University of Delaware	USA
Fang, Chun-Hsiung	National Kaohsiung University of Applied Sciences	Taiwan

Fang, Zhonghua	Shanghai Institute of Technical Physics	USA
Fisher, Wendy	The Open University	UK
Flores, Juan	University of Michoacan	Mexico
Florescu, Gabriela C.	National Institute for R&D in Informatics	Romania
Flynn, Don	Pacific Northwest National Laboratory	USA
Fogelholm, Mikael	Academy of Finland Communications	Finland
Fonseca Casas, Pau	Technical University of Catalonia	Spain
Fougeres, Alain-Jerome	University of Technology of Belfort-Montbéliard	France
Fox, Kelly	Texas Tech University	USA
Freney, S. J.	The University of New South Wales	Australia
Fu, Shih-Lung	National Taiwan University	Taiwan
Fu, Xiuju	Institute of High Performance Computing	Singapore
Fu, Yonggang	Shanghai Jiao Tong University	China
Fuhrer, Patrik	University of Fribourg	Switzerland
Fujikawa, Takemi	University of Western Sydney	Australia
Fukumi, Minoru	University of Tokushima	Japan
Fumizawa, Motoo	Shonan Institute of Technology	Japan
Fúster-Sabater, Amparo	Spanish Council for Scientific Research	Spain
Gagnon, Francois	School of High Technology	Canada
Ganapathi, Nanthini	Research Triangle Institute	USA
Ganchev, Ivan	University of Limerick	Ireland
García-Atanacio, César	National Nuclear Research Institute	Mexico
Gebre, Biruk A.	Stevens Institute of Technology	USA
Georgescu, Vasile	University of Craiova	Romania
Gesekus, Tim	Deutsche Flugsicherung	Germany
Geyer-Schulz, Andreas	University Karlsruhe	Germany
Goel, Gaurav	Netaji Subhas Institute of Technology	India
Gordon, Richard	University of KwaZulu Natal	South Africa
Goria, Stéphane	Université Nancy 2	France
Goriachkin, Oleg	Volga State Academy of Telecommunication and Informatics	Russian Federation
Goulding, Tom	Daniel Webster College	USA
Gregory, Mark A.	RMIT University	Australia
Guevara L., Miguel A.	University of Porto	Portugal
Guiasu, Radu Cornel	York University	Canada
Guiasu, Silviu	York University	Canada
Gulez, Kayhan	Yildiz Technical University	Turkey
Guo, Dagang	National University of Singapore	Singapore
Guo, Liwen	Stevens Institute of Technology	USA
Gupta, Neha	Maharaja Agrasen Institute of Technology	India
Gustavsson, Rune	Blekinge Institute of Technology	Sweden
Haba, C. G.	Technical University of Iasi	Romania
Haba, Cristian-Győző	Gheorghe Asachi Technical University of Iaşi	Romania
Hammond, Bruce R.	Saint Leo University	USA
Hamzah, Mohd Isa	National University of Malaysia	Malaysia
Hangai, Seiichiro	Tokyo University of Science	Japan
Hartono, Rachmat	BCL Technologies Inc.	USA
Hashimoto, Shigehiro	Osaka Institute of Technology	Japan
Hasnaoui, Salem	National Engineering School of Tunis	Tunisia
Heiserich, Gerd	University of Hannover	Germany
Hemmelman, Brian	South Dakota School of Mines & Technology	USA

Hendel, Russell Jay	Towson University	USA
Henninger, Michael	PH Weingarten	Germany
Herget, Josef	UAS	Switzerland
Herr, Stephan	Deutsche Flugsicherung	Germany
Hetzer, Dirk	T-Systems Nova Deutsche Telekom	Germany
Hlasny, Vladimir	Ewha Womans University	South Korea
Hochin, Teruhisa	Osaka Prefecture University	Japan
Hofmeister, Helge	BASF it Services	Belgium
Hoh, Carl	University California San Diego	USA
Holifield, David	University of Wales Institute Cardiff	UK
Hong, Wen-Hwa	Tatung University	Taiwan
Horne, Jeremy	Maricopa Community College	USA
Hovakimyan, Anna	Yerevan State University	Armenia
Hrušák, Josef	University of West Bohemia	Czech Republic
Hsu, Pai-Hui	National Taiwan University	Taiwan
Hsu, Shu-Mei	Tatung University	Taiwan
Hu, Kai-Wei	Tunghai University	Taiwan
Huang, Pin Chia	I-Shou University Kaohsiung	Taiwan
Huang, Ruhua	Wuhan University	China
Huang, Sheng He	University of Southern California	USA
Ibrahim, Noureddin	Minia University	Saudi Arabia
Iembo, Rosanna	University of Calabria	Italy
Ishikawa, Hiroshi	NUIS	Japan
Ismail, Amirah	National University of Malaysia	Malaysia
Ito, Akinori	Tohoku University	Japan
Ito, Kazumasa	Research Centre for Sustainable Communities	Japan
Iyyunni, Chakradhar	Larsen & Toubro Institute of Project Management	USA
Jablonski, Stefan	Association for Computing Machinery	USA
Jamsai Whyte, Suthida	Ubon Ratchathani University	Thailand
Jäntschi, Lorentz	Academic Direct Organization	Romania
Jenq, John	Monclair State University	USA
Jiménez R., Lourdes	University of Alcala	Spain
Jingchao, Chen	Donghua University	China
Johnson, Mark	Benet Laboratories	USA
Jones, Paul	University of Cincinnati	USA
Jonson, Mark	University of New Mexico	USA
Joy, Mike S.	University of Warwick	UK
Jung, Kyung Im	Samsung Electronics, Co. Ltd.	South Korea
Kadoch, Michel	School of High Technology	Canada
Kao, Chih-Yang	Yuan Ze University	Taiwan
Karamat, Parwaiz	The Open Polytechnic of New Zealand	New Zealand
Kasapoglu, Ercin	Hacettepe University	Turkey
Kaszubiak, J.	University Magdeburg	Germany
Katsikas, Sokratis	University of Piraeus	Greece
Katsuyama, Tomoo	Numazu National College of Technology	Japan
Kawaguchi, Masashi	Suzuka National College of Technology	Japan
Khalatov, Artem	Institute of Engineering Thermophysics	USA
Khatri, Anil	Bowie State University	USA
Khmelevsky, Youry	Okanagan College	Canada
Kim, Jeongdae	Information and Communications University	South Korea
Kim, Kyungdeok	Uiduk University	South Korea

Kim, Kyungwoo	University of Florida	USA
Kim, Seok-Man	Chungbuk National University	South Korea
Kim, Seong-Gon	University of Florida	USA
Kim, Yeo Jin	Samsung Electronics, Co. Ltd.	South Korea
Kim, Yong Hak	Korea Electric Power Research Institute	South Korea
Kim, Yong-Gi	Gyeongsang National University	South Korea
Kira, Dennis S.	Concordia University	Canada
Kirova, Vassilka	Alcatel-Lucent Technologies	USA
Kitagawa, Kazuhiro	World Wide Web Consortium Keio	Japan
Klapp, Jaime	National Nuclear Research Institute	Mexico
Kobal, Damjan	University of Ljubljana	Slovenia
Koleva, Maria	Bulgarian Academy of Sciences	Bulgaria
Kouki, A.	School of High Technology	Canada
Krakowska, Monika	Jagiellonian University	Poland
Kreisler, Alain	University of Paris-Sud	France
Kromrey, Jeffrey D.	University of South Florida	USA
Kronreif, Gernot	Austrian Research Centers	Austria
Kulkarni, Arun	University of Texas at Tyler	USA
Kumar, Aman	BCL Technologies Inc.	USA
Kung, C. M.	I-Shou University	Taiwan
Kurita, Takio	Hiroshima University	Japan
Kutter, Anna K.	University of Education Weingarten	Germany
La, Gaetano	University of Joensuu	Finland
Lahlouhi, Ammar	University of Biskra	Algeria
Laksmivarahan, S.	University of Oklahoma	USA
Lasmanis, Aivars	University of Latvia	Latvia
Latawiec, Krzysztof J.	Opole University of Technology	Poland
Latchman, Haniph A.	University of Florida	USA
Lau, Newman	Hong Kong Polytechnic University	Hong Kong
Lee, Hwajung	Radford University	USA
Lee, Jae-Suk	Gwangju Institute of Science and Technology	South Korea
Lee, Nam Ho	Korea Electric Power Research Institute	South Korea
Lee, Suk Bong	Samsung Electronics, Co. Ltd.	South Korea
Lee, Yih-Jiun	Chien Kuo Technology University	Taiwan
Letellier, T.	Victor Segalen Bordeaux 2 University	France
Leu, Fang-Yie	Tunghai University	Taiwan
Li, Longzhuang	Texas A & M	USA
Li, Xiaoping	National University of Singapore	Singapore
Lin, Feng-Tyan	National Taiwan University	Taiwan
Lin, Hong	University of Houston	USA
Lin, Huan Yu	National Chiao Tung University	Taiwan
Lin, Shu-Chiung	Tatung University	Taiwan
Lin, Wei-Kuo	Tatung University	Taiwan
Lin, Yi	Takming University	Taiwan
Linares, Oscar	Francisco Jose of Caldas District University	Colombia
Lind, Nancy	Illinois State University	USA
Litvin, Vladimir	California Institute of Technology	USA
Liu, Jun	University of Ulster	UK
Liu, Kuo-Shean	Tatung University	Taiwan
Liu, Shih-Chi	Tatung University	Taiwan
Livne, Nava L.	University of Utah	USA

Livne, Oren E.	University of Utah	USA
Lizano-Dimare, Maria	Sacred Heart University	USA
Lloret Mauri, Jaime	Polytechnic University of Valencia	Spain
Long, Changjiang	Huazhong University	China
López Román, Leobardo	University of Sonora	Mexico
López Zafra, Juan Manuel	Complutense University of Madrid	Spain
Lou, Shi Jer	NPUST	Taiwan
Loutfi, Mohamed	University of Sunderland	UK
Love C., Gloria	Dillard University	USA
Lowe, John	University of Bath	UK
Lowry, Pam	Lawrence Technological University	USA
Lyell, Margaret	Intelligent Automation, Inc.	USA
Lyuu, Yuh-Dauh	National Taiwan University	Taiwan
Ma, Yongqing	Victoria University of Wellington	New Zealand
Machotka, Jan	University of South Australia	Australia
Maggiore, Giulio	Telecom Italia	Italy
Mahendran, Francis	Motorola Software Group Singapore	Singapore
Mahgoub, Ahmed G.	Alexandria University	Egypt
Mak, Peng Un	University of Macau	Macau
Makaruk, Aleksander	Vienna University of Technology	Austria
Manley, Denis	Dublin Institute of Technology	Ireland
Mansikkamäki, Pauliina	Tampere University of Technology	Finland
Marino, Mark	Erie County Community College	USA
Marlowe, Thomas	Seton Hall University	USA
Marques Cardoso, A. J.	University of Coimbra	Portugal
Martínez, Lorenzo	Autonomous Technological Institute of Mexico	Mexico
Martínez, Manuel	Autonomous Technological Institute of Mexico	Mexico
Martínez, Pablo	Laboratory for Advanced Brain Signal Processing	Japan
Mascari, Jean-François	National Research Council	Italy
Masoumi, Nasser	University of Tehran	Iran
Masuko, Mayumi	Waseda University	Japan
Mathews, Brian	University of Bedfordshire	UK
Matsumoto, Kazunori	KDDI R&D Laboratories Inc.	Japan
Matsuno, Akira	Teikyo University	Japan
Mbobi, Aime Mokhoo	Centennial College	Canada
Mehrabian, Ali	University of Central Florida	USA
Mellouli, Sehl	Laval University	Canada
Michaelis, B.	University Magdeburg	Germany
Migliarese, Piero	University of Calabria	Italy
Miller, R.	The MITRE Corporation	USA
Minnaert, Eric	South Dakota School of Mines & Technology	USA
Mirza Sebzali, Yussef	Kuwait Institute for Scientific Research	Kuwait
Mitchell, Charles	Grambling State University	USA
Mitrovic, Frane	University of Split	Croatia
Mohtashami, Mojgan	Advanced Infrastructure Design- Inc.	USA
Moosavian, Ali A.	K. N. Toosi University of Technology	Iran
Moreau, Alban	University of Brest	France
Morii, Hisashi	Shizuoka University	Japan
Mostafaeipour, Ali	Yazd University	Iran
Mugellesi Dow, Roberta	European Space Agency	Germany
Munitic, Ante	University of Split	Croatia

Muñoz, Humberto	Southern University	USA
Mura, Gianluca	Polytechnic University of Milan	Italy
Muraleedharan, Rajani	Syracuse University	USA
Naddeo, Alessandro	University of Salerno	Italy
Nadworny, Margaret	Global Software Group, Motorola	USA
Nagalakshmi, Vadlamani	Gandhi Institute of Technology and Management	India
Nagaoka, Tomoyuki	Tokyo Metropolitan University	Japan
Nahmens, Isabelina	University of Central Florida	USA
Nair, V. S. Sukumaran	Southern Methodist University	USA
Nakashima, Takuya	Shizuoka University	Japan
Natriello, Gary	Columbia University	USA
Nave, Felecia M.	A&M University	USA
Nawawi, S. W.	University Technology of Malaysia	Malaysia
Nazmy, Taymoor M.	Ain Shams University	Egypt
Nedic, Zorica	University of South Australia	Australia
Nelwamondo, F. V.	University of the Witwatersrand	South Africa
Neo, Yoichiro	Shizuoka University	Japan
Nerguizian, Vahé	School of High Technology	Canada
Noh, Bong Nam	Chonnam National University	South Korea
Nousala, Susu	University of Melbourne	Australia
Novikov, Oleg	Tomko Inc.	Russian Federation
Nyan, M. N.	National University of Singapore	Singapore
Oberer, Birgit	University of Klagenfurt	Austria
Obrebski, Jan B.	Warsaw Polytechnic	Poland
O'Brien, Ann	Loughborough University	UK
Obwegeser, Nikolaus	Vienna University of Economics and Business	Austria
Ocelka, Tomas	Institute of Public Health Ostrava	Czech Republic
Oh, Yun Sang	Samsung Electronics, R. O. Korea	South Korea
Olla, Phillip	Madonna University	USA
Ong, Soh-Khim	National University of Singapore	Singapore
Ong, Vincent Koon	University of Bedfordshire	UK
Ophir, Dan	Ariel University Center of Samaria	Israel
Osadciw, Lisa Ann	Syracuse University	USA
Ostergaard, Soren Duus	IBM Europe	Denmark
Overmeyer, Ludger	University of Hannover	Germany
Oya, Hidetoshi	Shonan Institute of Technology	Japan
Ozdemir, Ahmet S.	Marmara University	Turkey
Palesi, Maurizio	University of Catania	Italy
Paré, Dwayne E.	University of Toronto Scarborough	Canada
Parks, Gary G.	National University	USA
Passos, Nelson	Midwestern State University	USA
Patel, Nishit	Stevens Institute of Technology	USA
Patil, Adarsh	University College Cork	Ireland
Pécatte, Jean-Marie	University Paul Sabatier-Toulouse 3	France
Pedrycz, Witold	University of Alberta	Canada
Peng, Jian	University of Saskatchewan	Canada
Pereira, Elisabeth	University of Aveiro	Portugal
Pfeifer, Michael	Technical University of Dortmund	Germany
Phillips, C. Dianne	NorthWest Arkansas Community College	USA
Phillips, James	Louisiana Tech University	USA
Pierce, Ryan M.	Arkansas State University	USA

Pieters, Cornelis P.	University for Humanistics	Netherlands
Pochiraju, Kishore	Stevens Institute of Technology	USA
Podaru, Vasile	Military Technical Academy	Romania
Polanski, Andrzej	Silesian University of Technology	Poland
Poobrasert, Onintra	NECTEC	Thailand
Poon, Gilbert	University of Waterloo	Canada
Postolache, Octavian	Institute of Telecommunications	Portugal
Potente, Till	RWTH Aachen University	Germany
Praus, Petr	Charles University in Prague	Czech Republic
Pulcher, Karen L.	University of Central Missouri	USA
Pun, Daniel	Central Queensland University	Australia
Putnam, Janice	University of Central Missouri	USA
Rachev, Boris	Technical University of Varna	USA
Rafaf, Mustapha	School of High Technology	Canada
Rahman, Anuar Abdul	Pusat Tenaga Malaysia	Malaysia
Rahman, Fuad	BCL Technologies Inc.	USA
Ramessur, Roshan	University of Mauritius	Mauritius
Repa, Václav	University of Economics in Prague	Czech Republic
Revetria, Roberto	University of Genoa	Italy
Reyes-G., Leonardo	Carnegie Mellon University	USA
Reyes-Méndez, Jorge Joel	Metropolitan Autonomous University	Mexico
Riaz Moghal, Mohammad	University College of Engineering and Technology	Pakistan
Ricquebourg, Vincent	University of Picardie Jules Verne	France
Rigaud, Bernard	Champollion University	France
Rodrigues, Joao	University of Madeira	Portugal
Rodríguez, María Dolores	University of Alcala	Spain
Rodríguez L., Gloria I.	National University of Colombia	Colombia
Rossignol, R.	Victor Segalen Bordeaux 2 University	France
Saad, Inès	Amiens School of Management	France
Saadane, Rachid	Institut Eurecom	France
Sahara, Tomohiro	Osaka Institute of Technology	Japan
Sala, Nicoletta	University of Italian Switzerland Largo Bernasconi	Switzerland
Salazar, Dora	Texas Tech University	USA
Saleh, Magda M.	Alexandria University	Egypt
Sanna, Andrea	Polytechnic University of Turin	Italy
Šaruckij, Mark	Cracow University of Economics	Poland
Sato, Tomoaki	Hirosaki University	Japan
Savva, Andreas	University of Nicosia	Cyprus
Sax, Eric	Mercedes-Benz Technology Gmbh	Germany
Schaeffer, Donna M.	Marymount University	USA
Schlette, Christian	RWTH Aachen University	Germany
Schrader, P. G.	University of Nevada	USA
Schulz, Lennart	Institute of Transport and Automation Technology	Germany
Schumacher, Jens	University of Applied Sciences Vorarlberg	Austria
Sert, Yasemin	University of South Florida	USA
Seyed-Abbassi, Behrooz	University of North Florida	USA
Shao, Shiyun	National University of Singapore	Singapore
Shaw, Jill	The Open University	UK
Shayeghi, Hossien	Iran University of Science	Iran
Shim, Eung Bo	Korea Electric Power Research Institute	South Korea
Shin, Jungpil	University of Aizu	Japan

Shiraishi, Yoshiaki	Nagoya Institute of Technology	Japan
Shum, Kwok	Stanford University	USA
Sim, Sang Gyoo	Samsung Electronics, Co. Ltd.	South Korea
Singh, Madhusudan	Netaji Subhas Institute of Technology	India
Sinkunas, Stasys	Kaunas University of Technology	Lithuania
Sleit, Azzam Talal	University of Jordan	USA
Smith III, Charles V.	University of South Alabama	USA
Soeiro, Alfredo	University of Porto	Portugal
Sokolov, Sergey	Keldysh Institute of Applied Mathematics	Russian Federation
Song, Hong Jun	University of Sydney	Australia
Srisawangrat, Songsiri	University of Regina	Canada
Srivastava, Sanjeev	Florida State University	USA
Srivastava, Smriti	Netaji Subhas Institute of Technology	India
Stasytyte, Viktorija	Vilnius Gediminas Technical University	Lithuania
Stix, Volker	Vienna University of Economics and Business	Austria
Stößlein, Martin	Dalian University of Technology	China
Su, J. L.	Quanzhou Normal University	China
Su, Wei	National University of Defense Technology	USA
Su, Wen-Ray	NCDR	Taiwan
Sulc, Bohumil	Czech Technical University in Prague	Czech Republic
Sulema, Yevgeniya	NTUU KPI	Ukraine
Sun, Baolin	Wuhan University	China
Suomi, Reima	Turku School of Economics	Finland
Szygenda, Stephen A.	Southern Methodist University	USA
Tadj, Chakib	University of Quebec at Montreal	Canada
Takemi, Fujikawa	University of Western Sydney	Australia
Tam, Wing K.	Swinburne University of Technology	Australia
Tarnikova, Yuliya	BCL Technologies Inc.	USA
Tay, Francis E. H.	National University of Singapore	Singapore
Taylor, Stephen	Sussex University	UK
Tchier, Fairouz	King Saud University	USA
Teng, Chia-Chi	Brigham Young University	USA
Teshigawara, Yoshimi	Soka University	Japan
Thede, Anke	University of Karlsruhe	Germany
Theologou, Michael	National Technical University of Athens	Greece
Thornton, Mitchell A.	Southern Methodist University	USA
Till, Robert	John Jay College	USA
Traum, Maria	Johannes Kepler University	Austria
Treven, Sonja	University of Maribor	Slovenia
Tsai, Li-Hung	Tatung University	Taiwan
Tsaur, Woei-Jiunn	Da-Yeh University	Taiwan
Tseng, Ko Ying	Tatung University	Taiwan
Turnitsa, Charles D.	Old Dominion University	USA
Tzeng, Jang-Ruey	Tatung University	Taiwan
Ucan, Osman N.	Istanbul University	Turkey
Väänänen, Kalervo	University of Turku	Finland
Vanka, Sita	Kakatiya University	India
Vasilache, Simona	University of Tsukuba	Japan
Vasinek, Vladimir	Technical University of Ostrava	Czech Republic
Venkataraman, Satyamurti	AIAMED	India
Verlinde, Patrick	Royal Military Academy	Belgium

Verma, Pramode	University of Oklahoma	USA
Viloria, Orlando H.	Simon Bolivar University	Venezuela
Vinod, D. S.	Sri Jayachamarajendra College of Engineering	India
Visram, Zabin	University of Warwick	UK
Vodovozov, Valery	Saint Petersburg State Electrotechnical University	Russian Federation
Volkov-Husovic, T.	University of Belgrade	Yugoslavia
Voss, Andreas	Dortmund University of Technology	Germany
Vrána, Stanislav	Czech Technical University in Prague	Czech Republic
Walter J., Ammann	SLF	Switzerland
Wang, Chuan-Ju	National Taiwan University	Taiwan
Wang, Lei	University of Houston	USA
Wang, Ling	University of Oklahoma	USA
Wei, Chao-Huang	Southern Taiwan University	Taiwan
Wei, Xinzhou	City University of New York	USA
Wells, Harvey	King's College London	UK
Welsch, Friedrich	Simon Bolivar University	Venezuela
Whiteley, Rick	Calabash Educational Software	Canada
Wilson, Ralph	Florida State University	USA
Woodhead, Steve	University of Greenwich	UK
Woodthorpe, John	The Open University	UK
Wu, Chun Yin	Tatung University	Taiwan
Wu, Shang-Yu	NCDR	Taiwan
Wu, Sheng-You	National Changhua University of Education	Taiwan
Xu, L.	National University of Singapore	Singapore
Yamacli, Serhan	Mersin University	Turkey
Yang, Hung Jen	National Kaohsiung Normal University	Taiwan
Yang, Sung Un	Syracuse University	USA
Yao, Wenbing	Brunel University	UK
Yap, K. L.	National University of Singapore	Singapore
Yasser, Muhammad	Chiba University	Japan
Yazawa, Toru	Tokyo Metropolitan University	Japan
Yilmaz, Levent	Auburn University	USA
Yoshida, Eri	Toyohashi University of Technology	Japan
Young, Jieh-Shian	National Changhua University of Education	Taiwan
Yu, Chong Ho	University of California at Berkeley	USA
Yu, Ke	National University of Singapore	Singapore
Yu, Xin	University of Bath	UK
Yuen, Fei Lung	University of Hong Kong	Hong Kong
Zadeh, Jeff	Virginia State University	USA
Zaretsky, Esther	Hebrew University	Israel
Zelinka, Tomas	Czech Technical University in Prague	Czech Republic
Zeller, Andrew J.	Norisbank GmbH	Germany
Zhang, Qingyu	Arkansas State University	USA
Zhang, Xiaozheng Jane	California Polytechnic State University	USA
Zheng, Huiyong	University of Michigan	USA
Zhu, Hui	Soochow University	China
Zobaa, Ahmed	Cairo University	Egypt
Zolotová, Iveta	Technical University of Košice	Slovakia

ADDITIONAL REVIEWERS FOR THE NON-BLIND REVIEWING

Blendler, Thomas	PlantaPhile	USA
De Oliveira Jr., Vani X.	Universidade Federal do ABC	Brazil
Fox, Roy	University of Missouri	USA
Hansen, David	Claremont McKenna College	USA
Homem De Mello, Paula	Universidade Federal do ABC	Brazil
Hyner, Gerald	Purdue University	USA
Irgens, Cristopher	University of Strathclyde	UK
Lasmanis, Aivars	University of Latvia	Latvia
Mariotti, Maria A.	Università di Siena	Italy
Michelini, Marisa	University of Udine	Italy
Molnar, Peter	Clark Atlanta University	Georgia
Orzari Ribeiro, Anderson	Universidade Federal do ABC	Brazil
Pani, Niranjan	Utkal University	India
Parfitt, Karen	Pomona College	USA
Reeves, Scott	University of California	USA
Revilla, Miguel	Universidad de Valladolid	Spain
Richta, Karel	Charles University	Czech Republic
Rucker, Sandra	Clark Atlanta University	Georgia
Schafer, Markus	University of Toronto	Canada

PROGRAM COMMITTEE

Chairs: Jorge Baralt (Venezuela) and
Hsing-Wei Chu (USA)

Adascalitei, Adrian	Gheorghe Asachi Technical University of Iaşi	Romania
Aguilar Torres, Fernando J.	University of Almeria	Spain
Ahmad, Ashraf	National Chiao Tung University	Taiwan
Al Obaidy, Mohaned	Gulf College	Oman
Al-Aomar, Raid	Jordan University of Science and Technology	Jordan
Alshara, Osama	Higher Colleges of Technology	UAE
Ariwa, Ezendu	London Metropolitan University	UK
Ayuga Téllez, Esperanza	Polytechnic University of Madrid	Spain
Barberà, Elena	Open University of Catalonia	Spain
Batos, Vedran	University of Dubrovnik	Croatia
Belyakov, V. A.	L. D. Landau Institute for Theoretical Physics	Russian Federation
Benbouziane, Mohamed	University of Tlemcen	Algeria
Bernardino, Jorge	Institute of Engineering of Coimbra	Portugal
Bönke, Dietmar	Reutlingen University	Germany
Bubnov, Alexey	Academy of Science of the Czech Republic	Czech Republic
Calenzo, Patrick	IM2NP	France
Carbon, Margarita	The Boeing Company	USA
Castaño-Moraga, Carlos	University of Las Palmas de Gran Canaria	Spain
Chang, Maiga	Athabasca University	Canada
Chang, Ruay-Shiung	National Dong Hwa University	Taiwan
Chang, Wen-Kui	Tunghai University	Taiwan
Chen, Yuhua	University of Houston	USA
Chen, Zhi	University of Kentucky	USA
Cheng, Xiaochun	University of Reading	UK
Chinesta, Francisco	Arts et Métiers ParisTech	France
Chiou, Yin-Wah	NanHua University	Taiwan
Clarke, Tim	University of Wales Institute Cardiff	UK
Conteh, Nabie Y.	Shenandoah University	USA
Curran, Kevin	University of Ulster	UK
D'Angelo, Anthony	University of South Florida	USA
Davies, B. J.	International Journal of Advanced Manufacturing Technology	UK
Davim, J. Paulo	University of Aveiro	Portugal
Dierneder, Stefan	Linz Center of Mechatronics GmbH	Austria
Dudek, Agnieszka	Wroclaw University of Environmental and Life Sciences	Poland
Ehmann, Kornel F.	Northwestern University	USA
Erkollar, Alptekin	University of Applied Sciences Wiener Neustadt	Austria
Estrems, Manuel	Polytechnic University of Cartagena	Spain
Faura, Félix	Polytechnic University of Cartagena	Spain
Franco, Patricio	Polytechnic University of Cartagena	Spain
Fu, Yonggang	Shanghai Jiao Tong University	China
Fúster-Sabater, Amparo	Spanish Council for Scientific Research	Spain

Gelbukh, Alexander	National Polytechnic Institute	Mexico
Gibson, Thomas L.	General Motors Research and Development Center	USA
Glovnea, Romeo P.	Brunel University	UK
Gomes, Samuel	University of Technology of Belfort-Montbéliard	France
Goriachkin, Oleg	Volga State University of Telecommunication and Informatics	Russian Federation
Greenberg, William	Virginia Polytechnic Institute and State University	USA
Grenmyr, Gustav	Chalmers University of Technology	Sweden
Gschweidl, Manfred	Vorarlberg University of Applied Sciences	Austria
Hepdogan, Seyhun	University of Central Florida	USA
Hetzer, Dirk	T-Systems International GmbH	Germany
Higashiyama, Yoichi	Ehime University	Japan
Hirschberg, Wolfgang	Graz University of Technology	Austria
Holifield, David	University of Wales Institute Cardiff	UK
Huang, Yo-Ping	Tatung University	Taiwan
Ioannides, Stathis	SFK Research & Development Company B. V.	Netherlands
Iovan, Stefan	Romanian Railway IT Company	Romania
Jazar, Reza N.	Royal Melbourne Institute of Technology University	Australia
Jiménez R., Lourdes	University of Alcala	Spain
Jingchao, Chen	University Donghua	China
Kalogiannis, Konstantinos	Brunel University	UK
Kaminski, Jacek	Chalmers University of Technology	Sweden
Keiski, Riitta L.	University of Oulu	Finland
Khamba, J. S.	Punjabi University	India
Kobielarz, Magdalena	Wroclaw University of Technology	Poland
Koeglmayr, Hans-Georg	Pforzheim University of Applied Sciences	Germany
Koleva, María	Bulgarian Academy of Sciences	Bulgaria
Kuravi, Sarada	University of South Florida	USA
Kuropka, Piotr	Wroclaw University of Environmental and Life Sciences	Poland
Lahlouhi, Ammar	University of Biskra	Algeria
Lamo, Yngve	Bergen University College	Norway
Lappas, Georgios	TEI	Greece
Li, Longzhuang	Texas A&M	USA
Li, Man-Sze	IC Focus Ltd.	UK
Linss, Gerhard	University of Technology Ilmenau	Germany
Litvin, Vladimir	California Institute of Technology	USA
Liu, Jun	University of Ulster	UK
Lloret Mauri, Jaime	Polytechnic University of Valencia	Spain
López Román, Leobardo	University of Sonora	Mexico
Lübbecke, Steffen	Steinbeis-Transferzentrum Qualitätssicherung & Bildverarbeitung	Germany
Luh, Guan Chun	Tatung University	Taiwan
Macke, Janaina	University of Caxias do Sul	Brazil
Mansikkamäki, Pauliina	Tampere University of Technology	Finland
Mares, Cristinel	Brunel University	UK
Masoumi, Nasser	University of Tehran	Iran
Masunov, Artëm E.	University of Central Florida	USA
Mbobi, Aime Mokhoo	Centennial College	Canada
Mehrabian, Ali	University of Central Florida	USA
Mikhaylov, Ivan A.	University of Central Florida	USA
Mosley, Pauline	Pace University	USA

Müller, Christoph Jo.	INCAP GmbH	Germany
Mura, Gianluca	Polytechnic University of Milan	Italy
Murugan, Natarajan	Coimbatore Institute of Technology	India
Nagai, Yasuo	Tokyo University of Information Sciences	Japan
Nagar, Atulya K.	Liverpool Hope University	UK
Nagl, Manfred	RWTH Aachen University	Germany
Narasimhan, V. Lakshmi	Western Kentucky University	USA
Newman, Stephen	University of Bath	UK
Nicolescu, Cornel Mihai	Royal Institute of Technology	Sweden
Nippa, Markus	Pforzheim University	Germany
Oberer, Birgit	University of Klagenfurt	Austria
Ojeda, Germán Y.	Colombian Institute of Petroleum	Colombia
Ong, Pangleen	University of Kentucky	USA
Ouwerkerk, David B.	General Motors Advanced Technology Center	USA
Paredes, Julia	Georgia Gwinnett College	USA
Pennington, Richard	Georgia Gwinnett College	USA
Pérez, Carlos A.	EAFIT University	Colombia
Petit, Frédéric	Polytechnic Montreal	Canada
Podaru, Vasile	Military Technical Academy	Romania
Praus, Petr	Charles University in Prague	Czech Republic
Pursell, David	Georgia Gwinnett College	USA
Rai, Bharatendra K.	University of Massachusetts Dartmouth	USA
Ramessur, Roshan	University of Mauritius	Mauritius
Remondino, Marco	University of Turin	Italy
Revetria, Roberto	University of Genova	Italy
Ribeiro, Cristina	University of Waterloo	Canada
Rieder, Mathias	Vorarlberg University of Applied Sciences	Austria
Rodger, P. M.	University of Warwick	UK
Rodríguez Lozano, Gloria.	National University of Colombia	Colombia
Rosenberger, Maik	University of Technology Ilmenau	Germany
Rossbacher, Patrick	Graz University of Technology	Austria
Rubim Sarate, João A.	Federal University of Rio Grande do Sul	Brazil
Sahli, Nabil	National Agency for Computer Security	Tunisia
Sala, Nicoletta	University of Italian Switzerland	Switzerland
Sauder, Deborah	Georgia Gwinnett College	USA
Schaeffer, Donna M.	Marymount University	USA
Schaetter, Alfred	Pforzheim University	Germany
Schellhorn, Mathias	Ilmenau University of Technology	Germany
Schumacher, Jens	Vorarlberg University of Applied Sciences	Austria
Schumann, Maik	Ilmenau University of Technology	Germany
Shiraishi, Yoshiaki	Nagoya Institute of Technology	Japan
Siddique, Mohammad	Fayetteville State University	USA
Singh, Harwinder	Guru Nanak Dev Engineering College	India
Spearot, James A.	General Motors Research and Development Center	USA
Srinivasan, Sesha	Tuskege University	USA
Staretu, Ionel	Transilvania University of Brasov	Romania
Stefanakos, Lee	University of South Florida	USA
Sulema, Yevgeniya	National Technical University of Ukraine	Ukraine
Sultanow, Eldar	University of Potsdam	Germany
Suomi, Reima	Turku School of Economics and Business Administration	Finland
Suzuki, Junichi	University of Massachusetts Boston	USA

Szotek, Sylwia	Wroclaw University of Economics	Poland
Tchier, Fairouz	King Saud University	USA
Toussaint, Luis	University of Technology of Belfort-Montbéliard	France
Trimble, Robert	Indiana University of Pennsylvania	USA
Tsoi, Mai Yin	Georgia Gwinnett College	USA
Valakevicius, Eimutis	Kaunas University of Technology	Lithuania
Vasinek, Vladimir	Technical University of Ostrava	Czech Republic
Venkataraman, Satyamurti	All India Association for Micro Enterprise Development	India
Vinod, D. S.	Sri Jayachamarajendra College of Engineering	India
Wallner, Daniel	Graz University of Technology	Austria
Wang, Lei	University of Houston	USA
Warwick, Jon	London South Bank University	UK
Weber, Edzard	University of Potsdam	Germany
Werner, Patrick	Ilmenau University of Technology	Germany
Whiteley, Rick	Calabash Educational Software	Canada
Yaghmaee, Mohammad H.	Ferdowsi University of Mashhad	Iran
Yanagisawa, Hideaki	Toyo University	Japan
Yingling, Yaroslava	North Carolina State University	USA
Yoon, Changwoo	Electronics and Telecommunications Research Institute	South Korea
Zalewski, Romuald I.	Poznan University of Economics	Poland
Zaretsky, Esther	Givat Washington College of Education	Israel
Zelinka, Tomas	Czech Technical University in Prague	Czech Republic
Zhu, Hui	Soochow University	China
Zobaa, Ahmed	University of Exeter	UK

ADDITIONAL REVIEWERS

Abd-Elmoaty, Sherif K.	University of King Faisal	Saudi Arabia
Acosta, J. A.	University of Seville	Spain
Andreev, Rumen	Bulgarian Academy of Sciences	Bulgaria
Balas, Valentina	Aurel Vlaicu University of Arad	Romania
Bandow, Gerhard	Fraunhofer Institute for Material Flow and Logistics	Germany
Barreiro, J.	University of Leon	Spain
Bayraktar, Seyfettin	Yildiz Technical University	Turkey
Cendon, Beatriz	Federal University of Minas Gerais	Brazil
Corcuera, Pedro	University of Cantabria	Spain
Dai, Fengzhi	Matsue College of Technology	China
de Pina Filho, Armando C.	Federal University of Rio de Janeiro	Brazil
Domínguez-Dorado, Manuel	University of Extremadura	Spain
Dutta, Sourav	IEEE Bombay	India
Ghosh, Sadhana	National Institute of Industrial Engineering	India
Glowacki, Miroslaw	AGH University of Science and Technology	Poland
Goel, Arun	National Institute of Technology	India
Hunek, Wojciech	Opole University of Technology	Poland
Ilkova, Tatiana	Bulgarian Academy of Sciences	Bulgaria
Kaminska, Aniela	Poznan University of Technology	Poland
Kässi, Tuomo	Lappeenranta University of Technology	Finland

Lee, Joohwan	Seoul National University	South Korea
Li, Yu-Chiang	Southern Taiwan University	Taiwan
Lin, Chih-Ting	National Taiwan University	Taiwan
Mahdoum, Ali	Centre for Development of Advanced Technologies	Algeria
Monat, André S.	Rio de Janeiro State University	Brazil
Nisar, Humaira	Universiti Tunku Abdul Rahman	Malaysia
Pedamallu, Chandra S.	Dana-Farber Cancer Institute	USA
Pieters, Cornelis (Kees)	University for Humanistics	Netherlands
Poltronieri, Palmiro	National Research Council	Italy
Resta, Marina	University of Genova	Italy
Salva, Sebastien	University of Auvergne	France
Schlue, Benjamin	University of Bremen	Germany
Sengul, Mahmut	Schlumberger Houston Technology Center	USA
Silva, Paulo Afonso	Military Institute of Engineering	Brazil
Sirjani, Mojtaba B.	Norfolk State University	USA
Song, Bo	University of Illinois at Chicago	USA
Šraml, Matjaž	University of Maribor	Slovenia
Sutherland, Trudy	Vaal University of Technology	South Africa
Tsinarakis, George. J.	Technical University of Crete	Greece
Turcu, Cristina Elena	The Stefan cel Mare University of Suceava	Romania
Yin, Jianwen	Dell Computer Corporation	USA
Zabierowski, Wojciech	Technical University of Lodz	Poland
Zeman, Klaus	Johannes Kepler University	Austria

ADDITIONAL REVIEWERS FOR THE NON-BLIND REVIEWING

Ahmad, Reyaz	Al Ain University of Science and Technology	UAE
Ahmadian, Mehdi	Virginia Tech	USA
Almeida, Maurício	Federal University of Minas Gerais	Brazil
Angelis, André	State University of Campinas	Brazil
Ansari, Jamshed	University of Allahabad	India
Arifin, Achmad	Institute of Technology Sepuluh Nopember	Indonesia
Bandow, Gerhard	Fraunhofer Institute for Material Flow and Logistics	Germany
Barbosa, Cátia	Federal University of Minas Gerais	Brazil
Beisel, Daniel	University of Braunschweig	Germany
Chen, YangQuan	Utah State University	USA
Christovao Pio M., Andre	Universidade Estadual Paulista Júlio de Mesquita Filho	Brazil
Dobrucky, Branislav	University of Zilina	Slovakia
Gregus, Jan	Comenius University	Slovakia
Hacioglu, Yuksel	Istanbul University	Turkey
Hiraishi, Hironori	Akita National College of Technology	Japan
Huo, Ran	Columbia University	USA
Lei, Ye	Northwestern Polytechnical University	China
Mizoguchi, Fumio	Tokyo University of Science	Japan
Motieifar, Alireza	University of Manitoba	Canada
Nishiyama, Hiroyuki	Tokyo University of Science	Japan

Patnaik, Amar	National Institute of Technology Hamirpur	India
Poncini, Gina	University of Milan	Italy
Ranasinghe, Arjuna	Alabama A&M University	USA
Schubert, Klaus	University of Münster	Germany
Sharma, Somesh	National Institute of Technology Hamirpur	India
Soomro, Tariq	Al Ain University of Science and Technology	UAE
Souza, Renato	Getulio Vargas Foundation	Brazil
Spinola, Mauro de M.	University of Sao Paulo	Brazil
Thomaz, Adriano	Anhanguera Eduacacional	Brazil
Wang, Bixiang	New Mexico State University	USA
Wang, Minqing	Northwestern Polytechnical University	China
Zhang, Xiaodong	Northwestern Polytechnical University	China
Zhang, Youmin	Concordia University	China

PROGRAM COMMITTEE CHAIRS
Jorge Baralt
Hsing-Wei Chu

GENERAL CHAIR
Nagib Callaos

ORGANIZING COMMITTEE CHAIRS
José Ferrer
Andrés Tremante

CONFERENCE PROGRAM MANAGER
Leonisol Callaos

HARDCOPY PROCEEDINGS PRODUCTION CHAIRS
María Sánchez
Dalia Sánchez

TECHNICAL CONSULTANT ON COMPUTING SYSTEM / CD PROCEEDINGS PRODUCTION CHAIR
Juan Manuel Pineda

SUBMISSION QUALITY CONTROL SUPPORT
Leonardo Contreras

META REVIEWERS SUPPORT
Maria Sánchez
Dalia Sánchez

SYSTEM DEVELOPMENT, MAINTENANCE AND DEPLOYMENT
Dalia Sánchez
Keyla Guédez
Nidimar Diaz
Bebzabeth García

OPERATIONAL ASSISTANTS
Marcela Briceño
Cindi Padilla

HELP DESK
Louis Barnes
Sean Barnes
José Aponte

IMCIC'12

Number of Papers Included in these Proceedings per Country

(The country of the first author was the one taken into account for these statistics)

Country	# Papers	%
TOTAL	68	100%
United States	23	33,82%
Brazil	9	13,24%
Japan	6	8,82%
India	4	5,88%
Russian Federation	3	4,41%
Slovakia	3	4,41%
Canada	2	2,94%
Czech Republic	2	2,94%
Italy	2	2,94%
Spain	2	2,94%
Australia	1	1,47%
China	1	1,47%
Colombia	1	1,47%
Iran	1	1,47%
Lebanon	1	1,47%
Malaysia	1	1,47%
Mexico	1	1,47%
New Zealand	1	1,47%
Nigeria	1	1,47%
Taiwan	1	1,47%
Germany	1	1,47%
United Arab Emirates	1	1,47%

Foreword

Our purpose in The 3rd International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC 2012) is to provide, in these increasingly related areas, a multi-disciplinary forum, to foster interdisciplinary communication among the participants, and to support the sharing process of diverse perspectives of the same transdisciplinary concepts and principles. Complexity, Cybernetics and Informatics are being increasingly related to each other in almost every scientific discipline, engineering area, and human activity. Their common transdisciplinarity characterizes and communicates them, generating strong relations among them and with other disciplines. They work together to create a whole new way of thinking and practice. This phenomenon persuaded the Organizing Committee to structure IMCIC 2012 as a multi-conference where participants may focus on one area, or on one discipline, while allowing them the possibility of attending events from other areas or disciplines. This systemic approach stimulates cross-fertilization among different disciplines, inspiring scholars, originating new hypothesis, supporting production of innovations and generating analogies; which is, after all, one of the very basic principles of the systems' movement and a fundamental aim in cybernetics.

IMCIC 2012 was organized and sponsored by the International Institute of Informatics and Systemics (IIIS), member of the International Federation for Systems Research (IFSR). IIIS is an organization dedicated to contribute to the development of the Systems Approach, Cybernetics, and Informatics potential, using both: knowledge and experience, thinking and action, for the:

a) identification of synergetic relationships among Systemics, Cybernetics and Informatics, and between them and society;
b) promotion of contacts among the different academic areas, through the transdisciplinarity of the systems approach;
c) identification and implementation of communication channels among the different professions;
d) supply of communication links between the academic and professional worlds, as well as between them and the business world, both public and private, political and cultural;
e) stimulus for the creation of integrative arrangements at different levels of society, as well as at the family and personal levels;
f) promotion of transdisciplinary research, both on theoretical issues and on applications to concrete problems.

On behalf of the Organizing Committee, I extend our heartfelt thanks to the Program Committee's members, to the 230 reviewers, from 81 countries, who made the **double-blinded** reviews, and to the 231 reviewers, from 47 countries, who made the **non-blind** reviews. (Some reviewers supported both: non-blind and double-blind reviewing for different submissions). A total of 1028 reviews made by 461 reviewers (who made at least one review) contributed to the quality achieved in IMCIC 2012. This means an average of 5.01 reviews per submission (205 submissions were received). Each registered author had access, via the conference web site, to the reviews that recommended the acceptance of their respective submissions. Each registered author could get information about: 1) the average of the reviewers evaluations according to 8 criteria, and the average of a global evaluation of his/her submission; and 2) the comments and the constructive feedback made by the reviewers, who recommended the acceptance of his/her submission, so the author would be able to improve the final version of the paper.

In the organizational process of IMCIC 2012, about 205 papers/abstracts were submitted. These pre-conference proceedings include about 68 papers, from 22 countries, that were accepted for presentation. The submissions were reviewed as carefully as time permitted; it is expected that most of them will appear in a more polished and complete form in scientific journals. This information about IMCIC 2012 is summarized in the following table, along with the other collocated events:

Conference	# of submissions received	# of reviewers that made at least one review	# of reviews made	Average of reviews per reviewer	Average of reviews per submission	# of papers included in the proceedings	% of submissions included in the proceedings
IMCIC 2012	205	461	1028	2.23	5.01	68	33.17%
ICSIT 2012	171	389	720	1.85	4.21	65	38.01%
CICIC 2012	65	193	357	1.84	5.49	20	30.17%
TOTAL	**441**	**1043**	**2105**	**2.02**	**4.77**	**153**	**34.69%**

We also extend our gratitude to the co-editors of these proceedings, for the hard work, energy and eagerness they displayed in their respective activities. We express our intense gratitude to Professor William Lesso for his wise and opportune tutoring, for his eternal energy, integrity, and continuous support and advice, as the Program Committee Chair of past conferences, organized by the International Institute of Informatics and Systemics, and as Honorary President of IMCIC 2012, as well as for being a very caring old friend and intellectual father to many of us. We also extend our gratitude to Professor Belkis Sánchez, who brilliantly managed the organizing process.

Special thanks to Dr. C. Dale Zinn and Professor José Ferrer for co-chairing IMCIC 2012 Program Committee, to Professors Jorge Baralt, Hsing-Wei Chu, and Michael J. Savoie for their General Co-chairmanship, and to Professor Belkis Sánchez for chairing its Organizing Committee. We also wish to thank all the authors for the quality of their papers, and to the 461 reviewers for the great job they did making the 1028 reviews that supported the acceptance process. We also extend our gratitude to María Sánchez, Juan Manuel Pineda, Leonisol Callaos, Dalia Sánchez, Keyla Guedez, Nidimar Díaz, Bebzabeth García, Marcela Briceño, Freddy Callaos, Abrahan Marin, Sean Barnes, Louis Barnes, José Aponte, Cindy Padilla, and Yulaime Padilla for their knowledgeable effort in supporting the organizational process and for producing the hard copy and CD versions of the proceedings.

Professor Nagib C. Callaos
IMCIC 2012 General Co-Chair

CONTENTS

(Post-Conference Edition)

Contents

Complexity, Informatics and Cybernetics

Computer Science and Engineering

Control Systems, Technologies and Applications

Cybernetics (Communications and Control)

ICT Applications in Education

ICT Applications in Management Engineering and Engineering Management

ICT Applications in Risk Management and Engineering

ICT Applications in Social and Political Systems

Informatics

Scientific and Engineering Computing

Software Engineering

Design and Modeling in Science, Education, and Technology

Design and Modeling Methods and Methodologies

Educational Research, Design, and Modeling

Hybrid Models and Modeling in Science and Engineering

Quantitative Models and Modeling in Science and Engineering

A Data Mining Method for the Medical Relationship between Diagnoses and Procedures – Vermont Hospital 2009

Khalid N. Alhayyan, Layinah Bouayad
{alhayyan, lbouayad}@usf.edu
Information Systems and Decision Sciences
College of Business
University of South Florida

ABSTRACT

The implementation of information systems in the medical field has created an abundance of valuable data. We use decision tress and association rules on real world hospital data to uncover interesting patterns. We build two different models linking diagnoses and procedures. We also use associative rules to explain the relationship between the principal payment source and the charges reported in our sample data. Our models could serve as a valuable tool for practitioners to increase decision confidence, as well as insurance companies to review their patients' cases at a considerably lower cost.

Keywords: Data Mining, Decision Tree, Association Rules, Medical Cases, Diagnosis, Procedure.

1. INTRODUCTION

Data mining techniques have been used in medical research for many years and have been known to be effective [1] [5]. While extracting a medical data pattern is a goal for this study, a good deal of effort has been exerted at understanding how such data is structured and distributed. It is not unusual, across reliable data mining studies, that a data mining study at this step would take a significant amount of time in order to see what a plausible link among dependent and independent variables, if any, that really worth to be studied [4]. Having accessed to large dataset, consisting of almost 54,000 data rows and 77 data columns brought many interesting research questions. That is, in order to solve problems such as misunderstanding and disapproving the relationship between a medical diagnosis and necessary medical procedures, faced by physicians, hence increasing costs on medical insurance companies, it seems reasonable to explore and build a hybrid methodology, combining data mining techniques such as association rules and classification trees. Without using data mining to drill down into medical diagnoses and procedures provided by physicians or involved practitioners, it will remain unknown just how much variability occurs, and what decisions are routinely made in a hospital. Our review indicates no prior research states clearly such rules governing this relationship. Additionally, it seems reasonable for any health-care stakeholder to explore how charges may differ across various patients who have been similarly diagnosed given different payment sources. Therefore, our second research question focuses on exploring the relationship that may exist between the payment party, or the principal payment source, and the corresponding diagnosis as independent variables and the total charges presented as the dependent variable.

The methodology is applied to real-world hospital data collected from a Vermont hospital web site (http://www.bishca.state.vt.us/health-care/vuhdds-2009-and-earlier) and is evaluated through contrasting the outcomes of different data mining techniques such as classification trees and association rules. The outcomes of this study are intended to help physicians, patients, and insurance companies to make a faster and more accurate understanding of the relationship between medical diagnosis and procedures as well as understanding of how the principal payment source and the corresponding diagnosis impact charges.

2. BACKGROUND

Most hospitals today have been involved with many different types of information systems for managing their healthcare data. These systems generate huge amount of data in various forms such as text, numbers, images, and charts [7], [8], [9], and [10]. The challenge is in finding intelligent ways that enable physicians and other healthcare stakeholders to make acute clinical decisions. Considering the abundance of data contents, scholars have focused on finding plausible techniques to extract useful information from different sources of data. Masuda et al [2] applied medical data mining to classify and predict prognosis and diagnosis of patients in specialized medical areas. Kononenko [3] argues that data mining methods can be applied for training physicians to solve a specific diagnosis problem. Chae et al [9] contrasted the predictive power of three different techniques (CHIAD and C5.0 data mining decision trees and logistic regression) in predicting health outcomes to provide policy information for hypertension management. Their findings indicated that the data mining techniques resulted in more accurate health outcomes. Many other researchers have employed different data mining techniques to report interesting medical-related knowledge discoveries. These studies investigated various medical-related researches [11] [12] [13] [14] [15] [16] [17] [18].

While there are several algorithms for classification and prediction tasks, decision trees and association rules are among the most frequently used techniques in the medical data mining area. While it is easy to find many situations to prove the decision tree and association rule to be useful in the business domain, the decision tree enables to predict procedures and diagnoses in the domain of medicine. Using tree-structured models in the form of 'IF condition-based-on attribute-values THEN outcome-value' to identify useful features of importance. These two features are argued to increase the readability and clarity of extracted findings. In addition, decision trees and association rules were found, in some cases, to outperform traditional statistical techniques in their predictive power [7] [9].

Decision trees are knowledge representation structures consisting of nodes and branches organized in the form of trees such that, every internal non-leaf node is labeled with values of the attributes [19]. Starting with the original dataset, it is constructed by splitting subsets of the dataset into two or more nodes repeatedly. The best split process is determined based on the significance of difference among the categories of the predictor variable with respect to the target variable. The goodness (prediction power) of an extracted-tree model is determined by the model accuracy, where the higher it is, the more powerful the prediction.

An association rule, on the other hand, gives an occurrence relationship among variables, where there are no predictors or target factors. It is intended to capture a certain type of dependence among variable values [9] [10]. In this paper for example, the association rule is used to identify the occurrence relationship between the medical diagnoses and procedures. For example, an inducted rule such as 'IF patient is diagnosed with Complications of pregnancy, childbirth, perperium" THEN the procedure is Obstetrical procedures' is evaluated by the frequent occurrences in data that provides an evidence of the relationship strength.

3. DATA DEVELOPMENT

3.1 Data Collection

This study uses secondary data collected from hospitals in Vermont in 2009. The data set is available for public use and is separated into three different files. Inpatient hospital data encloses records for patients admitted to the hospital and stayed for more than a twenty four hour period. Outpatient data refers to patients having procedures done at the hospital but are meant to leave before the end of the twenty four hour period. Emergency data contains records of patients admitted to the hospital through the emergency clinic. Even if encapsulated in Inpatient and Outpatient files, emergency data has been disjointed for ease of use. For this particular study, we have selected the use of Inpatient data since it allows for more interesting analyses such as the number of days stayed at the hospital in relationship to other factors such as admittance date or type of insurance.

Since the Vermont data was originally designed for analysis and mining, it was exceptionally easy to use compared to other public data sets. The data sets came with excel files detailing the description, length, and expected values for each column. Considering the complexity of terms used in the medical domain, these descriptor file were particularly helpful in making sense out of the collected data. Additionally, many experts who are blind to the study but closely engaged in the medical domain were consulted in order to increase our understanding of terms complexity, medical diagnoses and procedures, and medical payment sources. Out of the 77 columns in the data set, diagnoses and procedures were key variables. Twenty columns representing the different diagnoses a single patient might have gone through the hospital stay. Another twenty columns denoted the procedures performed on a single patient. Within the medical domain, it is well understood that any successive diagnosis or procedure requires that its predecessor be performed respectively. The different diagnoses and procedures were coded using ICD-9-CM available for use and interpretation online at ICD9.chrisenders.com. This particular websites describes codes used in 2009 data; which matches the dates our data was collected. In addition, the data set also contains diagnoses and procedure columns expressed in CCS codes (Clinical Classifications Software). These codes are also available for public use at the Healthcare Cost and Utilization Project (HCUP) website. In general, CCS is a tool for clustering patient diagnoses and procedures into a manageable number of clinically meaningful categories developed at the Agency for Healthcare Research and Quality (AHRQ, formerly known as the Agency for Health Care Policy and Research). CCS is used for grouping conditions and procedures without having to wade through thousands of codes. This "clinical grouper" makes it easier to quickly understand patterns of diagnoses and procedures so that health plans, policymakers, and researchers can analyze costs, utilization, and outcomes associated with particular illnesses and procedures. CCS collapses diagnosis and procedure codes from the International Classification of Diseases, 9th Revision, Clinical Modification (IDC-9-CM), which contains over 12,000 diagnosis codes and 3,500 procedure codes. Without CCS, the large number of ICD-9-CM codes poses difficulties in statistical analysis, data mining, and reporting.

Besides the diagnoses and procedures variables, three more columns were focused on for the purposes of answering our second research question. These columns are the principal payment source, the corresponding diagnosis, and the charges incurred by a patient.

3.2 Data Cleaning

Even though the Vermont data sets were primarily designed for mining purposes, some cleaning was necessary to allow for use of our selected data mining tool (Orange). In order to modify the file content, we first converted the text file available to an excel file for editing and filtering purposes. Then, redundant columns such as Charges_HCIA were removed. In addition, some column names were modified to reduce the length. Lengthy column names posed a problem while converting the excel file back to a tab-delimited text file. For this purpose, parts of names such as "_QTR", "GRP", and "CCS" were replaced by shorter syllables such as "Q", "G", and "N", respectively. To avoid other conversion issues, the charges column was modified to contain numbers with no decimal places.

After successfully converting the file into a format accepted by the chosen data mining tool, we needed to tackle issues generated during the loading process. Since the selected tool considers the rightmost column as being the dependent variable, it expects the last column in the dataset to be categorical. Therefore, we created a new column by discretizing the "CHARGES" column. The discretization process was made after careful analysis of the "CHARGES" distribution of values. Statistical properties of the column, as described below, show a variance between $219 and $725525, with an average of $17500.87, a median of 10,700.05, and a mode of 1331. The new discretized column included nine categorical values ("Extremely High", "Very High", "High", "Quite High", "Slightly High", "Middle", "Slightly Low", "Quite Low", and "Low").

3.3 Data Selection and Extraction

In terms of number of columns, only data relevant to our research question was kept for use. A pre-selected list of columns was kept for analysis.

Data selection was made based on two major factors. Looking at the number of columns representing the ICD-9-CM coded diagnoses and procedures (20 columns each), along with the number of possible values in each column (about 3000 different diagnoses available), we have decided to use the CCS coded columns. The CCS coded columns include two column for diagnoses (NDX and NDXG) and four columns representing procedures (NPX, NPXG, NPPX, and NPPXG). Having about 200 different possible values, CCS coded data was specifically designed to simplify mining of the data. As mentioned in several research studies in the medical realm, CCS coding is deemed more adequate for this type of analysis. In addition to these columns, we have included columns of principal payment source (PPAY), diagnoses group (NDXG), and charges (CHRGS) in order to be able to explore and answer the second research question.

3.4 Descriptive Statistics

Given the selected variables, visualization inspections have been conducted to ensure that the ranges of values across those variables are normally distributed. The primary diagnoses and procedure variables with values between 1 and 256 (in accordance with the CCS single-level codes), and the diagnoses and procedures groups coded with numbers ranging from 1 to 16 (representing the high level group of diagnosis or procedure). The secondary procedure and procedure group fall in the same ranges of values (since encoded using the same CCS standard), represent any additional group of procedure that might have been performed to the patient during his/her stay at the hospital.

The "Charges" column, as described earlier, was discretized using continuous charges values and includes values ranging from "LOW" to "Extremely High". Last but not least, the "PPAY" column (Principal Payment Source) varied from 0 to 12, and identifies the party responsible for paying the billed amount.

	Accuracy	Sensitivity	specificity
Model 1	42.89%	82.41%	69.78%
Model 2	47.24%	54.26%	91.32%
Model 3	69.59%	57.58%	91.24%

Table 1: Evaluation Results on Sample Data

4. RULE BUILDING AND EVALUATION

Focusing on exploring patterns in the data that would contribute to both understanding the behavior and improving the medical practice, we aimed at building a highly predictive model that addresses our two research questions.

First, we analyze the data using regression and smoothing splines algorithms. Because of the nature of data, where the same procedure can be linked to several diagnoses, these models all generated low predictive power. The adjusted R-square values, in each case, did not exceed 10%.

We then geared towards using data mining techniques. For practicality purposes, our first models used the Classification tree algorithm and were built and evaluated using a sample data of 22900 records. With 70% of the data used for training and 30% of used for testing, we were able to generate the following results:

Model 1: predicts single-level procedures (CCSPX), using single-level diagnoses (CCSDX),
Model 2: predicts single-level procedures (CCSPX), using group-level diagnoses (CCSDXGRP).
Model 3: predicts single-level procedures (CCSPXGRP), using group-level diagnoses (CCSDXGRP).

Once the models were identified, the final models were built using the complete data file, containing about 54,000 records. Using the diagnosis group, the model was able to predict with 60.43% accuracy the procedures needed to be administered.

	CA	Sensitivity	Spec	Precision
Classification Tree	0.60	0.8	0.72	0.5508

Table 2: Evaluation Results for Predicting Procedures

For verification and tracing purposes, we have also built a model predicting diagnoses from the procedures performed. This model can be used as review tool to be used by auditors and medical interns. The model generated an accuracy level of 42%.

	CA	Sensitivity	Spec	Precision
Classification Tree	0.42	0.00	1.00	-1.00

Table 3: Evaluation Results for Predicting Diagnoses

Sup %	Conf %	Rule	Description
10	99.3	**IF** NPXG = 13 **THEN** NDXG = 11	**IF** the procedure performed is "Obstetrical procedures" T**HEN** the diagnosis was "Complications of pregnancy, childbirth, perperium".
5	55	**IF** NPXG = 14 **THEN** NDXG = 13	**IF** the procedure performed is "Operations on the musculoskeletal system" **THEN** the diagnosis was "Musculoskeletal system and connective tissue"
5	70	**IF** NPXG = 7 **THEN** NDXG = 7	**IF** the procedure performed is "Operations on the cardiovascular system" **THEN** the diagnosis was "Diseases of the circulatory system".
4.9	64	**IF** NPXG = 9 **THEN** NDXG = 9	**IF** the procedure performed is "Operations on the digestive system" **THEN** the diagnosis was "Diseases of the digestive system".
3.4	85	**IF** NPXG = 11 **THEN** NDXG = 15	**IF** the procedure performed is "Operations on the male genital organs" **THEN** the diagnosis was "Liveborn"

Table 4: Association Rules (Procedures => Diagnoses)

In quest for a more predictive model, we have used the k nearest neighbors (with k=10) on sample data of 22900 records. Using 70% of the data for training and 30% for testing, we were able to get an accuracy of 69.95%, sensitivity of 47.17%, and specificity 90.37%. Even though the model provided a relatively high accuracy, the model did not show sufficient predictability to counteract the low explainability and traceability powers of the model. Because medical decisions can incur valuable consequences regarding patients' lives, the model being used needs to either have a tremendous predictive power, of a high degree of explainablity to support the decision being made.

Considering the fact the "clinical grouper" (NDXG and NPXG) provided higher accuracy, we focused on this model to extract and understand patterns. First, NPXG was made the DV while NDXG made the IV. We used association rules algorithm with variations across the minimum support and the minimum confidence percentages. Twenty experiments were conducted that involved fixing one factor and changing the other. The change step ranges from 0% to 100%. A set of rules were inducted shown in the following table:

Second, NDXG was made the DV while NPXG made the IV. Following the same procedure, we were able to extract a set of rules shown in the following table:

Sup %	**Conf %**	**Rule**	**Description**
10	93	**IF** NDXG = 11 **THEN** NPXG = 13	**IF** patient is diagnosed with "Complications of pregnancy, childbirth, perperium" **THEN** the procedure is "Obstetrical procedures".
6	61	**IF** NDXG = 8 **THEN** NPXG = None	**IF** patient is diagnosed with "Diseases of the respiratory system" **THEN** no procedure is given.
6	54	**IF** NDXG = 15 **THEN** NPXG = None	**IF** patient is diagnosed with "Liveborn" **THEN** no procedure is given.
5.4	80.3	**IF** NDXG = 13 **THEN** NPXG = 14	**IF** patient is diagnosed with "Musculoskeletal system and connective tissue" **THEN** the procedure is "Operations on the musculoskeletal system".
3.9	70	**IF** NDXG = 5 **THEN** NPXG = None	**IF** patient is diagnosed with "Mental disorders" **THEN** no procedure is given.
2.4	72	**IF** NDXG = 17 **THEN** NPXG = None	**IF** patient is diagnosed with "Symptoms, signs, and ill-defined conditions" **THEN** no procedure is given.
1.8	54	**IF** NDXG = 3 **THEN** NPXG = None	**IF** patient is diagnosed with "Endocrine, nutritional, metabolic, immunity" **THEN** no procedure is given.

Table 5: Association Rules (Diagnoses => Procedures)

Moving to answer our second research question, we analyzed the relationship between the charges and the principal payment source (PPS). Looking at some rules generated by the decision tree, we can conclude that charges differ depending on the payment party.

IF NDXG <=12.500 AND NDXG <=10.500 AND NDXG <=2.500 AND NDXG <=1.500 AND PPAY >1.500 AND PPAY >6.500 AND PPAY >7.500 AND PPAY <=11.500 AND **PPAY <=9.000** THEN **CHRGS = Q_High**	IF NDXG <=12.500 AND NDXG <=10.500 AND NDXG <=2.500 AND NDXG <=1.500 AND PPAY >1.500 AND PPAY >6.500 AND PPAY >7.500 AND PPAY <=11.500 AND **PPAY >9.000** THEN **CHRGS = E_High**

Table 6: Rules Relating Diagnoses and PPS to Charges -1

In other words, for the same diagnoses, the billed amount differs from "Quite High" for self-pay patients to "Extremely High" for other payment sources. This could probably be explained by discounts offered by hospitals for self-paying individuals.

Another rule depicted from the tree suggests that, for the same diagnoses, patients having Medicare are charged less than other patients. Further investigation needs to be performed to reveal whether the patients undergone the same procedures.

IF NDXG <=12.500 AND NDXG <=10.500 AND NDXG <=2.500 AND NDXG <=1.500 AND **PPAY <=1.500** THEN **CHRGS = High**	IF NDXG <=12.500 AND NDXG <=10.500 AND NDXG <=2.500 AND NDXG <=1.500 AND **PPAY >1.500** THEN **CHRGS = E_High**

Table 7: Rules Relating Diagnoses and PPS to Charges -2

Even though the conclusions made could be interesting in nature, it is important to recall that our model accuracy was low (34%). This suggests the need for further investigation.

	CA	**Sensitivity**	**Spec**	**Precision**
Classification Tree	0.34	0.05	0.98	0.24

Table 8: Evaluation Results of Model relating Principal Payment Source to Charges

5. BUSINESS APPLICATION

Most hospitals today have employed many different types of hospital information systems to manage their healthcare or patient data. These systems typically generate vast amounts of data in the form of number, text, chart, and image. How can healthcare practitioners turn that data into useful information that would enable to make intelligent clinical decisions? Considering the fast growth of data content, size, and diversity, researchers have focused on techniques to find useful information from collections of data. Although its application to medical data analysis has been relatively limited until recently, the term 'data mining' has been increasingly used in the medical literature over the past few years. The goal of predictive data mining in clinical medicine is to derive models that use patient information to support specific clinical decisions. Data mining models can be applied to building of decision-making procedures such as prognosis, diagnosis, and treatment planning, which once evaluated and verified, could then be embedded in clinical information systems.

Therefore, the purpose of this report is to focus on generating association rules that help physicians, hospital administrative, and may be insurance companies in two perspectives. First, understanding the link between patients' diagnoses and the corresponding procedures can help physicians make faster and more accurate decisions upon identifying the right diagnosis. Hospital administrative, on the other hand, will be able to follow up on un-necessary costs related to cases where diagnoses are the same but the procedures are different. The same application of such model would benefit insurance companies to cut their cost responsibilities. Second, the model that relates the payment source and the corresponding diagnosis from one side and charges from another side will be a great deal for insurance companies in order to keep track of their cost responsibilities.

6. FUTURE INVESTIGATION

While extracting a model that precisely determines the relationship between diagnoses and procedures seems a promising a data mining research question, we think that the methodology applied in this report is just an introduction to understand this relationship. Future research would require not only investigating this link on the group level, as we did in this report, but also, exploring if there is any pattern among the low-level groupings of both diagnoses and procedures. That is, we need to consider how the twenty groups of diagnoses are linked or related to the twenty groups of procedures. Answering this question would convey more accurate information to physicians and insurance companies.

7. CONCLUSION

The present study analyzes inpatient hospital data to build predictive data mining models. Predicting the procedures to be performed using doctors' diagnoses, the first model offered a great level of accuracy; Thus, providing a relatively reliable tool for doctors and insurance companies to review hospital actions at the very low cost of running the DM model. However, it is noteworthy that secondary data was used for this project. Therefore, conditions and methods used for the data collection are unknown to us.

The second model analyzed by the study predicts the patient's billed amounts using the diagnosis and the primary payment source. This model, especially designed for insurance use, suggests interesting findings with respect to the total amount charged to patients with the same diagnoses but with different types of insurances. However, for more reliability, we recommend selecting some of the cases used by the model to build a specific rule, and have them reviewed by professionals in the medical domain.

Overall, the Vermont hospital data gave us the opportunity to sharpen our data mining skills and provided us with some interesting insights about practices in the medical realm.

8. REFERENCES

[1] Sung Ho Ha and Seong Hyeon Joo, 2010, A Hybrid Data Mining Method for the Medical Classification of Chest Pain. World Academy of Science, Engineering and Technology, pp. 179-183.

[2] G. Masuda, N. Sakamoto, and R. Yamamoto, 2002, "A framework for dynamic evidence based medicine using data mining," In Proc. 15th IEEE Symposium on Computer-Based Medical Systems, IEEE press, pp. 117-122.

[3] Kononenko, 2001, "Machine learning for medical diagnosis: history, state of the art and perspective," Artificial Intelligence in Medicine, vol. 23, pp. 89-109.

[4] Michael Berry and Gordon Linoff, 2011, "Data Mining Techniques: For Marketing, Sales and Customer Relationship Management, "Third Edition, Text book, Wiley Computer Publishing.

[5] Fayyad, U. M., Piatetsky-Shapiro, G., and Smyth, P. (1996) "From data mining to knowledge discovery: an overview".

[6] Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R., "Advances in Knowledge Discovery and Data Mining", pages 1-34. AAAI Press / MIT Press, Menlo Park, CA.

[7] Kaur, H., & Wasan, S. K., 2006, "Empirical Study on Applications of Data Mining Techniques in Healthcare", Jamia Millia Islamia, New Delhi-110 025, India. Journal of Computer Science, 2(2), 194-200.

[8] Jonathan C. Prather, David F. Lobach, Linda K. Goodwin , Ph. D , Joseph W. Hales , Ph. D , Marvin L. Hage , W. Edward Hammond, 1997, "Medical data mining: Knowledge discovery in a clinical data warehouse", Annual Conference of the American Medical Informatics Association, pp 101-105.

[9] Chae, Y.M., Hee Ho, S., Won Cho, K., Lee, D., Ji, S., "Data Mining Approach to Policy Analysis in a Health Insurance Domain." International Journal of Medical Informatics 62, (2001): 103-11

[10] S Stilou, P D Bamidis, N Maglaveras, C Pappas, 2001, "Mining association rules from clinical databases: an intelligent diagnostic process in healthcare." Studies in Health Technology and Informatics 84: Pt 2.1399-1403.

[11] Bruce I. Reiner, "Uncovering and Improving Upon the Inherent Deficiencies of Radiology Reporting through Data Mining." J. Digital Imaging 23(2): 109-118 (2010).

[12] F. S. Khan, R. M. Anwer, O. Torgersson, and G. Falkman, "Data mining in oral medicine using decision trees," International Journal of Biological and Medical Sciences, vol. 4, pp. 156-161, 2009.

[13] Y. P. Yun, "Application and research of data mining based on C4.5 Algorithm," Master thesis, Haerbin University of Science and Technology, 2008.

[14] U. Abdullah, J. Ahmad, A. Ahmed, "Analysis of effectiveness of apriori algorithm in medical billing data mining," In Proc. 4th International Conference on Emerging Technologies, IEEE press, 2008, pp. 327-331.

[15] Y. Tan, G. F. Yin, G. B. Li, and J. Y. Chen, "Mining compatibility rules from irregular Chinese traditional medicine database by Apriori algorithm," Journal of Southwest JiaoTong University, vol. 15, 2007.

[16] R. Ceglowski, L. Churilov, and J. Wasserthiel, "Combining data mining and discrete event simulation for a value-added view of a hospital emergency department," Journal of the Operational Research Society, vol. 58, pp. 246-254, 2007.

[17] R. Delphine, M. Cuggia, A. Arnault, J. Bouget, and P. L. Beux, "Managing an emergency department by analyzing HIS medical data: a focus on elderly patient clinical pathways," Health Care Management Science, vol. 11, pp. 139-146, 2008.

[18] W. T. Lin, S. T. Wang, T. C. Chiang, Y. X. Shi, W. Y. Chen, and H. M. Chen, "Abnormal diagnosis of Emergency Department triage explored with data mining technology: An Emergency Department at a Medical Center in Taiwan taken as an example", Expert Systems with Applications, vol. 37, pp. 2733-2741, 2010.

[19] Han, J. and M. Kamber, 2001. Data Mining: Concepts and Techniques. San Francisco, Morgan Kauffmann Publishers.

Music Information Retrieval with Polyphonic Sounds and Timbre

Angelina A. TZACHEVA
Department of Informatics, University of South Carolina Upstate
Spartanburg, SC, 29303, USA

and

Keith J. BELL
Department of Informatics, University of South Carolina Upstate
Spartanburg, SC, 29303, USA

ABSTRACT

With the fast booming of online music repositories, the problem of building music recommendation systems is of great importance. There is an increasing need for content-based automatic indexing to help users find their favorite music objects. In this work, we propose a new method for automatic classification of musical instruments. We use a unique set of timbre related descriptors, extracted on polyphonic sounds (multiple distinct instruments) in combination with data mining classifiers. We report high classification accuracy.

Keywords: Music Information Retrieval, Data Mining, Timbre.

1. INTRODUCTION

To many people in many cultures music is an important part of their way of life. Greek philosophers and ancient Indian philosophers define music as horizontally ordered tones, or melodies and vertically ordered harmonies. Common sayings such as "the harmony of the spheres" and "it is music to my ears" point to the notion that music is often ordered and pleasant to listen to.

With 20th century music, there is a vast increase in music listening as the radio gained popularity and phonographs were used to replay and distribute music. The focus of art music is characterized by exploration of new rhythms, styles, and sounds.

Today, we hear music media in advertisements, in films, at parties, at the philharmonic, etc. One of the most important functions of music is its effect on humans. Certain pieces of music have a relaxing effect, while others stimulate us to act, and some cause a change in or emphasize our mood. Music is not only a great number of sounds arranged by a composer, it is also the emotion contained within these sounds [4].

The steep rise in music downloading over CD sales has created a major shift in the music industry away from physical media formats and towards Web-based (online) products and services. Music is one of the most popular types of online information and there are now hundreds of music streaming and download services operating on the World-Wide Web. Some of the music collections available are approaching the scale of ten million tracks and this has posed a major challenge for searching, retrieving, and organizing music content. Research efforts in music information retrieval have involved experts from music perception, cognition, musicology, engineering, and computer science engaged in truly interdisciplinary activity that has resulted in many proposed algorithmic and methodological solutions to music search using content-based methods [3].

This work contributes to solving the important problem of building music recommendation systems. Automatic recognition or classification of music sounds helps user to find favorite music objects, or be recommended objects of his/her liking, within large online music repositories. We focus on musical instrument recognition, which is a challenging problem in the domain.

Melody matching based on pitch detection technology has drawn much attention and many music information retrieval systems have been developed to fulfill this task. Fewer systems focus on timbre.

The original audio signals are a large volume of unstructured sequential values, which are not suitable for traditional data mining algorithms, while the higher level data representative of acoustical features are sometimes not sufficient for instrument recognition. This has stimulated the research on instrument classification and feature selection for content-based automatic music information retrieval.

We propose a new method for automatic classification of musical instruments with polyphonic sounds. We focus on timbre related features, in combination with data mining classifiers.

The rest of the paper is organized as follows: section 2 reviews related work, section 3 discusses timbre, section 4 describes features, section 5 illustrates our methodology, section 6 shows the experiment results, and finally section 7 concludes.

2. RELATED WORK

2.1 Classification of Monophonic Music Sound

Martin and Kim [8] employed the K-NN (k-nearest neighbor) algorithm to a hierarchical classification system with 31 features extracted from cochleagrams. With a database of 1023 sounds they achieved 87% of successful classifications at the family level and 61% at the instrument level when no hierarchy was used. Using the hierarchical procedure increased the accuracy at the instrument level to 79% but it degraded the performance at the family level (79%). Without including the hierarchical procedure performance figures were lower than the ones they obtained with a Bayesian classifier. The fact that the best accuracy figures are around 80% and that Martin and Kim [8] have settled into similar figures shows the limitations of the K-NN algorithm (provided that the feature selection has been optimized with genetic or other kind of techniques). Therefore, more powerful techniques should be explored.

Timbre classification systems have also been developed based on auditory processing and Kohonen self-organizing neural networks; where, data is preprocessed by peripheral transformations to extract perception features, then fed to the network to build a map. Such method did not quite well distinguish the instruments.

Bayes Decision Rules and Naive Bayes classifiers are simple probabilistic classifiers, by which the probabilities for the classes and the conditional probabilities for a given feature and a given class are estimated based on their frequencies over the training data. They are based on probability models that incorporate strong independence assumptions, which may, or may not have a bearing in reality, hence are naive. The resultant rule is formed by counting the frequency of various data instances, and can be used then to classify each new instance. Brown [2] applied this technique to 18 Mel-Cepstral coefficients by a K-means clustering algorithm and a set of Gaussian mixture models. Each model was used to estimate the probabilities that a coefficient belongs to a cluster. Then probabilities of all coefficients were multiplied together and were used to perform the likelihood ratio test. It then classified 27 short sounds of oboe and 31 short sounds of saxophone with an accuracy rate of 85% for oboe and 92% for saxophone.

Tzacheva and Bell [11] used timbre related features for classification of musical instruments with monophonic sounds. Authors proposed new features for preservation of temporal information. Since classifiers do not distinguish the order of the frames, they are not aware that frame t1 is closer to frame t2 than it is to frame t3 . Their proposed features allow for that distinction to be made. Authors achieved good classification accuracy.

2.2. Classification of Polyphonic Music Sound

One approach to address multi-timbre estimation in polyphonic sound is to apply sound separation techniques [12] along with the traditional classifiers. Each time when one classification label from a set is assigned, the sound separation module is applied to subtract the estimated timbre feature from the signal so that the signal of the single instrument is separate from the polyphonic sound signal. Then the classifier can be applied again on the residue of the signal to assign another label.

The sound separation process for each frame continues until the remnant of the signal is too weak to give any further timbre estimation. However, there is one problem in this method. After each sound separation process, the timbre information of the remaining instruments could be partially lost due to the overlap of multiple timbre signals, which make it difficult to further analyze the remnant of sound signal.

Instead of giving one classification label at a time, multi-label classification assigns multiple labels from a set, to the target estimation object. Some research on multi-label classification has been done in the text categorization area [7], [10], and in scene recognition [1]. Authors approached the problem by training a classifier with samples with multiple labels. A multi-label classification system for polyphonic sounds was proposed by Jiang at al. [6].

Typically a digital music recording, in form of a binary file, contains a header and a body. The header stores file information such as length, number of channels, sampling rate, etc. Unless it is manually labeled, a digital audio recording has no description of timbre or other perceptual properties. Also, it is a highly nontrivial task to label those perceptual properties for every piece of music based on its data content.

In music information retrieval area, a lot of research has been conducted in melody matching based on pitch identification, which usually involves detecting the fundamental frequency. Most content-based Music Information Retrieval (MIR) systems query by whistling/humming systems for melody retrieval. Few systems exist for timbre information retrieval, which indicates it as a nontrivial task [6]. In addition, typically, the timbre related methods use monophonic sounds (one distinct instrument). Such timbre estimation algorithms are rarely successfully applied to polyphonic sounds (multiple distinct instruments), which occur more often in the real music world. In this work, we propose a new method for automatic classification of musical instruments with polyphonic sounds. We use a unique set of timbre related descriptors in combination with data mining classifiers.

3. TIMBRE

The definition of timbre is: in acoustics and phonetics - the characteristic quality of a sound, independent of pitch and loudness, from which its source or manner of production can be inferred. Timbre depends on the relative strengths of its component frequencies; in music - the characteristic quality of sound produced by a particular instrument or voice; tone color. ANSI defines timbre as the attribute of auditory sensation, in terms of which a listener can judge that two sounds are different, though having the same loudness and pitch. It distinguishes different musical instruments playing the same note with the identical pitch and loudness. So it is the most important and relevant facet of music information. People discern timbre from speech and music in everyday life.

Musical instruments usually produce sound waves with frequencies, which are an integer (a whole number) multiples of each other. We call these frequencies harmonics, or harmonic partials. The lowest frequency is the fundamental frequency f0, which has close relation with pitch. The second and higher frequencies are called overtones. Along with fundamental frequency, these harmonic partials distinguish the timbre, which is also called tone color. The human aural distinction between musical instruments is based on the differences in timbre.

The body of a digital audio recording contains a large number of integers in a time-order sequence. For example, at a sampling rate 44,100Hz, a digital recording has 44,100 integers per second. This means, in a one-minute long digital recording, the total number of integers in the time-order sequence is 2,646,000, which makes it a very large data item. The size of the data, in addition to the fact that it is not in a well-structured form with semantic meaning, makes this type of data unsuitable for most traditional data mining algorithms.

Timbre is rather subjective quality and not of much use for automatic sound timbre classification. To compensate, musical sounds must be carefully parameterized to allow automatic timbre recognition.

4. FEATURE DESCRIPTIONS

Based on recent research, MPEG published a group of features for digital audio content data. They are either in the frequency domain or in the time domain. For those features in the frequency domain, a STFT (Short Time Fourier Transform) with Hamming window has been applied to the sample data. From each frame a set of instantaneous values is generated. We extract the following timbre-related features from MPEG-7:

Spectrum Centroid - describes the center-of-gravity of a log-frequency power spectrum. It economically indicates the pre-dominant frequency range. We use *Log Power Spectrum Centroid*, and *Harmonic Spectrum Centroid*.

Spectrum Spread - is the Root of Mean Square value of the deviation of the Log frequency power spectrum with respect to the gravity center in a frame. Like Spectrum Centroid, it is an economic way to describe the shape of the power spectrum. We use *Log Power Spectrum Spread*, and *Harmonic Spectrum Spread*.

Harmonic Peaks - is a sequence of local peaks of harmonics of each frame. We use the *Top 5 harmonic peaks - Frequency*, and *Top 5 Harmonic Peaks - Amplitude*.

In addition, we use the *Fundamental Frequency* as a feature in this study.

5. METHODOLOGY

We start with a set of musical instrument recordings containing two instruments mixed. One instrument is predominant. We divide each recording into frames of 40 milliseconds. Each frame overlaps the previous frame by 1/3. Next, we extract the above described features on each frame (object). Then, we train and test a classifier. We use a single-label approach and assign the predominant instrument as the correct class.

Formally defined as: let $S = \{X, F \cup L\}$ be the training dataset, where X is the set of instances, $L = \{l_1, l_2, \dots, l_n\}$ is the set of all class labels, and $F = \{f_1, f_2, \dots, f_m\}$ is the m-dimensional feature vector used to build a classifier C, which estimates the target object t = $\{v_1, v_2, \dots, v_m\}$. Then, the estimation result would be $C(t) = \{d: d \in L\}$.

We use two classifiers - a bayesian neural network and a J45 decision tree, and compare the results. Neural networks process information with a large number of highly interconnected neurons working in parallel to solve a specific problem. Neural networks learn by example. Decision trees represent a supervised approach to classification. It is a simple structure where non-terminal nodes represent tests on one or more attributes and terminal nodes reflect decision outcomes.

One class label with the highest confidence is assigned to the estimation object and the other candidate classes are simply ignored. Figure 1. illustrates our polyphonic sound classification method.

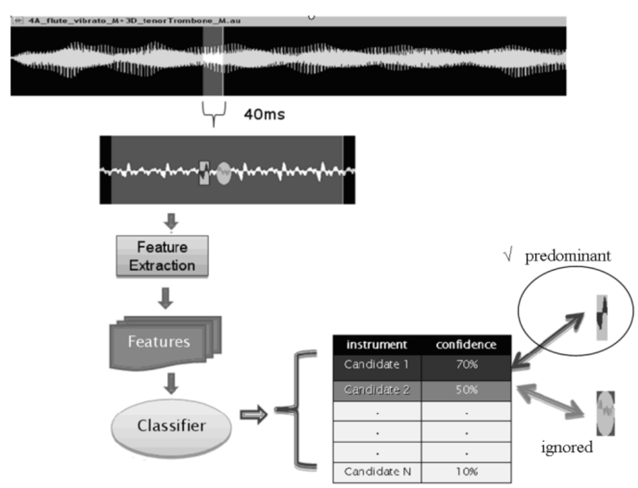


Figure 1. Music information retrieval with polyphonic sounds and timbre

6. EXPERIMENT

We have chosen 6 instruments: viola, cello, flute, english horn, piano, and clarinet to our experiments. All recordings originate from MUMS CD's by Opolko and Wapnick [9], which are used worldwide in similar tasks.

To create the analysis set a 0.2 second sample from 1.0 seconds to 1.2 seconds excerpt from the non-dominant instrument is mixed with the predominant instrument at every minute boundary (eg. 0.0s to 0.2s, 1.0s to 1.2s, 2.0s to 2.2s, etc.).

The mixed sound files used, with predominant instrument listed first, are: File1: Viola and Flute, File2: Flute and Englishhorn, File3: Viola and Englishhorn, File4: Cello and Piano, File5: Cello and Clarinet, File6: Clarinet and Piano.

Next, we split each recording into overlapping frames. Then, we extract the features described in the previous section 5. That produces a dataset with 1224 instances and 16 attributes.

We import the dataset into WEKA [5] data mining software for classification. We train two classifiers: Bayesian Neural Network and J45 Decision Tree. We test each one using 60% of samples for training and testing with remaining 40%, as well as testing using bootstrap. Results show Bayesian Neural Network has accuracy of 92.04% with 60/40 test and 94.28% via bootstrap testing. J45 decision tree has accuracy of 96.12% with 60/40 test and 98.93% via bootstrap testing. Summary results comparing the two classifiers are shown in Figure 2. The detailed results for Bayesian Neural Network and J45 Decision Tree are shown in Figure 3 and Figure 4 respectively.

	BayesNeuralNetwork	J48 DecisionTree
60/40 test	92.04 %	96.12 %
bootstrap test	94.28 %	98.93 %

Figure 2. Correctly classified instances % - classifier comparison

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
60/40 test	0.912	0.015	0.964	0.912	0.937	0.996	2C_cello_bowed
	0.99	0.031	0.89	0.99	0.937	0.999	3A#_bflatclarinet
	0.867	0.034	0.907	0.867	0.886	0.975	3C_viola_bowed
	0.936	0.026	0.912	0.936	0.924	0.989	4A#_flute_vibrato
bootstrap	0.937	0.012	0.973	0.937	0.955	0.998	2C_cello_bowed
	0.996	0.027	0.899	0.996	0.945	1	3A#_bflatclarinet
	0.904	0.02	0.944	0.904	0.924	0.988	3C_viola_bowed
	0.951	0.016	0.944	0.951	0.948	0.994	4A#_flute_vibrato

Figure 3. Bayesian neural network detailed accuracy by class

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
60/40 test	0.98	0.009	0.98	0.98	0.98	0.984	2C_cello_bowed
	0.98	0.008	0.97	0.98	0.975	0.983	3A#_bflatclarinet
	0.956	0.023	0.942	0.956	0.949	0.976	3C_viola_bowed
	0.927	0.013	0.953	0.927	0.94	0.963	4A#_flute_vibrato
bootstrap	0.992	0.004	0.992	0.992	0.992	0.998	2C_cello_bowed
	0.992	0.001	0.996	0.992	0.994	0.999	3A#_bflatclarinet
	0.982	0.004	0.988	0.982	0.985	0.998	3C_viola_bowed
	0.993	0.005	0.981	0.993	0.987	0.998	4A#_flute_vibrato

Figure 4. J45 decision tree detailed accuracy by class

7. CONCLUSIONS AND DIRECTIONS FOR THE FUTURE

We produce a music information retrieval system, which automatically classifies musical instruments. We use timbre related features. Our proposed methodology successfully classifies polyphonic music sounds (multiple distinct instruments). We report high classification accuracy. We find a decision tree classifier to be the most appropriate for this type of media. Our proposed method presents an improvement over previous timbre methods, which primarily focus on monophonic sounds.

This work contributes to solving the important problem of building music recommendation systems. Automatic recognition or classification of music sounds helps user to find favorite music objects within large online music repositories. It can also be applied to recommend musical media objects of user's liking. Directions for the future include automatic detection of emotions [4] contained in music files.

8. REFERENCES

[1] M.R. Boutell, J. Luo, X. Shen, and C.M. Brown. Learning multi-label scene classification, in Pattern Recognition, Vol. 37, No. 9, 2004, pp. 1757-1771.

[2] J.C. Brown. Musical instrument identification using pattern recognition with cepstral coefficients as features, Journal of Acousitcal society of America, 105:3, 1999, pp. 1933-1941

[3] M.A. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, M. Slaney. Content-Based Music Information Retrieval: Current Directions and Future Challenges. Proceedings of the IEEE, Vol. 96, Issue 4., 2008, pp. 668-696

[4] J. Grekow and Z.W. Ras. Detecting Emotion in Classical Music from MIDI Files, Foundations of Intelligent Systems, Proceedings of 18th International Symposium on Methodologies for Intelligent Systems (ISMIS'09), (Eds. J. Rauch et al), LNAI, Vol. 5722, Springer, Prague, Czech Republic, 2009, pp. 261-270.

[5] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten. The WEKA Data Mining Software: An

Update, SIGKDD Explorations, Vol. 11, Issue 1. New Zealand, 2009.

[6] W. Jiang, A. Cohen, and Z. W. Ras. Polyphonic music information retrieval based on multi-label cascade classification system, in Advances in Information and Intelligent Systems, Z.W. Ras, W. Ribarsky (Eds.), Studies in Computational Intelligence, Springer, Vol. 251, 2009, pp. 117-137.

[7] T. Joachims. Text categorisation with support vector machines: Learning with many relevant features, in Proceedings of Tenth European Conference on Machine Learning, 1998, pp. 137-142.

[8] K.D. Martin and Y.E. Kim. Musical instrument identification: A pattern recognition approach, in Proceedings of Meeting of the Acoustical Society of America, Norfolk, VA, 1998.

[9] F. Opolko and J. Wapnick. MUMS-McGillUniversityMasterSamples.CD's, 1987.

[10] R. Schapire and Y. Singer. BoosTexter: A boosting-based system for text ca-tegorization, in Machine Learning, Vol. 39, No. 2/3, 2000, pp. 135-168.

[11] A.A. Tzacheva and K.J. Bell. Music Information Retreival with Temporal Features and Timbre, in Proceedings of 6th International Conference on Active Media Technology (AMT 2010), Toronto, Canada, LNCS 6335, 2010, pp. 212-219.

[12] X. Zhang and Z.W. Ras. Sound isolation by harmonic peak partition for music instrument recognition, in the Special Issue on Knowledge Discovery, Fundamenta Informaticae Journal, IOS Press, Vol. 78, No. 4, 2007, pp. 613-628.

RHA(T)-System for Coding of Discrete Distributions and Their Alteration Processes

Tomas G. PETROV,
Faculty of Geology, Saint Petersburg University,
Saint Petersburg, Russia

and

Sergey V. MOSHKIN
Open Company "Infocomm" firm,
Saint Petersburg,
Russia

ABSTRACT

Brief description of ***RHA(T)*** information language – system designed for discrete distributions coding, analysis compositions of any nature, their ordering and description of their alternation processes is given. Here, ***R*** is rank formula of composition – a sequence of components by reduction of their value (contents); ***H*** is Shannon information entropy – complexity measure; ***A*** – anentropy – measure of purity; ***T*** – tolerance – measure of ultrapurity. The system was tested on objects, the components of which were chemical elements, molecules, minerals, population ages, nationalities etc. Lexicographic – after special alphabets – ordering of composition descriptions enables a linear hierarchic, periodic ordering of rank formulas. ***RHAT*** system in the chemical version of its utilization, being a System for Compositions representation, is a development of the Periodic System of Elements included by it as an alphabet. ***HAT*** characteristics are used as the most adequate ones for studying mixing and separation processes constantly taking place in nature and in engineering.

Keywords: coding, discrete distribution, entropy, anentropy, complexity, purity, mixing, separation.

1. INTRODUCTION

Natural, human, and social sciences have similar problems when describing discrete distributions (polycomponent compositions of objects), systematizing them and reflecting processes of their change. Among the problems absence of natural boundaries, indefiniteness and insignificance of names, instability of compositions in time and space. When nomenclature obsoletes, information on the objects which existed actually is lost along with the names. Large data arrays are poorly visible. Therefore, errors are undetected for a long time. Different forms of composition descriptions are used in different disciplines. When studying processes of compositions change, attention is usually focused on the behaviour of individual system components; this does not allow observing the overall trend in the system evolution. All this makes it difficult to implement new advances in the study of both the objects themselves and processes of their change, even in the adjacent areas of science. [1].

The system ***RHA(T)*** (***RHA*** in the initial version) [2–3]. removes a great number of difficulties in these areas, being a general-purpose method for conceptual coding of compositions of any nature, creation of unambiguously structured databases, as well as using a diagram for integral mapping of fundamental processes of composition changes; mixing and separation [4].

2. RANK FORMULA

Rank formula ***R***, a semi-quantitative characteristic of composition, is a sequence of event symbols, or elements in composition of an object, by decrease of their occurrence frequency (concentration) p_i. In case of contents equality, equal sign is put between them, and components are arranged according to the accepted alphabet. For chemical compositions, alphabet is represented by the sequence of elements in the Periodic System of Elements. Rank formula is written in a row. Rank formulas are situated in column under each other. ***R*** length denoted by ***n*** is taken for a ***detail measure*** of composition study and serves as ***R*** length standard when choosing types of events for further calculations. This is crucial, as it is often required to compare analyses of different lengths, with different component lists and with large differences in the content values. Thus, chemical analyses may include the content values differing by several of exponent. ***R*** is a name of sector in the composition diagram – simplex divided by median or hyperplanes median (Fig. 1).

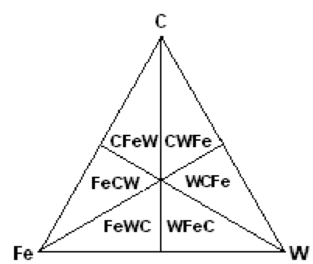


Fig.1. Rank formulas of carbon-tungsten-iron alloys as sector names of a triangle diagram partitioned by medians.

Ordering of ***R***`s assemblies is made using special alphabets. In this case, ***R*** is taken as a "word", where element (component) symbols are "letters". The Periodic System of Elements as the best fixed sequence of elements is taken for a chemical alphabet. Correspondingly, "H", "He", "Li" are the first three letters of this alphabet. For interval composition number axis is an alphabet. Vertically ordered alphabetic sequence of such "words" generates a linear hierarchic, periodic classification of composition rank formulas (Tab. 1).

Hierarchic character of classification is determined by two factors. The first one is that each rank formula with ***n*** detail belongs to a single rank formula with ***n***–1 detail. The second one is that the hierarchic character is shown in the rank formula itself. Thus, the first rank (***R***1) is occupied by the most widespread element in the given system, for example, oxygen O. Class of the second rank ***R***2 – OSi – silica-oxygenic enters into it. In its turn, ***R***3 class – OSiAl – alumino-silica-oxygenic enters into the latter class. System periodicity is shown in large enough assemblies in group arrangement of similar objects. In Tab.1, the system periodicity is exemplified by carbonatite arrangement.

From combinatorial analysis positions rank formulas are distribution from ***N*** alphabet symbols on ***n*** **.** Correspondingly, for example, for chemical compounds with ***n***=1, ***N*** equals 88. With ***n*** increasing, the number of rank formulas rises sharply and for ***n***=10 their number increases to $7.7 \cdot 10^{19}$. The actual number should be much smaller for cosmochemical and technological reasons. Generally, such a large number of possible ***R***`s also does not interfere with their use, as a similar number of words that can be generated combinatorially in natural languages does not interfere as well.

The text of ***R***-dictionary [5] was ordered according to the described above alphabet - in the collection of information about the theoretical chemical compositions of all known minerals.The collection included more than 4,500 records. ***R***-dictionary, as far as we know, is the first publication in which search characteristic is represented by strictly uniformly arranged data on the chemical composition of minerals. This method allows, after the available chemical formula (not only of minerals) is transformed into a rank formula, determination of the mineral name in 92% of cases. The rest 8% are polymorphs – minerals with the same chemical composition but different structure. Optical or X-ray methods are necessary for their determination. According to the R-dictionary materials, quantities of minerals with the given length of rank formulas were calculated (Fig. 2).

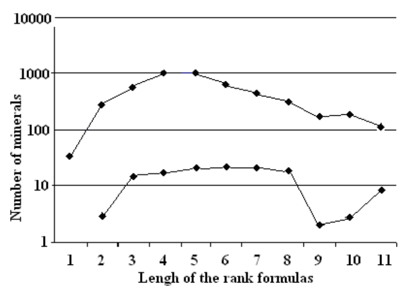


Fig.2 Distribution of minerals after the rank (chemical) formulas length

The upper curve - distribution of all minerals; the lower one, only "new" minerals discovered in recent years. When preserving the rules of new minerals distinguishing, curve shape should not change significantly.Therefore, we can, first, say that in conditions of accessible for studying part of the Earth's crust, maximum number of minerals forms as chemical compounds containing 4-5 chemical elements. We can surely say that statistics of minerals, for example moon ones, will be different. Second, a weak minimum on the upper curve and distinct one on the lower curve indicates that this phenomenon is clearly not a matter of nature. Apparently, this fact is an evidence of the initiated exhaustion of the minerals diversity for discovery of the new ones.

Rocks can be regarded as mixtures of minerals with different compositions. Problems of their unequivocal description and classification have long existed and have been solved separately for igneous, sedimentary, metasomatic, metamorphic rocks. Using sequences of mineral symbols after reduction of contents in rock as rank formulas, and R-dictionary as an alphabet, one can construct a common systematic of any rocks, for which mineral-molecular compositions exist (or can be estimated). Accordingly, chemical-molecular information on the technology, food, pharmaceutical products may be organized in such a way.

Rank formula is the first step for composition description. The second is introduction of quantitative characteristics.

3. QUANTITATIVE CHARACTERISTICS OF DISTRIBUTIONS

The ***1st quantitative*** characteristic of composition with ***n*** detail is C. Shannon ***information entropy*** as a composition ***complexity measure,*** $H = -\sum p_i \lg p_i$. Information entropy is an analogue of thermodynamic entropy of mixing [6[. Standardized to the interval 0÷1 entropy is calculated by formula $En = H/\lg n$. Isolines of standardized entropy in the three-component diagram are shown in Fig. 2.

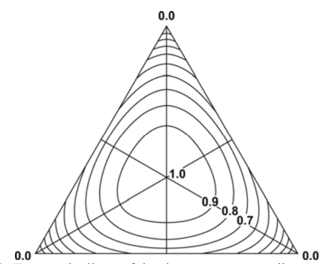


Fig. 2. Entropy isolines of the three-component diagram

When the system is split into two with composition change, entropy of at least one of the final products, as it was proved [7], is lower than the entropy of the original system. When two different compositions are mixed, entropy of the resulting composition is higher than that of at least one of the original ones. Conclusions of these theorems: during separation, complexity significantly decreases statistically; during mixing, composition complexity significantly increases statistically.

Entropy provides for a preferential consideration of large and medium in content components and is low sensitive to the small ones. Therefore, compositions differing strongly, especially in small components, can have the same ***H (En)*** value (Fig. 3). In addition, compositions of objects with the same ***R*** and ***En*** but relating to different types exist. In particular, this is the situation with some shales (rocks) and muscovites (mineral). This fact has led to the introduction of the second characteristics of ranked distribution of component compositions.

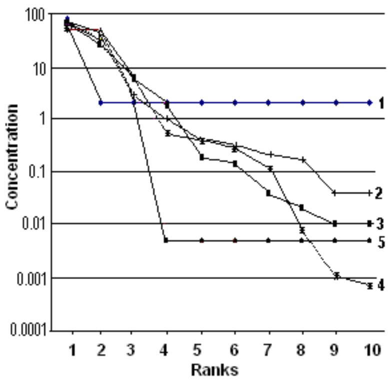


Fig. 3. Compositions variants at ***En*** = 0.376. 1 – theoretical composition – the first component 82%, the rest 2% each; 2 and 3 – rocks; 4 –- mineral of "simple" composition; 5 - theoretical composition - the first two components 47% each, third 5.96, and the rest 0.005% each.

For preferential estimation of ***small components smallness***, the ***purity measure*** of compositions, ***anentropy A*** is introduced [1–4]. In the simplest case, $A = -(\Sigma \lg p_i)/n - \lg n$. Anentropy values belong to the range 0 ÷ +∞. Minimum value $A = 0$ is obtained when $p_1 = p_2 = p_3 = \ldots = p_n = 1/n$. Maximum $A = +\infty$ is get if in the analysis at least one $p_i = 0$. For anentropy reduction to the interval 0÷1, anentropy is divided by ***A*** of the ***"analytically ideally pure system"***. Such composition consisting of ***n***–1 admixtures in contents equaling half sensitivity of a certain standard investigation method, and one component completing their sum to 1 is taken for such system. For chemical analyses of rocks and minerals, half sensitivity is taken for $p = 0.00005$. We get ***An*** independent of logarithm base. The name "anentropy" reflects its similarity to entropy and the opposite with respect to small components.

Fig. 4. shows a three-component diagram with plotted anentropy isolines. It is seen that ***An*** varies slightly at compositions commensurability and increases sharply when approaching the composition, in which one component dominates. In chemistry, it is an "elementary substance". Let us note that the most effective methods for improving purity of chemical elements leave in the matter dozens of elements – admixtures analyzed by modern methods.

Comparability of ***An*** values for analyses with different ***n*** is provided for by ***R*** length standardization. For identification of the overwhelming majority of geological objects and for solving of many genetic problems, ***n*** = 10 detail was accepted. This corresponds to consideration of 99.5-99.9..% of matter records in the system.

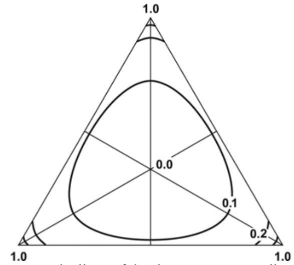


Fig. 4. Anentropy isolines of the three-component diagram

For describing compositions, in which taking into account of extra small component contents is necessary, one more characteristic – ***measure of ultrapurity*** – "***tolerance***" ***T*** was introduced [8]. It is calculated from the formula $\lg[(\Sigma 1/p_i)/n]$. In case of weak discrimination of compositions in ***EnAn*** diagram, one passes to ***EnT*** diagram.

All three characteristics ***at the given composition*** are independent of the phase state, temperature, pressure, and structure of an object; therefore, they allow reflecting processes of compositions change in systems of any nature. At the same time, values, especially ***An*** and ***T*** for the systems resulting from separation, the mechanism of which is generally selection, depend on the speed of this process. The higher its speed, the less perfect separation, the more difficult and less pure the composition becomes, the higher ***En*** values and the lower ***An*** values and ***T***, approaching the original composition.

Investments of individual concentration into ***H, A*** and ***T*** are shown in Fig.5.

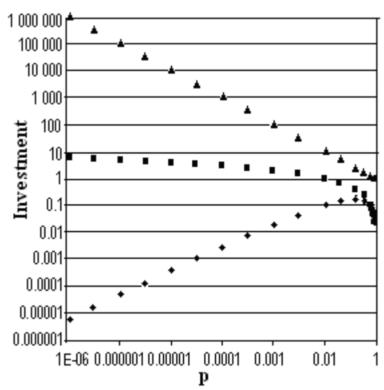


Fig. 5. Investments into functions: ◆ – ***En*** ($-p_i \cdot \lg p_i$), ■ – ***An*** ($-\lg p_i$); ▲ – ***T*** ($1/p_i$)

As one can see, maximum contributions to the integral characteristic is made by 1/***p***. That is, the use of tolerance will monitor the operational process when obtaining high purity matters. In other cases – to capture the effect of rare events on the overall behaviour of the system.

4. EnAn (EnT) DIAGRAMS

EnAn (or ***EnT***) diagrams are used:

1) for reflection of disordered datasets, features of point distinguished in this case, heterogeneities of the distribution density of points, the nature of arcs direction are revealed, manifestation of anomalous analyses;

2) for ordered ones, particularly, for processes reflection. ***En*** increase and ***An*** decrease correspond to ***mixing*** processes (mixture preparation: pharmaceutical, food etc.; exhaust gas emission into the atmosphere, river inflow into seas…). ***H*** decrease and ***A*** increase are peculiar to ***separation*** processes (any concentration, purification, crystallization etc.). Several typical trajectories of entropy characteristic changes under different processes of compositional changes are shown in Fig. 6.

Processes of minerals and rocks evolution shown here are crystallization processes, which are typical separation processes. They take place under the increment of content of the first components in ***R***`s and decrease of the last ones. Direction of trajectory of the national composition change in Turkmenistan is the same. Mixing processes are exemplified by the national composition change in Estonia and water dissolution of potassium nitrate.

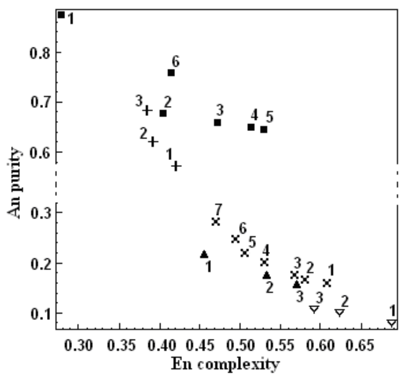


Fig. 6. Evolution of compositions, (calculations for 10 elements) ■ 1– water, 2,3,4,5 – water dissolubility of potassium nitrate at temperatures 0, 20, 40, 60^{0}C, 6 – potassium nitrate (KNO_3); + – mineral scheelite ($CaWO_4$) change in composition from the crystal center to the periphery [9], ✕ – magmatic rocks from gabbro to granite [10]; (calculations for 4 components): national composition [11]: ▽ –Turkmenistan (1959-1979), ▲ – Estonia (1959-1979).

Such diagrams enable comparison of the processes taking place in different systems. As we see, evolution of compositions of individual minerals in rocks and composition of rocks hosting them etc. may be traced.

5. RHAT IN GENERAL

RHAT assembly may be considered a "soft" system of coordinates specifying location of the composition ***area*** in the compositions diagram (in contrast to "hard" systems of coordinates describing ***point*** position).

From positions of semiotics ***RHAT*** it alphanumeric-sign symbol system of the description of compositions and their alteration **.**

From stand point of information theory this method is rank-entropy characterization of system composition on subset their dominant components

Entropy used in dozens of sciences, together with anentropy mathematically related with it, enabled construction of a diagram, which can be used to depict the processes of composition change without restrictions on the nature of changing systems.

The main advantage of using ***RHAT*** consists in simplicity and unity of the representation way and ordering the compositions of objects of any nature, which allows organizing quantitative analytical data as hierarchical, periodic systems. At the same time, systematizing of objects is performed regardless of whether they have names. The method was mainly developed on compositions of geological objects, where over twenty alternatives of its utilization are implemented [12]. Moreover, its potential is tested on age and national distributions of population, in linguistics, as well as on changes of personal income distributions, as a two-dimensional alternative to the one-dimensional Gini coefficient.

RHAT-method for chemistry is a development of the Periodic System of Elements. The method provided the possibility of a uniform ordering of all available (and pending) information on the chemical compositions of objects of any nature. Organization of the Hierarchical Periodic System of Compositions is correlated to the structure and information on the composition embedded into the Periodic System of Elements after 17 entries [13]. Main similarities and differences are shown.

For any chemical composition, there is a strictly defined place in the System of Compositions, as in a dictionary or encyclopedia. Properties of the Periodic System of Elements are inherited in the coding system: average atomic weights of compositions grow from the beginning to the end of the System of Composition.

RHAT system and PETROS was used for creation of a databank «Chemistry of Nature Objects». It includes ***chemical compositions*** of minerals, rocks, ores, waters, gases, meteorites, and other space bodies (46,000) records. Databank sampling is represented in Table.

Low ***En*** and high ***An*** values of the Sun, water, and minerals (pyrite) bear the evidence of strong differences of the first components in ***R*** from the last ones. High ***En*** and low ***An*** values correspond to a more uniform distribution of component contents (carbonatite, kimberlite, bone fishes).

In the Internet, there is an integrated catalogue of rocks and rock-forming minerals (4,300), minerals (micas – 1,600, pyroxenes – 900, tourmalines 800 etc.), as well as a catalogue of ***mineral compositions*** of alkalic rocks combined with the catalogue of chemical compositions (480 records).

Work on the method is provided by Petros 3 software package, the recent version of which allows using alphabets 100 in 100 symbols each [14].

Brief description of ***RHAT*** method, collections of compositions of different minerals and rocks as well as their codes in ***RHAT*** form are available at the web-site:

http://geology.spbu.ru/department/scientific/rha-language-method

ACKNOWLEDGMENT

The authors would like to thank N.I. Krasnova, I.S. Sedova, A.G. Bulakh, V.V. Gordienko and V.N. Dech for active assistance in the method development, as well as to thank all, whose active antagonism contributed to searching of the optimal substantiation and description.

Table. Data sampling from Databank «Chemistry of Nature Objects"

Rank formala										*En*	*An*	Object
H	He	O	C	Ne	N	Mg	Si=	Fe	S	0.138	0.837	Sun
H	O	C	N	Ca	P	K=	S	Na	Cl	0.428	0.434	human body
H	O	C	S	Na=	Si=	B	Fe	K	Ca	0.281	0.860	water H2S-CO3, Japan
H	Cl	S	B	Ca	Si	Na	Mg	K	F	0.412	0.425	fumarolic gas, Ebeko
N	O	Ar	H	C	Ne	He	S =	Kr	Xe	0.240	1.049	dry air
O	H	Mg	Si	Fe	Ca=	C	Al	Na	Ti	0.589	0.276	kimberlite, Ukraine
O	C=	Ca	Mg=	Si	Fe	Al	P	K	Na	0.579	0.172	carbonatite, Australia
O	Mg	Si	Fe	C	S	Ca	Cr=	Al	Mn	0.566	0.262	achondrite (meteorite)
O	Si	H =	Al=	K	Na=	Fe	Mg	Ca	C	0.478	0.234	granite, Murun Russia
O	Si	C =	Ca	H	Al=	K	Fe	Mn	Ti	0.609	0.159	carbonatite, Malavi
O	Si	Al	Na	H=	K	Ca	Fe	Mg	P	0.490	0.235	granite, Tuva Russia
O	Si	Al	Mg=	Ca	Na=	Fe	K	Ti	P	0.523	0.239	crust Earth bulk
O	Si	Al	Mg=	Ca	Fe	Na	Ti	S	K	0.558	0.217	Venus-14
O	Si	Al	Ca	Mg	Fe	Na	K	P	Ti	0.541	0.253	basalt, Moon
O	Si	Fe	H =	Al	K =	Ca=	Mg	C	Na	0.323	0.465	quarzite, Kazakhstan
O	Si	Fe	Mg=	Al=	Ca	Ti	Na	Mn	K	0.578	0.244	gabbro, Moon
O	Ca=	H	C	P	Na	Si=	Mg	S	K	0.646	0.220	bone fishes
O	Ca	C	H	Ba=	S	Fe	P	Si	F	0.630	0.142	carbonatite, Halutin Russia
O	Ca=	C	Si=	H	Mg	Al	F	Fe=	S	0.495	0.277	limestone, Ural Russia
S	Fe	As	Zn	Pb	Co	Ni	Se	Bi	Sb	0.283	0.911	pyrite, Sibaj Russia

REFERENCES

[1] T.G., Petrov, **"Information Language for Description of Multicomponent Object Compositions"**, Nauchn. Tekhn.Inform., Ser. 2, 2001, No. 3, pp. 8–18 (in Russian).

[2] T.G. Petrov, **"Ground of a Variant of the General Classification of Geochemical Systems",** Vestnik (Bulletin) of Leningrad State University. 1971, No. 18. Geol. Geogr. Iss 3. pp. 30-38 (in Russian).

[3] T.G. Petrov, O.I. Farafonova **Information-Component Analysis. RHA Method.** SPb. 2005. (in Russian**).**

[4] T.G. Petrov, **"Problem of separation and mixing in anorganic systems**", Vol.2. MSU. In 1995. pp.181-186. (in Russian) / / In.: Geology. Ed. V.T. Trofimov**.**

[5] T.G. Petrov, N.I. Krasnova**, R-Dictionary-catalogue of chemical compositions of minerals.** SPb "Nauka". (in Russian).

[6] E.A. Moelwyn-Hughes, **Physical Chemistry**, L-NY-P Perg. press, 1961.

[7] Yu.V. Shurubor, **"About One Property of Measure of Complication of Geochemical Systems"**, Dokl. Akad. Nauk SSSR, 1972, vol. 205, No. 2, pp. 453–456. (in Russian).

[8] T.G. Petrov, **"Range-Entropy Approach to the Description of the Geological Object Compositions and Their Changes"**, Obshch. i Prikladn. Cenologie., 2007, No. 5, pp. 27–33. (in Russian).

[9] E.M. Spiridonov, I.A., Baksheev, M.V., Seredkin, et al., **"Gumbeities of Ural and Conjugated Rock Mineralization"**, Geologiya Rudnykh Mestorozhdenii, 1998, vol. 40, No. 2, pp. 174—190. (in Russian)

[10] A.P. Sobolev, **Mezozoiskie granitoidy Severo-Vostoka SSSR i problemy ikh rudonosnosti** (Mesozoic Granitoids of Nort-East of USSR and Problems of Their Ore-Bearingness), Moscow: Nauka, 1989. (in Russian)

[11] **Demographic Encyclopedical Dictionary**, M. SE, 1985, (in Russian)

[12] T.G. Petrov, **"RHA method as a solution of systematization problem of analytical data on the material composition of geological objects"**, Otechestvennaya geologiya (National geology), 2008, No.4, pp. 98-105 (in Russian).

[13] T.G. Petrov **"Hierarchical periodical system of chemical compositions and its relation with the Periodical system of elements"**, Vestnik (Bulletin) of SPb State University, Ser. 7. 2009. No. 2, pp. 21-28 (in Russian).

[14] T.G. Petrov, S.V. Moshkin **"RHA method and its implementation in Petros-3 software system"**, Vychisleniya v geologii (Calculations in geology). 2011, No. 1. pp. 50-53 (in Russian).

Factors of Complexity for Managers of Projects of Information Systems

Ricardo YUGUE
School of Economics, Business and Accounting, University of Sao Paulo
Sao Paulo, SP 05508-010, Brazil

and

Antonio Cesar MAXIMIANO
School of Economics, Business and Accounting, University of Sao Paulo
Sao Paulo, SP 05508-010, Brazil

ABSTRACT

Even for the specific case of projects, the concept of complexity has no widely accepted definition. In the literature, this is associated to novelty, interdependencies, technology and to sets of variables as a whole. Some authors have an approach to complexity of projects based on subjective connotation.

This study aimed to identify factors related to the complexity of the projects in the area of information systems based on the perspective of their managers. It is classified as exploratory research and quantitative; and used a web-based survey with the participation of 140 Brazilian project managers of new information systems or of enhancements to existing systems.

The data analysis indicated that the complexity of projects managed by the participants can not be related to one factor or a specific group of factors, but possibly to different combinations of factors. Moreover, the scope of the project seems to be one of the most relevant contributors to the complexity of projects related to information systems.

Keywords: Project Complexity, Project Management and Factors of Complexity.

1. INTRODUCTION

The term "project" is used commonly with different meanings. Boutinet [1] refers to projects like figures ubiquitous in social life, defined by three core characteristics: uniqueness, materialization of the objective and identity. The PMI - Project Management Institute [2] adopts the concept that a project delivers products, services or results that are exclusive, non-repetitive or unique. Project is also temporary endeavor that starts with an idea and ends with the delivery of the materialization of this idea.

Archibald [3] states that the objective of a project, or the type of product that it generates, e.g. an information system, is one of the factors that defines how a project will be managed, differentiating it from others. However, even if different projects have similar goals, they will differ from each other depending on the context in which they are executed and on the degree of complexity it represents to their managers. According to Shenhar and Dvir [4], the context can be evaluated according to the complexity, uncertainty and the degree of familiarity with the organization's projects. Thus, among the characteristics of projects, complexity is one that has received increasing attention. Williams [5] affirms that, in general, there are two basic reasons for this phenomena: extensions or improvements of previous generations of the products (e.g. additional features and greater inter-relationships); and increasing projects restrictions in respect to deadlines and expectations about the delivery of the products. As a consequence, specifically related to projects of development of information systems, the research results of Xia and Lee [6] indicate that the complexity is associated with delays, cost overruns, restrictions of system functionalities and reduction on user satisfaction.

The concept of systems is used to define important components of projects complexity related to interconnections and interdependences of organizations and technologies [7], [8], [9], [10]. According to Shtub *et al* [11], complexity is at the root of the concept of projects, since the factors that determine the realization of projects include the complexity never seen before in the design, development and implementation of a new system. Remington *et al* [10] agree that complexity is important for the management of projects because of difficulties associated with decision-making and achievement of goals. However, they indicate a lack of operational definitions for the concept of complex projects. Geraldi [8], Adlbrecht and Geraldi [9] and Williams [5] believe that it is necessary to look at the project and its problems holistically, considering not only the specific components, but also its effects. In general, the concept of complexity on the projects has been explored based on different theories and perceptions, inclusively having different positions in relation to the concepts of difficulty, uncertainty and complexity. One must consider, however, that despite the established concept of complexity in other areas such as physics and biology, the findings of this study are restricted, in theory, to the knowledge area of management and, more specifically, of management of projects.

An important issue on studying the complexity of the projects is related to the factors that generates or influences it. Baccarini [7], Williams [5] and Fitsilis [12] proposed a classification based on two types of complexity: the organizational and technological, which are operationalized in terms of differentiation and interdependence. Tatikonda and Rosenthal

[13] refer to the concept of project complexity as relating to the novelty of the product, its development process and performance objectives; and its technological interdependence and difficulty.

Geraldi [8] and Geraldi & Adlbrecht [9] proposed an approach that considers that complexity management should consider not only the unique characteristics of complexity, but also its standard, and developed the concept of pattern of complexity. They defined three main types of complexity: Complexity of Faith (related to uncertainty), Complexity of Fact (referring to the amount of interdependent and concurrent information), and Complexity of Interaction (related to the interfaces between systems, people and places).

Based on the analysis and classification of the factors related to the complexity of the projects, Remington *et al* [10] defined two groups of categories operationally: dimension of complexity (characterizes the nature or origin of complexity) and factor of severity (in what extent is a problem). In an even broader way, Fitsilis [12], Baccarini [7], Geraldi & Adlbrecht [9] and Remington *et al* [10] believe in a subjective connotation, which implies in difficulty in understanding and deal with the projects' complexity.

Although there are many studies on the issue of complexity of projects, it was not possible to identify a model that could indicate which factors, in particular, are responsible for the complexity of certain types of projects, since they deal with the theme only conceptually, or are limited to very specific contexts. The Table 1 below shows the main factors of complexity found in the literature.

Table 1 – Factors related to the complexity of projects.

Factors of project complexity	References
- Uncertainty about the scope of the project	[7], [13], [14], [15].
- Uncertainty about the product of the project	[5], [10], [14], [15].
- Significant change in the scope of the project during its implementation	[5], [8], [9], [10], [12]; [15].
- High difficulty to achieve performance goals	[10].
- High number of stakeholders with influence on the project	[5], [10], [12].
- High interdependence between firms involved in the project.	[5], [7], [8]; [10].
- Novelty of the technology	[5], [7], [8], [10], [12], [13].
- High interdependence between the technologies	[5], [7], [8], [10], [12], [13].
- High multidisciplinarity	[7], [8], [9], [12], [16].
- Large number of different activities to be performed	[13].

2. METHOD

This research aimed to identify the factors that are related to the complexity of the projects, according to the managers of project of information systems. This research can be classified as non-experimental, conducted at a single time and having a quantitative approach [17]. It is predominantly exploratory [18].

This research had the participation of professionals working in Brazil as managers of information systems projects. There was no differentiation among the different industries. The sampling was non-random, for convenience, in conjunction with a snowball type sampling [19]. The contact with the project managers was made by using discussion groups on project management in social networks such as Yahoo!®, Google™ and LinkedIn®. The survey instrument was an electronic questionnaire applied through a website (QuestionPro™ - http://GP.questionpro.com) [20], [21]. Descriptive statistical techniques were used for the analysis of survey data. The Spearman correlation coefficient was employed to analyze the association between the level of complexity of projects managed and factors of complexity [17]. The quality of responses was verified by test-retest and calculating the Kappa index.

3. DATA ANALISYS

The application of electronic questionnaire obtained the participation of 313 professionals. Of this total, 140 professionals have managed projects of new information systems, or enhancements in existing systems. The profile of the participants is characterized by the experience in project management, since more than 90% of the 140 respondents have more than 2 years of experience in project management and more than 60% have more than 6 years of experience. In addition, more than 37% have some professional certification in project management.

The projects managed by the respondents in the last 12 months related to new information systems or enhancements in existing systems were distributed, in percentage terms, by the degree of complexity perceived, i.e., without complexity, low complexity, medium complexity, high complexity and very high complexity. Table 2 below shows average distribution of the projects managed, in percentage, among the different levels of complexity.

Table 2 - Descriptive statistics: distribution of the projects by level of complexity

	Level of complexity of the projects				
Statistics	Without	Low	Medium	High	Very high
N	140	140	140	140	140
Mean	4.5%	16.4%	31.6%	30.4%	17.2%
Median	0.0%	15.0%	30.0%	25.0%	10.0%
Mode	0%	0%	30%	20%	0%
Std. Deviation	8.83%	15.79%	19.85%	22.11%	21.49%
Minimum	0%	0%	0%	0%	0%
Maximum	60%	80%	100%	100%	100%

The data analysis presented in Table 2 shows the following:

- On average, 17.2% of the projects managed were classified as of "very high complexity." If added to the percentage of

projects rated as of "high complexity", they will represent together almost half of the sample, i.e. 47.6%.
- Relatively small portion of the projects managed (20.7% on average) were considered by respondents as of little or without complexity.

In order to analyze the respondent's understanding about the concept of complexity of projects, it was applied a question about situations that occur on a project and that are related to its complexity. These conditions were defined based on the factors related to the complexity that were found in the literature and by assigning values to them, in order to characterize how they contribute to the complexity of the projects in practice. In order to assess the contribution of each of the proposed situations to the complexity of the projects, it was used a Lykert type scale of five points, from "not contribute" to "contributes fully". The results of data analysis by using descriptive statistics are presented in the Figure 1 and the Table 3 below:

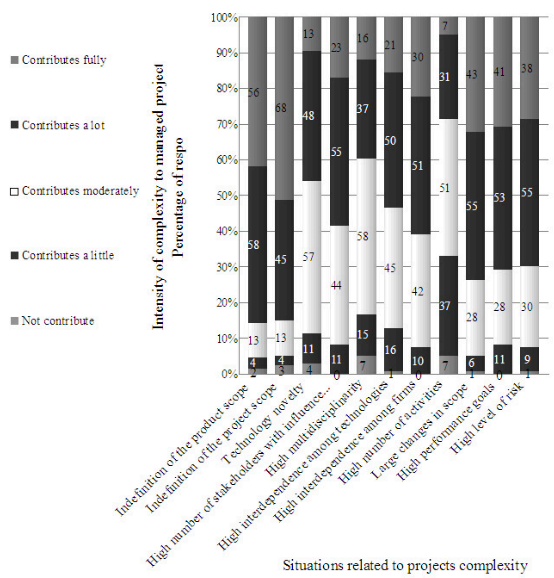


Figure 1 - Contribution of situations of projects to the complexity

Table 3 - Descriptive statistics: the contribution of situations of projects to the complexity

		Indefinition of the product scope	Indefinition of the project scope	Technology novelty	High number of stakeholders with influence on decisions	High multidisciplinarity	High interdependence among technologies	High interdependence among firms	High number of activities	Large changes in scope	High performance goals	High level of risk
N		140	140	140	140	140	140	140	140	140	140	140
Mode		4	5	3	4	3	4	4	3	4	4	4
Percen-tiles	25	4	4	3	3	3	3	3	2	3	3	3
	50	4	5	3	4	3	4	4	3	4	4	4
	75	5	5	4	4	4	4	4	4	5	5	5

Within the situations presented, which was considered with the greatest contribution to the complexity of projects of new systems, or enhancements to existing systems, was the vagueness or uncertainty of the project scope (mode = 5, median = 5). Authors such as Tatikonda and Rosenthal [13], Maximiano [14], and Turner and Cochrane [15] also indicate the uncertainty as important component of project complexity. Following, in descending order of degree of contribution, appears the uncertainty regarding the scope of the product of the project (mode = 4, median = 4). This factor is mentioned by Williams [5], Remington *et al* [10], Maximiano [14] and Turner and Cochrane [15] as a component of uncertainty of projects.

With a contribution a little smaller (mode = 4, median = 4) shows up the changes in the scope of the project during its implementation, the difficulty to achieve the goals of the project and the high level of risks, followed by the large number of stakeholders with influence on the project, the high interdependence among technologies and the high interdependence among the firms participating in the project. The change in project scope is also indicated as an important factor that contributes to the complexity of projects by Williams [5], Geraldi [8], Adlbrecht and Geraldi [9], Remington *et al* [10], Fitsilis [12] and Turner and Cochrane [15]. However, the performance goals related to the projects results, specifically in relation to meeting deadlines, cost, scope and quality, are cited only by a few studies. One of the authors is Remington *et al* [10] who includes high-level goals among the key themes of the complexity of the projects.

Regarding the high interdependence among the firms involved in the project, its relatively lower association with the complexity of projects was not expected, considering the importance given to it, most notably by Williams [5], Baccarini [7], Geraldi [8] and Remington *et al* [10]. This perception of the respondents may be related to the type of project they manage normally, which could be essentially internal, without considerable participation of other companies. The type of project managed by the respondents could have influenciated the answers and thus could be the reason for the relatively low importance given to the complexity of situations regarding the technologies (novelty and high interdependence between the technologies involved in the projects) if considering the high importance given by Williams [5], Baccarini [7], Geraldi [8], Remington *et al* [10], Fitsilis [12] and Tatikonda and Rosenthal [13] to this factor.

The high multidisciplinarity of the participants of the projects and the large number of different activities to be performed on the project are among the situations considered of the lowest degree of influence on the projects complexity. Contrary to the results of this research, the multidisciplinarity of the project participants is a factor of complexity highlighted by authors such as: Baccarini [7], Geraldi [8], Adlbrecht and Geraldi [9], Fitsilis [12] and Maximiano [16].

The number of project activities is one of the indicators used to evaluate the project size, especially in the case of new systems that can be sized based on the number of lines of programming. This low contribution to the complexity of projects is in accordance to the literature, since few authors consider it in the list of the most important factors.

The survey aimed to identify which factors were more complex in the projects managed by the respondents during the last 12 months of work. For this analysis it was used a Likert type scale of 5 points, from "without complexity" to "very high complexity". The results of the analysis of the responses received are presented descriptively in the following Figure 2 and Table 4.

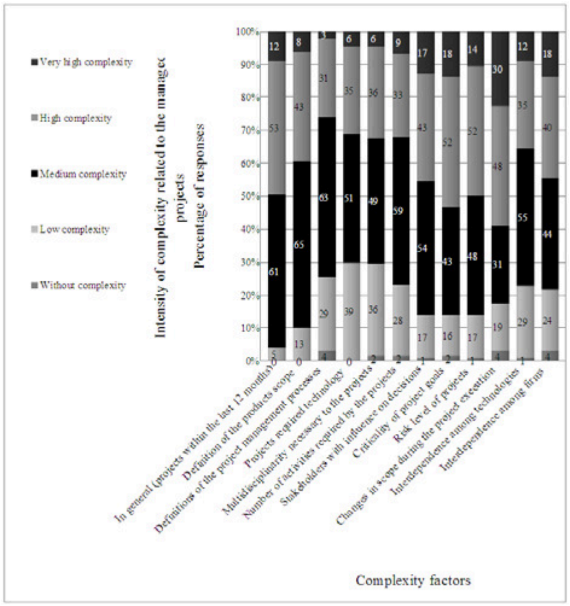


Figure 2 – Complexity factors and their intensity in the projects managed

Table 4 - Descriptive statistics: complexity factors – intensity in the projects managed

		In general (projects within the last 12 months)	Definition of the products scope	Definitions of the project management processes	Projects required technology	Multidisciplinarity necessary to the projects	Number of activities required by the projects	Stakeholders with influence on decisions	Criticality of project goals	Risk level of projects	Changes in scope during the project execution	Interdependence among technologies	Interdependence among firms
N		140	140	140	140	140	140	140	140	140	140	140	140
Mode		3	3	3	3	3	3	3	4	4	4	3	3
Percentiles	25	3	3	2	2	2	3	3	3	3	3	3	3
	50	3	3	3	3	3	3	3	4	3	4	3	3
	75	4	4	4	4	4	4	4	4	4	4	4	4

The criticality of the goals and changes in project scope (mode = 4, median = 4) were considered of the highest complexity in the projects managed by the research participants. More than 50% of projects were considered of high or very high complexity regarding these factors. It is plausible to consider that the criticality of the goals actually exerts great pressure on project managers if they are too restrictive to their performance, since they can generate uncertainty and risk. Scope changes during the execution of a project always result in redefinitions, replanning and insecurity with respect to the development of the project itself [2]. These two factors have some feedback relationship, since changes in project scope may result in pressure on the project goals. On the other hand, when the scope is changed there will be replaning and rework, which, in general, leads to increased costs and time, thus resulting in further pressure on the goals. Somehow, this relationship is related to the factor with the third highest intensity in complexity, which is the level of risk (mode = 4, median = 3).

On one level a little below appear the interdependence among firms and the influence of stakeholders. These factors seem to indicate that in practice the complexity is much related to factors that are out of the action of the project manager, in the sense that it depends on others, although there are also ways to influence them, as indicated by Cleland [22] in relation to stakeholders.

These last two groups of factors, except for the number of necessary activities, are opposite in relation to the higher importance given to these factors by other authors such as: Williams [5], Baccarini [7], Geraldi [8], Geraldi and Adlbrecht [9], Remington *et al* [10], Fitsilis [12], Tatikonda and Rosenthal [13] and Maximiano [16].

Figure 3 - Comparison among the situations and factors contributing to complexity.

Situation that contributes to the complexity of the projects (conceptual)	Contri-bution to com-plexity	Factors of complexity in the projects managed
- Indefinition about the project scope	HIGH	
- Indefinition about the scope of the product		
- Significant changes in the scope - Difficult to achieve the performance goals - High level of risks		
- High interdependence among the firms - High number of stakeholders with influence - High interdependence among the technologies		- Changes in the scope during the project - Criticality of the goals - Level of risks
- Novelty of the technologies - High multidisciplinarity		- Stakeholders with influence on the project - Interdependence among the firms
		- Interdependence among the technologies - Definition of the scope of the products of the project - Number of activities
- Large number of activities to be undertaken by the project	Low	- Multidisciplinarity - Technologies needed to develop the project - Definition of project management processes

Although these two groups of data did not use variables with identical descriptions, however there are correspondences between them. The differences between what is perceived conceptually and in practice in relation to the complexity of the projects are presented in the Figure 3 above, that aligns the degrees of contribution to the complexity of projects, comparatively higher or lower as the values of the modes and percentages presented in the Tables 3 and 4, and in the Figures 1 and 2. This comparison considers that, in general, the contribution to the complexity of projects is reached more frequently in higher scores. Thus, Figure 3 shows the factors with equivalent scores on the same lines.

The evaluation of the complexity of the managed projects have the average of the relative frequencies, in general, smaller than the average achieved by the relative frequency of the perception in conceptual terms, at the highest points on the ordinal scale. This difference was expected because, in the case of conceptual evaluation, the situations that were presented aimed to represent more clearly the occurrence of complexity source. Another reason is related to the level of complexity of projects managed, since not all of them were considered highly complex. However, this difference does not preclude making some comments regarding the order they appear in both analyses.

An important difference between the two scales is at the definition of the project scope and the products scope that are in the top of the list of situations that lead to complexity of projects in conceptual terms. In this research, the definition of project management processes is considered to be equivalent to the scope of the project, since it is defined as the work that needs to be done to deliver a product, service or result with the specified features and functions [2]. The low frequency of the contribution of the definition of management processes to the complexity of the projects managed indicates that, in practice and in general, there is weak relationship between the complexity of the projects and the project management processes, as indicated by Yugue [23].

Table 5 – Relation between the factors of complexity and the distribution of projects by degree of complexity.

	Distribution of the projects managed by degree of complexity					
Complexity factors	**Without**	**Low**	**Medium**	**High**	**Very high**	**High + Very high**
In general (for projects managed).	-0.312	-0.521	-0.467	0.258	0.540	0.681
Definition of the scope of the products..	-0.234	-0.369	-0.269	0.213	0.423	0.472
Definition of project management processes.	-0.057	-0.085	-0.020	0.046	0.127	0.120
Technologies required to develop the project.	0.116	-0.048	-0.047	0.023	0.071	0.036
Multidisciplinarity	-0.074	-0.176	-0.182	0.053	0.238	0.235
Number of activities necessary for the project.	-0.129	-0.179	-0.140	0.036	0.351	0.248
Stakeholders	-0.234	-0.184	-0.164	0.097	0.218	0.253
Criticality of the goals.	-0.152	-0.232	-0.165	0.130	0.245	0.290
Level of risks.	-0.079	-0.169*	-0.193	0.016	0.321	0.239
Changes in scope during the project execution	-0.066	-0.222	-0.234	0.098	0.264	0.273
Interdependence /of technologies.	0.030	-0.150	-0.239	-0.026	0.339	0.264
Interdependence among the firms.	-0.092	-0.135	-0.201	0.042	0.278	0.262

The correlation analysis (correlation coefficient Speaman - level of 0.05) between the degree of complexity of the factors in the projects managed and the distribution of projects (in percentage terms) among the different levels of complexity enabled the construction of Table 5.

An expectation of this study was that with a few and better defined variables, and for a specific type of projects, it would be possible to characterize the complexity of the projects. However, the data analysis found correlation between the percentage of projects managed with high or very high complexity and the complexity of projects in general (r = 0681), which seems to indicate that, probably, the complexity of the projects considered in this research is not due to a specific complexity factor, but rather a set of factors. This conclusion is in line with other authors such as: Williams [5], Geraldi [8], Geraldi and Adlbrecht [9] and Remington *et al* [10], who believe that one must look at the complexity of the projects holistically.

Another factor that presented positive correlation with respect to the high complexity of the projects was the definition of the scope of the products of the projects (r = 0.540). This datum indicates that the definition of scope is one of the biggest problems faced by project managers. The reason may lie in the concentration of many features of complexity in the planning phase of the project scope as it covers the definition of project objectives, identification of risks, the assumptions, constraints and goals to be achieved [2]. Also are considered at this stage: the necessary technologies, the interdependencies and the goals already set, which are characteristics related to complexity of the project too.

4. CONCLUSIONS

This research found that the projects performance goals are considered by respondents as of relatively high contribution to the complexity of the projects, but are rarely indicated by other authors. However, one must consider that this factor may be treated a factor of difficulty and not a factor of complexity. Anyway, this result indicates the need for greater attention from organizations for this factor. This research also indicated that changes in the scope of the project are also important factors of complexity, both conceptually and in relation to projects managed in practice.

It was hoped that, in general, the analysis of the factors of complexity of projects and the distribution of the projects managed by the level of complexity would indicate a more precise characterization of the complexity of the projects. However, the correlation analysis found only two relationships. The main of them points to a relation to the factors of complexity as a whole. Based on that, it is plausible to consider that this data is indicative that the complexity of the projects actually not due to a specific factor, but the set of factors, reinforcing the idea of the holistic approach.

The second relationship was found between the percentage of projects managed with high or very high complexity and the complexity of projects related to the definition of the scope of the products of the projects. In this case, it seems reasonable to assume that this relationship is indicative that yet during the definition of project scope it would be possible to perceive its level of complexity, since, in its process, are considered or analyzed most of the characteristics of complexity of a project.

Considering that this research is exploratory, not random, and thus, not possible to be generalized, the results of this research can be contributive to the formulation of better constructs suited to the study of the complexity of projects and their management. The conclusions, however, ask for further research on the criticality of the goals as a component of difficulty or complexity of projects and about the ways to approach the planning of the project scope in order to reduce its complexity and to facilitate its management.

5. REFERENCES

[1] J. P. Boutinet. **Psychologie des conduites à projet**. Paris: PUF, 1993.

[2] Project Management Institute – PMI. **PMOBOK. A guide to the project management body of knowledge**. 4th ed. Newton Square, PA: PMI, (PMBOK Guide), 2008.

[3] R. D. Archibald. **Managing high-technology programs and projects**. 3rd ed. New York: John Wily & Sons, 2003.

[4] A. J. Shenhar and D. Dvir. **Reinventing project management: the diamond an approach to successful growth and innovation**. Boston: Harvard Business School Press, 2007.

[5] T. M. Williams. "The need for new paradigms for complex projects". **International Journal of Project Management**. 17(5), 1999, pp. 269-273.

[6] W. Xia and G. Lee. "Grasping the complexity of is development projects". **Communications of the ACM**., Vol. 47, No. 5, May 2004, pp. 69-74.

[7] D. Baccarini. "The concept of project complexity: a review". **International Journal of Project Management**. 14(6), 1996, pp. 201-204.

[8] J. Geraldi. "Patterns of complexity: the thermometer of complexity". **Project Perspectives**. v. XXIX, pp. 4-9, 2007.

[9] J. Geraldi and G. Adlbrecht. "On faith, fact, and interaction in projects'. **Project Management Journal**. 38(1), 2007, pp. 32-43.

[10] K. Remington; R. Zolin and R. Turner. "A model of project complexity: distinguishing dimensions of complexity from severity". In: **International Research Network of Project Management Conference**. Berlin, 2009.

[11] A. Shtub; J. F. Bard and S. Globerson. **Project management: engineering, technology, and implementation**. Englewood Cliffs, NJ: Prentice Hall, 1994.

[12] P. Fitsilis. "Measuring the complexity of software projects". WRI World Congress on Computer Science and Information Engineering. **Computer Science and Information Engineering**. v. 7, 2009, pp. 644-648.

[13] M. V. Tatikonda and S. R. Rosenthal. "Technology novelty, project complexity, and product development project execution success: a deeper look at task uncertainty in product innovation". **IEEE Transactions on Engineering Management**. 47(1), 2000, pp. 74-87.

[14] A. C. A. Maximiano. "Apresentando os projetos". **001 Cap 01 RH Projetos – Projeto**. 2010. (Mimeo)

[15] J. R. Turner and R. A. Cochrane. 'Goals and methods matrix: coping with projects with ill-defined goals and/or methods of achieving them". **International Journal of Project Management**, 11(2), 1993, pp. 93–102.

[16] A. C. A. Maximiano. **Administração de projetos: como transformar idéias em resultados**. 2. ed. São Paulo: Atlas, 2002.

[17] R. H. Sampieri, C. F. Collado and P. B. Lucio. **Metodologia de pesquisa**. 3rd. ed. São Paulo: McGraw-Hill. 2006.

[18] C. Selltiz; M. Jahoda; M. Deutsch and S. W. Cook. **Métodos de pesquisa nas relações sociais**. São Paulo: EDUSP, 1974.

[19] L. P. L. Fávero; P. Belfiore; F. L. Silva and B. L. Chan. **Análise de dados: modelagem multivariada para tomada de decisões**. Rio de Janeiro: Elsevier, 2009.

[20] L. Vasconcellos and L. F. A. Guedes. "E-Surveys: vantagens e limitações dos questionários eletrônicos via internet no contexto da pesquisa científica". In: **X SemeAd – Seminário em Administração FEA/USP**. São Paulo, 2007.

[21] H. Kalantari D.; E. Kalantari D. and S. Maleki. "E-survey (surveys based on e-mail & web)". **Procedia Computer Science**, v. 3, pp. 935-941, 2011.

[22] D. I. Cleland. **Project management: strategic design and implementation**. 3rd ed. Singapura: McGraw Hill, 1999.

[23] R. T. Yugue. **Contribuição ao estudo dos processos de gerenciamento e da complexidade dos projetos**. Dissertation for master degree (Administration). Program of Graduation in Administration. São Paulo: School of Economics, Business and Accounting, University of Sao Paulo, 2011.

Extensions of One-Dimensional Gray-level Nonlinear Image Processing Filters to Three-Dimensional Color Space

Orlando HERNANDEZ and Richard KNOWLES
Department Electrical and Computer Engineering, The College of New Jersey
Ewing, New Jersey 08628-0718, USA

ABSTRACT

This paper describes the development of computer simulation models to evaluate the filter algorithms in different color spaces. The objective of this project is to study the extension of filters from a one-dimensional gray-level space to a three-dimensional color space. The principal task of this project was to conduct and evaluate mathematical extensions of filters and determine which combination of filters and color spaces presents the best image quality. Results have shown that the use of the Rank Order filter with the YIQ color spaces provides the highest signal-to-noise-ratio. Therefore, the use of Rank Order filters in YIQ color spaces would be the most effective in the noise smoothing of images.

Keywords: color image smoothing, nonlinear imaging filters, color image processing

1. BACKGROUND

Color is that attribute of light-energy which is related to the wavelength. It is well known that color carries a very important part of information regarding objects of interest in an image [1]. The attention to the representation of color (color-models) and color-based segmentation of images has increased over the recent past few years. Applications include (i) wound-image processing, (ii) TB-screening, and (iii) Segmentation of Fascia.

Smoothing processing is one of the common methods in image enhancement that clears up the image noise, and its theories are all aimed at monochrome image. Using color space transformation of color image can not only gain more useful information, but also use the processing methods for a monochrome image. Using color space transformation based on L-H-S, the experiment for color image smoothing processing is made. The experimental results have shown that the method can availably cleans up the noise infection of color image and improve the imaging quality [2].

Other techniques have explored the theory of vector bundles over Riemannian manifolds in order to smooth multi-valued images [3]. And these frameworks have considered standard PDE's used in image processing as generalized heat equations, related to the geometries of the base manifold, given by its metric and the subsequent Levi-Cevita connection and of the vector bundle, given by a connection. As a consequence, the smoothing is made through a convolution with a 2D kernel, generalizing Gaussian, Beltrami and oriented kernel. In particular, extensions of the oriented kernel have been constructed, and illustrated with applications to color image smoothing.

Additionally, simplification and merging techniques based on an initial region partition of color images: a region adjacency graph (RAG) have been proposed; as is a general scheme that can be employed for both color image simplification and/or segmentation [4]. A new filtering algorithm of RAG was presented and included within a merging algorithm. The region models associated to each node are simplified and moreover the RAG is simplified by merging similar nodes.

In this work, the development of computer simulation models to evaluate the filter algorithms in different color spaces was undertaken. The objective of this project was to study the extension of filters from a one-dimensional gray-level space to a three-dimensional color space. The principal task was to conduct and evaluate mathematical extensions of filters and determine which combination of filters and color spaces presents the best image quality. Results have shown, and discussion will support, that the use of the Rank Order filter with the YIQ color spaces provides the highest signal-to-noise-ratio. Therefore, the use of Rank Order filters in YIQ color spaces would be the most effective in the smoothing of images.

2. INTRODUCTION

In this project, the development of computer simulation models to evaluate the filter algorithms in different color spaces was executed. The objective was to study the extension of filters from a one-dimensional gray-level space to a three-dimensional color space. The principal task was to conduct and evaluate mathematical extensions of filters and determine which combination of filters and color spaces yields the best image quality results.

The filters that were investigated are listed below:

- Weighted order statistics
- Stack
- Spatial-rank order selection
- Nonlinear mean
- Teager
- Polynomial
- Rational

The three-dimensional color spaces that were evaluated are listed below:

- RGB – Red Green Blue
- HSV (HSB) – Hue Saturation Brightness
- XYZ (CIE) – International Commission on Illumination (French)
- YIQ – Luma In-Phase Quadrature

The signal to noise ratio is used to compare and contrast the performance of each combination of filters with color spaces. Every possible combination of filters and color spaces were tried along with matrix multiplications (dot product and cross

product). The performance of each combination was contrasted with the gray-level counterparts.
All the filters used in this investigation are nonlinear filters. A nonlinear filter is an operation where each filtered pixel y(m,n) is a nonlinear function of x(m,n) and its neighbors. In this investigation, a window of 3 x 3 is used to determine the neighbors of each pixel.

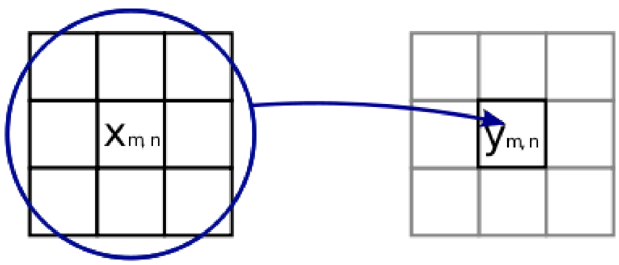


Figure 1: 3 x 3 Window (neighboring pixels)

The equation used for the output of the Weighted Order Statistic (WOS) filter is:

```
Y(m,n)=
median{
   (w1♦x(m-1, n-1), (w2♦x(m, n-1),
   (w3♦x(m+1, n-1), (w4♦x(m-1, n),
   (w5♦x(m, n), (w6♦x(m+1, n),
   (w7♦x(m-1, n+1), (w8♦x(m, n+1),
   (w9♦x(m-1, n+1)
}                                        (1.0)
```

In terms of computing the output pixel value, the median of all the pixels in the window is selected where certain pixel values are weighted or repeated. The pixels at the four corners have a weight of 1. The center pixel has a weight of 3 while the others have a weight of 2. The median value is then selected as the output. The stack filter is the same as the WOS filter except that none of the pixel values are weighted. Therefore, it takes the median value of the 9 pixels.

The equation used for the output of the Nonlinear mean filter is:

```
Y(m,n) = {x(m-1, n-1) + x(m, n-1) +
  x(m+1, n-1) + x(m-1, n) + x(m, n) +
  x(m+1, n) + x(m-1, n+1) + x(m, n+1) +
  x(m-1, n+1)}/9                         (2.0)
```

The output is equal to the sum of its neighboring pixels divided by the total number of neighbors. Because the window is a 3x3 window, the number of neighbors is 9. Therefore, the output is the mean of the pixel values within the window.

The equation used for the Teager filter is:

```
Y(m,n) = {3x^2(m,n) - ½[x(m+1,n+1) *
  x(m-1,n-1)] - ½[x(m+1,n-1) *
  x(m-1,n+1)] - x(m+1,n) * x(m-1,n) -
  x(m, n+1) * x(m, n-1)}                 (3.0)
```

In this filter, the center pixel is squared and multiplied by three. It is then subtracted by four values: (1) half of upper left corner pixel value multiplied by the lower right corner pixel value; (2) half of the upper right pixel value multiplied by the lower left pixel value; (3) the middle pixel value in the top row multiplied by the middle pixel value in the last row; and (4) the middle pixel value in the first column multiplied by the middle pixel value in the last column.

The equation used for the polynomial filter is:

```
y(m,n) = x(m,n)+λx(m,n)                  (4.0)
y(m,n) = [x(m-1,n)-x(m+1,n)]^2 *
  [2x(m,n)-x(m-1,n)-x(m+1,n)] +
  [x(m,n-1)-x(m,n+1)] ^2 *
  [2x(m,n)-x(m,n-1)-x(m,n+1)]            (4.1)
λ = intensity of enhancement             (4.2)
```

This filter is a polynomial function which uses multiplication and powers to remove the noise in an image. This type of filter is more of an edge detection filter whereas; more emphasis is put on the edges in the image.

The equations used for the output of a rational filter are:

```
y(m,n) = x(m,n) + λ[vn1(m,n) * cn1(m,n) -
  vn2(m,n) - cn2(m,n)]                   (5.0)
vn1(m,n) = 2x(m,n) - x(m,n-1) - x(m,n+1) (5.1)
vn2(m,n) = 2x(m,n) - x(m-1,n) - x(m+1,n) (5.2)
cn1(m,n) = gn1(m,n) / [k gn1^2(m,n) + h] (5.3)
cn2(m,n) = gn2(m,n) / [k gn2^2(m,n) + h] (5.4)
gn1(m,n) = [x(m,n+1) - x(m,n-1)]^2       (5.5)
gn1(m,n) = [x(m+1,n) - x(m-1,n)]^2       (5.6)
k = 1/(2 g0)                             (5.7)
h = g0 / 2                               (5.8)
```

This filter embodies a complex algorithm. There are several equations that are required in order to implement this filter. The value of `g0`, the position of resonance was set to 1.

These filters are applied to images of different color spaces such as YIQ, RGB, XYZ, and HSV. A color space is an abstract mathematical model which describes the way that the colors in the image can be represented as tuples of numbers, usually 3 color components. In this investigation, only the color spaces with 3 components are used. There are color spaces with four color components such as CMYK. RGB is an additive color model where red, green, and blue light are added together to reproduce a broad array of colors. RGB is the most commonly used color space in technology. Most computer monitors and televisions use RGB. YIQ is the color space used by the NTSC color TV system. The Y stands for luma, I for in-phase, and the Q stands for quadrature. There are very few televisions today that perform true YIQ decoding. The XYZ color space was created by the International Commission on Illumination (CIE). This color space is very similar to RGB in that its tri-stimulus values X,Y,Z are roughly red, green, and blue but are not observed as red, green, or blue. Instead, they are thought of as derived parameters from the red, green, and blue colors. The HSV color space stands for Hue – Saturation – Value. It is also known as HSB where the B stands for brightness. HSV is one of the most common cylindrical coordinate representations of points in an RGB color model where the geometry of RGB is rearranged in an attempt to make it more perceptually relevant than the Cartesian representation. This color model is common in computer vision applications.
As one can see, the use of different filters with different color spaces and types of mathematical computations (magnitude, cross product, dot product) can improve or decrease the quality

of an image and remove some noise. In this project, the principal task is to find the combination that produces an image with the best quality. In order to do so, one must use the signal to noise ratio. The signal to noise ratio quantifies how much a signal has been corrupted by noise. It is defined as the ratio of signal power to the noise power corrupting the signal. In this project, the gray-level image that is filtered is compared to the filtered color to grey level image using the signal to noise ratio. The alternate definition of signal to noise ratio was used which is the ratio of mean to standard deviation of a signal. The equation for SNR in dB is:

$$SNR_{dB} = 20\log\frac{\mu}{\sigma} \tag{6.0}$$

Therefore, the numerator inside the log function is the mean of the signal, meaning the mean of the grey-level filtered pixel values. The denominator inside the log function is the standard deviation of the filtered color to gray image pixel values subtracted by the filtered gray-level image pixel values. The result will help determine the quality, which will help conclude which combination of filters and color spaces produces the best image in terms of quality and reduction in noise.

3. PROCEDURE

The model is shown in the diagram below:

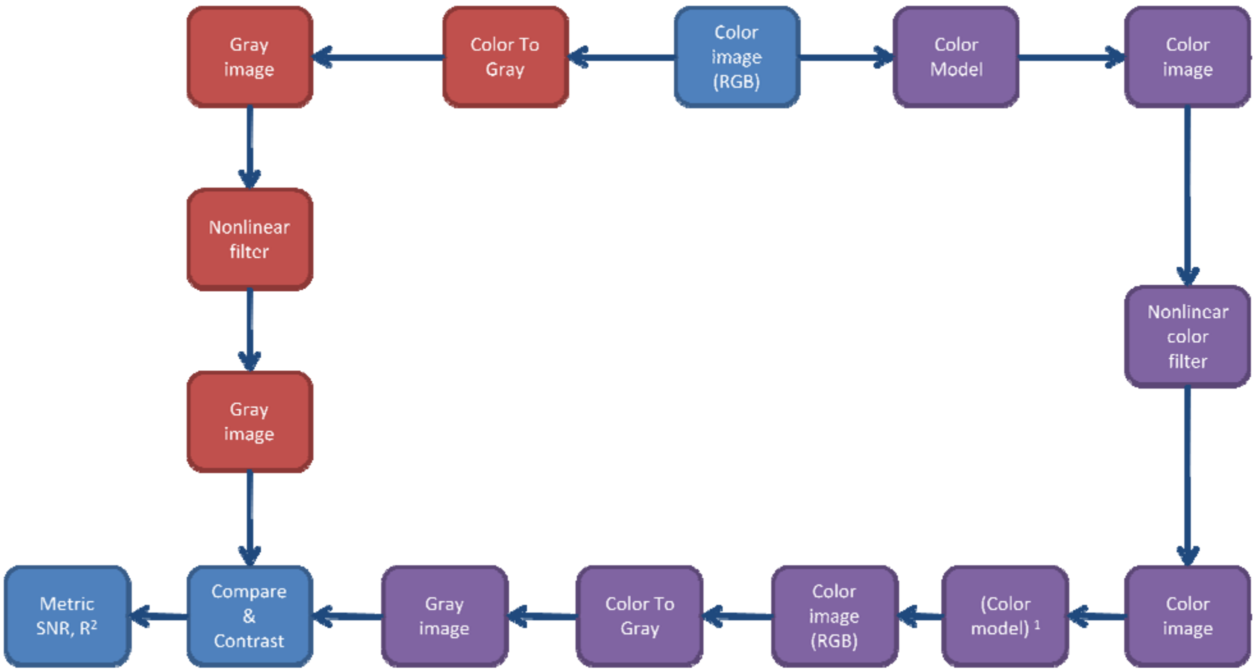


Figure 2: Software Architecture Used

The procedure for each filter was as follows:

1. Start with a noisy color image in RGB color space.
2. Use a defined equation to convert RGB image to grey-level image.
3. Apply nonlinear filter to the image.
4. Take the original color image in RGB color space and use algorithms to convert the image into another color space (YIQ, HSV, XYZ).
5. Apply filter to the color image.
6. Convert the image back to RGB.
7. Use the same equation used in STEP 2 to convert the RGB image to gray-level.
8. Compare and contrast the grey-level images in STEPS 3 and 7.
9. Apply the signal to noise ratio to determine the image quality.

The filter algorithms were implemented using MATLAB which includes an image processing toolbox. The program is very functional and user-friendly. The MATLAB Image processing toolbox also has functions that aid in the conversion of color spaces. MATLAB is very useful for mathematical computations such as the use of matrices and vectors. There is functionality provided for matrix multiplications. The use of matrix multiplications is complex in image processing due to the use of vectors. Therefore, the magnitude, cross product, or the dot product can be used when multiplying pixel values in the equations of the filters. The use of matrix and vector multiplication is addressed in the **DISCUSSION** section.

4. RESULTS

The following table shows the Signal to Noise Ratio (SNR) for each combination.

Table 1: **RGB**

	WOS	Stack	Rank Order	Nonlinear Mean	Teager	Polynomial	Rational
Magnitude	20.8228	20.1962	14.0901	20.0767	N/A	N/A	N/A
Cross Product	N/A	N/A	N/A	N/A	7.13618	-9.05561	13.9935
Dot Product	N/A	N/A	N/A	N/A	8.70002	5.5154	9.73282

Table 2: **HSV**

	WOS	Stack	Rank Order	Nonlinear Mean	Teager	Polynomial	Rational
Magnitude	1.84821	1.8566	11.8685	1.16047	N/A	N/A	N/A
Cross Product	N/A	N/A	N/A	N/A	3.7356	-19.3505	0.73193
Dot Product	N/A	N/A	N/A	N/A	-10.8557	4.03196	-1.2355

Table 3: **YIQ**

	WOS	Stack	Rank Order	Nonlinear Mean	Teager	Polynomial	Rational
Magnitude	29.9345	29.7441	33.566	30.3338	N/A	N/A	N/A
Cross Product	N/A	N/A	N/A	N/A	7.282	-16.2347	2.9111
Dot Product	N/A	N/A	N/A	N/A	5.04037	10.4187	2.59794

Table 4: **XYZ**

	WOS	Stack	Rank Order	Nonlinear Mean	Teager	Polynomial	Rational
Magnitude	12.3291	11.967	14.4147	11.9524	N/A	N/A	N/A
Cross Product	N/A	N/A	N/A	N/A	1.14778	-5.62941	12.8802
Dot Product	N/A	N/A	N/A	N/A	0.606417	12.6513	11.5968

The following figures show the two images with the 2nd highest SNR result which is the Nonlinear Mean filter using YIQ and magnitude equation.

Figure 3: Output of Gray-level image for Rank Order filter using YIQ and magnitude

Figure 4: Output of color to Gray-level image for Rank Order filter using YIQ and magnitude

Figure 5: Output of Gray-level image Nonlinear mean filter using YIQ and magnitude

Figure 6: Output of color to Gray-level image for Nonlinear mean filter using YIQ and magnitude

5. DISCUSSION

The equations used for the filters were obtained from [5]. The output equations were usually given for 2-Dimensional images. It was very simple to apply these equations to the gray level images. However, the algorithms became much more complex when it came down to applying the filters on 3-Dimensional images. One had to be careful traversing through the array of pixels since the size of the original noisy RGB image was 324 x 432 x 3. The two dimensional image size was simply 324 x 432. The 3D image has 3 channels rather than one like a 2-Dimensional image. A group of three for loops was used to manipulate the 3-Dimensional image. The loops simply computed the output using the algorithms for each of the three channels.

The multiplication of two vectors requires the use of cross product or dot product. The cross product of two vectors with 3 elements produces an output vector with three elements. However, the dot product of two vectors with 3 elements produces a scalar output, where as one value is returned. Therefore, the use of the dot product required to expand the vector to three elements in order to complete the filter algorithm. Matrix multiplication was only used for the Teager, Polynomial, and Rational filters. The Weighted Order Statistics filter, stack filter, rank order, and the Nonlinear Mean filter don't require the use of cross or dot product. In these filters, the magnitude equation was used to add and subtract the pixel values for the filter algorithms.

The magnitude equation for the RGB images was simply

$$\sqrt{R^2 + G^2 + B^2} \tag{7.0}$$

The same method was applied for the XYZ, YIQ, and HSV color spaces.

MATLAB provides a function that computes the cross product or dot product of two vectors. These functions were used in the program code. These functions were simply used whenever two pixel values were multiplied. For example, if the computation is x(m,n) * x(m+1, n +1), then the cross product and dot product functions are used separately to multiply the two values. However, the cross and dot product was not used whenever a pixel value was to a power or squared such as in the polynomial function. Instead, the power function was used because the cross product of a vector multiplied by itself is equal to zero. Therefore, using the power function would produce better results.

Nonlinear methods and techniques are intensive topics of research in image processing. The best way to find the best combination of filters with color spaces and matrix multiplication is to develop a table which has been done. From the table, one can see that the Rank order filter used on YIQ color space and magnitude for computations produced the highest Signal to Noise Ratio of 33.566. The Nonlinear mean filter used on YIQ and magnitude for computations produced the 2nd highest Signal to Noise Ratio of 30.3338. Therefore, these two combinations would be best for improving the quality of an image. From using eyesight, one can see that these two images from both combinations have great quality. From looking at the comparison table, one can see that the YIQ color space has the highest mean of the SNR when a WOS, Stack, Rank Order, or Nonlinear Mean filters are used. Again, in this case neither dot products nor cross products are used. From the table, one can conclude that the polynomial filter produces the worse SNRs when a cross product is used. The SNRs are all large negative numbers. However, the use of the dot product resulted in high positive SNRs. The highest SNRs for the three filters using cross product or dot product are above 10 dB. The use of the rational filter has the most occurrences of high SNRs. By observing the table, one can see that in general, the RGB and the XYZ color spaces produce the best results. The HSV color space clearly returns the worst results. This color space contains the most occurrences of negative SNRs.

6. CONCLUSIONS

From looking at the output of each filter, it was interesting to see the effects of edge detection and noise reduction. For example, the Teager filter seemed to be a great edge detection filter. The gray level image was pitch black with white outlines. By looking at the Signal to noise ratios from each combination, the Rank order filter on a YIQ image resulted in the highest SNR of approximately 34. The worse SNR was found in the polynomial filters in any color space with the use of cross product. Also, one can conclude that the Rank order filter has the highest mean of SNRs, which means that it is the best nonlinear filter to use compared to the others. It is noticeable that the use of the filters that use the magnitude equation produces a much larger SNR than those filters that require the use of cross or dot products.

7. REFERENCES

[1] Galigekere, R. R.; "Color-image processing: An introduction with some medical application-examples," **2010 International Conference on Systems in Medicine and Biology (ICSMB),** pp. 3, 16-18 Dec. 2010.
[2] Zhu Shangdong; Ai Zhibin; Chen Xiyuan; "On color image smoothing processing based on color space transformation," **Proceedings of the IEEE International Vehicle Electronics Conference, *(IVEC '99),*** pp. 271-274 vol. 1, 1999.
[3] Batard, T.; Berthier, M.; "Heat kernels of generalized laplacians-application to color image smoothing," **16th IEEE International Conference on Image Processing (ICIP),** pp. 461-464, 7-10 Nov. 2009.
[4] Lezoray, O.; Elmoataz, A.; "Graph based smoothing and segmentation of color images," **Proceedings. Seventh International Symposium on Signal Processing and Its Applications,** pp. 517- 520 vol.1, 1-4 July 2003.
[5] Sicuranza G. and Mitra S.; "Nonlinear Image Processing," **Communications, Networking and Multimedia,** 19 September 2000.

Urban Environment Monitoring: System and Technology Issues

Mario LA MANNA
Selex Sistemi Integrati
Via Tiburtina km. 12.400,
Roma 00131 Italy

ABSTRACT

Urban environment monitoring needs a complex wireless sensor network architecture, which is capable to ensure maximum connectivity between the nodes of the network, to provide adequate throughput and to allow reduced response time. System and technology issues of such a network have to respond to specific guidelines, in order to be compliant with scenario requirements and to offer maximum performance in a number of different operating situations. URBEM (URBan Environment Monitoring) is a new concept idea, which encompasses both design and simulation of a heterogeneous wireless network for urban environment monitoring, with the goal to respond to modern urban scenario requirements.

Keywords: Wireless sensor network, urban environment, clustering, data fusion, performance simulation.

1. INTRODUCTION

URBEM (Urban Environment Monitoring) is a heterogeneous wireless sensor network for application in a complex urban environment. Nodes in URBEM are deployed and left unattended, with the goal to survive as much as possible in a hard environment. In that context, energy consumption is sensibly reduced to increase network lifetime. This leads also to specific topology control strategies. In URBEM, the sources of information are the sensors. Each sensor explores the physical environment and reports the available information to one or more observers, or sinks, through wireless communications. Two different kinds of traffic exist in URBEM: phenomenon driven and continuous. The former type of traffic happens when a specific event has been detected by some sensors, leading to information transmitted toward the sink. The latter kind of traffic is positional information, such as information being regularly generated and emitted towards the sink. Both kinds of traffic require high network throughput and low latency. Moreover, as information from multiple sensors is merged to improve sensible target localisation, recognition and tracking, suitable distributed data fusion algorithms are employed. For the above reasons, efficient simulation methodologies are adopted during the design phase, in order to assess correctness and quality of design. Simulation is based on modular tools and friendly interface, in order to allow easy and fast collection of results and to apply trial and error strategies.

2. TECHNOLOGY ISSUES IN URBEM

The URBEM architecture (Fig. 1) is based on nodes of local sensors, tightly connected through clusters (Tier 1). Inside clusters, local information coming from sensors is merged and forwarded to the upper level (Tier 2). Cluster configuration is managed dynamically, in order to insure autonomous re-configurability, node mobility, resistance to node fault or destruction and node dynamic deployment.

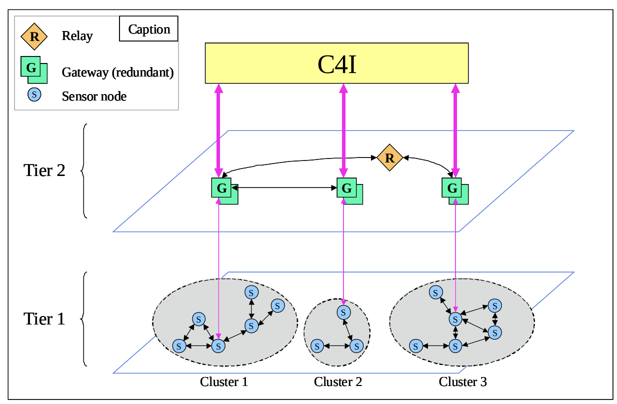


Fig. 1. URBEM Architecture.

Communications inside a cluster are organized by means of multi-hop packet exchange and suitable routing algorithms.
In the upper level, namely Tier 2, connection between clusters is provided by means of highly dependable (at least duplicated) gateways and wide bandwidth high speed links, featuring high robustness and resistance to cyber attacks. These links are based on direct ground based inter cluster connections. When a direct connection is not available, due to obstacles posed by the environment, poor quality of existing infrastructure, node failure, etc., the link can be implemented through suitable relays.

3. DATA FUSION STRATEGIES

Data fusion inside URBEM is organized in a tree structured fashion, by first merging information coming from separate sensors at platform level (inside Tier 1) and then collecting and centralizing the merged data (inside Tier 2) and forwarding the final data to the central station.

Due to the high density of sensors present in URBEM and to the need to centralize a huge amount of information, data have to be necessarily merged in stages, to the extent possible at each stage. The first stage is the one existing at platform level, upper level stages are at node and gateway level, and so on, until arriving to the top central point in the network. Independently from the network topology, the data fusion process can be represented as in Fig. 2.

The process in Fig. 2 represents K data fusion threads, where each thread corresponds to data fusion of a specific piece of picture in the scenario and the K different picture pieces are assumed to be uncorrelated and adaptive to the scenario characteristics.

As an example, the K different picture pieces can be related to a scenario where K different target objects have been discovered and are being tracked by K different data fusion threads. In the picture, it is assumed that each data source is contributing in the whole structure with the same R data rate and that, for each thread, there are Q data fusion steps, each of which merging data coming from N different nodes.

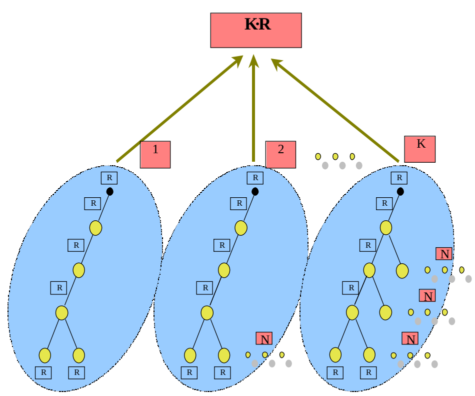


Fig. 2. Data Fusion Process.

4. PERFORMANCE METRICS IN URBEM

The main performance parameters, to be evaluated by simulation, in a wireless sensor network, with reference to the URBEM architecture, are the following.

1) Connectivity,
2) Response time,
3) Throughput,
4) Energy usage,
5) Congestion control,
6) Node density management,
7) Network and data centricity,
8) Mobility and self configurability,
9) Distributed signal/data capability,
10) Scalability,
11) Dependability and quality of service.

Performance parameters can be evaluated by analysis and/or simulation. Analysis can be performed at the top level of the network, while simulation can be performed at different levels of abstraction and by adopting scalable strategies. The basic goal is to evaluate network performance, in presence of bursty traffic (alarm storms).

A typical analysis approach consists of the following steps.

1) Definition of the high level network architecture,
2) Modelling of the network,
3) Performance analysis of the high level building blocks.

The analysis approach is generally applied at higher system level (see Fig.3).

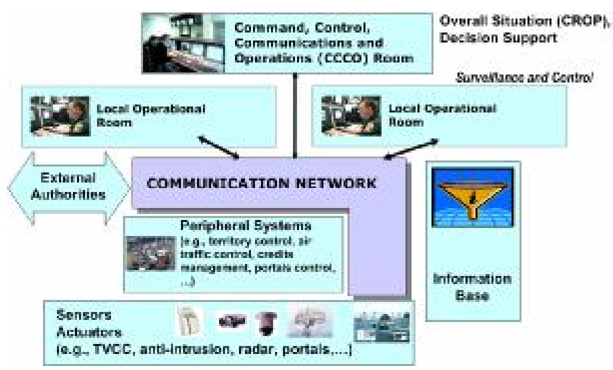


Fig. 3. High Level Analysis Approach.

A typical simulation approach consists of the following steps (see Fig. 4).

1) Definition of the lower level network architecture,
2) Modelling and emulation of the network,
3) Definition of typical scenarios,
4) Simulation runs and collection of results (Monte Carlo Trials).

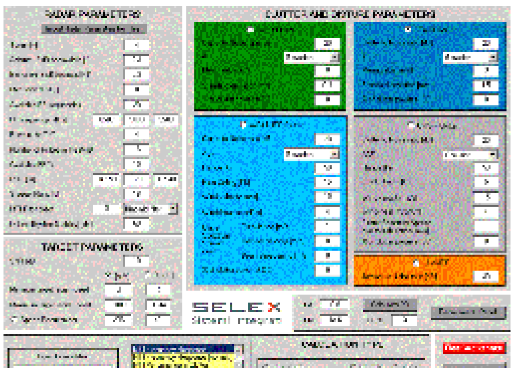


Fig. 4. Simulation Approach.

The URBEM performance simulation approach benefits from advantages coming from analysis and simulation and from the principle of the “divide and conquer”. Analysis is performed at the top level of the network, while simulation is performed at the lower levels, by adopting scalable strategies. The overall results, in terms of parametric formulas, graphics, etc. (see some result examples in Fig. 5) are obtained in a short time and are practically usable in the stage of feasibility assessment.

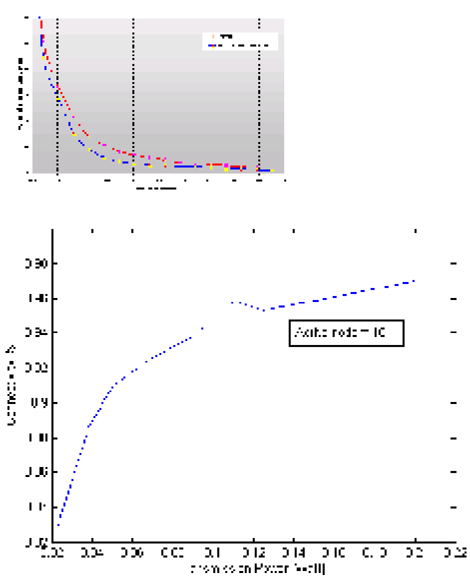


Fig. 5. Result examples from Performance Simulation.

5. PERFORMANCE SIMULATION PROCESS

A typical performance simulation process of a multi cluster network is composed of cluster simulation (lower level) and network analysis (higher level).
Cluster simulation allows to evaluate, by means of modelling and emulation of the network clusters and by using the process of Monte Carlo trials, the critical parameters relative to the network at cluster level.
Typical main cluster parameters are the following.
1) Cluster Connectivity = CConn
2) Routing Latency Time = CRou
3) Data Fusion Latency Time = CDatf

Network simulation allows to evaluate, by means of analysis, the critical parameters at network level. Typical main network parameters are the following.
1) Network Connectivity = NConn
2) Routing Latency Time = NRou
3) Data Fusion Latency Time = NDatf

The overall performance simulation allows to evaluate, by means of combined analysis and simulation, the critical parameters at global level.
The global parameters derived from the previous ones are the following.
1) Global Connectivity = Cconn * NConn
2) Established Track Response Time = CRou + CDatf + NRou + NDatf

Further parameters can be evaluated, by combining homogeneous results produced by cluster simulation and network analysis, by using specific formulas, like in the above example, spreadsheets or different calculation methods.

6. OPEN ISSUES AND FURTHER WORK

The URBEM (URBan Environment Monitoring) architecture concept implements a heterogeneous wireless network for urban environment monitoring. The practical use of the URBEM architecture needs the use of new technologies, some of them already in an advanced stage of development. These technologies span from the hardware of miniaturised sensor nodes, with low power consumption, to the definition of advanced communication protocols for cross-layer searching and to the definition of efficient data fusion methodologies, based on adaptive fusion.

7. REFERENCES

[1] Paul J. Croft et al. " UrbanNet: Urban Environment Monitoring and Modeling with a Wireless Sensor Network", 90th American Meteorological Society Annual Meeting, Atlanta, 2010.

[2] U. Lee et al. " Mobeyes: Smart mobs for urban monitoring with a vehicular sensor network",Wireless Communications IEEE, Vol.3, Issue 15, October 2006.

[3] B. A. Warneke, et al., "Smart Dust: Communicating with a Cubic-Millimeter Computer", Computer Magazine, Jan. 2001, pp. 44-51.

[4] B. A. Warneke, K. S. J. Pister, "Exploring the Limits of System Integration with Smart Dust", ASME International Mechanical Engineering Congress & Exposition, 2002-

[5] Charles E. Perkins, Pravin Bhagwat: DSDV Routing over a Multihop Wireless Network of Mobile Computers, Charles E Perkin (Editor) Ad Hoc Networking, pp. 53-74, Addison-Wesley, 2001

[6] Charles E. Perkins, Elisabeth M Royer: The Ad Hoc On-Demand Distance-Vector Protocol, Charles E Perkin (Editor) Ad Hoc Networking, pp. 173-219, Addison-Wesley, 2001

[7] C. Mbarushimana: Comparative Study of Reactive and Proactive Routing Protocols Performance in Mobile Ad Hoc Networks, Advanced Information Networking and Applications Workshops, 2007

[8] David B. Johnson, David A. Maltz, Josh Broch: DSR: The Dynamic Source Routing Protocol for Multihop Wireless Ad Hoc Networks, Charles E Perkin (Editor): Ad Hoc Networking, pp. 139-172, Addison-Wesley, 2001

[9] G. Ding, J. Vicente, S. Rungta, D. Krishnaswamy, W. Chan, K. Miao. "Overlays on Wireless Mesh Networks: Implementation and Cross-Layer Searching" 9th IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services, Dublin, 2006.

[10] A. Roecker, and C.D. McGillem, "Comparison of Two-Sensor Tracking Methods Based on State Vector Fusion and Measurement Fusion", *IEEE Trans. On Aerospace and Electronic Systems*, vol. 24, no. 4, July 1988, pp. 447-449.

[11] A. Farina, and F. Studer, *Radar Data Processing. Advanced Topics and Applications* (vol. II), Research Studies Press (England), John Wiley & Sons (USA), May 1985.

[12] A. Farina, and F. Studer, "Radar Data Processing. Introduction and Tracking (vol. I)", Research Studies Press (England), John Wiley & Sons (USA), May 1985.

[13] Y. Bar-Shalom, X.R.Li, Multitarget-Multisensor Tracking: Principles and Techniques. YBS Publishing, Storrs, CT, 1995.

[14] N. G. Wah, Y. Rong , "Comparison of Decentralized Tracking Algorithms", *ISIF 2003*, pp. 107-113

[15] Y. Bar-Shalom, "On the Track-to-Track Correlation Problem", IEEE Trans. On Automatic Control, vol. 26. no 2, April 1981.

[16] C. Mbarushimana: Comparative Study of Reactive and Proactive Routing Protocols Performance in Mobile Ad Hoc Networks, Advanced Information Networking and Applications Workshops, 2007

[17] Abtin Keshavarzian, Huang Lee, Lakshmi Venkatraman, Dhananjay Lal, Krishna Chintalapudi, and Bhaskar Srinivasan: Wakeup scheduling in wireless sensor networks. Proc. of the ACM MobiHoc'06, Florence, Italy, May 2006.

[18] Fabian Kuhn, Thomas Moscibroda, Roger Wattenhofer: Fault-tolerant clustering in ad hoc and sensor networks. Proc. of the 26th IEEE International Conference on Distributed Computing Systems (ICDCS 06), Lisboa, Portugal, July 2006.

[19] R. Montemanni, L. M. Gambardella, A. K. Das:. Mathematical models and exact algorithms for the min-power symmetric connectivity problem: an overview. Handbook on Theoretical and Algorithmic Aspects of Sensor, Ad Hoc Wireless, and Peer-to-Peer Networks, pages 133–146, Auerbach Publications, 2006.

[20] M. Cardei, D. Du: Improving wireless sensor network lifetime through power aware organization. ACM Journal of Wireless Networks, 11(3):333–340, May 2005.

[21] P. Santi, The critical transmitting range for connectivity in mobile ad hoc networks. IEEE Transactions on Mobile Computing, 4(3), May-June 2005.

[22] Hai Liu, P. Wan, C.-W. Yi, Xiaohua Jia, S. Makki, and N. Pissinou. Maximal lifetime scheduling in sensor surveillance networks. Proc. of the 2005 IEEE Infocom, volume 4, pages 2482–2491, Miami, Florida, March 2005.

Web Services and Ontologies for Building Data

Deborah L. MacPherson
Cannon Design, Specifications and Research
National Information Exchange Model (NIEM) Business Architecture Committee (NBAC)
National Building Information Modeling Standard (NBIMS-US) Technical Subcommittee
OmniClass Development Committee
dmacpherson@cannondesign.com

ABSTRACT

Building Information Modeling (BIM) needs to be placed in a larger architectural context with data modeling, geography, web infrastructures, public safety, and real estate. A common reference system is needed to ensure fundamental characteristics of building interiors are understood in affiliated software and services. Only a few data need hooks, for example footprints, entrances, floors, circulation, occupancy, spaces and specified contents. Implementation of successful BIM web services will be a symphony of the following: a Data Model (isolate certain International Foundation Classes); a Dictionary (APIs with the buildingSMART Data Dictionary); classification and taxonomies (OmniClass); and Process Descriptions (Information Delivery Manuals and Model View Definitions). A kit of parts, including metadata organization and rules for what is discoverable in registries and repositories, needs to conform to Simple Object Access Protocol (SOAP), Representational State Transfer (REST), Web Services Description Language (WSDL), and new Internet technologies as they are invented. The whole system could benefit from a collection of Tiny Ontologies All Stitched Together (TOAST) branching out from a common core. The US National Information Exchange Model (NIEM) is an ideal home for elements related to justice, emergency management, health and family services - many people could benefit from this work.

Keywords
BIM, GIS, NIEM, OmniClass, Data Structures

INTRODUCTION

As timely and practical as Building Information Modeling (BIM) is, the greatest use may be via web services. For example, to achieve President Obama's 2011 goal for "...a firefighter to download the design of a burning building onto a handheld device" several applications, networks, policies, and Internet layers need to work together **[Figure 1]**. While this might be possible in a closed environment today, for any firefighter to access up-to-date information, BIMs need to be placed in a larger architectural context with data modeling, geography, web infrastructures, public safety, and real estate.

Current BIMs are not scalable to thousands of simultaneous users focused on specific buildings during an emergency, or to seamlessly navigate indoor and outdoor locations. A common reference system is needed to ensure fundamental characteristics of building interiors are understood in affiliated software and services. Only a few pieces of data need hooks, for example footprints, entrances, floors, circulation, occupancy, spaces and specified contents.

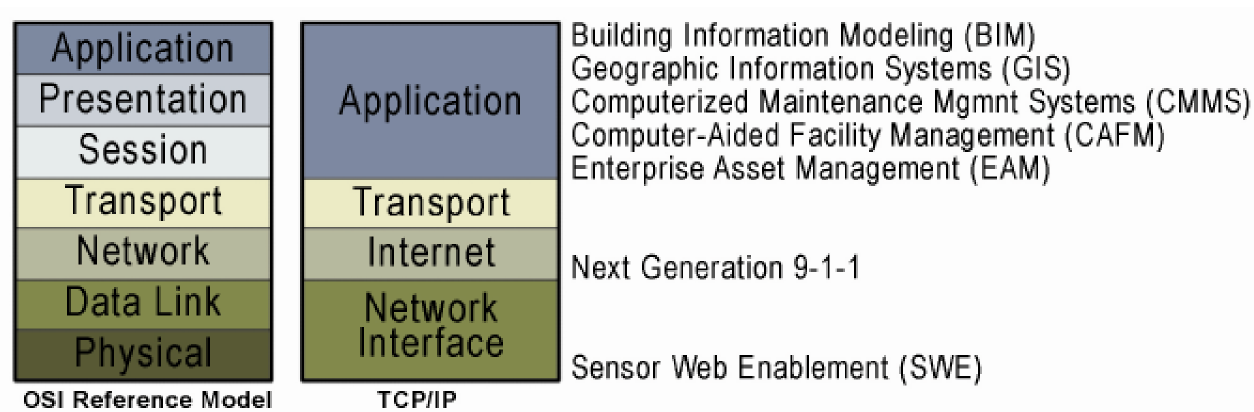


Figure 1- Software, Networks, and Internet Architecture

OPEN ARCHITECTURE

Standards Development Organizations (SDOs) and vendors should have a summit to understand what everyone else is doing relative to themselves. Significant results could be achieved quickly. Everyone is in a better position to do this now, most Memorandums of Understanding (MOUs) are already in place. Shared concepts, software interfaces, service interfaces, and data structures can be extended in more detail by each stakeholder to understand interrelationships of shared building information through their own data, systems, and viewpoints.

Different parts of an open architecture for BIM and related technologies will originate, and need to be maintained, by diverse professional organizations. Senders and receivers need to agree on exactly which information to exchange and what it should look like. The whole system needs to be dynamic and cyclical, updating with continuous improvements and strict version control. Defining where each part belongs in the whole world's digital infrastructure will ensure each part is worth doing, will be sustainable, published openly, and owned by people that care about what the information means. The results will not be duplication - but harmonization and increased customer satisfaction.

WEB SERVICES

Web Services are automated machine-to-machine exchanges and data integration processes using open standards over the Internet to simplify complex data into messages.

Successful web services for BIM requires the following: a Data Model (isolate certain International Foundation Classes); a Dictionary (APIs with the International Framework for Dictionaries); classification and taxonomies (OmniClass); and Process Descriptions (Information Delivery Manuals and Model View Definitions). A kit of parts, including metadata organization and rules for what is discoverable in registries and repositories, needs to conform to Simple Object Access Protocol (SOAP), Representational State Transfer (REST), Web Services Description Language (WSDL) and new Internet technologies as they are invented.

Leveraging existing work by others will require specified sets of building data to stay bound together in messages. Message types will define the order data are presented, beginning with the most important sets first, then compiling more detail depending on the contracted web service.

ONTOLOGIES

Ontologies are often used for search, decision support, and software design. Matthew West defines ontologies simply as "The things there are and the rules that govern them." BIM ontologies need to be web-based to support workflow functions, provide overviews, relationship structures, necessary attributes, reminders about information to include, and consistency across integration models used with master and reference data. Ronald Reck suggests a collection of Tiny Ontologies All Stitched Together (TOAST) branching out from a common core. The US National Information Exchange Model (NIEM), a data model for data exchanges using a common core, is an ideal home for facility data elements related to emergencies, justice, health, and family services - many people could benefit from a focused collaborative effort.

Large complex models like NIEM and the Open Geospatial Consortium (OGC)'s CityGML are made for different reasons. Ontologies drawing from them can be leaner and fit for purpose. Maintaining linkages between and among data assets, keeping specific data types bound together, and resolving multiple records describing the same assets needs ontologies to tell machines how to interpret the information, what to do with it, and what level of detail to store in each environment without restating the same instructions for each web service or affecting other data models internal structures.

LINKED DATA

Changing a published building layout or occupancy should be able to tap back into open floor plans, linked databases, and public records. Exporting heavy files like typical BIMs are today sends more information than needed and too many copies are generated. Using a Service Oriented Architecture (SOA) approach instead, data remains with one application for reference or modification in other applications by maintaining links and mutually supportive artifacts. Shared elements and a messaging structure created for BIM going outwards could also benefit BIM going deeper into the design process at Architecture/Engineering firms. Some typical software such as BIM and specification writing programs can maintain rolling associations, but they are only two of many **[Figure 2]**. If web services and ontologies could be used to establish semantic connections rather than just software connections, archival flows could extend to knowledge management systems so a critical mass can be reached within firms and across the industry. The number of known answers to repeated questions

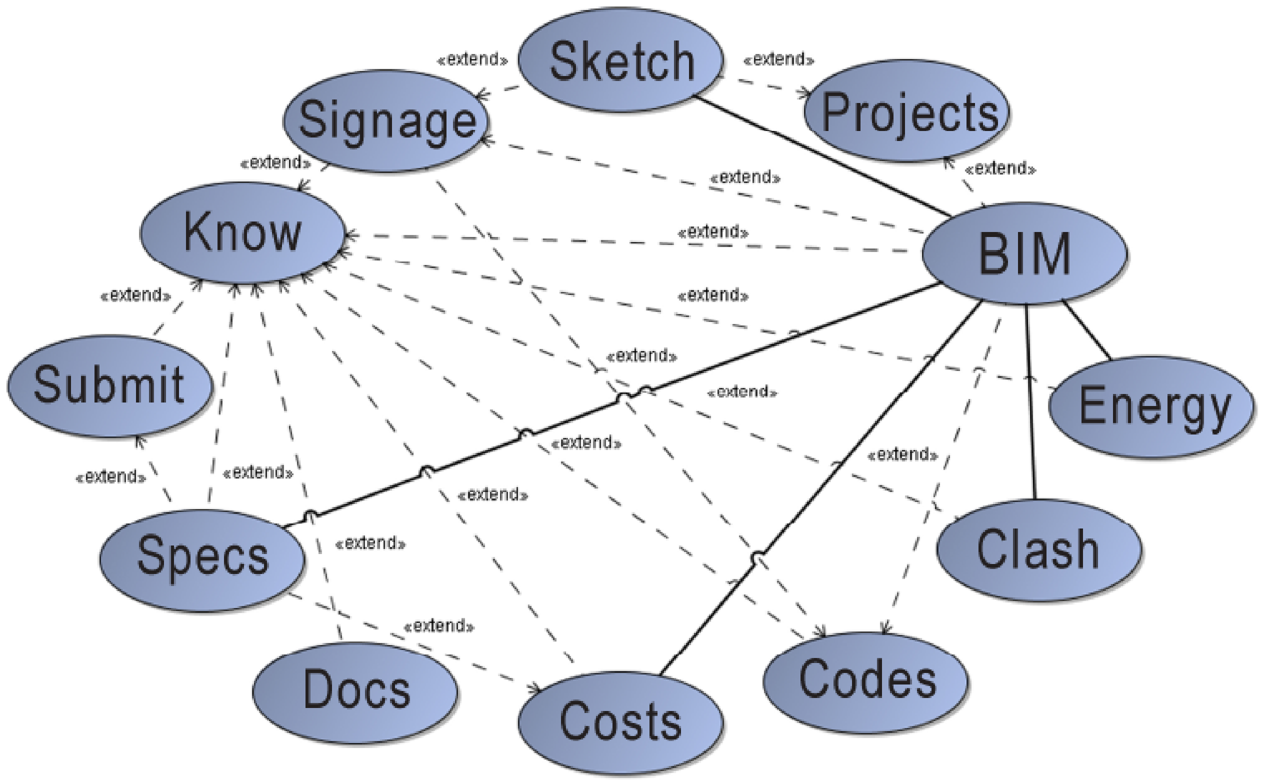


Figure 2 - Typical A/E Software

and combinatorial complexity of geometric objects can be reduced using thematic and logical restrictions. When an OmniClass type usually has a limited range of typically associated components, it will get easier to link back and forth between applications and resources. Repeated instances of OmniClass types, and formulaic buildings like dormitories, will be especially adaptable.

BENEFITS

Live data could be automatically captured in compliance with policy constraints to flow through to first responders, maintenance contractors, real estate agents, energy auditors, financial agents, and other authorized users without duplication and potential human error. Real-time information compared to historical trends will let machines keep track of minutiae and fastest routes to improve scheduling, perform analysis, and assess scenarios.

Several initiatives need to converge so the results will begin showing up in software. Defining a core set of elements with relationships between them and the outside world will let users, or machines on their behalf, periodically submit questions such as "have the regulations for chlorine storage changed in Arlington County?" and get an answer from the local jurisdiction or other authoritative data source. Each user and their machine would only receive the information they need, or are allowed to access, rather than an unwieldy batch of data to extract an answer. The extraction is performed via web services. Invasion of privacy and terrorist planning can be prevented by limiting distribution to public information or authorized web services with elemental traceability. Some web services and ontologies will be free for the general public, some will cost a lot of money. The first and most collaborative versions should be for public safety in public buildings such as schools, hospitals, courthouses, stadiums, airports, hotels, and shopping malls. These are also buildings where anyone may want to look up their current location on a handheld device so consumer demand and Internet service providers can help speed development.

CONCLUSION

Web services and ontologies are progressive; there will never be a final version, finite set of questions and answers, or set of dots to connect. A shared reference architecture needs to celebrate complexity in the world, not dull it down, or the results will not be realistic and effective.

Acknowledgements: Louis Hecht, Ron Reck, Matthew West, Joseph Talamantez, and Christopher Shoup.

In Praise of Wavelets – 3 Disparate Case Studies

Apkar Salatian, Francis Adepoju
School of Information Technology and Communications
American University of Nigeria,
Yola Bypass, PMB 2250,
Yola, Nigeria
Email: {apkar.salatian, francis.adepoju}@aun.edu.ng

Abstract

Wavelets are a mathematical tool that can be used to extract information from many different kinds of data, Sets of wavelets are generally needed to analyze the data fully. In this paper we will show how wavelets analysis can be applied to 3 disparate case studies from our own research. In our first case study we will show that wavelet analysis can be used to facilitate clinical decision support from Intensive Care Unit data. In our second case study we will show how wavelets can create conceptual models from building data. In our final case study we will show how wavelet analysis can help in telecommunications by reducing transmission costs and improve the quality of service of a network. We conclude that wavelets are a suitable tool for processing large and noisy data sets for higher level processes to interpret.

Keywords – Wavelet analysis, medicine, building, telecommunication

I. INTRODUCTION

Real-world data tend to be incomplete, noisy, inconsistent, high dimensional and multi-sensory and are not suitable for further processing without reasonable data preprocessing. One approach to overcome this problem is to use wavelets.

Wavelets are functions that satisfy certain mathematical requirements and are used in representing data or other functions. The wavelet transform is a synthesis of ideas that emerged over many years from different fields. Generally speaking, the wavelet transform is a tool that partitions data, functions, or operators into different frequency components and then studies each component with a resolution matched to its scale [1]. Therefore, it can provide economical and informative mathematical representation of many objects of interest [2].

In this paper we will show from our previous research that the results of wavelet analysis can have many applications. Using real data, we will present 3 disparate cases to demonstrate how higher level processes can make use of wavelets. In our first case study we will show that wavelet analysis can be used to facilitate clinical decision support from data generated by the bedside monitors in an Intensive Care Unit. In our second case study we will show how wavelets can create conceptual models from data generated by the environmental sensors in a building. In our final case study we will show how wavelet analysis can help in telecommunications by reducing transmission costs for restricted bandwidth and improve the quality of service in a network.

The structure of this paper is as follows. Section 2 describes the wavelet analysis algorithm. Section 3 presents results of applying wavelet analysis to a real data set. Our discussion of how wavelet analysis can be used in 3 disparate case studies from our own research is given in section 4 and final conclusions are given in section 5.

II. WAVELET ANALYSIS

Wavelets are a mathematical tool they can be used to extract information from many different kinds of data, A wavelet is a wave-like oscillation with amplitude that starts out at zero, increases, and then decreases back to zero.

At first glance, wavelet transforms are very much the same as Fourier transforms except they have different bases. The wavelet transform is capable of providing time and frequency localizations simultaneously while Fourier transforms only provide frequency representations. Fourier transforms are designed for stationary signals because they are expanded as sine and cosine waves which extend in time forever - if the representation has a certain frequency content at one time, it will have the same content for all time. Hence the Fourier transform is not suitable for a non-stationary signal which has time varying frequency [3].

In wavelet analysis, the scale that we use to look at data plays a special role. Wavelet algorithms divide a given function or continuous-time signal into different scale components. One can assign a frequency range to each scale component. Each scale component can then be studied with a resolution that matches its scale. If we look at a signal with a small window, we would notice small features. Similarly, if we look at a signal with a large window, we would notice gross features. There has been a requirement for more appropriate functions than the sines and cosines that comprise the bases of Fourier analysis, to approximate choppy signals. Wavelets have been shown to be well-suited for approximating data with sharp discontinuities [4] – this feature was common in our data from our case studies.

Generally, Wavelet transform of signal f using wavelet Ψ is given by:

$$W_{\psi}(f)(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\psi\Big(\frac{t-b}{a}\Big)dt \quad (1)$$

where the variable a is the dilation factor, variable b is the translation factor.

The wavelet analysis procedure is to adopt a wavelet prototype function, called an analyzing wavelet or mother wavelet. Temporal analysis is performed with a contracted, high-frequency version of the prototype wavelet, while frequency analysis is performed with a dilated, low-frequency version of the same wavelet. We will now describe the wavelet method.

Assume that $Y(t)$ is the value of an observable time series at time t, where t can take on a continuum of values. $Y(t)$ consists of two quite different unobservable parts: a so-called trend $T(t)$ and a stochastic component $X(t)$ (sometimes called the noise process) such that

$$Y(t) = T(t) + X(t) \quad (2)$$

where it is assumed that the expected value of $X(t)$ is zero. There is no commonly accepted precise definition for a trend, but it is usually spoken of as a nonrandom (deterministic) smooth function representing long-term movement or systematic variations in a series. Kendall [5] asserted that the essential idea of a trend is that it shall be smooth while Priestly [6] refers to a trend as a tendency to increase (or decrease) steadily over time or to fluctuate in a periodic manner. The problem of testing for or extracting a trend in the presence of noise is thus somewhat different from the closely related problem of estimating a function or signal $S(t)$ buried in noise. While the model $Y(t) = S(t)+X(t)$ has the same form as equation (2), in general $S(t)$ is not constrained to be smooth and thus can very well have discontinuities and/or rapid variations.

The detection and estimation of trend in the presence of stochastic noise arises in the data from our case studies. A wavelet analysis is a transformation of $Y(t)$ in which we obtain two types of coefficients: wavelet coefficients and scaling coefficients - these are sometimes referred to as the *mother* and *father wavelet coefficients* respectively. The wavelets are scaled and translated copies (*father wavelets*) of a finite-length or fast-decaying oscillating waveform (*mother wavelet*).

The mother and father wavelets coefficients are fully equivalent to the original time series because we can use them to reconstruct $Y(t)$. Wavelet coefficients are related to changes of averages over specific scales, whereas scaling coefficients can be associated with averages on a specified scale. The information that these coefficients capture agrees well with the notion of a trend because the scale that is associated with the scaling coefficients is usually fairly large. Trend analysis with wavelets is to associate the scaling coefficients with the trend $T(t)$ and the wavelet coefficients (particularly those at the smallest scales) with the noise component $X(t)$. A more interesting situation arises when we observe trends with correlated noise. Under certain models and choice of wavelet function, the wavelet transform de-correlates the noise process and allows us to simplify the statistical analysis involved.

There are numerous mother and father wavelet functions which have specific uses. For example, Chan and Fu have shown that the Haar transformation is well suited for retrieving time series data [7]. Likewise, Nilsson and Funk [8] show that the Daubechies D4 Discrete Wavelet Transformations is more suited for biomedical retrieval because it overcomes Haar's inability to detect specific forms of oscillating sequences. For further discussion on other wavelet functions the reader is advised to read [9].

III. RESULTS

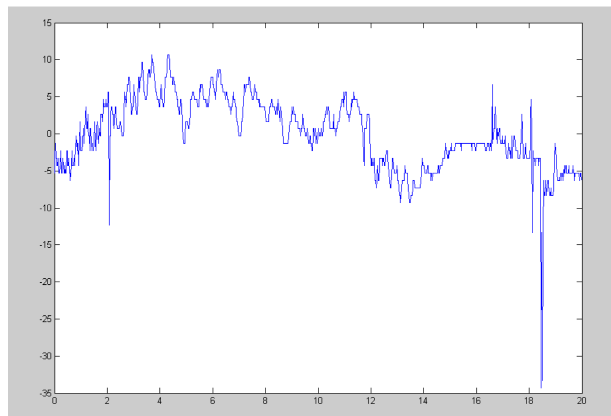


Figure 1: Original Data

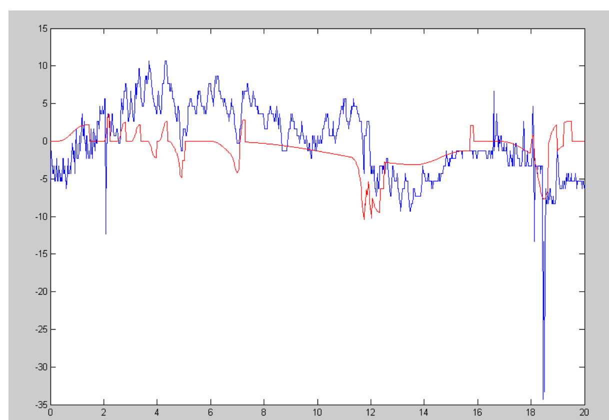


Figure 2: Final Trends

Figure 1 shows the original waveform of a Blood Pressure signal recorded using a catheter at a Neonatal Intensive Care Unit in the UK. The frequency of the data is 15 - 20Hz and 1000 data items were recorded. It can be seen that the data is noisy due to clinically insignificant events such as taking blood samples from the patient.

The final trends are captured in figure 2 which clearly mark the data trend when the data was analyzed with Shannon and Daubechies wavelets - it is this compressed data (trends), rather than the large original data set, that can be processed by higher level functions.

IV. DISCUSSION

Using our previous research, we will discuss how wavelet analysis can be used in 3 disparate case studies from our own research. We will show in our first case study that wavelet analysis can be used to facilitate clinical decision support from Intensive Care Unit monitor data. In our second case study we will show how wavelets can create conceptual models from building data collected from environmental sensors. In our final case study we will show how wavelet analysis can help in telecommunications by reducing transmission costs to make better use of restricted bandwidth and improve the quality of service of a network. We shall look at each in turn.

A. Clinical Data Support

The Intensive Care Unit (ICU) bedside monitors present the medical staff with large amounts of continuous data such as the heart rate and blood pressure. The frequency of the data can be higher than one value every second which creates information overload for medical staff. Indeed, monitoring systems have become increasingly complex, and the data rate is so high that all of the data cannot be utilised fully by medical staff, whose main function is to take care of the patient and not just to observe all of the information provided by the equipment [10].

In [11, 12] we described how wavelets can derive trends from the ICU monitor data to reduce the information overload to medical staff. In [13] we described how the generation of these trends allows *interval-based (qualitative) reasoning* - this has many applications in ICU monitoring to enhance clinical decision support.

Trends allow Temporal Reasoning - this involves 'reasoning about points and intervals of time, quantitative and qualitative relations, common temporal scales, temporal relations with respect to the present and alternate temporal hypotheses' [14]. To reason temporally one can consider a time line - it is assumed that all times ultimately map onto a real number line. Events occur either at *points* or within *intervals* on this time line. Using these *points* and *intervals*, one can reason temporally on the time line. For example, when considering the time point *now* everything to the left of this point is in the *past* i.e. certain actions or events have *occurred* while everything to the right of this point is the *future*. Hence one can use past events to draw expectations of what will happen in the future relative to *now*. Using the *past* one can, say, change plans for actions in the future by consolidating what have already occurred earlier.

Trends allow the identification of clinical conditions and the outcome of therapies for clinical decision support and clinically insignificant events for removal [15] – these are achieved by reasoning about the temporal relationships between intervals based on their endpoints (see [16]). Clinical conditions can be identified for medical audit e.g. the clinical condition shock can be identified as an increasing heart rate and decreasing blood pressure – this is achieved by identifying patterns in overlapping trends of particular signals**.** The outcomes of therapies are determined by comparing future trends to the trend when the therapy was administered to see if an expectation was met (or not). Clinically insignificant events which could not be removed by standard filtering are identified by associational reasoning of meeting intervals of a single signal or overlapping meeting intervals of multiple signals.

Interval-based reasoning hence removes the burden and complexity of reasoning on a point to point basis.

B. Conceptual Modeling of Buildings

Building operators are confronted with large volumes of continuous and noisy time series data from multiple environmental sensors which require interpretation. To assist with interpretation, building operators could benefit from a conceptual model derived from the data. In [17] we used the wavelets approach to derive trends in environmental monitor data to allow conceptual models to be created – these will determine how well the building is operating.

Using wavelet analysis to generate trends allowed us to develop a conceptual model of a complex system (an occupied building) in sufficient detail to enable the building operators to make informed decisions about how to improve the buildings energy performance. Such a model would allow temporal and interval-based (qualitative) reasoning which has many applications in environmental monitoring.

Using temporal reasoning, a building operator can determine if expectations of particular events were met e.g from the conceptual model, does sticking a notice on office doors asking staff to switch of lights every time they leave the office have any effect on electric load for lighting?

Using interval-based reasoning the building operator could derive from the conceptual model that a wall is storing heat when the wall internal surface temperature increases faster than the brick interior temperature. Faults can also be identified e.g a rise in air temperature in a supply duct which then remains constant when the external air temperature continues rising above the set point might indicate that the valve controlling hot water flow to the air heater battery has stuck in the fully open position. A building operator can reason about temporal relationships between intervals based on their endpoints – this allows a building operator to find associations between different signals with or without the presence of an external event.

C. *Telecommunications*

We have shown that wavelets analysis can be used as a lossy data compression technique for the transfer of large data sets to make better use of restricted bandwidth in our paper [18]. Wavelets analysis allow for more efficient use of network resources – the resulting compressed data reduces storage requirements and makes better use of bandwidth since smaller files take up less room on the access pipe and are therefore faster to transfer over a network. From our results, we have also shown that our approach serves to remove redundancy in the data such as noise.

Wavelets analysis is a lossy compression technique that can be adjusted to different quality levels, gaining higher accuracy in exchange for less effective compression. A set of *complementary* wavelets will deconstruct data without gaps or overlap so that the deconstruction process is mathematically reversible – this is useful for many applications because the receiver of the compressed data can perform decompression to obtain the original signal.

We have also shown that wavelets analysis can be used to improve the Quality of Service (QoS) of a telecommunication network in our paper [19]. QoS refers to the set of technologies and techniques for managing network traffic with the goal of providing a certain level of performance to a data flow in a network.

The QoS issues that concern telecommunication systems are delay, jitter, loss rate, throughput and network resource availability. We shall look at how wavelet analysis can affect each in turn.

Delay is the elapsed time for a packet to traverse the network from the source to the destination. Since the packet is made smaller by the application of wavelet analysis, the delay will decrease because the data will require smaller bandwidth – this is beneficial for telecommunications.

Jitter is defined as the variation in delay encountered by similar packets following the same route through the network. The jitter requirement only affects real-time streaming applications because this QoS requirement arises from the continuous traffic characteristics of this class of applications. Services intolerant of delay variation will usually try to reduce the delay variation by means of buffering. In time critical systems late data arrivals can make the data useless because the data is no longer considered synchronous – this would result in receiver buffer underflow; likewise early arrival can lead to receiver buffer overflow. Indeed large delay variation (jitter) degrades the performance of the data stream buffer in the receiver and the smoothness of the data flow. Since the packet is made smaller by the application of wavelet analysis, the jitter will decrease because the delay will decrease – this is beneficial for telecommunications.

Loss Rate refers to the percentage of data lost among all the delivered data in a given transmission time interval. Any packet loss and packet delay can degrade the quality of the data at the receiver. Large packet delay is equivalent to packet loss because in real-time applications new data overwrites old data. In real-time applications we might tolerate a strict amount of data loss. Since data loss is not a function of packet size, wavelet analysis would have no serious affect on data loss and would not necessarily benefit telecommunication.

Throughput is defined as the rate at which packets are transmitted in a network. Since packets are made smaller by the application of wavelet analysis the throughput will increase – this is beneficial for telecommunications.

Network resource availability is the infrastructure associated with the transmission of data e.g equipment, power, etc. In some systems it is absolutely imperative to have good network resource availability because the generated traffic may be critical. In some applications the speedy delivery of data can be an extremely important issue as well as reliability in terms of data delivery. Since network resource availability is not a function of packet size, wavelet analysis would have no affect on network resource availability and would not necessarily benefit telecommunications.

V. CONCLUSIONS

In this paper we described a time series analysis technique called data wavelets to derive trends in the data – this acts as a form of lossy data compression. Wavelet transforms have advantages over traditional Fourier transforms for representing functions that have discontinuities and sharp peaks, and for accurately deconstructing and reconstructing finite, non-periodic and/or non-stationary signals.

We have shown from our previous research that wavelets are a suitable tool for processing large and noisy data sets for higher level processes to make use of.

We showed that wavelet analysis facilitates clinical decision support in the form of qualitative reasoning for patient state assessment.

Using data wavelets to generate trends allowed us to develop a conceptual model of a building in sufficient detail to enable building operators to make informed decisions about how to improve the buildings energy performance.

Wavelets analysis compressed data to allow for more efficient use of network resources such as storage and bandwidth since smaller files take up less space and are faster to transmit over a network. Wavelets also facilitate improvements in QoS.

REFERENCES

[1] Daubechies, I. (1992) Ten Lectures on Wavelets**.** Capital City Press, Montpelier, Vermont, 1992.

[2] Abramovich, F., Bailey, T., Sapatinas, T. (2000), Wavelet analysis and its statistical applications. JRSSD, (48): 1-30, 2000.

[3] Polikar, R. The wavelet tutorial. Internet Resources: http://engineering.rowan.edu/polikar/WAVELETS/WTtutorial.html, Accessed 14 October 2011.

[4] Amara Graps, Jnl of Computational Science and Engineering, vol. 2, num. 2, published by the IEEE Computer Society, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720, USA, 1995.

[5] Kendall, M., Time Series. London: Charles Griffin, 1973.

[6] Priestley, M.B, Spectral Analysis and Time Series. (Vol. 1): Univariate Series. London: Academic Press, 1981.

[7] Chan, K-P. Fu, A. W-C. (1999). Efficient time series matching by wavelets. In ICDE, pages 126–133.

[8] Nilsson, M., Funk, P. (2004). A Case-Based Classification of Respiratory Sinus Arrhythmia, pages 673–685, ECCBR'04.

[9] Singhal A., Singh, R.P., 2, Tenguria, M., Comparison of Different Wavelets for Watermarking of Colored Images, 2011 IEEE 3rd International Conference on Electronics Computer Technology (ICECT 2011), Kanyakumari, India pages 187 – 191, 2011.

[10] T. Sukuvaara, M. Sydänmaa, H. Nieminen, A. Heikelä, E. M. J. Koski, Object-Oriented Implementation of an Architecture for Patient Monitoring, IEEE Engineering in Medicine and Biology, pages 69-81, December 1993.

[11] Salatian, A., Adepoju, F., Odinma, A. (2010) 'A Data Wavelets Approach to Deriving Trends in Historical ICU Monitor Data', Proceedings of the 2010 IEEE Sensors Applications Symposium, pages 162 – 165, IEEE Catalogue Number: CFP10SAS-CDR, ISBN 978-1-4244-2787-1/09, Limerick, Ireland, 23-25 February, 2010.

[12] Salatian, A., Adepoju, F., Oriogun, P. (2011) 'Using Data Wavelets for Data Mining to Enhance ICU Clinical Decision Support', Proceedings of 1st International Conference on Intelligent Systems & Data Processing (ICISD 2011), pages 373-378, ISBN 978-1-6123-3002-0, Vallabh Vidya Nagar, India, January 24 - 25, 2011.

[13] Salatian, A. & Adepoju, F. (2011) 'Towards an ICU Clinical Decision Support System Using Data Wavelets', International Journal of Computer Applications, Number 1, Article 8, pages 37 – 43, ISBN: 978-93-80749-02-6, 2011, April 2011.

[14] Kohane I. S., Temporal Reasoning in Medical Expert Systems, Boston University School of Medicine, Boston, MA 02118 U.S.A, 1986.

[15] Salatian A. (2003) Interpreting Historical ICU Data Using Associational and Temporal Reasoning, 15th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2003), Sacramento, California, USA, pages 442-450, ISBN 0-7695-2038-3, ISSN 1082-3409, 3-5 November 2003.

[16] Allen J. F. (1981) Maintaining Knowledge about Temporal Intervals, Department of Computer Science, University of Rochester, Rochester, NY 14627, TR 86, 1981.

[17] Salatian, A., Adepoju, F., Taylor, B., Odinma, A. (2011) 'Using Wavelets to Create Conceptual Models from Building Monitor Data', Proceedings of the 2nd International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC 2011), Volume 1, pages 235-240, ISBN 9781-93633-8-20-7, Orlando, Florida USA, March 27th - 30th, 2011.

[18] Salatian, A., Adepoju, F., Oborkhale, L., Data Compression for ICU Telemedicine to Reduce Transmission Costs and allow Qualitative Reasoning, Proceedings of the 2nd International Conference on Software Technology and Engineering (ICSTE 2010), Volume 1, pages 242-246, IEEE Catalogue Number: CFP1030L-PR, ISBN: 978-1-4244-8665-6, San Juan, Puerto Rico, USA, 3-5 October 2010.

[19] Salatian, A., Adepoju, F., Oborkhale, L., Using Wavelets to Improve Quality of Service for Telemedicine, International Journal of Advanced Science and Technology, Volume 27, pages 27 – 34, ISSN: 2005-4238, February 2011.

Physics of Open Systems: Generation of System Knowledge

Boris F. Fomin, Tamara L. Kachanova
Department of Automation & Control Processes, Saint-Petersburg State Electrotechnical University "LETI"
Saint-Petersburg, Russia

ABSTRACT

Information technologies of Physics of Open Systems automatically generate reliable, theoretical, system knowledge using the data and which are collected by the empirical science.
Descriptive technologies begin the cycle of production of system knowledge. They create an empirical base of generation of system knowledge. Application of these technologies requires participation of subject area experts.
Projective technologies finish the cycle. Their task is transferring the system knowledge into application sphere. The specialists of subject area work with projective technologies.
Technologies of analytical kernel and constructive technologies of Physics of Open Systems generate directly the system knowledge. The knowledge generation is performed automatically, without experts.
Technologies of analytical kernel generate knowledge on the level of system ontology. This knowledge is about organization of the space of senses of the system, semantic organization and semantic activity of the system, semantic forms of all qualitative determinacies of the system whole, completeness, significance, reliability, applicability of the obtained theoretical knowledge, and about constructive definition of the system in each its qualitative determinacy and each actual state.
Constructive technologies generate knowledge on the level of subject ontology. This system knowledge forms solution resources of the system problems: cognitive schematic descriptions of local and global states of the system; cognitive schematic descriptions of intrasystem mechanisms; analytical descriptions of dependencies disclosing relations in the inner world of the system; constructive formal definitions and rational explanations of emergent properties of the system.

Keywords: system knowledge, ontological modeling, communicative modeling, states modeling, solutions resources, system ontology.

1. INTRODUCTION

Open systems interactively exchange with environment by substance, energy and information. The fundamental laws, still unknown to science, define structure, behavior, states and properties of open systems. Complexity of open systems is tightly related with growth of their scale and with heterogeneity of arising structures. The interdependence of heterogeneous components becomes the main problem for understanding the complexity of open systems.
Natural, social and anthropogenic systems must be considered as open systems. The new scientific ideas, new mathematical structures, new technologies for scientific understanding, which are aimed at the reconstruction of global system behavior and rational explanation of both regularities and mechanisms of formation of system properties and states, are needed to overcome the complexity of such systems. *The key purpose is production of scientific knowledge about open systems.* Solving future global problems depends on the success on this research direction.
As initial presentation of open natural, social and anthropogenic systems their empirical descriptions are being used. In the empirical description the system is given in its actual states. A complete representative empirical description of open system is a unique source of objective information about its natural scale and complexity.
For generating system knowledge on the basis of the empirical descriptions the scientific theory is needed. The Physics of Open Systems is such theory. Scientific understanding and rational explanation of the essence of open systems defined by empirical descriptions are supported by the technologies of Physics of Open Systems.
The Physics of Open Systems is a post-cybernetic paradigm of systemology that proposes a new approach to solving problems of cognition, scientific understanding and rational explanation of the complexity phenomenon of open systems [1]. It considers open systems in their natural scale and complexity. It has the deep methodological foundations, adequate metatechnology and its own theoretical apparatus. Its ideas, approaches and methods are implemented in information technologies providing automatic generation of reliable theoretical knowledge on open systems.
The systemological conception and main scientific statements of the Physics of Open Systems were developed in St.-Petersburg State Electrotechnical University "LETI" in 1992-2003 [2-7].
The project "Physics of Systems" has been stated in 2003, and the consortium "Institute of Strategic Developments" (http://www.isd-consortium.ru) is carrying out works on this project. Authors, developers of technologies and participants of applied approbations of Physics of Open Systems are the consortium members. The results go through approbation on six directions: computational toxicology, genomics, system biology [8, 9]; theoretical medicine [10-12]; the solar-terrestrial physics [13-16]; security [17-18]; generation of scientific system knowledge [19]; knowledge management.
A general review of systemological conception of the Physics of Open Systems is given in the present paper. Within this conception, the technologies of Physics of Open Systems are considered in context of the problems of generation of scientific knowledge about open systems in accordance with their empirical descriptions. On the basis of information technologies of Physics of Open Systems, a process of automatic generation of the complete, reliable, theoretical, system knowledge on the levels of system and subject ontology is described.

2. SYSTEMOLOGICAL CONCEPTION OF PHYSICS OF OPEN SYSTEMS

The Physics of Open Systems has four levels of organization. *Methodological foundations* define conception of cognition paradigm of open systems in the form of logically complete system of concepts, disclosing senses of systemogenesis.
Metatechnology realized conception of cognition paradigm of open systems at the normative language of constructive expression of their senses.
Constructive theory offered methods for generation of concepts of systems' normative language, considering these concepts as formal objects.
Information technology was realized in algorithmic systemology, and its formal objects obtained adequate computable representations.

Methodological foundations
Physics of Open Systems is represented in the methodological models, defining a system on levels of vision, cognition, understanding, and explanation of system senses [4-6], fig.1.

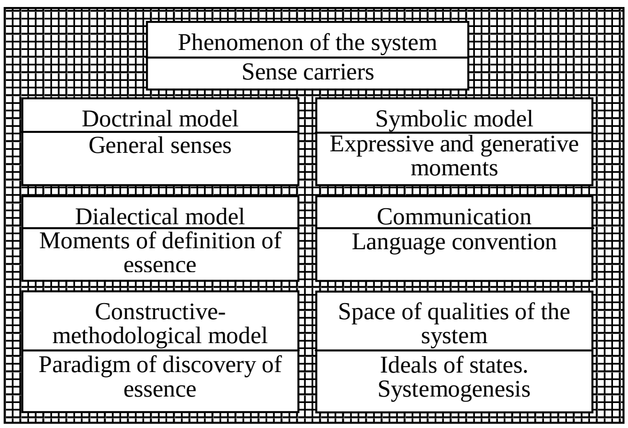


Fig. 1. Methodological foundations

The Physics of Open Systems, entirely based on the world of experience, builds a philosophical system of doctrines and fundamental concepts about senses and relationship between senses of the system.
A doctrinal model defines a concept of the system through its representations in complete abstract forms of the system senses.
A dialectical model defines doctrines of cognition of the system through basic concepts related by dialectical triads, which form the unified, holistic and hierarchically arranged system of concepts.
A constructive-methodological model sets the sequence of steps of understanding the senses of the system on the basis of measure category and universal principle of symmetrization. It creates structural images of the system.
A symbolic model introduces a set of sense relations which are transferring characteristic intrasystem regularities by way of generative and expressive moments.
Communication creates a space of concepts of the system, where the scientific knowledge about the system is expressed by the words of language of systems. Triads of the concepts, which are stating dialectical relations reflecting an organization of the system in the world of its senses, disclose contents of the words of language of systems.
Space of qualities of the system represents the complete space of the system senses and explains all possible actual and potential manifestation forms of phenomenon of the system. The configuration of the space of qualities of the system is established by mechanisms of systemogenesis which form ideals of the system's qualitative determinacies and states of the system.

Metatechnology
The main purpose of the metatechnology is the realization of both ideas and principles of the Physics of Open Systems in adequate scientific apparatus organized into unified schema of cognition, understanding and explanation of general mechanisms of systemogenesis [3], fig.2.

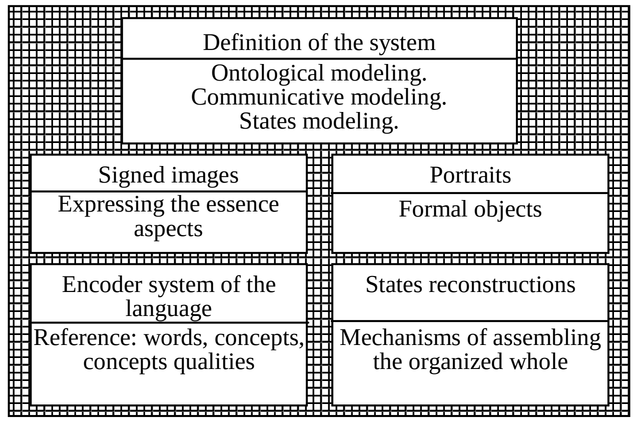


Fig. 2. Metatechnology of Physics of Open Systems

In the metatechnology of ontological modeling, a pure sense of the system, represented by its symbolic model, is transferred into signed images, for which the formalized concepts expressing different aspects of the system essence, serve as means of expression. The sense of the sign transfers into its subject value embodied in formal objects of portraits of the system.
In the metatechnology of communicative modeling, the language of systems turns into a holistic theoretical system of scientific knowledge, obtains the status of an encoder system defining a complete set of semantic associations between words and concepts of the language of systems, which are introduced at the level of communication. Lexical structure of the language is enriched at the level of reference by both the qualities of concepts and their meaningful estimations generating the constructive definitions of concepts through relations with the objects of portrait images of the system.
In the metatechnology of states modeling, the models of states of the system, where the system is represented as an organized whole, are introduced; the models of rational explanation of the properties conditioned by the whole system are being built; the models of mechanisms responsible for formation of the global system properties are being created; the models of properties of every parameter in each concrete actual state are being defined.

Constructive theory
There are three theoretical chapters in Physics of Open Systems [2, 20-21]. These are presented below.
A theory of ontological modeling creates formal models defining the system in its qualitative features, properties and organization of the space of its states. The basis of this theory is being formed by the axioms of the systems and by the principles of systemogenesis which create the ideal objects with specific symmetries of forms of system organization. The task of this

theory is to establish regularities revealing relations of the ideal objects. These ideal objects and established regularities are implicitly applied to the description of the empirical reality of concrete systems, fig. 3.

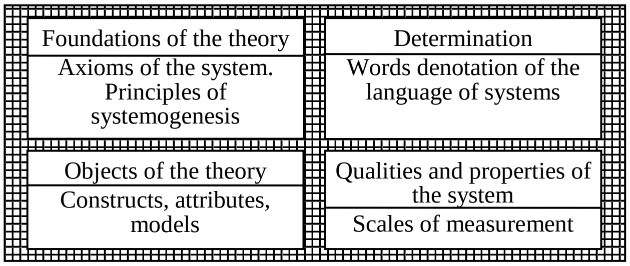


Fig. 3. Constructive theory

A theory of communicative modeling develops the language of systems at the level of determination by introduction of measures in semantic space of the system. The language gets ability to distinguish and explain the properties of concrete system, to express the scientific knowledge about the system and to estimate the value and utility of this knowledge.
A theory of states modeling explores the models of states of the system; mechanisms of assembling the states; classes of the states; emergent properties of the system; mechanisms forming a variability of both the properties and values of parameters of the states; attributes of elements of the system organization. For the purpose of measuring objects of the theory of states modeling, the measures establishing rules of their mapping onto the special qualitative and quantitative scales are created.
Constructibility of the theory provides computability of all objects, elements, concepts, qualities and properties, introduced by the metatechnology and proves consistency of the procedures of their calculation.

Information technology
Processes of cognition, understanding and explanation of the essence of concrete systems are realized in the information technology [2], fig. 4.

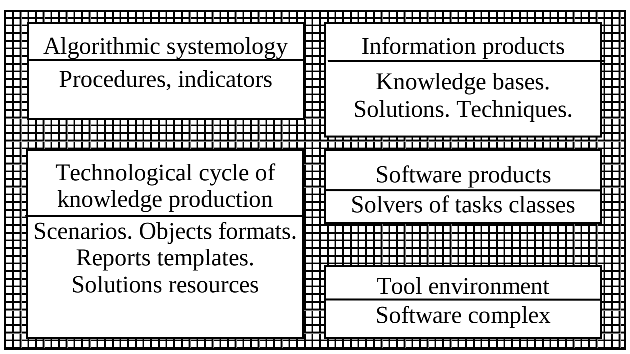


Fig. 4. Information technology

An algorithmic systemology provides computability of all objects of the metatechnology and formal language of the system. It provides a basis for creation of the information technology with the special mathematical methods, effective computational procedures and apparatus of technological indicators. Such technology is created.
A technological cycle of knowledge production establishes use of procedures for the informational technology in accordance with the universal scenario of system knowledge generation. Stages and steps of the scenario create formal objects of the technology which are mapped in normative formats. The technology defines regulations for representation of the obtained knowledge in the normative documentary reports. Resources of applied tasks' solutions on the basis of the system knowledge are the outcome of the technological cycle.
A tool environment in automatic mode generates the solution of all tasks of producing, formatting and representing the knowledge, creating and providing the solution resources for users. Information technology creates information and software products. It generates the complete system knowledge about applied problems, as well as forms and represents the knowledge bases, models and techniques for solving system problems on the basis of the knowledge. For automatic solving of typical system tasks, the technology creates software tools, namely tasks solvers.

3. TECHNOLOGIES OF PHYSICS OF OPEN SYSTEMS

The technologies of the Physics of Open Systems are represented by the analytical kernel, descriptive, constructive and projective components, described below.
An analytical kernel is the basis of the technologies of the Physics of Open Systems. It comprises the ideas, approaches and methods of ontological modeling, communicative modeling and states modeling. Environment of the analytical kernel solves two tasks: 1 – description of the system problems in empirical data and concepts of subject area (*descriptive component*); 2 – representation of the system knowledge for solving these problems in forms appropriate for users (*constructive component*) and application of the obtained knowledge for development and estimation of variants for solving the applied problems (*projective component*), fig. 5.

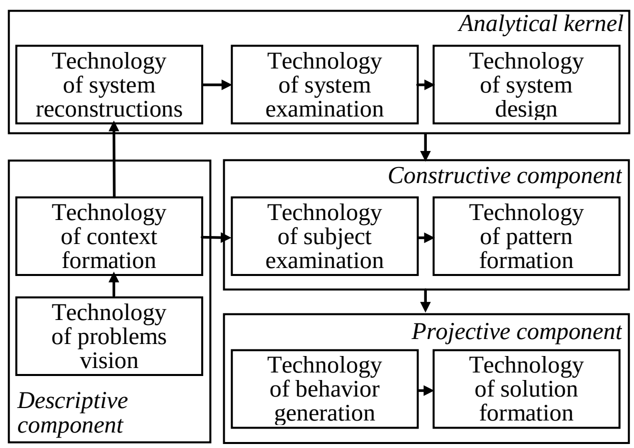


Fig. 5. Components of technologies of Physics of Open Systems

Analytical kernel
The technologies of analytical kernel serve as an "intellectual machine" for generation of the system knowledge. It consists from technology of system reconstructions, technology of system examination and technology of system design [1, 19-21].
A technology of system reconstructions generates, organizes, forms and represents an intellectual resource of the system knowledge.
A technology of system examination executes the analysis of sense of the intellectual resource (estimates the scientific system

knowledge from the perspective of its reliability, completeness, applicability, significance and actuality) and creates a cognitive resource of the system knowledge.

A technology of system design synthesizes the adequate models of states of the system, researches the emergent properties of the system and generates, organizes, forms and represents a technological resource of the system knowledge.

Descriptive component

The technologies of descriptive component are connected with the analytical kernel by the abstraction channel, in which the subject representation about the system, in its natural scale and complexity, is transferred onto the system level [2, 19].

A technology of problems vision provides creation and application of interfaces for the description of problems in the subject area in the form of system projects. In the subject area the description of each problem includes its isolation, interdisciplinary verbal description, structuring, stratification and organization of monitoring. The problem is represented as a system project to production of the system knowledge. The problem representation in the form of a system project is related with justification of applicability of the system approach for solving the problem, with estimation of both scale and complexity of the problem, with definition of the size of empirical data which can be represented for solving the problem, and with regulation of data delivery.

A technology of context formation is responsible for transformation of the problem description in the system project into an interpreted normative initial representation of the system (as the object of research); selection and description of measures of the system measurement; and formation of data repository about the system.

Constructive component

The technologies of constructive component are linked to the analytical kernel by the channel of concretization, in which the system knowledge is transferred onto the subject level [8, 19]. The constructive component operates the obtained resource of knowledge. It transforms system knowledge generated by the technologies of analytical kernel into the informational, intellectual, cognitive and technological resources of the solutions of applied problems.

An informational resource of solutions is a knowledge, which appears a product of the system analysis and of understanding of the empirical fact (e. g. defects and quality estimations of empirical description, level of parameters significance, relevance of both parameters and objects of observations in relation to the tasks being solved).

An intellectual resource of solutions is the families of formal models, creating a cognitive potential for research activity (e. g. system models, models of interaction, estimations of both entirety and completeness of the system knowledge).

A cognitive resource of solutions is a knowledge meant for reasoning and action. It has the translation potential, and enables creation of universal conceptual ways for scientific communication (e. g. models, objects, schemas, language of systems).

A technological resource of solutions is an objective knowledge about the system in the whole and in the parts. It provides a rational explanation of states of the system and mechanisms of its variability (states, states space).

A technology of subject examination realizes a transformation process of knowledge about states and mechanisms of the system which are expressed by the language of systems, into unified schemas of subject ontology of the system. Knowledge, generated by the technologies of analytical kernel about states and mechanisms of the system is the knowledge about the system's inner world, which does not have the subject format. Translating this knowledge into the subject formats requires the application of expression tools enabling to link the system understanding of mechanisms and states with concepts and representations related to both, mechanisms and states in the subject area.

A technology of pattern formation uses knowledge resources for choice of the knowledge elements needed for solving applied tasks. It reduces the knowledge elements to the formats taking into account the specifics of the subject description of the problem at the level of both, data and conditions for their obtaining. It offers the formalized methods for solving the problems and also the templates for grapho-analytical presentation of the results.

Projective component

The technologies of projective component use the resources of the solutions in order to create the subject interface [19].

A technology of behavior generation is responsible for

1) the construction of an objective cognitive model of the problem based on its subject ontology and quantitative forms of the system solutions;

2) the application of this model for generation of behavioral portraits disclosing the system properties through demonstration of its variability in the events, states, space and time.

A technology of solution formation forms the libraries of typical schemas for solving of applied tasks, develops and uses the service-oriented solvers for classes of the applied problems.

4. GENERATION OF SYSTEM KNOWLEDGE

The cycle of generation of the system knowledge contains four stages. On the first stage, the technologies of descriptive component create the base for generation of the system knowledge, and shape the available *empirical knowledge* about the problem (i.e. experiments protocols, scientific facts, empirical terms and dependencies).

On the second stage, the technologies of analytical kernel generate system knowledge. *The system knowledge* includes theoretical models which disclose the essence of the multi-qualitative system and explain its complexity.

On the third stage, the technologies of constructive component generate (on the basis of obtained theoretical models) the system knowledge about regularities, inner organization of systems mechanisms, and also about attributes, properties and cognitive schemas of these mechanisms. The stage is completed by construction of the solutions' resources, empirical interpretation of the system knowledge, and formation of the subject interface.

On the fourth stage, the technologies of projective component use the system knowledge to prepare a solution of the system problem.

The system knowledge is generated by the technologies of analytical kernel and constructive component. The techno-cubes, having three dimensions, serve as images of technologies of analytical kernel.

The first dimension sets the representation of the system (in the whole, in parts and in elements).

The second dimension defines the tasks (cognition, understanding and explanation) and subjects (parameters, relations structures, states and mechanisms) of technologies.

The third dimension discloses steps and key moments of solving the tasks of technologies. The spaces of techno-cubes are filled with *knowledge elements*, for which the normative formats of representation are established.

Techno-cub of system reconstructions

The dimensions of the techno-cube of system reconstructions are set by coordinates "Representation", "Cognition" and "Expression" [2, 6-7, 19], fig. 6.

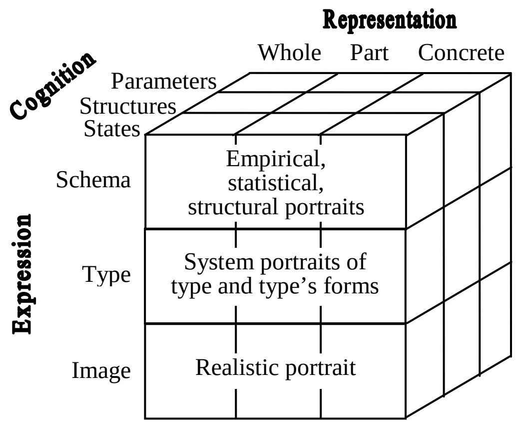


Fig. 6. Techno-cube of system reconstructions

In the "Schema" position, the system is represented at the level of both empirical fact and structures of binary relations. In the "Type" position, the system is given in semantic forms of all its qualitative determinacies. In the "Image" position, the potential of transferring semantic forms of qualitative determinacies of the system onto empirical fact is disclosed.

The technology of system reconstruction automatically produces system knowledge based on empirical description of the system. The empirical description is then transformed into abstract representation of the system in a form of signed connections graph. The graph vertices are parameters of state of the system and its environment. The graph edges are statistically significant binary relationships between parameters. The structure of the binary relationships represents multiplicity of intra-system correlations. The signs of the binary relationships define different forms of behavior of the system through variability of parameters of its state.

The first axiom of Physics of Open Systems states what the changes in all system parameters are harmonic. Out-of-balance condition of the connection graph shows heterogeneity of the system and its complexity. Connection graph with signs out-of-balance serves as a base for an automatic generation of complete set of *system models* and *models of interaction*.

Each system model determines the whole system in one of its qualitative determinacies, formed by the special system-forming mechanism. Complete set of system models determines all qualities of the system. Generation of system models begins with the finding of all *unbalanced triangles* in connection graphs. Resolving lack of balance in the connection graph is realized by finding symmetries of structures of relationships – singletons with the ability to harmonize connections between parameters. A *singleton* is an unbalanced triangle with main axial symmetry and system roles of vertices. One vertex is *special* and identifies one characteristic quality of the system. Two other vertices serve as carriers of system-forming *two-factor interaction*. All singletons with the same special vertex form a *kernel of the system model* with preservation of the axial symmetry and two-factor relationships of these singletons. The kernel determines a single quality of the system. System model with such a kernel represents the system as a whole in its one quality. The system as a whole in all its qualities is represented by the complete set of system models. This complete set discloses *complexity* inherent in the system.

The models of interaction (*doublets* and *triplets*) are generated from a variety of singletons. They define all types of structural and behavioral invariants explaining the unity of the multi-qualitative system. Higher symmetries of multi-factorial intrasystem interactions are manifested through models of interaction.

The result of the technology is knowledge about space of qualities of the system, which consists of images of family of abstract system models. In this space each system model matches a region, in which the type of qualitative determinacy of the system (particular quality of the system) is assigned. Each region covers all variety of manifestations of the assigned type of qualitative determinacy. Conceptual borders, in which this type is manifested in different forms and with different intensity, determine the structure of the region.

The technology of system reconstruction represents elements of the obtained system knowledge in six normative formats: *empirical, statistical, structural, two system portraits and realistic portrait*.

Techno-cube of system examination

The dimensions of the techno-cube of system examination are set by coordinates "Representation", "Understanding" and "Communication" [19-21], fig. 7.

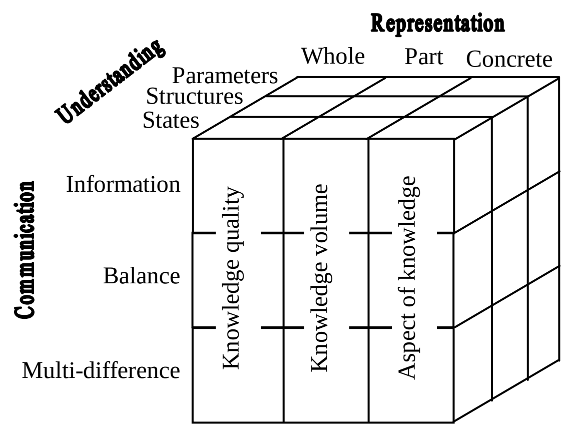


Fig. 7. Techno-cube of system examination

The position "Information" characterizes the empirical fact in its ability to generate a complete reliable knowledge about the system. The position "Balance" explains system models from the perspective of their form, completeness, homogeneity, contrast of idea expression of the system whole in each of its quality. The position "Multi-difference" estimates completeness of actualization of all types and forms of qualitative determinacies of the system.

The technology of system examination assesses generated system knowledge and constructs (based on system models) a *complete set of ideal states of the system*. It also maps each region of the space of qualities into the space of attributes and determines a set of objects with quality characteristics for the given region. The technology works with words, concepts and assessments of the *language of systems*. It uses these objects for expressing properties of concrete systems using generated system knowledge about these systems.

The technology works with different forms of representation of the system: empirical description; complete set of system models; system model of each quality; and complete family of condensed triangles.

The empirical description of the system is assessed based on its sufficiency for generation of complete system knowledge. In the complete system knowledge, the heterogeneity of the system is completely revealed i.e. unbalances are resolved, and changes in all parameters are explained by system mechanisms.

The family of system models is assessed by its ability to express completely the organization of the space of system qualities (i.e., to define all regions of the space through compact structural invariants isolating the system in each of its quality, and to give an explanation to all forms of representation of all qualities of the system through mechanisms determining variability of its states).

The *condensed triangle* is the ultimate concentrated image of the system in a quality expressed by one concrete system model. The condensed triangle serves as an instrument that maps a region of the space of qualities into the space of attributes of the system.

The main purpose of technology of system examination is a transformation of the family of system models into a set of models of the ideal states of the system. Main axial symmetry of a system model allows only two ways for concordance of its signs in agreement with the *first axiom of systems*. Each alternative gives rise to a *model of stereotype of behavior* of the system. The model of each stereotype is transformed into two models of ideal states of the system in accordance with the *individuation axiom*. This axiom establishes existence of a unique border between high and low values. A complete set of models of ideal states determines the system as a whole with all its possible qualities and all ways of manifestation of these qualities in reality.

The direct mapping of regions of the space of system qualities into its space of attributes is achieved by mapping of the set of models of ideal states on the empirical description of the system. This mapping is achieved by using condensed triangles and special scales of numerical forms for the levels of parameters values. The technology constructs scales for each parameter in each ideal model. The set of all quantitative assessments of parameters determines a *region of ideal* in the space of attributes. This region contains set of objects whose state corresponds to a concrete ideal of the system with different intensity of manifestation of qualities of the ideal in reality. A set of such objects forms a *cluster of observed objects*.

A joint set of singletons, system models, and models of ideal states forms a *complete layout of the system* where senses of the system are revealed in abstract representations. The result of the technology of system examination is knowledge on the quality of the empirical description, the quality of all system and ideal models, and the quality of mappings of regions of the qualities space into the space of attributes of the system.

Technology of system examination represents elements of system knowledge in three normative formats (*quality, volume, aspect of knowledge*).

Techno-cube of system design

The dimensions of the techno-cube of system design are set by coordinates "Representation", "Explanation" and "Attributes" [19], fig. 8.

The position "Qualities" characterizes the system as the whole (in its structural invariants, reconstructions of states, forms of system regularities). The position "Properties" discloses organization of regions of the ideals in space of attributes of the system, and also rules of conjugacy of the ideals and dominants of states of the system. The position "Differences" gives a complete explanation for each actual state of the system and its states space as a whole.

The technology of system design applies a set of clusters of observed objects for construction of *models of actual states of the system*. Each ideal state of the system is realized in different observed objects with different intensity. On the basis of each set of clusters, the technology generates *models of implementation forms of the ideal*. Each such model includes cluster of observed objects and assessments of degree of implementation of the ideal in these objects.

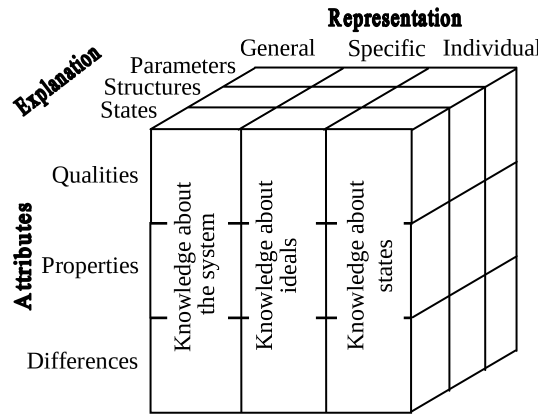


Fig. 8. Techno-cube of system design

The main purpose of the technology of system design is an automatic generation of reconstructions of actual system states which are represented in the system empirical description by states of observed objects. The *reconstruction of the actual state* arises as a result of "assembling" all the models of implementation forms of the ideals which are corresponding to this concrete state. In a model of the ideal state, the system has one quality, generated by the two-factor interaction which forms the kernel of the system model from a singleton with a common axial symmetry. In the reconstruction of the implemented state the system is multi-qualitative (as a result of that assembly) and is generated by interactions that form the kernel of the model of the reconstruction from singletons of the ideal models.

Each reconstruction acts as a carrier of knowledge about a state of the system, considered as the whole, and about *emergent properties* of the system in this state. The states of the system are revealed in reconstructions by parameters and mechanisms which characterize and determine these states. Each parameter has a set of attributes that are assessed from the position of the system as a whole by special quantitative and qualitative scales. These attributes characterize each parameter by assessments of the level of its value, predetermination of this level, importance, mobility and roughness of parameter.

At a reconstruction of a concrete actual state each particular system mechanism assists in confirmation or changes this state. The concrete role of each particular mechanism in the determination of this state is done by the reconstruction of the observed state. The complete set of reconstructions contains the knowledge of the system as a whole, as well as its emergent properties. Thus, it represents the knowledge on limitations and patterns of conjugacy on different qualities of the system in their observed states.

The results of the technology are the models for the rational explanation of the properties of each parameter in each concrete state, the properties of the system as a whole, the properties of the observed states of the system, and the mechanisms that form changes of each parameter and of the global system properties.
The technology of system design represents elements of the system knowledge in three normative formats: *knowledge about system, ideals and states*.

Generation of solutions resources

The technologies of analytical kernel form a level of the *system ontology* [8, 19] (see fig.5). Their task is production of the system knowledge. The system knowledge represents an organized set of formal constructs and their attributes disclosing the essence of the system. This knowledge has an abstract form distracted from concrete reality. It serves as a rational basis for the production of solutions' resources of the system problems. The *solutions' resources* are the system knowledge designed for explanation of the system in interpreted concepts, forms and relations.
The level of *subject ontology* is formed by the technology of context formation and the technologies of constructive component. At this level, the knowledge about the system is given in the forms of the empirical and the system knowledge, fig. 9.

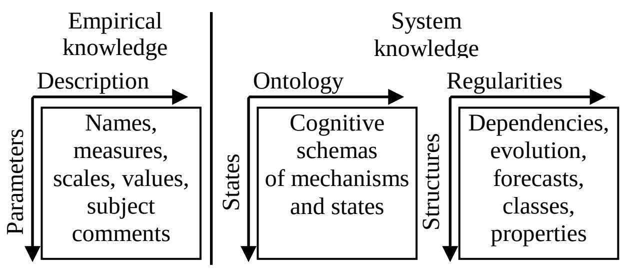


Fig. 9. Production of solutions resources

Along the "Representation" coordinate of the techno-cubes of system ontology, the system is disclosed in whole, as well as in parts and elements. On the level of the subject ontology this coordinate is unfolded into three coordinates "Description", "Ontology", and "Regularities". Along the "Description" coordinate an empirical knowledge about the researched system is fixed. Along the "Ontology" coordinate, on the basis of the system knowledge, the essential world of the system is disclosed in the key concepts and relations and expressed in *cognitive schemas*. Along the "Regularities" coordinate the most significant, necessary and stable relations between parameters, characteristics of their variability and actual and potential states of the system are represented. These relations are the external *analytical descriptions* of the essential world of the system.
The "Parameters", "States", and "Structures" categories characterize knowledge about the system in the external, inner and external-inner forms. Through the "Parameters" category, the *description of the system* at the level of empirical knowledge *is introduced*. At this level the system and its states are defined as hypotheses. Through the "States" category, the *system receives a constructive definition* in terms of the system ontology. At the level of subject ontology the *definition of the system is transformed into descriptions of its mechanisms and states*, and filled with facts of the empirical knowledge. It becomes the concrete-subject representation. The "Structures" category establishes rules, by which the development of regularities determined by the inner mechanisms is being carried out on the basis of the obtained descriptions of the inner world of the system.

5. CONCLUSION

The Physics of Open Systems in its methodological foundations relies on the following statements:

1. The system is a multi-qualitative unity of the whole;
2. The system in each of its quality is defined in some locality being a part of the whole and simultaneously being a unified whole equipped by this quality in the context of the given part;
3. Behavior of the system whole in each of its locality is dominant;
4. The system in each locality has a two-factor structure;
5. The factors of the system within locality are homogeneous and each of these factors is formed by the unique system mechanism;
6. The system in each of its qualitative determinacy is defined by the unique mechanism of two-factor interaction;
7. The system regularities are conditioned by the action of intrasystem mechanisms establishing relations between the parts and elements of the system whole.

The first four statements are directly implemented in the constructs of the technologies of the analytical kernel. These technologies produce the system knowledge about the organization of the system considered as a unified whole. They define constructs of this whole and explain their roles in forming the whole, its parts and states.
The other three statements obtain the concrete form and content in the technologies of constructive component. The constructs of these technologies express the given statements in the cognitive schemas of mechanisms and states, and also in the form of formal dependencies, properties and assessments. The solutions' resources are the end product of the technologies of the analytical kernel and constructive component. The elements of the resources form reliable system knowledge (understood, and checked), having the rational explanation, which allows a subject interpretation.
The solutions' resources are transferred to the technologies of the projective component (level of applications). At this level, the experts in the subject area work with the solutions' resources. Their task is to use these resources in solving concrete, applied problems.

Paper is prepared with financial support from ISTC within project #3476p "Unified Method of State Space Modeling of Biological Systems".

References

1. B.F. Fomin, T.L. Kachanova, "Physics of Systems is a postcybernetic paradigm of systemology" // Intern. Sympos. **"Science 2.0 and Expansion of Science: S2ES"** in the context of The 14th World-Multi-Conference "WMSCI 2010", June 29th - July 2nd, 2010 – Orlando, Florida,USA. – pp. 244-249.
2. T.L. Kachanova, B.F. Fomin, **"Technology of system reconstructions"**, "Politechnika", St. Petersburg, 2003. – 148 p. (In: Problems of innovation development, issue 2), (in Russian). ISBN 5-7325-0772-8.

3. T.L. Kachanova, B.F. Fomin, **“Meta-technology of system reconstructions”**. ETU (“LETI”) Publishing Center, St. Petersburg, 2002. – 336 p., (in Russian). ISBN 5-7629-0439-3.
4. T.L. Kachanova, B.F. Fomin, **“Foundations of the systemology of phenomena”**. ETU (“LETI”) Publishing Center, St. Petersburg, 1999. – 180 p., (in Russian). ISBN 5-7629-0293-5.
5. T.L. Kachanova, B.F. Fomin, **“A new methodological platform of systemology”**. ETU (“LETI”) Publishing Center, St. Petersburg, 1998. – 41 p., (in Russian).
6. T.L. Kachanova, B.F. Fomin, **“Symmetries, interactions in localities, behavior components of complex systems”**. ETU (“LETI”) Publishing Center, St. Petersburg, 1998. – 126 p., (in Russian).
7. T.L. Kachanova, B.F. Fomin, **“Reconstruction of complex systems behavior using experimental data”**, ETU (“LETI”) Publishing Center, St. Petersburg, 1997. – 67 p., (in Russian).
8. Ageev V., Fomin B., Fomin O., Kachanova T., Kopylev L., Spassova M., Chen C. “Physics of Open Systems: a new approach to use genomics data in risk assessment”/In **“Risk Assessment”**. Book 2 //Ed. M.G. Tyshenko. “InTech”, 2012. ISBN 979-953-307-894-5
9. V. Ageev, B. Fomin, O. Fomin, T. Kachanova, S. Shirshov, K. Turalchuk, L. Kopylev, C. Chen, “Technologies of Physics of Systems will help to realize ToxCast mission” // **The First ToxCast™ Data Analysis Summit** Hosted by U.S. EPA”s National Center for Computational Toxicology EPA Campus, Research Triangle Park NC May 14-15, 2009.
10. V.I. Nemzov, N.K. Belisheva, T.L. Kachanova, “Dependence of functional capacity in patients with bronchial asthma on some geocosmic agent fluctuations” /**Scientific proceedings, Saint-Petersburg State Pavlov Medical University**, V. VIII. №1, SMU Publishing Center, St.Petersburg. 2001. – P. 67-72. (in Russian).
11. **Inflammation mechanisms in bronchi and lungs. Anti-inflammatory therapy** //Ed. G.B. Fedoseev. Nordmed-Izdat, St.Petersburg, 1998. – P. 335-362. (in Russian). ISBN 5-93114-004-2.
12. V.I. Nemzov, T.L. Kachanova, “Mechanisms of the inflammation forming in bronchi and lungs”/ In **“Bronchial asthma”**. V.2. //Ed. G.B. Fedoseev. Medical informational agency, St.Petersburg, 1996. – P. 109-119, (in Russian). ISBN 5-85619-083-1.
13. V.O. Ageev, A.V. Araslanov, T.L. Kachanova, V.O. Samoilov, K.A. Turalchuk, B.N. Filatov, B.F. Fomin, O.B. Fomin, C.A. Shirshov, “System analysis of working conditions' influence onto personnel health status of hazardous chemical production”// Proc. of 4th “PACO’2008” Intern. Conf. **“Parallel Computations and Control Problems”**, Moscow, Russia, October 27-29, 2008. V. A. Trapeznikov Institute of Control Sciences. – P. 1-22, (in Russian).
14. T.L. Kachanova, A.A. Semipolez, B.F. Fomin, M.L. Khodachenko, “States reconstructions of the system “Sun – Interplanetary medium – Earth” // Proc. of 11th Intern. Conf. “**System analysis in engineering and management**”, St. Petersburg, Russia, June 27-29, 2007, SPU Publishing Center, St. Petersburg, P. 19-28, (in Russian).
15. B.F. Fomin, T.L. Kachanova, M.L. Khodachenko, N.K. Belisheva, H. Lammer, A. Hanslmeier, H.K. Biernat, H.O. Rucker, “Global system reconstructions of the models of solar activity and related geospheric and biospheric effects”// In Proc. of 39th ESLAB Symposium “**Trends in Space Science and Cosmic Vision 2020**”, ESTEC, Noordwijk, The Netherlands, 19-21 Apr. 2005, Eds.: F. Favata, J. Sanz-Forcada, A. Gimenez, SP-588, 2006. P. 381-384.
16. “Prediction of Solar Flaring and CME Activity by Means of COnceptual MODelling (COMOD) Technology for Reconstruction of Complex Systems”/ B. Fomin, T. Kachanova, M. Khodachenko, N. Belisheva, H. Lammer, A. Hanslmeier, H. Biernat, H. Rucker// “**CITSA-2004”. Communications, Information & Control Systems, Technologies & Applications**, 2004. P. 161-166.
17. V.O. Ageev, A.V. Araslanov, T.L. Kachanova, B.F. Fomin, O.B. Fomin, “Global reconstruction of states and behavior of open systems: social tensity in the districts and regions of Russian Federation” // Proc. of 6th “SICPRO” Intern. Conf. **“System Identification and Control Problems”**, Moscow, Russia, January 29 to February 1, 2007. V. A. Trapeznikov Institute of Control Sciences. – P. 1-17, (in Russian).
18. V.O. Ageev, T.L. Kachanova, B.F. Fomin, K.A. Turalchuk, “The accident analysis in urban water supply networks on the basis of system knowledge”/ Scientific and technical reports of SPSTU, №3 (121) 2011, SPSTU Publishing Center, St.Petersburg. 2011. – P. 297-303 (in Russian)
19. T.L. Kachanova, B.F. Fomin, “**The methods and technologies of system knowledge generation**”, ETU (“LETI”) Publishing Center, St. Petersburg, 2012. – 132 p., (in Russian). In print
20. T.L. Kachanova, B.F. Fomin, **“Introduction into language of systems”**. “Nauka”, St. Petersburg, 2009. – 340 p., (in Russian). ISBN 978-5-02-025360-5.
21. V.O. Ageev, T.L. Kachanova, B.F. Fomin, O.B. Fomin, “Language of open systems and expertise of system knowledge”// Proc. of 9th “SICPRO’09” Intern. Conf. “**System Identification and Control Problems**”. Moscow, Russia, January 26-30, 2009. V. A. Trapeznikov Institute of Control Sciences. P. 1-46, (in Russian).

Estimating Complexity of Algorithms as a Black-Box Problem: A Normalized Time Index

Pablo REVUELTA SANZ,
Electronic Technology Dpt. Carlos III University of Madrid.
Leganés, 28918, Madrid, Spain.

Belén RUIZ MEZCUA
Computer Science Dpt. Carlos III University of Madrid.
Leganés, 28918, Madrid, Spain.

and

José M. SÁNCHEZ PENA
Electronic Technology Dpt. Carlos III University of Madrid.
Leganés, 28918, Madrid, Spain.

ABSTRACT

The comparison of algorithms presents some important problems, which are critical when it is done over different hardware. In this study, we have proposed a behavioral method to reduce machine dependency, keeping the advantages of the time consumption method regarding exactness. This solution is based on a normalized time index. We provide in this study some experiments and results of image processing algorithms to validate this method.

Keywords: Machine-independent complexity, Complexity measures, Measurements and Relations among complexity measures.

1. INTRODUCTION

To measure the complexity of an algorithm, independently of the hardware on which such algorithm is running, we have developed a new method, estimating the number of operations needed to perform a given task. In computer science, measuring or estimating the complexity is a very important matter, because it allows comparing different algorithms working on the same problem, ranking them as more or less efficient. When dealing with simple algorithms, it is easy to calculate or estimate the exact number of operations. For example, to compute the factorial of *n*, we need *n-2* products. If we only take into account the dependency to the input data size and not the proportionality of performances among algorithms is disregarded, then we denote the order. The order *o(f(size of input data))* is a function of the input data size related to the number of basic operations needed to perform the evaluated task. Hence, the order is an estimation of the exact number of such elementary operations, no matter the proportional differences. For example, taking $o(kn^3)$ running time [1]. This fact allows, as said, the addition of two n^2 algorithms works as a new n^2 algorithm. In these terms, the order is an analytical but approximated estimation of the algorithms' complexity.

However, this information has some important disadvantages. On the one hand, we may find two algorithms with the same order, but consuming one of them the double of time (or computational cost) to achieve the same result and accuracy. Another problem is found in some complex processes, where it is not so evident to estimate the order. As stated in [2], "their theoretical complexity is not quite clear because they are iterative". Sometimes, the order is dependent on some extra information, or even specifically on the input data internal organization, and not only its size. This is the case of semantic dependencies, e.g., speech recognition, compressing algorithms, and others.

As every algorithm works over a finite frequency machine, the order is thus related to computing time and cost. But when the order is difficult to be estimated, some other solutions for this evaluation are proposed. The computation time is a generalized solution to measure the efficiency of the algorithms. The main counterpart of this option is the dependency with the machine. The usual way to solve this problem is to specify the computing conditions: "X ms at a 2.4 GHz Pentium IV processor, with 512 MB of DDR memory," and so on. We can observe similar proposals in image processing where efficiency is given, e.g., in frames per second over a specific hardware [3]. The advantage of this option against the order is the proportionality of the measure of the algorithm complexity. This advantage becomes a drawback when we want to compare algorithms from different studies, machines, or even years, when the computing architectures have changed.

Comparisons of the speeds of different algorithms can be found in studies, such as that by Kim et al. [4]. However, the danger of "historical slant" is evident. For example, Kim et al. used a Pentium IV 2.4 GHz PC. The processing time of their algorithm cannot be compared with that of Wei and Quan [5], who employed a Pentium IV 1.9 GHz PC, one year earlier. Year after year, the speed of computers increases and time comparisons become increasingly difficult. Even with the help of explicit computing conditions, using the running time is a very limited way to solve the problem of ranking algorithms. An interesting clue to solve these problems can be found in some studies comparing the efficiencies of different algorithms in the same machine. For example, we found the following results: "The FH-9 detector is more than five times faster than DoG and ten times faster than Hessian-Laplace" [6]. Similar comparisons can be found, e.g., in [7]. It will be useful, thus, to propose an index for both complex algorithms or situations where the order is difficult to be estimated. This index should allow inter-machine and asynchronic comparisons along the years, languages, and technologies. The solution can be found in generalizing a relative comparison that can be reproduced in any computer and at any computing conditions, remaining invariant to these parameters, without sacrificing the proportional measure of the number of operations. Thus, the main goal of this study is to propose and validate a measure of the complexity and the number of operations by means of a machine-independent method.

This paper is structured as follows. Following the Introduction Section, the proposal is described in Section II. Section III shows some tests and results regarding how this proposal works to evaluate complexities, which will be discussed in the final section. This final section also summarizes the conclusions of this study.

2. INDEX SPECIFICATION

Before brain imagery and other “hi-tech” solutions for psychologists and linguistic researchers, behaviorism was the only experimental approach to psychology. This approach treated the human being, as well as rats or other “input-output (I-O) systems,” as a black-box problem [8]. A scheme of this model is shown in Figure 1.

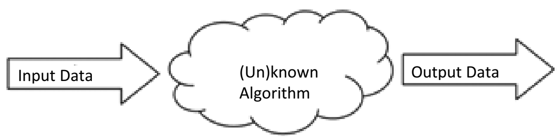


Fig. 1. Simplified behavioral scheme.

Furthermore, in psychology, some “problems” related to an underestimation of the brain complexity and freewill have been observed. However, the problems encountered when trying to reduce the human being to an “I-O system” vanishes when treating with deterministic systems, as it is the case of computers.

As shown earlier in the Introduction section, estimating the complexity is a generalized problem, and it is a crucial task to compare different algorithms.

Generalization of the behavioral scheme allows us to tackle problems where the complexity of the algorithm varies with the input or even with decisions in the middle, depending on some new input data. Moreover, some external inputs, depending on some intermediate data outputs and/or human decisions, can make the problem of complexity computation unaffordable from a classical point of view.

As stated earlier, we have assumed having a finite number of steps as a condition. However, this constraint should be removed, because there are algorithms that are theoretically able to run forever in an infinite loop depending on a periodic (or at least repeatable) input of information, retrieving periodic (or at least repeatable) output data. Examples of this kind of algorithms are information servers, feedback-controlled algorithms, IIR filters, etc. In all these cases, the concept of complexity can be applied, because we can presume (and design) the algorithms to solve these problems in many different ways and, hence, with different complexities and running times. Therefore, we require an approach to estimate complexity, which can deal with all these scenarios.

The standard notion of complexity is related to the number of operations needed to achieve a specific computation point from a starting one. We assume that the number of steps to process this path is related to the time they take to be processed. This assumption is true if the processing time is constant and the machine has a serial processor, as it will be discussed later. However, it is evident that the processing time varies with the architecture of the machine and other parameters. The consequence of this fact is the difficulty to compare similar algorithms whose complexities are given in terms of the input data size, time and specific computing conditions on which they are run, etc.

2.1 Normalization and properties

The relative differences between the machines can be eliminated by normalizing the total time (dependent on the specific hardware and task) to a standard and basic operation (supposed to depend in the same way of the specific hardware). Accordingly, we propose the following index:

$$NT = t_{task}/t_{ref} \qquad (1)$$

In the Eq. (1), the “Normalized Time” (NT) is given as the relation of the time required to achieve a specific task in terms of the time required to compute a reference and basic task.
This index presents the following properties:

- This index is machine independent (see “Assumptions” subsection).
- We have a degree of freedom to choose the reference operation, which can be useful to define a basic operation as simple as possible, with some consequences. For example, if we chose the reset of a byte, then the NT obtained for signal-filtering algorithms is dependent on the size of the input file (useful to obtain the order of the algorithm). On the other hand, for the same case, if we chose the reset of the input file, then we would obtain a NT that is independent of the size.
- Chopping algorithms. Complex algorithms perform different tasks and have some well-defined intermediate steps. In these cases, it is possible to locally measure the complexity of any definable piece of the algorithm. Moreover, we could assume that the total complexity is the addition of each partial complexity. If human interaction is needed, then this process is mandatory. However, time-computing must be stopped during human interaction, because the complexity is obviously independent of the time required for human decisions.
- Real-time systems characterization. The complexity estimation of real-time systems, as we saw, is sometimes an unaffordable task. However, if we apply the previous property, every real-time algorithm can be expressed as a variable loop of operations. Thus, each loop should be characterized with an NT_i. With these data, we can estimate instantaneous complexity, average complexity, etc.

2.2 Interpolation and order

The order is a very intuitive way to explain how the system deals with the input information. However, it has some restrictions, which has already been explained in the Introduction section. Nevertheless, it is interesting to find a path to compute the order from the NT. We are interested in this “single direction” path because the original idea is to deal with algorithms whose complexities and, hence, orders, are unknown. Moreover, the order has less information than the NT (the order is a truncated transformation of the NT). Thus, it is not possible to obtain the NT index from the order. However, we can estimate the order from the NT through interpolation. By computing different NT indices for several input data sizes and taking t_{ref} from an operation independent of the input data size, we can obtain a set of points that can be matched to an analytical curve which represents the order, as it will be shown later.

2.3 Assumptions

We must take some assumptions into account to apply the measure of the complexity proposed in this paper. In general terms, we assume that the hardware that implements the algorithm is a general purpose system. In other words:

- Serial and mono-task processing. The reference operation should be computed in some basic steps. In this case, we can assume that it is done serially. If the hardware has parallel structures to compute some kind of operations, then the normalized time ratio of the algorithm will not be the same than that of a serial architecture. Moreover, it will depend on the specific hardware structures. In different machines, the reference times may be proportional, but not in the case of the global time. Thus, we are forced to assume a serial processing structure. In serial multi-task systems, the measure of the NT of some portions of the algorithm can be distorted because of other concurrent processes. The same problem can happen with the measure of the reference time. Thus, when we are using a standard PC, it is important to average several measures, as it will be shown in the Validation section.
- No hardware acceleration. The NT is distorted not only in parallel architectures. Some hardware has kernels to accelerate specific and complex tasks. Its performance is increased when complex tasks are carried out, but is not necessarily increased when basic tasks are requested; thus, we will find distortions in the NT if the hardware has this kind of modules.

These constrictions can be summarized in a simple concept: the Strictly Serial Machine (SSM) approximation. In this paradigm, the system is supposed to have a single processing unit which, at every clock event, only processes the algorithm instructions. This approximation will be discussed in Section 4. Finally, let us clarify that our proposal is an empirical method to estimate the complexity. This constraint forces us to take some general assumptions about the hardware used to estimate the complexity (such as the SSM assumption). If these assumptions are taken into account, then this index can express a generalized "merit" for any kind of algorithm.

3. VALIDATION

In this section, we have implemented a set of tests to validate some relevant aspects of the proposed method over image processing algorithms. Every result will be discussed in the next section.

3.1 Convergence to stable value

Irrespective of the supposed independency with the hardware, modern operative systems are designed to provide pseudo-parallel computation with a concurrent procedure to select tasks that use the CPU at every time. As a result, some little variations can be found when we estimate the NT of an algorithm. Depending on the average load of the CPU, both the measurements, namely, reference and total time, may be slightly distorted.

Fig. 2 shows the stabilization of the standard deviation of an NT measurement of an algorithm (it is irrelevant at this point how does this algorithm work), with a mean value of 8.25. A total number of 20 measurements were carried out.

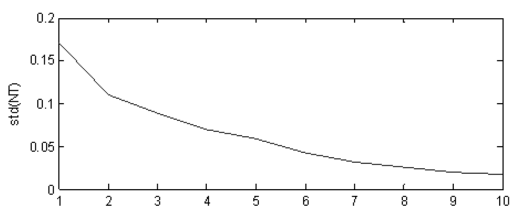


Fig. 2. Stabilization of the standard deviation (std) of the NT index against the average number of iterations.

In fig. 2, the abscissa represents the number "x" of samples averaged over the 20 samples available. The ordinate axis represents the standard deviation of the averaged "x" samples.

3.2 Order estimation

The dependency between order and computation time is well known. However, in our proposal, we can decide whether we want to obtain such dependency or not, because we can observe different references: if the order must be estimated, the reference time must be constant, and subsequently, independent of the input data size.

Fig. 3 shows the NT (reference: "for" loop of $250 \bullet 10^6$ iterations) of an image-processing algorithm against the input data size.

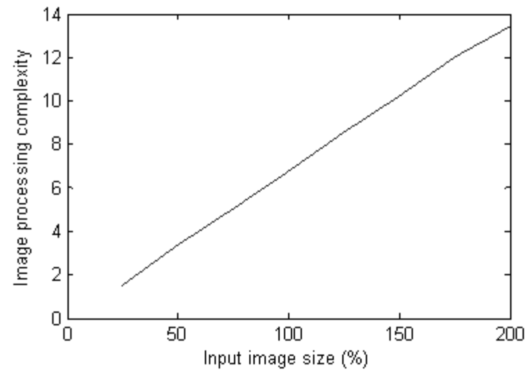


Fig. 3. Image-processing algorithm NT index. The size percentage is applied to each dimension from 25 to 200% (25% step).

Although this kind of algorithm is used to deal with N×M images, the order of this specific algorithm is linear, depending only of one dimension. In this case, the complexity only depends of the number of columns.

Fig. 4 shows the NT of a laplacian filter with 3×3 kernel applied to the same set of images, computed with a constant reference.

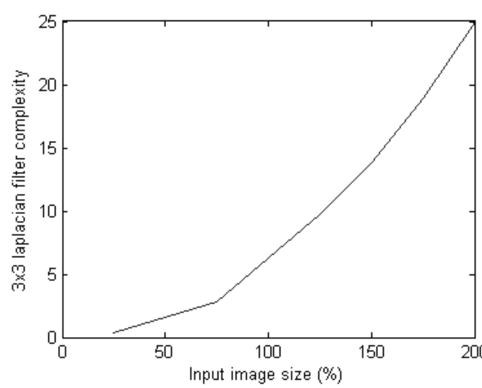


Fig. 4. NT index of a 3×3 kernel laplacian filter computed over a constant reference.

In Fig. 4, the curvature yields to an order of $o(n^2)$ for a constant kernel, which corresponds to the mathematical estimation of the order.

3.4 Same task, different algorithms with equal order

Let us consider the example given in Fig. 4. A laplacian filter can be implemented with a k×k matrix, which is convoluted with the original image of size n×n. This process yields an order of $o(n^2k^2)$. The same filter can be implemented dimension independent, and its order remains the same, but the computation time is lower. This difference is shown in Fig. 5.

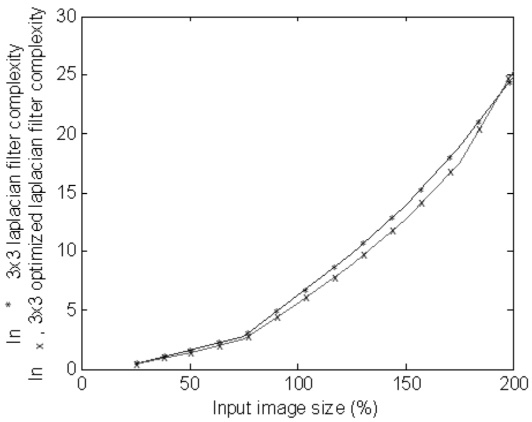


Fig. 5. Two algorithms with the same order and different efficiencies. The reference is constant (“for” loop of $250 \cdot 10^3$ iterations).

3.5 Input data size dependency or independency

By definition, the order represents the input data size dependency. Nevertheless, we may be interested only in the relative differences among a set of algorithms independently from the input data size. This option measures the “vacuum efficiency” of the algorithms, their “merit”: how well do they manage information. In other words, we measure infinitesimal efficiencies.

Fig. 6 shows the results of the NT (averaged for 5 iterations for each measure) of the algorithms shown in Fig. 5, but with reference to the time to reset the input image.

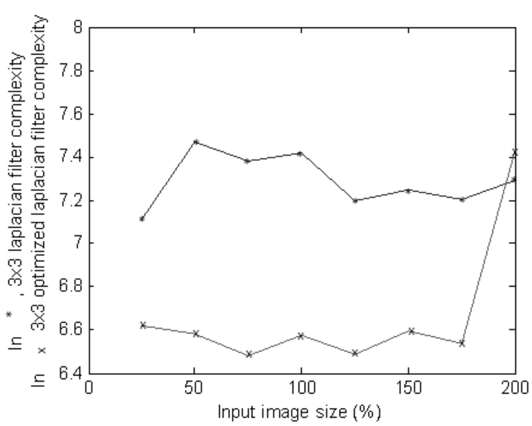


Fig. 6. Two algorithms with the same order, but different efficiencies. The reference is relative to the input data size (time to reset the input image).

Fig. 6 shows how faster is the dimension independent algorithm (in almost every case) with respect to the other one. The order for both the algorithms is the same because both lines are almost constant. Furthermore, both NTs are mostly constant because the reference time is dependent on each data input. In this representation, the NT shows that the laplacian filtering of an image with a 3×3 kernel takes around 7.2 times the operations (i.e. time) of resetting the same image, while the optimized one reduces this relation to 6.5. Thus, the first one is 10.7% faster while their orders are the same.

3.6 Different machines comparisons

The final goal of the proposal is to allow inter-machines comparisons. For that purpose, we have implemented and tested two algorithms over two different references and six machines (OS is Windows XP in every case) (Table 1). These algorithms perform two different filters over one dimensional array of 100000 random and integer values. It is irrelevant the exact functioning of the algorithms because we are only interested in relative differences among machines.

In Table 1, 20 complexity measurements are averaged for each result, and outsider results are highlighted in bold.

4. DISCUSSION

The relative comparison among the algorithms has revealed some problems. The measurement of the time, for both reference and processing time, has not been found to be completely constant. This variation comes from the SSM approximation, which is not exactly appropriated in a standard PC running over a commercial OS. This problem can be minimized averaging several samples of the same measurement, as shown in Fig. 2. In these terms, it is important to have a reference long enough for one simple reason: the error (standard deviation) of the measure of the reference time is supposed to be constant and dependent of the OS architecture, PC load, etc., and independent of the size of this reference. Thus, the error in a small reference (short-time reference) is relatively higher than that in a long reference. Finally, it can yield to total errors that are not negligible, because the reference is the divisor in the NT expression. If the reference is long enough, then the error becomes smaller even in the worst case; in the same test, the mean value was 8.25 with a maximum standard deviation of 0.17, which represents a maximum absolute error of 0.036 for this simulation.

The choice of the reference has been found to be critical to the results obtained for more than one reason. The NT can measure different things depending of the chosen reference. If we need to estimate the order, the reference has to be constant, as shown in Fig. 6. It is important to notice that the algorithm’s complexity should be described by a number or a function, representing how much time it takes to perform a task, when compared with how much time the same machine takes to perform a standard task (dependent or independent of the same input). This is the main reason why we have to specify the reference, and how it was obtained. This relativity is its strongest point, because it allows making this measure theoretically independent of the hardware, thus comparing only the software performance. The reference should then be as simple as possible but long enough to be easily implemented in any language and knowledge level. For example, we considered for that purpose resets and “for” loops; however, any other basic operation would be valid. It has been demonstrated, however, that there is an intrinsic hardware dependency. However, this aspect is not critical, as shown in Fig. 2 and Table 1, for almost all computers (except second and third). In these cases, the results for different machines (or same machine and different moments) are close enough to be regarded as the same. Nevertheless, it must be considered when using hardware accelerators or parallel systems (such as GPUs, FPGAs, DSPs, etc.) or even some standard PC architectures. It can be observed in table 1 that the second and third computers show an important independency from the other results. Likewise, the difference between them is irrelevant. The conclusion is that second and third machines present different architectures than the other ones. We can, hence, compare algorithms among similar machines (regarding their architecture), even if their computing speed is different, but not among architecturally different ones. This is the main limitation concerning the SSM assumption.

Processor	Clock	RAM	Case 1: Algorithm 1; Ref: "reset"	Case 2: Algorithm 1; Ref: "for"	Case 3: Algorithm 2; Ref: "reset"	Case 4: Algorithm 2; Ref: "for"
Intel Core 2 T5200	1.6 GHz	990 MB	23	1.9	9.1	7.8
Intel Pentium D	2.8 GHz	960 MB	**15**	**3.4**	**6**	**14**
Intel Pentium D	2.8 GHz	960 MB	**15**	**3.5**	**6**	**14**
Intel Core 2 6600	2.4 GHz	2 GB	26	1.7	9	7.7
Intel Core 2 6600	2.4 GHz	2 GB	23	2	9.1	7.8
Intel Core 2 Duo E7300	2.6 GHz	2 GB	22	1.8	8	7.4

Table 1. Hardware comparison of two different algorithms.

Indeed, this decollation works in the opposite way depending on the algorithm; in the former one ("reset" as reference), it shows a decrease in the measured complexity than that in other computers, but demonstrates an increase in the latter ("for" reference). This distortion is caused by the different speeds of processing basic and measured operations among different computers. In the same table, one can observe how the second and third computers could perform a faster operation with regard to the "reset" (−35% in the complexity measurement), but a slower operation with regard to the "for" loop reference (70% more complexity measured). This difference is constant among the machines, and the order remains the same in any one of them and just changes proportionally. This effect is shown in fig. 5 for two different algorithms performing the same task in the same machine. The proposed method arrives to measure differences that could achieve, in absolute time measurements, around milliseconds. The discussion about if differences of milliseconds are relevant or not, depends of the final application and the accumulation of these small delays. Thus, our proposal regards to relative differences, independently of the context. Even if the order is constant and independent of the input data, the final differences in the processing time can be much more relevant than those obtained from the order itself for different sizes of the input data.

We also found other sources of variations in the NT given by the algorithms, such as the internal structure of the data, in addition to its "external" properties, such as the size. It is a normal situation to find algorithms which deal with information faster if it is organized in a proper way. Examples of such algorithms were presented in the Validation section.

Regarding the order, from simple examples (Fig.s 3 and 4), we can observe how the order takes into account only the dependency on the size of the input data. In the first example, the order is $o(n)$, and thus, linear. In Fig. 4, the order is $o(n^2)$. However, there are infinite lines and curves that match with theses orders, having different running times, even with different orders of magnitude. Constant and proportional differences are hidden in the order estimation. For a linear dependency, the NT can be expressed as NT=a•n+c. The order will always be $o(n)$, irrespective of a and c and, hence, it is not possible to obtain the NT index from the order.

5. CONCLUSIONS

Measuring the complexity or efficiency of a specific procedure can be a very difficult and even an impossible task. Therefore, one can find different proposals with pros and contras. The two main proposals are the order and the computation time (or frames per seconds or Hz in some fields).

The order presents some important problems:

- With the order, the relation of the operations of the algorithms is given in terms of the size of the input data, but information about how much time or number of operations are needed is not presented. As a result, although two algorithms can have the same order, one can take much more time than the other.
- In very complex algorithms, the order may be impossible to estimate.
- Finally, in semantic-dependent algorithms, the data itself can be as important as the data size.

Furthermore, the time measure also shows important problems. Time comparisons need specific computing conditions to give an "absolute" idea of the complexity of the algorithm. Thus, time measurements do not allow direct comparisons among different researches. This drawback is heightened by the different machines on which the algorithms to be compared are implemented.

In this paper, we have proposed a method to estimate the complexity, without calculating the number of operations. We estimated the complexity using the normalized time (NT) and discussed the way to estimate the algorithm order from the NT. With this estimation, we can evaluate and compare the complexity of different algorithms in several different machines.

This method works by taking into account some assumptions about the hardware architecture, which allow decoupling or at least minimizing hardware and software dependencies. Finally, it can be applied as an empirical and generalized method to measure the merit of any "finite steps" process to solve a given problem. However, we found some inter-machine problems that should be analyzed in detail in future works.

6. REFERENCES

[1] L. Boxer, "Point set pattern matching in 3-D", **Pattern recognition**, vol. 17, no. 12, 2010, pp. 1293-1297.

[2] L. di Stefano, M. Marchionni, S. Mattoccia, "A fast area-based stereo matching algorithm", **Image and Vision Computing**, vol. 22, 2004, pp. 983-1005.

[3] J.Ch. Kim, K.M. Lee, B.T. Choi, S.U. Lee, "A dense stereo matching using two-pass dynamic programming with generalized ground control points", **Computer Vision and Pattern Recognition**, in 2005 IEEE Computer Society Conf. on, vol. 2, 2005, pp. 1075-1082.

[4] Y. Wei, L. Quan, "Region-based progressive stereo matching", in 2004 Proc. Conf. on **Computer Vision and Pattern Recognition**, vol. 1, 2004, pp. 106-113.

[5] H. Bay, A. Ess, T. Tuytelaars, L. van Gool, "Speeded-Up Robust Features (SURF)", **Computer Vision and Image Understanding**, vol. 110, 2008, pp. 346-359.

[6] V. Lepetit, J. Pilet, P. Fua, "Point matching as a classification problem for fast and robust object pose estimation", in **Computer Vision and Pattern Recogni-tion**, Proc. of the 2004 IEEE Computer Society Conf. on, vol. 2, 2004, pp. II-244-II-250.

[7] B.F. Skinner, **The Behavior of Organisms**, Copley Pub Group. ed. 1991.

[8] Middlebury Database, http://vision.middlebury.edu/stereo/ data/ 2011.

Assessing Resilience of Complex Socio-ecological Systems through Loop Dominance Analysis

Newton Paulo Bueno, Federal University of Viçosa, Dep. of Economics, Viçosa, MG, Brazil, 36570-000, phone number: 553138991566, e.mail: npbueno@ufv.br

ABSTRACT

Although resilience is an important concept to assess the vulnerability of ecological systems, literature is still unclear on how to measure it in more complex social-ecological systems. This paper presents a practical procedure to assess the degree of resilience of fisheries which builds on system dynamics analysis, specifically on the recent field of research known as loop dominance analysis. It shows that such systems may lose resilience when they exceed certain thresholds beyond which shifts in dominance of feedback loops take place and suggests how to identify those thresholds.

Keywords: socio-ecological systems, resilience indicator, system dynamics, loop dominance analysis

1. INTRODUCTION

Although the field is considered by some to be fragmented, an emerging consensus on the critical importance of the concept of ecological resilience for assessing the vulnerability of social-ecological systems (SESs) promises even greater relevancy and utility of that concept for decision makers in the near future (Eakin and Luers [2]). As defined by Holling [8], ecological resilience refers to the ability to absorb change and disturbance and still maintain the same relationships that control a system's behavior. However, questions regarding to operationalizing the concept of resilience for the analysis of complex social systems have been raised. This paper aims to help bridge that gap by proposing a practical procedure to assess the degree of resilience of SESs which builds on system dynamics analysis, specifically on the recent field of research known as loop dominance analysis. We suggest that SESs lose resilience when they cross tipping points, that is threshold conditions that, when crossed, shift the dominance of feedback loops that control the processes. For illustrating the argument, we use a very stylized system dynamics model developed by Bueno e Basurto [1] for a small Mexican fishery.

2. METHODOLOGY

2.1 System dynamics methodology and system's behavior patterns

A central message of system dynamics methodology is that structure drives behavior; system dynamics methodology explains how exactly the former drives the behavior of variables of interest in a particular system (Forrester [7]) . There are only three unique behavior patterns based on the net rate of change, or atomic linear behavior, of a variable of interest, say fish population. The first is linear behavior, when the variable grows or declines steadly. The second atomic behavior is exponential growth or decay, when the variable moves away from its initial value faster over time. The last pattern is logarithmic growth or decay, when the variable moves away from its initial condition at a slower rate over time. Thus, atomic behavior can be described by the second time derivative of the values of the variable of interest: a second derivative equal to zero indicates a constant rate of change , and positive and negative second derivatives indicate, respectively, increasing and decreasing rates of change. The three atomic behaviors (or combinations among them) can describe most behavior presented by systems. For example, a positive second derivative of the level variable fish population indicates that the fishery is in a collapse trajectory, since its dinamics is driven by a positive feedback loop, in which smaller populations generate smaller regeneration rates of the system. This allow us to define loop dominance as follows (Ford [6], p.8):

"A fedback loop dominates the behavior of a variable during a time interval in a given structure and set of system conditions when the loop determines the atomic pattern (the second derivative path) of that variable`s behavior."

2 .2 Definition of loop dominance

System dynamicists have been developing a number of new techniques for understanding complex systems' behavior. One of the more iluminating of existing studies is the classical Richardson`s [12] paper on loop dominance which we shall use in this paper to study the problem of assessing resilience in fisheries. Our conjecture is that the loss of resilience of social-ecological systems like fisheries can be seen as a bifurcation point in their dynamics through time caused by shifts in loop dominance. More precisely, we shall argue that systems lose resilience when their dominant polarity shifts from negative to positive. The development is in terms of continuous systems but a similar development holds for processes expressed in discret terms.

The polarity of a single feedback loop involving a single level x and an inflow rate $\dot{x}$ = dx/dt is defined by sign $(\frac{d\dot{x}}{dx})$, which is consistent with a more intuitive characterization as follows. The denominator of the fraction – dx – can be thought of as a small change in x, for instance a small change in fish caught in a particular fishery, which is traced around the loop until it results in a small change - $d\dot{x}$ – in the inflow rate, say in the regeneration rate of the system, $\dot{x} = \frac{dx}{dt}$. If the change in the rate, $d\dot{x}$, is in the same direction as the change in the level, dx, then they have the same sign. As $\dot{x}$ is an inflow rate and thus is added to the level, the loop is a positive one and hence reinforces the initial change In such case sign $\left(\frac{d\dot{x}}{dx}\right)$ is positive and will be negative if the polarity of the loop is negative, that is if the resulting change in the inflow rate is in the oposite direction to the change dx. If $\dot{x}$ is an outflow rate, all we have to do to extend the above definition for loop polarity is to attach a negative sign to the expression for $\dot{x}$, since variation in the same direction in the outflow, e.g in the death rate, and in the level, e.g. in the fish population means that the loop polarity is negative.

The basic mathematics of loop dominance analysis is as follows[1].

Taking the derivative of the input flow into the state of variable of interest, $\dot{x}_k$, with respect to the state of the state variable of interest, x_k, yields:

$$\frac{d\dot{x}_k}{dx_k} = \frac{\partial f_k}{\partial x_1}\frac{dx_1}{dx_k} + \frac{\partial f_k}{\partial x_2}\frac{dx_2}{dx_k} + \cdots \frac{\partial f_k}{\partial x_k}\frac{dx_k}{dx_k} + \cdots + \frac{\partial f_k}{\partial x_n}\frac{dx_n}{dx_k}$$

Which simplifies to:

$$\frac{d\dot{x}_k}{dx_k} = \sum_{i=1}^{n} \frac{\partial f_k}{\partial x_i}\frac{\dot{x}_i}{\dot{x}_k} \quad \text{for } \dot{x}_k \neq 0 \qquad (1)$$

Each term in equation (1) represents all minor feedback loops (or pathways) leaving the ith state variable and coming into the variable of interest x_k. We can decompose the effect of each minor feedback coming into the state variable x_k by doing:

$$\frac{d\dot{x}_k}{dx_k} = \sum_{i=1}^{n} \sum_{j=1}^{m(i)} \frac{\partial f_k^j}{\partial x_i}\frac{\dot{x}_i}{\dot{x}_k} \qquad (2)$$

Where m(i) is the number of minor loops that leave the ith state variable and come into the kth state variable, $\frac{\partial f_k^j}{\partial x_i}$ is the polarity of the minor feedback loop and the ratio $\frac{\dot{x}_i}{\dot{x}_k}$ represents the net changes in the ith state variable and the net changes in the kth state variable.

Thus, the system is dominated by negative feedback loops, that is presents an equilibrium behavior pattern, only if $\frac{d\dot{x}_k}{dx_k}<0$. Furthermore, if the sign of $\frac{d\dot{x}_k}{dx_k}$ shifts from negative to positive this indicates it has crossed a tipping point beyond which self-reinforcing loops start to control the process and the system loses resilience, as we have remarked before.

These ideas suggest a practical procedure for assessing resilience in fisheries which can be summarized in the following steps:

1) Identify the variable of interest (x) that will determine feedback loop dominance in the major loop of a system dynamics model, that is the loop that drives the dynamics x→dx → dx/dt → x, and compute the ratio $\frac{d\dot{x}}{dx}$ for the actual observed conditions of the system and over a chosen reference time interval, attaching a negative sign to $d\dot{x}$ if $\frac{dx}{dt}$ is an outflow rate.
2) Identify a control parameter which affects an auxiliary variable and can vary the gain of the major loop . Use the parameter to vary the gain of the major loop until $\frac{d\dot{x}}{dx}$ changes sign in the reference time interval.
3) Compute the value of the variable of interest at the point where $\frac{d\dot{x}}{d\,x}$ changes sign and compare this value to the equilibrium value obtained in the last run before $\frac{d\dot{x}}{d\,x}$ changed sign; tipping point of the variable of interest is in this interval.
4) Repeat steps 2 through 5 for the other loops
5) Select the larger among the critical values of the variable of interest as the critical resilience level of the system and compare this to the actual observed level (or the simulated level at the actual observed conditions of the fishery); resilience degree is computed as:

$$\frac{\textit{actual observed level} - \textit{critical resilience level}}{\textit{critical resilience level}} \times 100$$

In section 3, we apply this procedure to a system dynamics model developed for a small Mexican fishery.

2.3 The model for the Seri fishery

The Seri fishery is one of many Mexican small-scales fisheries located in the Gulf of California. It exploits the Callos de hacha (CDH), a sessile bivalve mollusk that lives buried in sandy bottoms. The entire Seri CDH fishery takes place inside the Infiernillo Channel, a long, narrow, and shallow body of water flanked on the east by the Mexican state of Sonora. The basic model for Seri fishery is formalized as the simple stock-flow structure presented in Figure 1. The equations of the complete VENSIM model and other information regarding parameters description are provided in Bueno and Basurto (op. cit).

[1] See Mojtahedzadeh et al. [9].

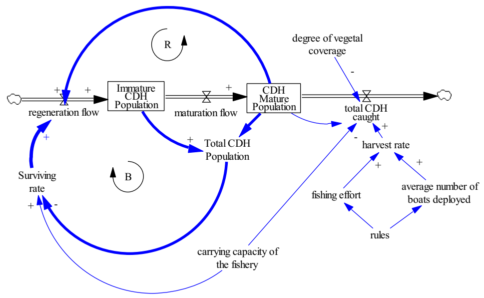


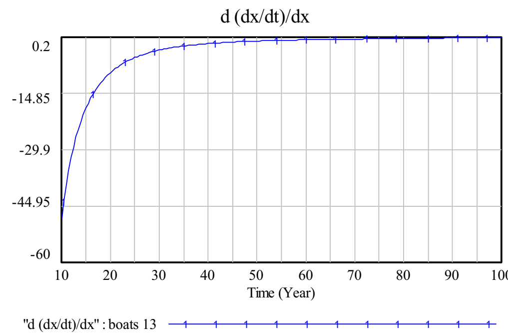
Figure 1. A simple model for CDH population dynamics. Variables inside the boxes are the state variables of the system while the others are either parameters, like the carrying capacity of the fishery, estimated by extrapolating observation in a sample of the total area, or auxiliary variables, like the harvest rate. The harvest rate is given by institutional rules which determine fishing effort, that is the number the days fishermen go to sea, and by the capture capacity, given by the average number of boats deployed in the fishery, which also depends on the specific institutional rules adopted by the community. Positive signs indicate there is a direct relationship between the variables and negative signs, an inverse one. Thus total amount of CDH caught increases with the harvest rate and with CDH density in the channel and decreases when the size of the area of the channel covered by marine vegetation rises. The R mark indicates the presence of a positive or self-reinforcing loop. Hence, increases (decreases) in CDH Mature Population increase (decrease) regeneration rate and thus the CDH Mature population next period. The B mark indicates the presence of a negative or balancing loop. Regeneration rate, thus, depends on the fertility rate of female CDH and on the survival rate given by the Beverton-Holt equation: the larger the total CDH population (CDH) in relation to carrying capacity (CC), in the x axis, the lower the survival rate, in the y axis, and the smaller the CDH population next period. The sign of the loop, that is whether it is of self-reinforcing of balancing type, is obtained by multiplying the signs of the relationships included in the loop. The system may collapse if the flow of total CDH caught exceeds a safety level beyond which the regeneration rate next periods will be lower than total harvests, which leads the self-reinforcing loop (in its collapse mode, in which decreases in CDH population lead to further decreases) to dominate the systems dynamics.

3. APPLICATION OF THE LOOP DOMINANCE PROCEDURE TO ASSESS THE RESILIENCE OF THE SERI FISHERY

For assessing resilience of the Seri fishery, apply the procedure proposed in section 2 as follows.

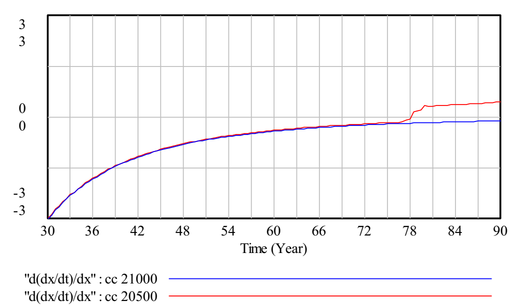
1) Variable of interest is the Mature CDH Population, which is a key variable for the system regeneration; the ratio $\frac{d\dot{x}}{dx}$ over the chosen reference time interval is shown in Figure 2.

Figure 2: ratio $\frac{d\dot{x}}{dx}$ for a 13 boats fleet size



2) Parameter *carrying capacity* is chosen, assuming that this parameter influences positively the regeneration capacity of the fishery; the evolution of the ratio $\frac{d\dot{x}}{dx}$ is shown in Figure 3.

Figure 3: Loop dominance shift caused by variation in the carrying capacity



3) System undergoes a loop dominance shift somewhere in the interval cc 20500- 21000 millions of CDHs, and the tipping point (at cc = 20500), that is the point beyond which $\frac{d\dot{x}}{d\,x}$ changes sign from negative to positive, is at time 78. Critical Mature Population is somewhere in the interval equilibrium population at cc=21000 and the Mature Population at the tipping point. Thus, a estimative for the critical mature population based on the carrying capacity is 3282 < critical mature population <4252 .
4) Procedure is repeated for the parameters average boats deployed, productivity per man and fishing effort. Results are as follows. Tipping points: fleet size: year 89; productivity: year 67, and fishing effort: year 82. Critical mature CDH population forecasts: fleet size:

15,9 to 16 boats → 4081 Mature CDHs < Xc < 4255 Mature CDHs; productivity: 660 CDH/man/day and 670 CDH/man/day → 4048 Mature CDHs < Xc < 4282 Mature CDHs; fishing effort: 50,5 to 51% → 4075 Mature CDHs < Xc < 4308 Mature CDHs.

5) resilience degree is computed as:

$$\frac{observed\ level - critical\ resilience\ level}{critical\ resilience\ level} \text{ x } 100$$

The critical resilience level of the fishery (the larger one) is 4308 million of mature CDH and, hence, the resilience degree of the system is :

$$\frac{\text{equilibrium observed level} - \text{critical resilience level}}{critical\ resilience\ level} \text{ x } 100$$

$$= \frac{8900-4308}{4308} = \sim 107\% .$$

Which means that the observed value in the Seri fishery is a little more than twice the critical resilience value.

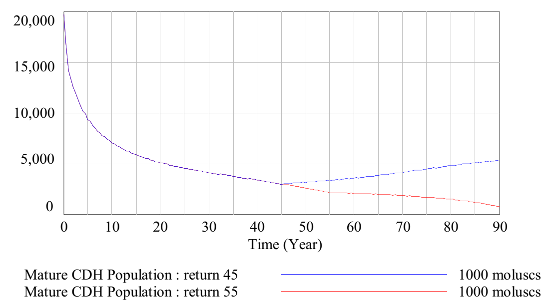
4. DISCUSSION

Formal loop dominance analysis tools are still in their formative phase. One of the main criticisms raised against these approaches is that they do not provide a unified theory that automatically provide modelers with dominant structures. While much progress is expected to be reached in the field, this is likely to remain as a permanent shortcoming of the system dynamics methodology, given the analytical intractability of non-linear high order actual social-ecological systems. So it seems that at least in part loop dominance analysis is likely to remain as much as science as art.

Despite its shortcomings, however, simple analyses such as the one presented in this paper might be very helpful for resource users.

For instance, assume that fleet has grown from a size nearly below the tipping point of the system, say the one obtained at the fleet of 15 boats, to a size nearly above that point, in which 17 boats are deployed. Assume also that fishermen, after realizing that the fleet is too large, decide to reduce it to the original level after 45 (return 45) or 55 years (return 55) as depicted in Figure 4. In the second scenario, mollusk population has fallen below the system´s tipping point (4308000 mature CDHs in the Seri fishery), which has led the fishery entering into an endogenously driven downward spiral. Once system is stuck in that path it will remain there, even if the fleet size returns to the *status quo ante*, that is to the previous set of practices and institutional rules. Thus, in order to prevent their system to collapse, fishermen would have to agree on a new set of rules restricting fishing effort until the mollusk population has been restored to its sustainable levels

Figure 4: Effect of fishermen's institutional responses near the tipping point



Several recent studies using the system dynamics approach have attempted to apply the concept of resilience to study processes of dominance shifts in different complex systems. A representative sample of these works is as follows. Rudolph and Repenning [13] show that the accumulation of routine events can shift organizations from a resilient, self-regulating regime which off-sets the accumulation of interruptions in to existing plans and procedures, to a fragile self-escalating regime that amplifies them. Taylor and Ford [16] explain why the accumulation of tasks in the development phase of new projects in an organization can generate ripple effects during the completion of these projects and thus lead them to failure. Sengupta et al. [14], in another context, show how an irrigation systems can lose resilience due to insufficient maintenance of equipment. Ford [5], finally, suggests that electric companies may be trapped in a spiral of losses for expanding capacity ahead of demand beyond a certain critical level. The basic idea of all these works is that systems lose resilience when they exceed certain thresholds in which shifts in dominance of feedback loops take place. Given the commonalities among the processes it seems plausible, while this is still an unexplored field by economists, that approach might also help to better understand regime shifts in economic processes as long wave cyclical movements in investments and output (Sterman [15]) and perhaps structural breaks in economic activity, as the recent financial crisis.

5. CONCLUSION

Usually one assumes that social-ecological systems respond to gradual change in a smooth way, but sometimes there are drastic shifts – bifurcation points - which are typically hard to predict in advance, and expensive or even impossible to reverse (Folke et al. [3]). Mostly because measurement or predictions of thresholds in SESs have low precision, the precise meaning of resilience and its identification remain a subject of debate. This work has attempted to bridge that gap by providing a relatively simple procedure to assess resilience in SESs based on loop dominance analysis.

Once a system dynamics model has been calibrated for a particular SES, the proposed procedure allow to identify critical intervals for the population which would prevent systems entering into downward trajectories as typically happens in loss of resilience processes. Those values thus could then be considered as benchmarks by users, who might assess resilience by forecasting actual population directly from collected samples in different parts of the systems.

Computing such benchmark values is a potentially useful tool because it might help to develop indicators of gradual change and early warning signals of loss of ecosystem resilience by monitoring only (at most) a handful of key ecosystem variables. To be aware of those signals is important since, as literature has related fishermen around the world prefer, simply and pragmatically, to grow the fleet until there is compelling evidence to stop (Morecroft, [10], p. 341). However, as experience that people have before crossing the tipping point is likely to misguide them when their systems start to be driven by positive downward feedback loops (Moxnes [11]), the level of fish population that they intuitively consider as a compelling evidencefor reducing their fishing effort may indeed be quite bellow from the critical0 safe level of resource use.

6. ACKNOWLEDGMENTS

I acknowledge the CNPq and the FAPEMIG for their financial support.

7. REFERENCES

[1] Bueno, N.P. and X. Basurto (2009) Resilience and collapse of artisanal fisheries: a system dynamics analysis of a shellfish fishery in the Gulf of California, Mexico. *Sustainability Science*, vol. 4 (2).

[2] Eakin, H. and A. Luers (2006) Assessing the vulnerability of social-environmental systems. *Annual Review of Environment and Resources*. 31 (365-94).

[3] Folke, C; Carpenter,, S.; Elmqvist, T, Gunderson, L.; Holling, C.S., and B. Walker (2002) Resilience and Sustainable Development: Building Adaptive Capacity in a World of Transformations. *Ambio*, Vol. 31, No. 5 (Aug., 2002), pp. 437-440.

[5] Ford, A. (1996) System Dynamics and the electric power industry. *System Dynamics Review*, vol. 13 (1), Spring.

[6] Ford, D. (1999) A behavioral approach to feedback loop dominance analysis. *System Dynamics Review*, vol. 15 (1), Spring.

[7] Forrester, J. W.(1961) *Industrial Dynamics*. Portland, OR.: Productivity Press.

[8] Holling,C.S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecological Systems*, 4: 1-23.

[9] Mojtahedzadeh, M., Andersen, D., and G. Richardson (2004) Using digest to implement the pathway participation method for detecting influential system structure. *System Dynamics Review*, vol. 20 (1), Spring.

[10] Morecroft, J. (2007) *Strategic modeling and business dynamics – a feedback systems approach*. Chichester, England: John Wiley & Sons Ltd.

[11] Moxnes, E. (2000) Not only the tragedy of the commons: misperceptions of feedbacks and policies for sustainable development. *System Dynamics Review*, 16 (4).

[12] Richardson, G. P. (1995) Loop polarity, loop dominance, and the concept of dominant polarity. *System Dynamics Review*, vol. 11 (1), Spring.

[13] Rudolph, J.W. and N. Repenning. 2002. Disaster dynamics: understanding the role of quantity in organizational collapse. *Administrative Science Quarterly*, vol. 47 (1), March.

[14] Sengupta, N.; Swati, S; and E. Ostrom. 2001 Sustainability, equity, and efficiency of irrigation infrastructure. In: Constanza, R. *Institutions, ecosystems, and sustainability*. Boca Raton: Lewis Publishers.

[15] Sterman, J. (1985) A behavior model of the economic long wave. *Journal of Economic Behavior and Organization*, 6 (1), pp.17-53.

[16] Taylor, T.; Ford, D. and A. Ford (2006) Tipping point failure and robustness in single development projects. *System Dynamics Review*, vol. 22 (1), Spring

Relevance of Glucose Level Variation Trend in Diabetic Patient's Blood and Onset of Diabetic Retinopathy Problems

Durgesh KUMAR
MJK Retina Services, Sri Ram Janki Netralaya, Betiahata,
Gorakhpur, Uttar Pradesh 273001, India

Pavan KUMAR
Department of Biotechnology, Indian Iinstitute of Technology Roorkee
Roorkee, Uttarakhand 247667, India

ABSTRACT

Majority of the world population is affected with diabetes. Retina of eye is found to involve in one third to one fourth of diabetics as a complication called diabetic retinopathy which is one of the most common causes of blindness in diabetes. It has been found that duration of diabetes with characteristic rise and fall is major cause for damage to retina. Various spans of remarkably elevated blood sugar levels often results into changes in perfusion of retinal capillaries bed which leads to onset of diabetic retinopathy. Many associated physiological dynamics have revealed the various modes of clustering of events affecting this condition. Computer simulation based methodologies have been used to solve various issues primarily related to bio-medical complications. Our effort in this work is to check the significant pattern of glucose level dynamics with some data trend responsible for causing diabetic retinopathy condition. The characteristic span profile of the dynamics is found to be held responsible for critical conditions. The obtained data set is trained meanwhile using neural network to predict the possibility of occurrence of such situation with relevant statistical inferences.

Keywords: Retina, Diabetic Retinopathy, Data Trend, Span Profile and Physiological Dynamics.

1. INTRODUCTION

Computational approaches are now being developed to resolve and predict the occurrence of several most common health disorders. Diabetes is a common physiological disease affecting a majority of world population. It is reported that diabetes mellitus affects 4 per cent of the world's population and about half of that have some sort of diabetic retinopathy problem at some stage [1]. Diabetic retinopathy is a common complication in diabetes and is a most prevalent cause of legal blindness between ages of 20 and 65 years. Diabetic retinopathy is more common in type I (insulin dependent or early onset) diabetes (in 40% cases) than in type II (non-insulin dependent or late onset) diabetes (in 20% cases). Such retinopathy occurs in type I as well as type II diabetes mellitus cases and it has also been found that most of type I and about three-fourth of type II sugar patients will develop retinopathy after one and half decade of their diabetic span [2]. Duration of diabetes is a more important risk factor than its metabolic control. In type I diabetics, retinopathy rarely develops within five years of onset of diabetes or before puberty, but 5% of type II diabetics have retinopathy at first presentation of diabetes. It is observed that very strict glycemic monitoring and control is required to cure surgical and cardiac patients, likewise precautions methodology are devised to avoid the occurrence of retinopathy problems [6,7]. Prolonged spans of elevated blood sugar levels often results into remarkable changes in perfusion of retinal blood capillaries with loss of pericytes and increased adhesiveness of blood cells and platelets in retinal vasculature of the eye leads to onset of diabetic retinopathy. Many associated physiological dynamics pertaining to anemia, protienurea, dyslipidemia and hypertention affect this condition.

The stipulation can be avoided to a certain extent by predicting the risk which arises only after sustenance or variation of blood sugar level in a characteristic way by using the data trend of a certain time span. The long term profile of glucose level can be used in computing the exact time of the approaching risk when the patient must get proper regulation of blood sugar level to avoid any significant damage to retina.

The modelling of time span profile for diabetic retinopathy is done in order to compute the exact time of falling into risk. The screening policies are also to be implemented for the early detection of retinopathy in patients with even non-insulin dependent diabetes [3]. The insulin dosage is carefully administered to diabetic patients with higher level of complications and on the other hand there are chances to develop retinopathy problems in meanwhile stage [4,5]. Using the profile one can go directly to consult his ophthalmologist in the initial stage of such a condition and need not wait for having complaints in the vision to make such a move. Model equations are found to be of great help in depicting and characterizing the dynamics and hence concluding some decisions based on numerical outcome of the putative model. The outcome of the model is iteratively being incorporated to the neural network and each numerical value is optimized using genetic algorithm to achieve more accuracy. It is found that neural network approach can be used for making predictions based on training given to the collected data. Each input data is first normalized and make to a certain scale then the values are trained, tested and validated to setup a neural network model. The temporal based massive data is available to support data mining research and working on models this effort is reasonably utilized in field of medical care [8].

The dynamics of other physiological factors are to be assimilated into the baseline sugar trend models in type I and type II diabetes to optimize the goal. Moreover statistical analysis has been made to identify the precise risk zone. Some co-incidental roles of indispensable genetic factors have been incorporated in the model to substantially develop a worthy approach. Although a lot of efforts workers did in characterizing diabetic retinopathy but a very inadequate noticeable work can be found which is dedicated particularly to resolve its cause. Digital chips based on this approach may be devised to monitor glucose level variation trends particularly for the patients falling in a defined risk zone. The accurate predictions based on these approaches would be of great help to identify the meticulous inception of such eye disorders.

2. A TYPICAL DIABETIC RETINOPATHY

Diabetic retinopathy is the most common diabetic disease of eye found to be a leading cause of blindness in most of the adult population. It is caused by adverse changes in the blood vessels of the retina. Generally in diabetic retinopathy, blood vessels may swell and leak fluid while in some cases abnormal new blood vessels grow on the surface of the retina, as shown in **Fig**. **(1)**, the retina is the light-sensitive tissue at the back of the eye required for vision. Diabetic retinopathy involve retina in five stages viz. mild, moderate, severe and very severe non-proliferative retinopathy and proliferative retinopathy. Former types (non-proliferative) are associated somehow with varying degree of obstruction in blood supply to retina while later stage (proliferative) is characterized by growth of new blood vessels with fragile wall along the interface of retina and vitreous i.e. neovascularisation. In such retinopathy the abnormal new blood vessels being fragile get damaged and result in blurring or loss of vision due to leaking of blood on macula or into vitreous cavity. The fluid and proteins may leak into the macula and its centre i.e. fovea where vision is sharpest, leading to blurring of vision. Such condition is defined as macular edema. It may usually occur at any stage of diabetic retinopathy but often tends to occur with the progress of disease. Most of the patients with proliferative retinopathy have macular edema.

To protect vision, every person including young and mature people with type I or II diabetes should have a comprehensive dilated eye examination as soon as possible after detection of diabetes and also get it done at regular intervals as required. Symptoms like floating spots are observed associated to haemorrhages that tend to happen more than once. If left untreated, proliferative retinopathy may result into severe vision loss and even total blindness.

Diabetic retinopathy characterized by macular edema is detected during a comprehensive eye tests as, visual acuity test, dilated eye exam, tonometry, fluorescein angiogram etc. To prevent progression of diabetic retinopathy effectively patient with diabetes is required to control his blood pressure, and blood cholesterol level additionally to blood sugar. Anaemia and nephropathy should also be investigated for and treated. Focal laser surgery is required for treating macular edema. Procedure called scatter laser treatment and surgical vitrectomy are quite effective in treating proliferative retinopathy. Now researchers are testing drugs that may have potential to stop retina from sending chemical signals (growth factors) to promote growth of new blood vessels. And also to block and render tem ineffective if released. These drugs might be of help in reducing the need for laser surgery to treat diabetic retinopathy problems.

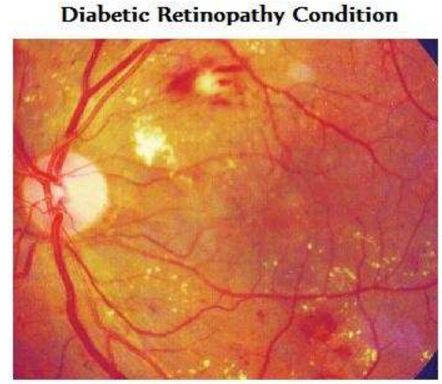


Figure 1. Perfusion of blood capillaries associated to diabetic retinopathy condition.

The clinical findings diagnosis at several stages over a vast time phase resulted into database for some numerical based study and risk characterization.

3. ROLE OF GLUCOSE LEVEL DYNAMICS

Glucose level dynamics is used to govern several physiological key factors and their interaction dynamics at each and every instance. In our study of diabetic retinopathy the impact of its higher level sustenance and frequent variable peaks are considered to be matter for prime research. The focus here is to allocate numerical weight factor that may incorporate several considerations included, at any point of time. The training of such data using neural network etc provide some clue in identifying the risk zone using statistical base and would become the matter of clinical attentiveness. Various other altered physiological states mentioned above like anemia, dyslipidemia, albuminurea and hypertention etc. along with genetic variants may interplay with glycemic status dynamics and if all could be computed may provide a vivid and continuous shade of expression in prediction of diabetic retinopathy.

Observed Glucose Dynamics at long term basis

Data of blood sugar level (fasting) has been assumed to configure a model framework which is used in critical study of its subjective dynamics. An observed data-set for a certain length of period has randomly selected to utilize for our putative model, **Table (1)**.

Blood Glucose level (mg/dL)	108.1	104.3	79.2	…	135.2	122.1
Time-Period (months)	1	2	3	…	69	70

Table 1. Observation time in months and recorded blood sugar level in mg/dL for respective duration

Utilizing the above data trend we would be able to utilize approach of data mining which have been generally used to apply over the population of data to screen out the data of interest in order to conduct a relevant statistical analysis.

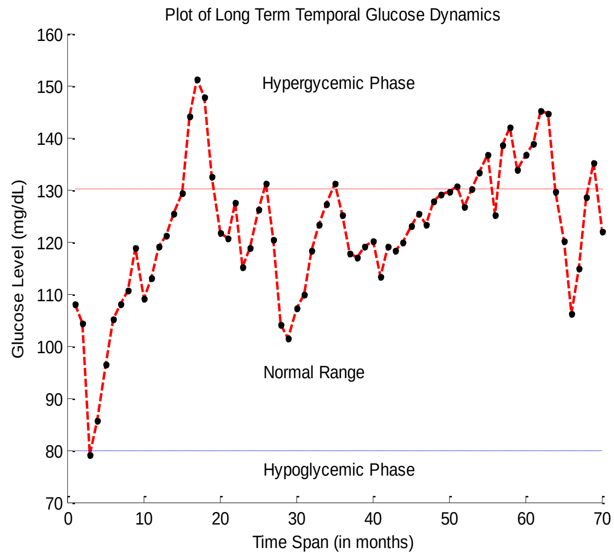


Figure 2. Observed trend of general variation in glucose level with selected time span in diabetes

4. ALGORITHM AND DATA ANALYSIS

To identify the Risk Prone Zone (RPZ) computationally from the entire observed data is the forth most task. Generally glucose level shows abrupt variations over a period of time in a diabetic patient **Fig. (2)**.

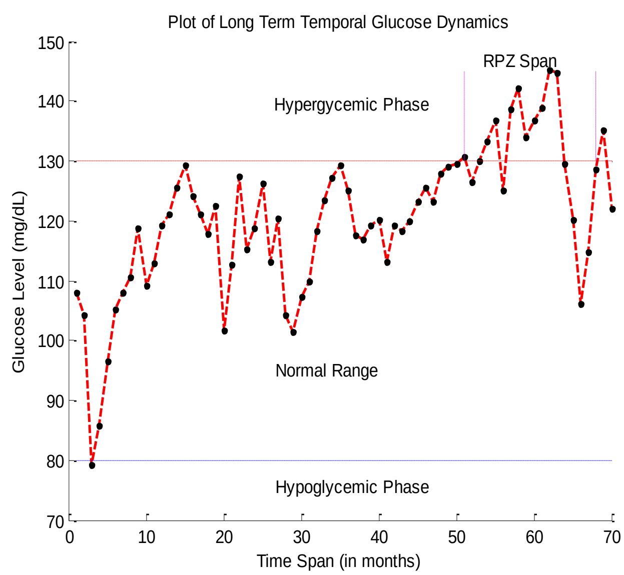


Figure 3. Identification of Risk Prone Zone from glucose level profile at certain period of span

Relevance of Statistical Analysis

According to the assumptions of our model the factors that are to be taken into account to compute the risk factor are;

- Span length of rise in blood sugar level above the normal
- Magnitude of different elevations in blood sugar during a particular span
- Frequency of such abnormal rise in any short duration of time length

Allocating weight to the above different parameters a simple model can be configured. We may compute risk factor variable (R_f) of diabetes to cause retinopathy at any point of time within any risk prone zone, shown **Fig. (3)**.

$R_f = wgl_t * wt_i * X_{ps}$ .....(1)

$X_{ps} = Rs_t * Rf_{i-1}$ .....(2)

i – Denotes the index for particular span; t – Time step

X_{ps} Weight of the previous high level glucose span

wgl_t Weight of the glucose level magnitude

wt_i Weight of the time factor or temporal issue

Rs_t Reliability value at a particular time instance

Rf_{i-1} The risk factor computed for previous span

In a risk prone zone span, depicted in **Eq. (1)** & **(2)** where parameter R_f follows a sigmoidal path while X_{ps} shows a decay with time based reliability factor (Rs_t) in between the different elevated sugar level spans. The other associated physiological factors governing glucose level and vision loss are not incorporated in overall study. Also the impact of treatment and dosage regulation not included separately to signify their role in shaping the data. The effort has been made just to incorporate the general data into the model to understand the vision related issue.

The first assumption envisages over the fact that any functional unit has initially maximum reliability and it decreases on basis of interaction of several subunits in a specific topological fashion with variable dynamics, **Fig. (4)**. A varying consequence of transformation in topology can be taken into account to depict the observed level of functionality at any point over a period of time. The general reliability term expression is considered here, **Eq. (3)**.

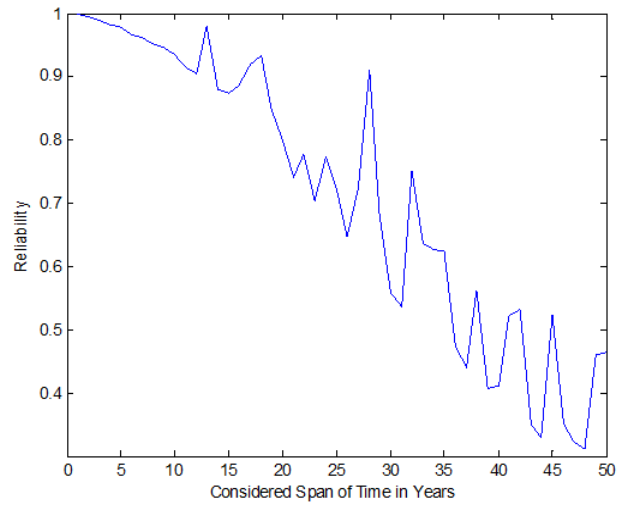


Figure 4. Stochastically decreasing reliability factor for visual system in diabetic situation.

Reliability at a time t is given by;

$Rt = e^{-lambda * t}$ .....(3)

The variable lambda – Denoted the chances of failure
e – Euler's number, having numerical value 2.7183

More effectively we can write for a an interactive small frame where two subunits are interacting sequentially while other three are also working in sequential harmony to each other, but the group of two and three are interacting in parallel fashion. The reliability of this small system with a small value for chance of failure (lambda of respective subunit i – lbdi) can be computed in respect to span of time, given by the equation **Eq. (4)**;

$Rs(t) = e^{(-(lbd1 + lbd2) * t)} + e^{(-(lbd3 + lbd4 + lbd5) * t)}$ .....(4)

It is preferable that large population size data has to be processed to design it for statistical test. Subsequently p-value is to be taken into account to signify the existence of risk zone over any band of time scale. Although reliability has very undersized impact over the biological system model but it always impart the property of decaying strength over time. So the strength of risk prone band has significant influence of this factor. Here stochastic assumption among functional subunits is made that may result into feasible functional outcome of any system as in our case visual framework.

Neural Network Applications

The simulation of glucose dynamics has been done in Matlab 8.0 to numerically evaluate and predict the process behaviour using neural network approach.

The data is first normalized using upper and lower bound of the numerical range and then the normalized data is used to serve as input for the neural network model. The procedure is repeated for train, test and validation data. Constituting the hidden layer a multilayer perceptron is designed to process the input together with weight. Using activation function as sigmoidal type the output is obtained. The given input is data set of sugar patient for entire observable time duration. After optimizing the number of layers to ensure optimal memory, ten hidden layers of neurons in the network are assumed. Inbuilt tool box is employed to build the neural network and to obtain the output of statistical significance.

The idea is to predict the Risk Prone Zone over the period of time from the given data, once it has been trained earlier with requisite data set. The model validation for a specific data profile is required to train and test the data after normalization, since each input data is first normalized and make to a certain scale then the values are trained, tested and validated to setup a neural network model. So adjusting the weight strength and minimizing the error more optimal predictions may be obtained regarding the temporal based inception of problems. Meanwhile data is optimized using genetic algorithm operation applying an objective function to ensure more optimized input data for neural network after its operation.

5. CONCLUSIONS

It was observed that span of glucose elevation in the blood is also important in addition to the level of glucose in the blood. So three things are collectively assumed to be relevant to reach any statistical conclusion i.e., times for which glucose level went high, magnitude of its elevation and span for which high frequency of glucose elevation or sustained increase in level is recorded. This strategy is to be applied over a long term observation which would give some inference of identifying the chances of onset of risk prone zone. It is also to be noticed that the study of physiological dynamics and its impact over retinopathy studied has still have to incorporate the significant influence of plenty of other factors and ailment conditions to make the approach more realistic and applicable.

6. REFERENCES

[1] L.P. Aiello, T.W. Gardner, G.L.King, G. Blankenship, J.D. Cavallerano, F.L. Ferris, et al, “Diabetic retinopathy”, **Diabetes Care**, Vol. 21, 1998, pp. 143-156.

[2] R. Klein, B.E. Klein, S.E. Moss, M.D. Davis and D.L. De Mets, “The Wisconsin Epidemiologic Study of Diabetic Retinopathy II: Prevalence and risk of diabetic retinopathy when age at diagnosis is less than 30 years”, **Arch. Ophthalmology**, Vol. 102, 1984, pp. 520-526.

[3] S.C. Brailsford, R. Davies, C.R. Canning and P. Roderick, “Evaluating screening policies for the early detection of retinopathy in patients with non-insulin-dependent diabetes”, **Health Care Management Science**, Vol. 1, 1998, pp. 1-10.

[4] G.V.D. Berghe, P. Wouters, F. Weekers, et al, “Intensive insulin therapy in critically ill patients”, **In. N. Engl. J. Med.**, Vol. 345, No. 19, 2001, pp. 1359-1367.

[5] G.V.D. Berghe, A. Wilmer, G. Hermans, et al, “Intensive insulin therapy in the medical ICU”, **Int. N. Engl. J. Med.**, Vol. 354, No. 5, 2006, pp. 449-461.

[6] J.M. Carr, F.W. Selke, M. Fey, et al, “Implementing tight glucose control after coronary artery bypass surgery”, **Int. Ann Thorac. Surg**., Vol. 80, 2005, pp. 902-909.

[7] A. Bhatia, B. Cadman and I. Mackenzies, “Hypoglycemia and cardiac arrest in a critically ill patient on strict glycemic control”, **Int. Anesth. Analg.**, Vol. 102, 2006, pp. 549-551.

[8] B.A. Janz, M. Saeed, G.D. Clifford, R.G. Mark, et al, “MIMIC II: A massive temporal database to support research in integrating data, models, and reasoning in critical care”, **Int. Proceedings of AMIA 2005 Fall Symposium**, Washington DC, USA, Oct 2005.

CRM PERFORMANCE ANALYSIS USING THE CONCEPT OF GENETIC FUZZIMETRIC TECHNIQUE

Issam Kouatli,
School of Business - MIS Division, Lebanese American University
LAU, Beirut Lebanon

ABSTRACT

Fuzzy logic and Genetic Algorithm combination is becoming popular intelligent technique among researchers to investigate the decision making process in different application fields. This paper proposes a methodology of performance analysis of CRM systems based on such hybridization of fuzzy logic and genetic Algorithm. The technique is referred to as Genetic Fuzzimetric Technique. It is based on a concept of Fuzzimetric Arcs that defines the initial selections of fuzzy sets and allowing the mutation and cross-over of sets using Genetic algorithm principles and hence optimizing the output. CRM –performance analysis was used as an example to explain the operation of GFT where the objective was to monitor the performance level of CRM. Based on GFT principle, a tool was built to demonstrate the mechanism of GFT to CRM performance analysis example.

Keywords Fuzzy systems, Decision Support Systems, DSS, CRM performance analysis.

1. INTRODUCTION

Management of information systems today are much more complex due to the variety of socio-technical elements introduced as the system evolve. Fuzzy logic (Zadeh 1965 & 1973) [1,2] is one type of theory that can be used to help achieving a decision given uncertain (fuzzy) inputs and as such it can provide a translation of the qualitative abilities of the human brain into quantitative functions. Uncertainty in the decision making is becoming more and more needed as most decisions are becoming too fluid and unstructured. A tool that can help managers to take a decision under uncertainty is becoming highly demanded. Many systems uses fuzzy inference engine as embedded features of their perspective systems, for example, Cheng et al. [3] proposed an e-marketplace negotiation system that is based upon fuzzy inference system. Dweiri et.al. [4] proposed to use fuzzy decision making systems to measure the efficiency based on three criteria, project cost, project time and project quality. Most mathematical modeling of decision support system may lack the existence of accuracy element due to the fact that either because the variables are actually "fuzzy" (not accurate) or due to the fact that the human individual's interpretation of the inputs may vary. This fact has led a number of researchers in the field to adopt the mixture of accurate mathematical formulae with some elements of fuzziness. Jimenez et.al. [5] for example proposed a method for solving linear programming problems using fuzzy parameters and Wang et al. [6] proposed a fuzzy modification to AHP (Analytical Hierarchy processing) to aid decision maker in optimization of maintenance strategies.

This paper introduces a new user-friendly fuzzy inference tool that can be utilized by managers or by researchers to aid them in decision making without the need to understand the mechanism of fuzzy logic/system. It is basically an inference engine (shell) using the concept of GFT and hence termed as "**FIE**" (Fuzzy Inference Engine). **FIE** was developed on the basis of the concept of Fuzzimetric Arcs for fuzzy set choice and selection introduced by Kouatli et al. [7] as well as then concept of multivariable system introduced by Kouatli

[8]. Methodology of definition and selection of fuzzy sets (Fuzzy variables) using the concept of Fuzzimetric Arcs in conjunction with Genetic algorithm was also proposed by Kouatli [9]. The concept principle of Fuzzimetric Arcs and its combinations with Genetic Algorithm is beyond the scope of this paper. However, interested researchers can view relevant articles about this subject [7, 8, and 9]. JAVA language was used to implement the concept of FIE.

2-Genetic Algorithm & Fuzzimetric Arcs Principles

Genetic algorithm can help in establishing the rule-set values based on combinations of multiple rules or "genes". Each gene can be regarded as one rule towards modeling the system where the antecedents (If- part) and the consequents (THEN-Part) can be found using "crossover" operator between all possible combinations of fuzzy variables between the input universe and the output universe. For example, if the input universe is represented by four linguistic terms (variables) (Tiny, Low, Medium, High) and the output universe is represented by three linguistic terms (Low, Average and High), and then the total possible crossover of input/output would be composed of a total of 12 possible genes (Rules). Some of these genes are not relevant and will die as the fitness factor/function becomes very low. This will leave only the healthy "good" genes in the pool of rule-set genes. The remaining healthy genes form the chromosomes of the rules which are basically the resulted Knowledge base.

The final output performance of fuzzy system is highly dependent on the fuzzy set shape. Genetic algorithm may help in selection and mutation of the fuzzy set shape in order to optimize the system to the desired output. The main components of genetic fuzzy system are illustrated in figure 1, and consist of:

1-selection of good "genes" of the knowledgebase

2- Selection of good "genes" of fuzzy variables (fuzzy set shapes)

3- The input & justification process

4-The output and the De-fuzzification process

5- The fuzzy inference engine

6- Feedback link that it can be either in training mode or a feedback of the output as part of normal production of the system (i.e. output change treated as an input to the system as well.).

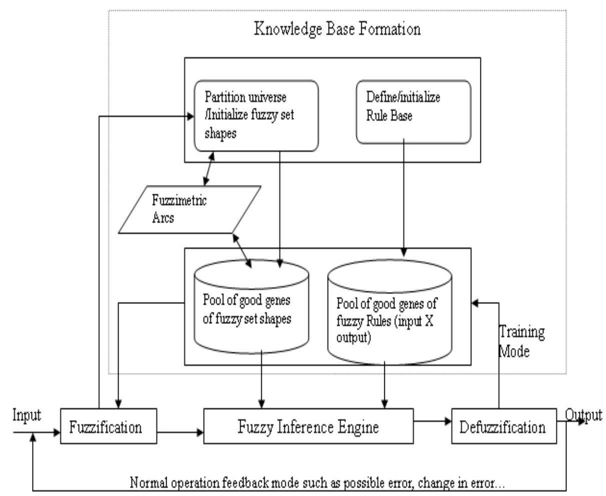


Figure 1: General structure of Genetic Fuzzy algorithm.

JAVA language was used to construct the FIE structure to have a maximum of 10 inputs and a maximum of 10 outputs. As an example, Figure 2 shows a snapshot of the 3 simple steps in creating a simple rule-set composed of one-input-one output (SISO) system

Step1- Specify the number of inputs and number of outputs using drop –down menu, then specify if the required relationship to be either directly proportional, inversely-proportional or advanced (used for customization of linear or nonlinear-systems).

Step 2- Specify the level of each input – in this example 9 levels specified (other than the 0 level) for each of the input and output

Step3- run the test which is basically calls the program and does the fuzzy inference and display the de-fuzzified value as shown in Fig. 2

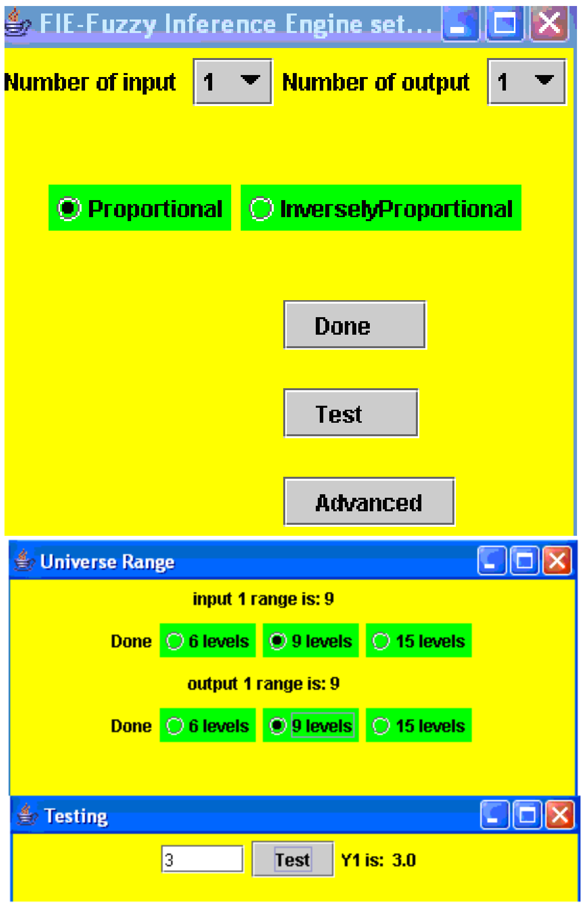


Fig. 2 Snapshot(s) of FIE simple user-interface

If advanced button has been chosen, then the system displays a rule-set option menu where you can specify each rule using the drop-down menu as shown in Figure 3 where R11 stands for rule-set connecting the first input with first output and hence labeled as R11 which is composed of the four rules described above and in this example shows a linear relation ship between the input and the output, i.e. the figure 6 states the definition of directly proportional relationship represented by the four rules as:

RULE1 If input=PO Then Output=PO
RULE2 If input=PS Then Output=PS
RULE3 If input=PM Then Output=PM
RULE4 If input=PL Then Output=PL

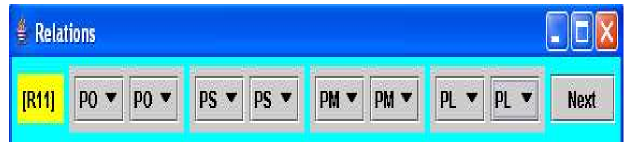


Fig. 3- Snapshot of Rule-set Definition R11 using Drop-down menus (Advanced)

3- Fuzzy CRM performance analysis - A Root-Cause Example

The use of Root-Cause diagram is mainly to segregates the possible causes of problem in a manner that identifies the root cause (and hence it is termed as root-cause Analysis). Each cause reflects the characteristic required to improve the performance of the fish-bone analysis. The measurement of inputs are usually qualitative rather than quantitative, hence, the input level of each input can be represented qualitatively as Low, High….etc.

Our interest in diagram is in the utilization of causes (inputs labeled as X_1 till X_{10}) taken as fuzzy variable (see figure 4). Fuzzy inference in this case can be used to analyze the CRM performance level (the only output of the system labeled as Y_1). Each one of these input ill have a specific influence on the output (directly proportional inversely proportional or non-linear) as it will be explained in table 1.

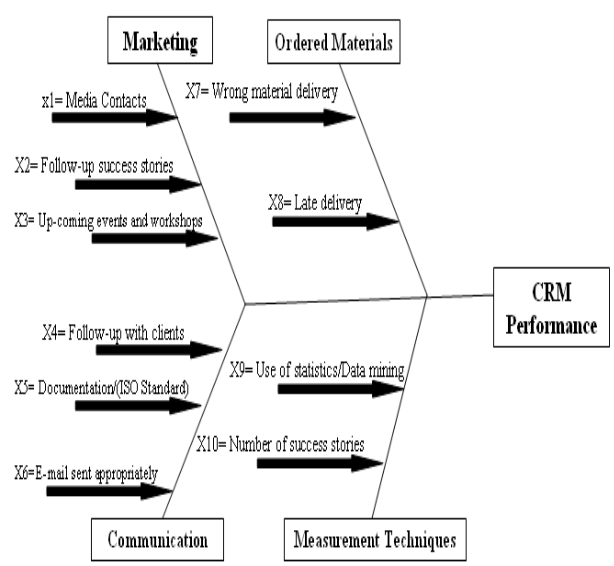


Fig. 4- CRM performance analysis in a form of root-cause diagram.

Running the FIE application would need to identify the number of inputs and outputs. Figure 5 shows a snapshot for such definition to our CRM performance analysis example.

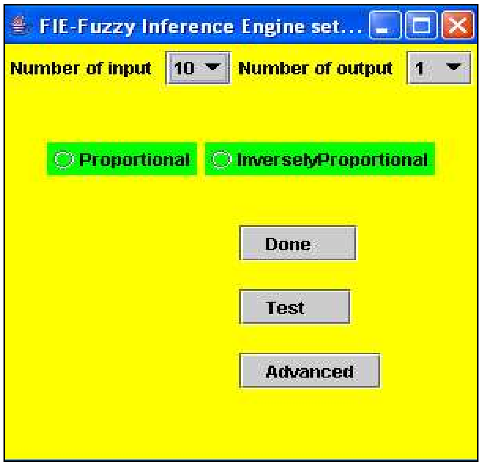


Fig. 5- Snapshot of the first menu of FIE Number of Input/output required.

There are number of factors that will be needed to specify when running the Fuzzy Inference Engine (FIE). These factors are explained below and summarized in table1:

Table 1: Details of chosen values used in the Fuzzy-CRM performance example

Input Description	Chosen levels	Relation type
X1=Media Contacts	9	Directly Proportional
X2=Follow-up success stories	9	Directly Proportional
X3=Up-coming events and workshops	9	Directly Proportional
X4=Follow-up with clients	9	Directly Proportional
X5=Documentation/ISO standards	6	Advanced – Non linear
X6=Email sent appropriately	6	Advanced – Non linear
X7=Wrong Material delivery	9	Inversely Proportional
X8= Late Delivery	9	Inversely Proportional
X9=Use of statistics/Data mining	15	Inversely Proportional
X10= Number of success stories	15	Inversely Proportional

After defining the number of inputs and outputs (10 inputs and one output as shown in Fig.5), the steps involved finding the final output (CRM performance level) can be described in 3 main steps. These are:

Step1:-Fuzzification: Decision needs to be made about the fuzzification strategy for each input must be made. Only number of levels required to input without the need to understand the structure of Fuzzimetric Arcs (angles required for each option). These are:
X1 to X4 assumed to have 9 levels ($\alpha = 30^0$),
X5 and X6 assumed to have 6 level ($\alpha=45^0$),
X7 and X8 assumed to have 9 levels ($\alpha = 30^0$), and
X9 and X10 assumed to have 15 levels ($\alpha = 18^0$)
Figure 6 shows a snapshot of the definitions of fuzzification levels.

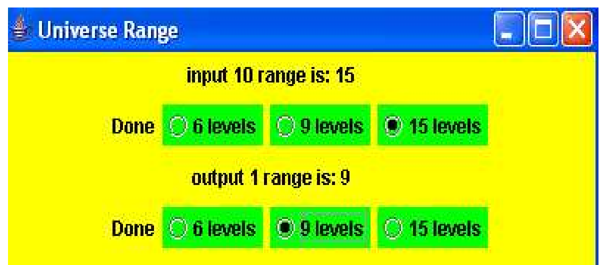


Figure 6 – Snapshot of fuzzification levels definitions using FIE

Note that, in general, each one of the inputs may have more or less influence to the output that the other (weight). For simplicity, FIE assumes that all inputs carries equal weight, i.e. in our example, of 10 inputs, the weights for any one of the inputs would be (1/10=0.1)

Step2: Define Knowledge/Rule-set(s): Each one of the inputs will have Rule-set relationship with the final output would be directly proportional, inversely proportional or non-linear relationship. For the sake of this example, X1 till X4 assumed to be directly proportional with the output Y1, X5 and X6 assumed to have non-linear rule-set with the outputs and X7 to X10 assumed to be inversely proportional with the output Y1. Figure 7 shows a snapshot of rule-set(s) definition using advanced button for both linear and non-linear rules.

Fig. 7. Definition of the 10 rule-sets using advanced button of FIE

Step3:- Inference and De-Fuzzification: after setting-up the system as shown in the above 3 steps then FIE can be called as a method from another program and hence FIE can be used as the inference part of a specific application package on CRM performance. For testing purposes, FIE designed to have test button that allow the user to view the output (Y). Figure 8 shows the snapshot of the test of Fuzzy-CRM measurement using an example of input values of X1 to X10. Note that as the output fuzzified levels was chosen as 9 levels (see figure6), then the final output shown in the snapshot (Y1=4.68) is dependent on that chosen levels, i.e. the Fuzzy-CRM performance measurement Y1=4.68 out of 9 levels

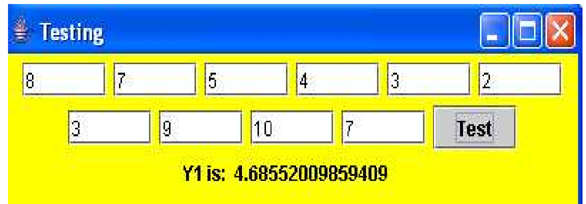


Fig. 8- Snapshot of testing Button (De-fuzzification step)

4- Conclusion

CRM performance analysis was used in this paper as a vehicle to demonstrate the structure and steps used in FIE -A tool for Decision Making based on Genetic Fuzzimetric Technique Principle. The advantages of this tool are:

a- Systematic approach of selection and definition of Fuzzy sets

b- Simplified process of fuzzification

c- Better optimized decision making principle based on rule-type definition (directly or inversely proportional).

d- Can be used by developer as part of closed-loop system where the Inference/Defuzzification component may be called from an external software package to develop a specific area of management systems.

e- May also be used by managers for generic decision making problems.

f- Definition of rule-set(s) was also made simple by defining the rules in drop-down menus.

As the strength of fuzzy systems is that it can deal effectively with uncertainty in systems, and as the inputs in most management problems are not absolutely accurate, the proposed tool (FIE) would be ideal in such situation where the manager has to deal with such uncertainty. FIE is still at its prototype stage, where the purpose of this paper is just to explore and identify the concept. Further enhancement would be required to identify the mutation/crossover mechanisms as well as the tuning algorithm to achieve optimum performance.

References

[1} Zadeh, L.A. (1965) Fuzzy sets Information Control 8, 338-53.

[2] Zadeh, L.A (1973) Outline of a new approach to the analysis of complex systems and decision processes. IEEE Transactions on Systems, Man and Cybernetics 3, 28-44.

[3] Chi-Ben Cheng, Chu-Chai Henry Chan, Kun-Cheng Lin, "Intelligent agents for e-marketplace: Negotiation with issue trade-offs by fuzzy inference systems". Decision Support Systems NOV 2006 vol. 42 issue 2 pages: 626

[4] Dweiri,F.T. And Kabalan. M (2006) Using fuzzy decision making for the evaluation of the project management internal efficiency. Decision Support Systems. Vol. 42, pg. 712

[5] Jimenez, M. And Arenas, M, and Bilbao, A. And Rodriguez, M. (2007) Linear programming with fuzzy parameters: An interactive method resolution. European Journal of Operational Research. Amesterdam. Vol. 177, Iss. 3; pg. 1599

[6] Wang, L. And Chu, J. And Wu, J. (2007) Selection of optimum maintenance strategies based on a fuzzy analytical hierarchy process. International Journal of Production Economics. Amesterdam. Vol.107, Iss.1; pg. 151

[7] Kouatli, I. And Jones, B. (1990) An improved design procedure for fuzzy control systems. International Journal of Machine Tool and Manufacure,

[8] Kouatli, I.(1994), A simplified fuzzy multivariable structure in a manufacturing environment. Journal of Intelligent Manufacturing. vol. 5, page 365-387.

[9] Kouatli, I.(2008) Definition and Selection of fuzzy sets in genetic-fuzzy systems using the concept of fuzzimetric arcs. Kybernetes, Vol. 37 Issue 1

[10] Chou, T. And Chou, S., and Tzeng, G. (2006) evaluating IT/IS investments: A fuzzy multi-criteria decision model approach. European Journal of Operational Research. Amesterdam. Vol.173, Iss. 3; pg. 1026

[11] Gunasekaran, N. And Rathesh, S. And Arunachalam, S., and Koh, S. (2006) Optimizing supply chain management using fuzzy approach. Journal of Manufacturing Technology Management. Bradford. Vol. 17, Iss. 6; pg. 737

Operation of Marine Diesel Engine Based on Differential Equations

Frane MITROVIC*, Goran Kovacevic*, Dean SUMIC*

***Faculty of Maritime Studies, University of Split**
Zrinsko-Frankopanska 38, 21000 Split, Croatia

ABSTRACT

The purpose of this paper is to present the efficiency of the application of the system dynamics simulation modelling in investigating the behaviour dynamics of the diesel engine complex system. The marine diesel engine is defined by a set of non-linear differential equations, i.e. by a continuous simulation model of a higher order and the so-called equations of state. At the same time, the simulation model is discrete as it strictly satisfies the chosen value of the fundamental integration time step DT. The paper presents a mathematical model of the system consisting of a marine diesel engine model, which forms the basis for designing a system dynamic qualitative model (mental-verbal, structural and schematic model) and a quantitative model (system dynamic mathematical and information simulation model). A scenario of mixed propulsion states has been presented. Parameters obtained with the aid of the hardware simulator Kongsberg ERS-L11 MAN B&W-5L90MC-VLCC Version MC90-IV have been used in conducting the simulation.

Key words: simulation, system dynamics, modelling, diesel engine, simulator, heuristic optimisation.

1. INTRODUCTION

As there are a large number of various diesel engines in use, it appears that designing a universal model applicable to each individual type of engine would be very useful and effective, particularly in the case of designing complex mathematical models of diesel engines that include all essential sub-processes:

- description of the working medium (fuel, air),
- mechanical system dynamics,
- cooling system,
- lubrication system,

as well as their mutual interactions.

However, such a complex model is not suitable for the computer-supported simulation modelling of a diesel engine. Other sub-processes of the universal model are indirectly involved through the values of the various parameters that have been defined for given stationary states.

Some of the features of the diesel engine operation process have a discrete quality (injection process, ignition, burning and combustion of the air and fuel mixture) so that a discrete model of the engine is undoubtedly more exact. Since in high-speed diesel engines the time of discrete event simulation (injection period) is much shorter than the dominant time constants of the very process, the discrete model of the engine closely approaches the continuous one, so that only the latter will be analysed in this paper.

Figure 1. shows a generic scheme of a marine turbocharged diesel engine. The following basic functional units can be noticed:

- engine mechanism (M),
- exhaust gas receiver (KIP),
- turbine (T) and the compressor (K) of the turbocharger,
- scavenge air receiver (KZ),
- high-pressure fuel pump (VTSg).

Each of the above functional units has been allocated input and output variables, variables of states and disorders, which are relevant for the model. They connect the units into a unique multivariable system.

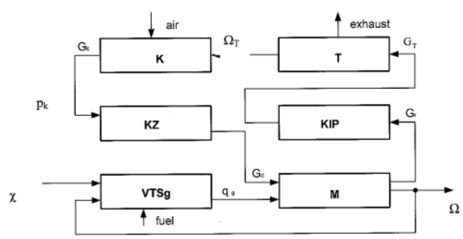


Figure 1: Generic scheme of a marine diesel engine

Where:

M	- engine,
KIP	- exhaust gas receiver,
T	- turbine,
K	- compressor,
KZ	- scavenge air receiver,
VTSg	- high-pressure fuel pump,
Ω	- angular speed of the engine crankshaft (KV),
G_r	- amount of the exhaust gas from the engine, entering the KIP,
p_r	- exhaust gas pressure in the KIP,
ρ_r	- exhaust gas density,
G_T	- amount of the exhaust gas flowing through the turbine,
Ω_T	- angular speed of the turbocharger,
G_k	- amount of air that the turbocharger supplies to the KZ,
p_k	- air pressure in the receiver,

ρ_k - specific air mass in the KZ,
χ - regulator action,
q_g - amount of fuel supplied to the engine by VTS.

2. EQUATION OF STATE OF A MARINE TURBOCHARGED DIESEL ENGINE

Dynamic features of a diesel turbocharged engine are determined by the dynamic features of the components (Figure 1.), therefore the differential equations of these units have to be considered together. However, in some cases at low volumes in the inlet and exhaust pipelines, and when there is no gas-dynamic supercharging, the impact of the pipeline volume on the engine dynamic features is negligible.

This enables the simplification of its equations. As a result, it may be assumed that $V_B \approx V_r \approx O$ and, accordingly, $T_B = T_r = T_h = O$. Taking this condition into consideration, the system of the engine unit equations takes the following form:

$$\begin{aligned} &d_D(p)\varphi = \chi + \theta_k\rho - \theta_D\alpha_D, \\ &d_T(p)\varphi_T = \xi + \theta_{T1}\chi - \theta_{T2}\rho, \\ &k_B\rho = \varphi_T - \theta_B\varphi, \\ &k_r\xi = \varphi + \theta_r\rho - k_h\chi. \end{aligned} \tag{1}$$

Since the paper discusses the system dynamics modelling from the aspect of revolution speed, it is necessary to analyse changes in angular speed of the diesel engine crankshaft. That is why the resulting equation can be presented in the following form:

$$(\det A)\varphi = D(\varphi)$$

where the determinants of the equation system (1) take the form:

$$\det A = \begin{vmatrix} d_D(p) & 0 & -\theta_k & 0 \\ 0 & d_T(p) & \theta_{T2} & -1 \\ \theta_B & -1 & k_B & 0 \\ -1 & 0 & -\theta_r & k_r \end{vmatrix},$$

$$D(\varphi) = \begin{vmatrix} \chi - \theta_D\alpha_D & 0 & -\theta_k & 0 \\ \theta_{T1}\chi & d_T(p) & \theta_{T2} & -1 \\ 0 & -1 & k_B & 0 \\ -k_h\chi & 0 & -\theta_r & k_r \end{vmatrix}.$$

By solving these determinants

$$\det A = \begin{vmatrix} d_D(p) & 0 & -\theta_k & 0 \\ 0 & d_T(p) & \theta_{T2} & -1 \\ \theta_B & -1 & k_B & 0 \\ -1 & 0 & -\theta_r & k_r \end{vmatrix} =$$

$$= \begin{vmatrix} d_D(p) & 0 & -\theta_k & 0 \\ \theta_B d_T(p) & 0 & k_B d_T(p) + \theta_{T2} & -1 \\ \theta_B & -1 & k_B & 0 \\ -1 & 0 & -\theta_r & k_r \end{vmatrix} =$$

$$= \begin{vmatrix} d_D(p) & -\theta_k & 0 \\ \theta_B d_T(p) & k_B d_T(p) + \theta_{T2} & -1 \\ -1 & -\theta_r & k_r \end{vmatrix} =$$

$$= \begin{vmatrix} d_D(p) & -\theta_k & 0 \\ \theta_B d_T(p) & k_B d_T(p) + \theta_{T2} & -1 \\ k_r\theta_B d_T(p) - 1 & k_B k_r d_T(p) + k_r\theta_{T2} - \theta_r & 0 \end{vmatrix} =$$

$$= \begin{vmatrix} d_D(p) & -\theta_k \\ k_r\theta_B d_T(p) - 1 & k_B k_r d_T(p) + k_r\theta_{T2} - \theta_r \end{vmatrix} =$$

$$= k_B k_r d_D(p) d_T(p) + (k_r\theta_{T2} - \theta_r) d_D(p) + k_r\theta_B\theta_k d_T(p) - \theta_k$$

$$D(\varphi) = \begin{vmatrix} \chi - \theta_D\alpha_D & 0 & -\theta_k & 0 \\ \theta_{T1}\chi & d_T(p) & \theta_{T2} & -1 \\ 0 & -1 & k_B & 0 \\ -k_h\chi & 0 & -\theta_r & k_r \end{vmatrix} =$$

$$= \begin{vmatrix} \chi - \theta_D\alpha_D & 0 & -\theta_k & 0 \\ \theta_{T1}\chi & 0 & k_B d_T(p) + \theta_{T2} & -1 \\ 0 & -1 & k_B & 0 \\ -k_h\chi & 0 & -\theta_r & k_r \end{vmatrix} =$$

$$= \begin{vmatrix} \chi - \theta_D\alpha_D & -\theta_k & 0 \\ \theta_{T1}\chi & k_B d_T(p) + \theta_{T2} & -1 \\ -k_h\chi & -\theta_r & k_r \end{vmatrix} =$$

$$= \begin{vmatrix} \chi - \theta_D\alpha_D & -\theta_k & 0 \\ \theta_{T1}\chi & k_B d_T(p) + \theta_{T2} & -1 \\ (k_r\theta_{T1} - k_h)\chi & k_B k_r d_T(p) + k_r\theta_{T2} - \theta_r & 0 \end{vmatrix} =$$

$$= \begin{vmatrix} \chi - \theta_D\alpha_D & -\theta_k \\ (k_r\theta_{T1} - k_h)\chi & k_B k_r d_T(p) + k_r\theta_{T2} - \theta_r \end{vmatrix} =$$

$$\begin{aligned} &= k_B k_r d_T(p)\chi + k_r\theta_{T2}\chi - \theta_r\chi - \\ &-k_B k_r\theta_D d_T(p)\alpha_D -_r \theta_D\theta_{T2}\alpha_D + \theta_D\theta_r\alpha_D + k_r\theta_k\theta_{T1}\chi - k_h\theta_k\chi \end{aligned}$$

and taking these expressions into account:

$$\begin{aligned} d_D(p) &= T_D p + k_D \\ d_T(p) &= T_T p + k_T \end{aligned} \quad \text{and}$$

we obtain the possibility to present the equation of a turbocharged diesel engine, as follows:

$$d_{DH}(p)\varphi = S(p)\chi - u(p)\alpha_D, \tag{2}$$

where:

$d_{DH}(p) = T_H^2 p^2 + T_{DH} p + k_{DH}$

- related operator of the turbocharged engine,

$S(p) = T_s p + k_s$

- operator of action from the aspect of fuel supply control,

$u(p) = T_u p + k_u$

- operator of action from the aspect of consumer adjusting.

$$\left.\begin{aligned} d_D(p) &= T_D p + k_D \\ d_T(p) &= T_T p + k_T \end{aligned}\right\} \Rightarrow d_D(p) d_T(p) =$$

$$= (T_D p + k_D)\cdot(T_T p + k_T) =$$

$$= T_D T_T p^2 + (k_T T_D + k_D T_T)p + k_D k_T$$

$$\det A = \underbrace{k_B k_r T_D T_r}_{=T_H^2} p^2 +$$

$$+\underbrace{\left[k_B k_r\left(k_T T_D + k_D T_T\right) + \left(k_r \theta_{T2} - \theta_r\right)T_D + k_r \theta_B \theta_k T_T\right]}_{=T_{DH}} p +$$

$$+\ \underbrace{k_B k_D k_r k_T + \left(k_r \theta_{T2} - \theta_r\right)k_D + k_r k_T \theta_B \theta_k - \theta_k}_{=k_{DH}};$$

$$D(\varphi) = \underbrace{k_B k_r T_T}_{=T_s} p\chi + \underbrace{(k_B k_r k_T + k_r \theta_{T2} - \theta_r + k_r \theta_k \theta_{T1} - k_h \theta_k)}_{=k_s}\chi -$$

$$-\underbrace{k_B k_r \theta_D T_T}_{=T_u} p\alpha_D - \underbrace{(k_B k_r k_T \theta_D + k_r \theta_D \theta_{T2} - \theta_D \theta_r)}_{=k_u}\alpha_D$$

$$(\det A)\varphi = D(\varphi) \Rightarrow \underbrace{(T_H^2 p^2 + T_{DH} p + k_{DH})}_{d_{DH}(p)}\varphi = \underbrace{(T_s p + k_s)}_{=S(p)}\chi - \underbrace{(T_u p + k_u)}_{=u(p)}\alpha_D \cdot$$

This is the form taken by the dependence of the coefficients of the above expressions on the parameters of the elements that are comprised within the engine system:

$$\begin{aligned}
T_H^2 &= T_D T_T k_B k_r, \\
T_{DH} &= \left[T_D\left(k_B k_r + \theta_{T2}\right) + T_T\left(k_D k_B + \theta_B \theta_k\right)\right]k_r - \theta_r T_D, \\
k_{DH} &= k_D k_r\left(k_B k_T + \theta_{T2}\right) - \theta_k\left(1 - k_T k_r \theta_B\right) - \theta_r k_D, \\
T_s &= k_B k_r T_T, \\
T_u &= T_T \theta_D k_B k_r, \\
k_s &= k_r\left(k_B k_T + \theta_k \theta_{T1} + \theta_{T2}\right) - \theta_k k_h - \theta_r, \\
k_u &= \theta_D k_r\left(k_B k_T + \theta_{T2}\right) - \theta_r \theta_D.
\end{aligned}$$

Eq (2), in a differential form, can be formulated:

$$T_H^2 \frac{d^2\varphi}{dt^2} + T_{DH}\frac{d\varphi}{dt} + \varphi k_{DH} = T_s \frac{d\chi}{dt} + \chi k_s - T_u \frac{d\alpha_D}{dt} - k_u \alpha_D \qquad (3)$$

$$p = \frac{d}{dt},\ \ p^2 = \frac{d^2}{dt^2};$$

$$T_H^2 \frac{d^2\varphi}{dt^2} + T_{DH}\frac{d\varphi}{dt} + k_{DH}\varphi = T_s \frac{d\chi}{dt} + k_s \chi - T_u \frac{d\alpha_D}{dt} - k_u \alpha_D.$$

If all elements of Eq (2) are divided by the related operator, this results in:

$$\varphi = Y_{DH}^{\chi}(p)\chi + Y_{DH}^{\alpha}(p)\alpha_D,$$

where the transfer functions:

$$Y_{DH}^{\chi}(p) = \frac{S(p)}{d_{DH}(p)} = \frac{T_s p + k_s}{T_H^2 p^2 + T_{DH} p + k_{DH}},$$

$$Y_{DH}^{\alpha_D}(p) = -\frac{u(p)}{d_{DH}(p)} = -\frac{T_u p + k_u}{T_H^2 p^2 + T_{DH} p + k_{DH}}$$

allow for designing a turbocharged engine scheme which is more suitable than the one in Figure 2.

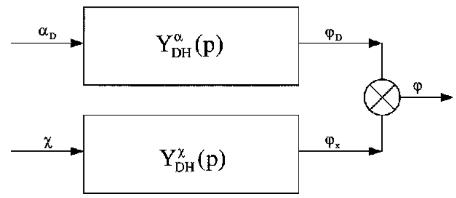


Figure 2. Simplified structural scheme of a turbocharged diesel engine

3. COMPUTER SIMULATION MODEL FOR A MARINE DIESEL ENGINE

3.1. System dynamic mathematical model for a marine diesel engine

The system dynamic mathematical model for a turbocharged diesel engine can be defined by the expression (the explicit form of the Eq (3):

$$\frac{d^2\varphi}{dt^2} = -\frac{d\varphi}{dt}\frac{T_{DH}}{T_H^2} - \varphi\frac{k_{DH}}{T_H^2} + \frac{d\chi}{dt}\frac{T_S}{T_H^2} + \chi\frac{k_S}{T_H^2} - \frac{d\alpha_D}{dt}\frac{T_u}{T_H^2} - \alpha_D\frac{k_u}{T_H^2} \quad (4)$$

where:

$d^2\varphi = D2FI$

- acceleration of the angular speed of the engine crankshaft (KV),

$d\varphi = D1FI$

- angular speed change rate of the engine crankshaft (KV) (speed gradient),

$T_{DH} = TDH$

- time characterising the sluggishness of the engine as a regulated object,

$T_H^2 = TH2$

-square of the time characterising the sluggishness of the engine as a regulated object,

$T_s = TS$

- time characterising the engine sluggishness,

$d\chi = DKAPA$

- speed of moving the rack lever high - fuel pressure pump (VTS),

$\chi = KAPA$

- relative movement of the rack lever of the fuel pressure pump (VTS),

$k_s = KS$

- coefficient of the engine gain,

$d\alpha_D = DALFAD$

- change rate of the relative consumer load,

$\alpha_D = ALFAD$

- relative change of the consumer load,

$T_u = TU$

- time characterising the generator sluggishness,

$k_u = KU$

- coefficient of the gain of the external engine load,

$\varphi = FI$

- relative change of the angular speed of the KV,

$k_{DH} = KDH$

- factor depending on the self-equalisation coefficient and engine gain coefficient,

ALFAD

- external engine load and

SLOPE

- subroutine of the first derivation (KAPA and ALFAD).

On the basis of the given mathematical model for a marine diesel engine, it is possible to design system-dynamic simulation models.

4. SIMULATION SCENARIOS OF THE MARINE DIESEL ENGINE MODEL

4.1. Simulation scenario of the start-up of an idling marine diesel engine – starting

The marine diesel engine is started at TIME= 0 at zero load.

The diesel engine system is fitted with an electronic universal (UNIREG_DISK1) regulator. After a series of scenarios, we have obtained coefficients of the electronic universal PID regulator for the start-up of the idling marine diesel engine. The coefficients have been obtained with the aid of the heuristic optimisation. The latter implies the use of all familiar methods, including the ones that cannot be expressed in a mathematically exact way, so that expertise in modelling is necessary. Heuristic optimisation in system dynamics involves the "retry and error“ method, a method that could be described as trying to obtain results gradually, i.e. "step by step“.

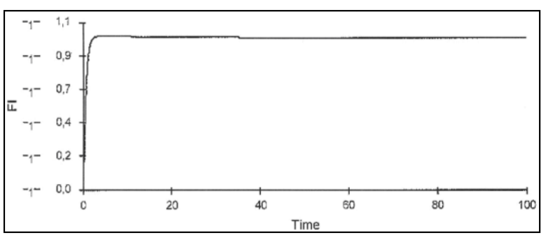


Figure 3. FI=Relative change of the crankshaft angular speed

KPP=20, KPI=0.8, KPD=1.5

Comments on the obtained results from the Scenario I:

The behaviour dynamics of an idling marine diesel engine is entirely in accordance with its model, i.e. with the second order differential equation. The revolution speed regulation has been performed by the electronic universal PID regulator whose elements coefficients have been changing. The heuristic optimisation of the PID regulator parameters has been carried out, resulting in the following combination of coefficients:

KPP=20, KPI=0.8, KPD=1.5

This combination entirely meets the criteria of the revolution speed of a marine diesel engine. The same results have been obtained in the Kongsberg hardware simulator at the University of Split.

5. DISCUSSION

Although simulation data is provided by mathematical model, results of the simulation were similar to those provided by certified Kongsberg's simulator. The fact is that even the real marine diesel engine is entirely in accordance with its model due to PID regulator that is the very same one as it is in real engine and in marine engine simulator. The revolution speed regulation has been performed by the electronic universal PID regulator whose elements coefficients have been changing. The heuristic optimisation of the PID regulator parameters has been carried out by means of very same manner as it is done in real engine since ball governor is not used any more in marine diesel engine.

6. CONCLUSION

The presentation and the results of the simulation of the system-dynamic models for a marine diesel engine lead to the conclusion that applying System Dynamics, a modem scientific discipline, in the research of behaviour dynamics of complex ship systems is an exceptionally effective, rational, and prospective methodology. It is extraordinarily useful in education of both engineering faculty students and graduate engineers of all profiles as it provides a cost-effective, fast and accurate method of acquiring new insights and skills in the field of engineering systems and processes. This brief presentation gives to an expert all the necessary data and the opportunity to collect further information on the same system using a fast and scientific method of investigation of a complex system.

REFERENCES

[1] Jurčević, M., Mitrović, F., Nadrljanski, M.: **System Dynamics and Theory of Chaos in Freight Rate Forming in Shipping**, Promet Traffic & Transportation, Vol.22, 2010, pp 433-438

[2] Munitić, A., Antonić, R., Dvornik, J.:**System Dynamics Simulation Modeling of Ship Gas-turbine generator**, ICCC'03, International Carpathian Control Conference, 26-29 May, 2003, KOŠIĆE, SLOVAK, pp. 357-360

[3] Munitić, A., Milić, L., Milić, I.:System **Dynamics Simulation Model of the Marine Diesel Engine**, Naše More, , No. 43(3-4)/96, Faculty of Maritime Studies, University of Dubrovnik, Croatia, pp. 97-102

[4] Munitić, A., Milić, L., Milković, M.: **System Dynamics Computer simulation model of the Marine Diesel-drive generating set automatic,** 15 IMACS, Berlin, August 1997, Proceedings, Volume 5, Systems Engineering, , Wissenschaft & Technik Verlag, Germany, pp. 245-250

[5] Munitić, A., Milić, L.:**System Dynamics Simulation Modelling of the Diesel-drive Generating Set**,, 9 European Simulation Symposium ESS 97, , Simulation in Energy Systems I, Passau, Germany, October 19-23, pp. 679-684

[6] Munitić, A., Oršulić M., Dvornik, J.: **Computer Simulation of Complex Ship System "Gas turbine-Synchronous generator"**, ISC 2003, The Industrial Simulation Conference, 9-12 June, 2003, VALENCIA, SPAIN, ISBN: 90-77381-03-1, pp. 192-197

LITERATURE

[1] Dvornik, J.: **Sustavno dinamičko simulacijsko modeliranje brodskog Diesel motora pod opterećenjem**, Master thesis, Faculty of Electrical

Engineering, Mechanical Engineering and Naval Architecture, University of Split, Croatia, 2004

[2] Isakov, L.I., Kutljin, L.I.: **Kompleksnaja avtomatizacija sudovljih dizeljnih i gazoturbinmljih ustanovok**, in Russian, Leningrad, Sudostreonnie, 1984

[3] Medica, V.: **Simulacija dinamičkih uvjeta rada dizel motora s prednabijanjem kod pogona električnog generatora**, Doctoral dissertation, Faculty of Engineering, University of Rijeka, Croatia, 1988

[4] Milić, L.: **Dinamičko kontinuirani simulacijski model brodskog dizelmotornog porivnog sustava**, Doctoral dissertation, Faculty of Maritime Studies, University of Rijeka, Croatia, 1996

[5] Munitić, A., Kuzmanić, I., Krčum, M.: **System Dynamics Simulation Modelling Of the Marine Synchronous Generator Set**, International Conference: Modelling And Simulation, May 13-16, 1998, Pittsburgh, USA, 1998

[6] Munitić, A., Oršulić M., Dvornik, J.: **Continuous Computer Simulation Model of the Marine Gas Turbine**, 15th Annual European Simulation Symposium, ESS 2003, Delft, The Netherlands, October 26-29, 2003

[7] Munitić,A.: **Kompjuterska simulacija uz pomoć Sistemske Dinamike**, BRODOSPLIT, BIS Split, Croatia, 1989

[8] Nalepin, R. A., Demeenko, O.P.: **Automatizacija sudovljih energetskih ustanovok**, Sudostroennie, in Russian, 1975

[9] Suprun G.F.: **Sintezsistem elektroenergetiki sudov**, in Russian; Leningrad, Sudostroenie, 1972

[10] Šretner, T.: **Brodski motori s unutarnjim izgaranjem**, Faculty of Engineering and Naval Architecture, University of Zagreb, Croatia, 1970

[11] Veretenikov, L.P.: **Isledovanie procesov v sudovljih elektro-energetieesldh sistemah – teorija i metodiji**, in Russian, Sudostroennie, 1975

Distributed Control in Multi-vehicle Systems

Paul A. Avery
Southwest Research Institute
San Antonio, Texas, 78228, USA
and
Richard Garcia, Ph.D
Southwest Research Institute
San Antonio, Texas, 78228, USA

ABSTRACT

The Southwest Research Institute (SwRI) Mobile Autonomous Robotics Technology Initiative (MARTI) program has enabled the development of fully-autonomous passenger-sized commercial vehicles and military tactical vehicles, as well as the development of cooperative vehicle behaviors, such as cooperative sensor sharing and cooperative convoy operations. The program has also developed behaviors to interface intelligent vehicles with intelligent road-side devices. The development of intelligent vehicle behaviors cannot be approached as stand-alone phenomena; rather, they must be understood within a context of the broader traffic system dynamics. The study of other complex systems has shown that system-level behaviors emerge as a result of the spatio-temporal dynamics within a system's constituent parts. The design of such systems must therefore account for both the system-level emergent behavior, as well as behaviors of individuals within the system. It has also become clear over the past several years, for both of these domains, that human trust in the behavior of individual vehicles is paramount to broader technology adoption. This paper examines the interplay between individual vehicle capabilities, vehicle connectivity, and emergent system behaviors, and presents some considerations for a distributed control paradigm in a multi-vehicle system.

Key Words: Cooperative Vehicle Systems, Cooperative System Dynamics, Intelligent Vehicles

1. INTRODUCTION

Vehicles are becoming more intelligent with the proliferation of in-vehicle technologies and advancements in perception, reasoning, and actuation technologies. The "Intelligent Vehicle" domain is being driven by a number of factors, not the least of which is consumer demand for more driver-assist and active-safety capabilities in their vehicles. Vehicles are also becoming more connected through communication technologies sponsored by the USDOT and the Connected Vehicle program, which seeks to develop standards and technology for in-vehicle, and roadside, radio devices. These devices allow vehicles to communicate with each other (V2V), and roadside devices (V2I) in a variety of ways, for a variety of purposes. The convergence of these two domains however is creating a third type of application domain, the cooperative vehicle system (CVS). A CVS is altogether a different technology than either the Intelligent or Connected vehicle, and yet even as these two constituent technologies race ahead, little work is being done to understand the broader system effects that will emerge when these technologies reach a critical mass, or a tipping point in the jargon of complex systems.

SwRI is an applied research and development non-profit organization, and the Intelligent Systems Department at SwRI has been conducting internally- and client-funded R&D in the areas of "intelligent vehicles", "connected vehicles", and "cooperative vehicle systems", for a number of years. This paper will discuss SwRI's research in these areas, and will focus on the design considerations for cooperative vehicle systems with respect to the broader system dynamics that can emerge. As the constituent technologies are more broadly adopted, the interaction between devices will have an increasing effect on system-wide behaviors, which will then have a feedback effect on the behaviors of individual devices.

The SAE Dedicated Short Range Communications (DSRC) Traffic Information Group Sub-Committee is developing standards for traveler information and vehicle teaming. Efforts are underway to standardize transportation-related information messages displayed for the driver, and SwRI actively participates in these efforts. Research conducted at SwRI has implemented the emerging SAE message set on a three-vehicle platoon, one of which was the MARTI-1 autonomous commercial vehicle (the MARTI software has recently been ported to a military HMMWV). Simulations of larger platoons have also been created, and have provided valuable insight into the effect of vehicle-to-vehicle communication on overall system stability.

Traffic system dynamics have been thoroughly examined in the literature for decades [1] - [4], through both empirical and formal methods. As vehicle population and density increases, and our urban centers become choked, economic and environmental factors become critical for decisions related to urban planning and infrastructure; however, so long as the individual vehicles are controlled by individual drivers, the overall efficiency of dense traffic systems will remain low. This is a property of the system that emerges as a result of interactions between vehicles within their environment. Emerging intelligent vehicle technology, from active safety systems, to inter-vehicle communication, to autonomous capabilities, is enabling a higher degree of connectedness and cooperation among vehicles, which will affect the system-level behavior and performance. However, if a systems approach is not taken in the design of these systems, particularly as they relate to their interaction with other vehicles or roadside devices, undesired system-level effects may emerge.

2. Intelligent Vehicles

The concept of intelligent vehicles encompasses a wide variety of paradigms and technologies, but simply stated are vehicles that possess hardware and software, which enable them to in some way sense their environment, reason on that information,and provide some feedback to the environment, which a typical observer would deem "appropriate" [9]. Thus intelligent vehicles encompass not only the enabling hardware and software, but a design philosophy that this machine will perform an action in a human-relevant environment, usually for the benefit of the human occupants. The means with which this is accomplished is varied; however, stand-alone intelligent vehicle technologies have been driven primarily by automobile manufacturers (OEMs), and sluggish consumer demand. So long as the human still has control of the vehicle's functions, there hasn't been a large demand for the vehicle to take over these functions. Cruise control is a nice feature, which is an intelligent vehicle technology that has been around for a while, but adaptive cruise control, where the vehicle will also maintain a safe following distance from vehicles in front of you, has just recently been introduced, and is not yet widespread [5]. Parallel parking assist is another good example of an intelligent vehicle technology building block. This system uses cameras and software to interface with drive-by-wire capabilities within the vehicle, to self-park. Although useful, a self-parking car by itself is still relatively isolated in its impact the intelligent vehicle domain. Automatic braking systems and traction control are other examples where humans have relinquished control of vehicle functions to hardware/software systems, and the vehicles make decisions within tens or hundreds of milliseconds, and then take "appropriate" action.

These technologies are often developed for safety, comfort, or both. But their development is leading towards a vehicle that is very "aware" of its internal state and surroundings, and with the right software, is able to make decisions quickly, and make physical adjustments to the vehicle's speed and trajectory. The natural progression of this technology is towards fully autonomous vehicles. The challenge in this progression, however, manifests as more human control is relinquished. More trust has to be given over to the system that it will sense the correct parameters in the environment, processes that data appropriately, asses potentially complicated situations, and then make the "right" decision. However, as researchers in this technology, it's easy to see why there is hesitation to relinquish this trust. Human drivers do not always make the "right" decision, and yet we will hold our technology to a much higher standard. And one of our most significant challenges currently is defining the bounds with which our technology will operate to a sufficiently high degree of certainty to pass the "trust test".

3. Connected Vehicles

The concept of connected vehicles is similarly fairly simple, with a more complex application. According to the Research and Innovative Technology Administration (RITA), connected vehicles are vehicles that are equipped with communications technology, which enable the communication between other vehicles (V2V) and with properly equipped roadside devices (V2I). The communication technology in the United States is called Dedicated Short Range Communications (DSRC), and is based on the 5.9Ghz frequency. However, simply allowing vehicles to communicate with each other or roadside devices, does not necessarily make them "connected". The connection comes when a message is successfully passed between two devices. And even then, there may be no further relevance if, for example, the message is invalid. When researchers talk about connected vehicles, again they are not only referring to the hardware and software enabling technology, but to a design philosophy. A connected vehicle is a vehicle with access to information that has been gathered and processed by a completely separate entity. This means the information has the potential to expand the vehicle's awareness of its surroundings in a way on-board sensors could never do. Connected vehicles have the ability to share information about other vehicles, or about their local, or non-local environments. The applications for this are quite expansive, and are the basis for the USDOT Connected Vehicle program.

SwRI has been conducting R&D in connected vehicles for a number of states DOT, as well as commercial vehicle OEMs. Some of the applications are safety-related, such as merge assist, where vehicles communicate to avoid a collision at a highway on-ramp, or in signal phase and timing (SPAT) applications, where vehicles and roadway devices can warn of an impending red-light violator. Other applications are for efficient operations, such as in fleet management, where roadside and vehicle data can be used to assist in re-routing a commercial vehicle to avoid a long and costly delay. Still other applications are targeted for traffic management systems operators, where semi-real-time aggregate data from vehicles can calculate travel times, help uncover developing congestion, or pinpoint the location of an accident. In this case, the previous example is relevant, and other vehicles could be warned and re-routed to avoid the affected route.

One aspect of connected vehicles that begins to hint at the potential of cooperative vehicle behaviors is that of "vehicle teaming". The SAE DSRC Traffic Information Group Sub-Committee is developing standards for traveler information and vehicle teaming. Efforts are underway to standardize transportation-related information messages displayed for the driver, and SwRI actively participates in these efforts. Research conducted at SwRI has implemented the emerging SAE message set on a three-vehicle team, one of which was the autonomous MARTI vehicle. Simulations of larger vehicle team have also been created, and have provided valuable insight into the effect of vehicle-to-vehicle communication on the overall team's string stability, which can also lead to effects of greater fuel efficiency and reduced carbon emission for gas-powered vehicles [4],[6],[7].

4. Cooperative Vehicle Systems

Cooperative vehicle systems are an even more abstract concept than intelligent or connected vehicles. CVS are comprised of individual devices, such as the vehicles and road-side devices we've been discussing, which function together as a cohesive system using their own independent control systems and sets of objectives toward one or more collective goals. Often-times the behavior of a cooperative system, whether it's a vehicle system or not, can seem non-intuitive, and can relatively quickly change from one point of apparent equilibrium to another. These rapid transitions have been studied extensively in the literature for a wide variety of complex, interconnected systems, and are prevalent in dynamic systems where a forcing function is present. Traffic systems in general are good examples of complex systems, which form self-organizing patterns, and can rapidly change states based on a minor perturbation.

Cooperative vehicle systems can thus be approached from different scales. A single team of vehicles moving down the roadway may be considered a CVS depending on how they are teaming. An entire urban network of traffic could also be a CVS. SwRI has developed several example technologies that can be used to one extent or another as part of a CVS. These technologies build on both the intelligent and connected vehicle technologies, but take the next step of enabling vehicles to seamlessly integrate information obtained from other sources into their own world model. Specifically, SwRI has developed several systems for demonstrating the benefit of CVS.

One system enables vehicles to cooperatively share sensor data concerning objects of interest in their environment with other vehicles that are nearby. The use-case SwRI highlighted was pedestrian collisions on roadways due to pedestrians crossing a road illegally. This was demonstrated using small research vehicles with our partners in France, and on the full-size MARTI vehicle in New York City at the ITS World Congress in 2008. A second system we developed is the cooperative convoy system (CCS). This system has a passive and an active component. The passive component is an algorithm that keeps track of which vehicles are in the team and which physical position they occupy, and can provide this data to a user interface for driver information. This passive system was demonstrated at the ITSA Annual Meeting in Washington, DC in 2009. The active component is geared towards vehicles with greater autonomy, and enables autonomous vehicles to send commands to other vehicles in the team to, for example, rearrange the team order. This was demonstrated at the Robotics Rodeo at Ft. Hood in 2009 and at Ft. Benning in 2010, where the MARTI vehicle and two other vehicles were shown to maintain a convoy, and the vehicles could rearrange, or even leave and rejoin the convoy based on the required mission. SwRI is currently working to expand the capabilities of these systems based on commercial and military requirements.

SwRI has also conducted research into the dynamics of large-scale urban traffic systems, and how the technologies of intelligent and connected vehicles might affect these dynamics [8]. Using an agent-based modeling approach, SwRI investigated the effect of increasing the vehicle population that is enabled with connected vehicle technology. The percentage of vehicles equipped with the technology was varied from 0% to 100%, and the model enabled the collection of numerous system-level parameters, such as congestion. In May of 2006, the U.S. Department of Transportation (USDOT) outlined a multi-tiered effort known as the "National Strategy to Reduce Congestion on America's Transportation Network" [23]. This is commonly known as the "Congestion Initiative." The first phase is focused on relieving urban congestion. In conjunction with this effort, RITA is continuing its initiative to improve safety and mobility on the nation's roadways by supporting efforts to integrate standardized traffic management communication infrastructure with vehicle systems. The efforts by RITA and others, including work done by SwRI, to develop infrastructure and vehicle-to-vehicle (V2V) technologies is critical to advancing the intelligent traffic system model. SwRI's research using the agent-based model approach was targeted at answering some of the non-intuitive questions that arise with large-scale traffic systems. One such question is what level of deployment, or market penetration, is required before system-level benefits can be observed. Many other complex systems exhibit the ability to change phases, shift from one state of equilibrium to another, when between 5% and 10% of the constituent parts are affected [13],[14].

If traffic systems are similar, this would bode well for both the commercial OEM interests and the Government agencies looking to deploy technology, because the current thinking in some circles is that 100% deployment is necessary. SwRI's model found that congestion was positively affected when at least 10% of the vehicles were able to communicate with other devices, and that a level of 25% created even lower congestion. The model's limitations however, such as restricted route choices for congested vehicles, created an effect of increasing congestion when 50% to 100% of the vehicles were equipped. This was due to the vehicles having limited routes to choose from, and ultimately selecting the same routes as many other vehicles. SwRI is working to expand the capabilities of this model to include more realistic route choices on the scale of a large urban traffic network, but we feel confident in the vehicle agent behavior model, and that it will scale well. Some of the perception and intelligence algorithms used on the MARTI autonomous vehicle were modified and included as part of the vehicle agents' algorithms. Each agent was also created with slightly different characteristics, for example, individual propensities to travel slightly faster or slower than the posted speed limit. Small random fluctuations in some vehicle parameters were also included to introduce stochastic noise into the vehicle's behavior.

This model was also used to show how teams of vehicles could travel together using V2V communications with the result of increased string stability. This has a significant impact on fuel economy and carbon emissions since vehicles moving at more constant speeds, rather than the oscillating speeds found in unstable vehicle strings, are more efficient. Again, the ABM and simulation allows us to collect data such as fuel usage and carbon emission from individual vehicle agents, which can then be aggregated to the larger traffic system.

In contrast to centrally planned traffic management systems, none of the individual devices in a CVS need contain an understanding of global events or objectives, and emergent behavior plays a bigger role in the overall system's dynamics. Development of cooperative vehicle systems requires an understanding of how the behaviors at the device, or vehicle, level will affect the overall system behavior, and conversely, how that will then feedback to the device level.

5. Distributed Control

The ability to share information among devices plays a critical role in the emergent behavior of decentralized cooperative systems [8],[10], and subtleties such as communications latency, message content, and density of devices will significantly affect the overall performance of the system. All of these are critical components to consider in the development of a CVS. However, those components provide the structure of the system, but the behavior of the system will emerge as a result of how that structure is used. In other words, the behavior of the devices, in this case the vehicles, within the context of this information-rich environment, and within the confines of the physical traffic infrastructure will determine the emergent system properties. Control of such a system cannot be managed as a centralized system with master and slave devices. The "intelligence" in the system cannot be maintained at a central server and then dispatched to devices for them to mindlessly carry out the new instruction. This control paradigm will fail, and has been shown to fail in other domains [11],[12],[19]. One of the greatest weaknesses, for example, in the US power grid is its centralized command and control nature, and is why the "SmartGrid", and other decentralization efforts have been underway for a number of years [20],[21][22]. Similarly, a complex, interconnected traffic system will be robust and fault-tolerant if it is under decentralized control.

The SwRI cooperative convoy system was developed with this paradigm. In the SwRI CCS, the same set of algorithms run on every vehicle in the team. There is no "leader" vehicle, which determines and keeps track of the make-up of the team, which is then "pushed out" to the team members. The CCS developed and demonstrated by SwRI uses a distributed control architecture that is robust and fault-tolerant. Every vehicle in the team runs the same CCS software, which at a rate of about 20hz determines the position, speed, heading, and team position of every vehicle in the team, including the self-vehicle. The algorithm uses nothing other than the vehicle heartbeat message from other vehicles, which contains various parameters including the GPS position and speed of the vehicle. This message is based on the SAE J2735 standard for Dedicated Short Range Communications (DSRC) Message Set Dictionary, and a simple calculation allows the vehicles to properly sort the vehicles in the team. This is important because the CCS was developed such that active vehicle sensors are not necessary, such as LIDAR or cameras. With this calculation, a vehicle can determine which vehicle it is directly following, and can adjust its speed accordingly to maintain a desired following distance. The team ordering also allows vehicles to "sense" speed changes of any vehicle in the team, as soon as they are reported by that vehicle as part of its heartbeat message. With this type of information, vehicles can react to speed changes in vehicles that are not directly in front of them, which means, the entire team can now react as a single cohesive unit if, for example, the front vehicle slows suddenly.

Human drivers can be slow to react, or may overreact to a vehicle slowing in front of them, and this reaction continues for each new vehicle entering the situation. This is how a small perturbation in dense traffic can be amplified into major, long-lasting traffic jams. Vehicle teams that are communicating position and speed changes numerous times a second can be less susceptible to this effect through a more parallel vehicle reaction, rather than each vehicle reacting in series. Since each vehicle in the team maintains its own accounting of the team structure, the loss of a vehicle from the team is not catastrophic. Autonomous convoy scenarios usually involve the concept of a "leader vehicle", which the others follow, which inevitably leads to questions about what happens to the follower vehicles if the leader vehicle is "taken out". In a military sense, this could occur due to hostile action or equipment failure. In a commercial sense the convoy fails if the leader vehicle leaves the convoy for some reason.

The CCS software was developed for both commercial and military scenarios. The military convoy scenario contains a second message set that is sent only when triggered, and contains a command structure, which is a new instruction and is targeted at a specific vehicle. This command may be for a specific vehicle to increase or decrease its following distance, to follow at a lateral offset, or even to change its position within the team structure. There is still no "leader" vehicle in this scenario; however, a single vehicle is given command authority to issue this command message.

Information is a crucial aspect in the behavior of a cooperative system. Depending on the communication structure of a cooperative system, individuals may use direct and/or indirect (stigmergic) forms of communication. Typically, stigmergic communication takes the form of one individual modifying the environment, and another individual sensing that change and reacting to it. In a cooperative vehicle system, this may take many different forms. But one form in particular is relevant to the connected vehicle technology we've been discussing, and that is communication via an infrastructure-based devices, or Roadside Equipment (RSE), which is essentially a DSRC radio. The function of this type of device is to provide a communication bridge between nearby DSRC-equipped vehicles and other vehicles, which are too far away to be communicated with directly. Information collected by an RSE is usually sent to a central traffic management system, which can then distribute the information to other RSE devices. The net effect of this communication structure is that vehicles have access to information that is difficult or impossible to collect themselves, but persists within the CVS environment due to the storage and broadcast capabilities of the RSE devices.

Indirect communication introduces large latencies of course, but the information transmitted in this way is not meant for real-time use, but is used for more deliberative planning. For example, a vehicle may directly sense heavy congestion, which can be relayed to all other DSRC devices within the limitations of the DSRC equipment. This may include other vehicles and RSE devices, but eventually the information that a specific segment of road is experiencing significant congestion makes its way into the larger system repository of information. Further upstream of the congestion, a separate RSE device receives the information and begins locally broadcasting it, where nearby vehicles can receive it, and then use it for route planning purposes. In an effort to better understand the issue of communications latency, and the resultant emergent behavior on CVS, SwRI has developed and demonstrated a number of connected vehicle scenarios, most recently at the ITS World Congress in Orlando, FL.

SwRI is currently conducting research into cooperative vehicle behaviors for autonomous vehicles within military-relevant scenarios using the two MARTI vehicles, one commercial and one military, as well as other vehicles in SwRI's fleet that are equipped with the connected vehicle device technology. Again, the approach with these scenarios is to enable the intelligence to reside at the vehicle level, but to approach the design of the vehicle interaction such that the collection of vehicles accomplishes the mission. A mission of perimeter patrol for example might have several vehicles, with different perception and locomotion capabilities. If the goal of the system is to patrol a perimeter, and presumably either alert on or pursue a detected anomaly, a centralized control paradigm would require a leader, either a leader vehicle or human operator, to coordinate the activities of all the vehicles, assimilate their data, assess their situations, and then send out new commands. A distributed control paradigm would enable each individual vehicle to sense and react to its own environment, while sharing "relevant" information with the other vehicles. What information is relevant is determined by the developers of the system, as are other parameters such as how much information is shared, and how often.

Developers must also determine what action should be taken. This is the behavioral design aspect of the system, and is not necessarily intuitive. For example, if all of the vehicles have a simple behavior that causes them to navigate to a location when another vehicle detects something, the system can be fooled by a false detection, or by purposely creating a diversion, which attracts all of the perimeter patrol vehicles to one location, leaving the rest of the perimeter unobserved. And depending on how many vehicles there are and how big they are, etc, they may all get stuck trying to get to the location. Alternatively, if the vehicles have a slightly more sophisticated behavior that enables only the closest vehicles to the location to go inspect the report, and then, only to a max number of vehicles, this diversionary, or false-positive, effect can be avoided. In this case, no "leader" vehicle told which vehicle to go where, the vehicles collectively decided who goes and who stays, and the decision is largely based on the circumstances of the event, the environment, and their individual capabilities. This scenario can be enhanced further through a simple feedback mechanism. If the first vehicle on-scene confirms the initial report, the "weight" of the report is increased. The vehicles could have a behavior modifier that causes them to react more quickly and decisively as the "report weight" increases. It's not likely a false-detection, although it could still be a diversion.

This feedback mechanism is often seen in the response of social insects [14] - [18]. Ant species will use this to find the best food source or nest site, and bees use this mechanism in the defense of their hives. When a bee stings it releases a pheromone, which attracts nearby bees to sting, which continues in a reinforcing feedback loop [15]. The net effect is that the entire hive is quickly engaged after the first sting has occurred, and the threat is quickly overrun with stings. This effect is considered a form of collective intelligence because the collective acted as a cohesive unit to sense and react to something in the environment, which no single individual could have fully sensed or reacted to. When an ant column selects one path to a food source over another, it is using the intelligence of the collective to do so, since no single ant has explored all possible paths, calculated the speed for each, and determined a winner. Rather, faster ants come back quicker, laying chemical signals faster, which are reinforced by other returning ants. The best trails simply get more pheromone, and the ants are predisposed to follow the strongest chemical trail.

One of us (Garcia) has also conducted research in the cooperative control of unmanned aerial systems (UAS) using rotorcraft. A simulation was developed where a manned aircraft was teamed with a small group of smaller, unmanned aircraft (UA). The specific goal was to have the UA dynamically assemble into standard flight formations with the manned aircraft during flight maneuvers [25]. The small group of UA were developed with two main behaviors: Obstacle avoidance and desired position. Obstacle avoidance was triggered when a UA flew to within a minimum radius, and was implemented by calculating non-linear repulsive potential fields for each aircraft, which of course assumes a 360 degree situational awareness for the aircraft. The behavior for maintaining a "desired position" simply attempted to maneuver the vehicle to within a desired offset location of all other vehicles, again using the potential field method.

The desired offset was calculated by summing the repulsive potential fields from the obstacle avoidance behavior with an attractive potential field generated from a goal position, which was implemented as a dynamic variable, dependent on the formation type and local lead position [24]. The unmanned aircraft were specifically restricted to state information about their local lead aircraft, which similarly to the CVS was the vehicle directly in front of the aircraft. These local lead aircraft were dynamically selected by the order in which they joined the formation. Using the formation type, each UA could determine its constant offset from its local lead, and combining this information with the local lead's state data allowed each UA to determine the current goal position in the absolute coordinate frame [26].

Although this method allowed the group to successfully form correct formations, several severe and sometimes dominant emergent behaviors developed. Specifically, experimentation showed that dynamic motion of one vehicle would cause perturbations within the position of the other aircraft, which due to the control techniques used were not immediately damped. Each aircraft was essentially responding only to the movement of its nearest neighbor, without any understanding of the motion of the group as a whole. This is an identical result as we observed with the string stability of vehicle systems. In the UA system, additional complexity is seen in the formation stability since it can encompass three spatial dimensions, whereas vehicle systems typically only deal with a single dimension, forward, except in the case where a vehicle is following at a lateral offset, or wingman configuration, as SwRI has also demonstrated. The formation of aircraft systems can easily two or three dimensional aspects, and the reaction of a rotor craft to, for example, avoid another aircraft, can certainly take advantage of multiple dimensions.

In an effort to suppress the emergent formation instability, each vehicle also had a third behavior for predicting the intent of other aircraft. Prediction of future motion was based on the orientation of other vehicles, which was possible because the unmanned aircraft were physically identical, and thus would have similar performance characteristics. This orientation cue provided a small decrease in the latency of communicated intent from other aircraft, which allowed the system perturbations to be damped more quickly. Further damping could have been realized by enabling a direct communication link among aircraft, where the intended travel vector could be communicated quickly to the other vehicles in the system, and the system of aircraft could move together as a single cohesive unit.

6. Conclusions

The development and deployment of cooperative vehicle systems provide new challenges for researchers, device manufacturers, and policy makers because the behavior of large-scale cooperative systems emerges as a result of spatio-temporal dynamics within the system's constituent parts, and is heavily influenced by the environment in which it operates. A systems approach must be taken, and adopted as early in the development process as possible, to ensure the performance of such systems operates within agreed upon constraints. Commercial and military applications alike will require this kind of assurance for broad acceptance to occur. Southwest Research Institute is actively working in the areas of intelligent vehicles, connected vehicles, and cooperative vehicle systems to assist in the transition of these technologies to commercialization and broad acceptance. Central to the successful deployment of these technologies, however, is the broader issue of distributed control in multi-vehicle systems, especially when the system consists of mixed-type vehicles, such as when the vehicles in a system range from fully manned, to semi-autonomous, to fully-autonomous. This control strategy cannot take the form of deterministic heuristics; rather, it must be approached from the standpoint of a decentralized system, where individuals behavior deterministically for a given input, but the system as a whole exhibits emergent behavior.

This paper has discussed some of the research performed by SwRI, and examined the interplay between individual vehicle capabilities, vehicle connectivity, and emergent system behaviors. Considerations for a distributed control paradigm in a multi-vehicle system were also presented within the context of recent and active work being performed at SwRI.

REFERENCES

[1] B. D. Greenshields, "A study of traffic capacity," **Proc. Highway Research Board**, vol. 14, pp. 448-477, 1935.

[2] K. C. Chu, "Decentralized control of high-speed vehicular strings," **Transportation Science**, vol. 8, no. 4, pp. 361 – 384, 1974.

[3] Caudill, R. J. and Garrard, W. L.: Vehicle-Follower Longitudinal Control for Automated Transit Vehicles, **J. of Dynamic Systems, Measurement, and Control**, December, pp. 241-248 (1977).

[4] P. Seiler, A. Pant, and K. Hedrick, "Disturbance propagation in vehicle strings," **IEEE Transactions on Automatic Control**, vol. 49, no. 10, pp. 1835 – 1841, 2004.

[5] J. Zhou and H. Peng, " Range Policy of Adaptive Cruise Control Vehicles forImproved Flow Stability and String Stability," in **IEEE Transactions On Intelligent Transportation Systems**", Vol. 6, Issue 2, Pages 229-237, June 2005

[6] P. Barooah and J. Hespanha, "Error amplification and disturbance propagation in vehicle strings with decentralized linear control," in **Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference** 2005, 2005, pp. 4964–4969.

[7] J. Zhou and H. Peng, "String stability conditions of adaptive cruise control algorithms," **in IFAC Symposium**. on "Advances in Automotive Control", Salerno, Italy, 2004.

[8] Avery, P., "Intelligent Vehicle Behavior and Emergent System Stability", **Proceedings of the Society of Automotive Engineers of Japan**, JSAE, May, 2010, Yokohama, Japan

[9] Avery, Paul, A., Brown, Michael (2008), Intelligent Control of Unmanned Systems as an Emergent Property of Situational Awareness, **AUVSI Unmanned Systems North America**,

San Diego, CA

[10] Smith, R., Hadaegh, F., (2007), Closed Closed-Loop Dynamics of Cooperative Vehicle Formations With Parallel Estimators and Communication, **IEEE Transactions on Automatic Control**, Vol. 52, No. 8.

[11] Callaway, D., Newman, M., Strogatz, S., and Watts, D., 2000, Network Robustness and Fragility: Percolation on Random Graphs, **Physical Review Letters**, v. 85, 5468–5471

[12] Carlson, J., M., Doyle, J., 2002, Complexity and robustness, **Proceedings Of The National Academy Of Sciences** Of The United States Of America, v99, suppl. 1, 2538-2545

[13] Beekman, M., Sumpter, D., and Ratnieks, F., 2001, Phase transition between disordered and ordered foraging in Pharaoh's ants, **Proceedings Of The National Academy Of Sciences** Of The United States Of America, v. 98, 9703-9706

[14] De Vries, H., Biesmeijer, J., 2002, Self-organization in collective honeybee foraging: emergence of symmetry breaking, cross inhibition and equal harvest-rate distribution, **Behavioral Ecology and Sociobiology**, v. 51, 557-569

[15] Millor, J., Pham-Delegue, M., Deneubourg, Jl., Camazine, S., 1999, Self- Organized Defensive Behavior In Honeybees, **Proceedings Of The National Academy Of Sciences** Of The United States Of America, v. 96, 12611-12615

[16] Bonabeau, E., Theraulaz, G., Deneubourg, Jl., Aron, S., And Camazine, S.,1997, Self-Organization In Social Insects, **Trends In Ecology & Evolution**, V.12, 188-193

[17] Mallon, EB., Pratt, SC., and Franks, NR., 2001, Individual and collective decision-making during nest site selection by the ant Leptothorax albipennis, **Behavioral Ecology And Sociobiology**, v. 50, 352-359

[18] Roces, F., 2002, Individual Complexity and Self-Organization in Foraging by Leaf-Cutting Ants, **Biological Bulletin**, v. 202, 306-313

[19] North American Reliability Council Steering Group, 2004, Technical Analysis of the August 14, 2003, Blackout: What Happened, Why, and What Did We Learn?, **ftp://www.nerc.com/pub/sys/all_updl/docs/blackout/NERC_Final_Blackout_Report_07_13_04.pdf**

[20] North American Reliability Council, 2004, 2004 Long-Term Reliability Assessment: The Reliability of Bulk Electric Systems in North America, **ftp://www.nerc.com/pub/sys/all_updl/docs/pubs/LTRA2004.pdf**

[21] CERTS, The CERTS Microgrid, Distributed Energy Resources Integration, **http://certs.lbl.gov/DER_H1.html**

[22] CERTS, 2003, Integration of Distributed Energy Resources The CERTS MicroGrid Concept, CEC Consultant Report, P500-03-089F, **http://certs.lbl.gov/pdf/LBNL_50829_app.pdf**

[23] National Strategy to Reduce Congestion on America's Transportation Network, http://isddc.dot.gov/OLPFiles/OST/012988.pdf

[24] L. Barnes, M. A. Fields, and K. Valavanis, "Swarm Formation Control Utilizing Elliptical Surfaces and Limiting Functions," in **IEEE Transactions on Systems, Man, and Cybernetics**--Part B, vol. 39, 2009.

[25] M. A. Fields, E. Haas, S. Hill, C. Stachowiak, and L. Barnes, "Effective robot team control methodologies for battlefield applications," in **IEEE/RSJ International Conference on Intelligent Robots and Systems**, 2009, pp. 5862-5867.

[26] R. D. Garcia, K. P. Valavanis, and A. Kandel, "Fuzzy Logic Based Autonomous Unmanned Helicopter Navigation With A Tail Rotor Failure," in **15th Mediterranean Conference on Control and Automation** Athens, Greece.

Optimization of efficiency of power electronic converter suited for electroplating

Pavol SPANIK[1] Peter DRGONA[1] Michal FRIVALDSKY[1] Roman RADVAN[1] Peter HURTUK[1]

Jozef BUDAY[2]

1) Department of mechatronics and electronics, University of Zilina
Zilina, 01026, Slovakia

2) EVPU (Electronic research and testing institute),
Nova Dubnica, 01851, Slovakia

ABSTRACT

This paper deals with optimization of full bridge converter with synchronous rectifier, used for electroplating. Main focus of the paper is on efficiency increase of power stage through reduction of conduction and switching losses on the secondary side of converter. First, losses and efficiency of the diode rectifier and synchronous rectifier are compared. Next the possible optimization steps are provided and experimentally verificated through the experimental measurements. Several results show that efficiency increased when proper optimization steps were utilized.

1. INTRODUCTION

Switching power supplies for electroplating processes are characterized by low output voltage (tens of volts), which is comparable to the voltage drop of rectifying elements and large current output (thousands of amps). Permissible output current ripple is 2-5% of the average current. Ripple of the output voltage is 1%. Significant losses are caused by current flow through semiconductor structure for required parameters.

The goal of research is minimizing the switching power supply losses using of synchronous rectifier. It is necessary to use a transistor with low resistance in the conducting state of structure. It is also necessary to select an appropriate topology of the synchronous rectifier circuit with respect to the number of transistors. Very important part of topology is control of MOSFET transistors, which enables the transistor to conduct current from zero voltage UDS. Commutation failure is not acceptable. The paper describes the issue of power converter for electroplating processes, its control and also the principle of synchronous rectifier. It brings also experimental verification of the synchronous rectifier, efficiency comparison of power converter with synchronous rectifier and with diode rectifier, where body diode of MOSFET transistor is used for rectification.[1]

2. REQUIREMENTS FOR POWER CONVERTER

All technological processes take place in an environment of galvanic baths of various sizes and designs. Their activities require (with some exceptions) power supplies of small DC voltage with high output currents. In the aspect of required performance we can divide resources for electrochemical applications to several categories. Power supplies with minimal performance (up to 1 kVA) are less important and used for technological processes like degreasing. Medium performance (up to 40 kVA) is typical for the metallization and small refinery links and big performance supplies (from 100 kVA) are used for large refining links and production of chlorine and aluminum.

Supplies specific characteristics depend on the character of electroplating techniques. It is a particular need for low output voltage, very high currents. The fact that the bath is a source of reverse voltage of variable level is important, too. For some technological processes, it is also necessary to allow change the polarity of the output voltage of power supply. Output current with minimal ripple allows achieve superior surface in most cases, ie. homogeneous layer of deposited metal. In the case of minimal current ripple of metalized material surface, liquid products are removed worse and we must remove it with additive surface washing. The optimal value of current ripple is topic of debates for years, but acceptable level of current ripple is set to 2-5% of current average value [1].

3. SYNCHRONOUS RECTIFIER

In synchronous rectifier, diodes are replaced with MOSFET transistors. Diodes must be engaged with respect to the fact, that in presence of driving signal to the gate of transistor, this works in the third quadrant of the static output characteristics (Fig. 1).

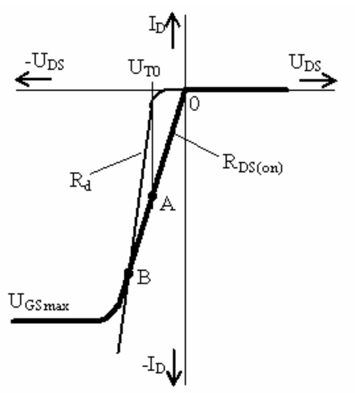


Fig. 1 V-A characteristic of MOSFET transistor and the body diode

In the third quadrant, the body diode is conductive, because it is polarized in forward direction. If the supply output voltage is polarized reverse, transistor works in the first quadrant of its static characteristics and transistor is closed. We can also see that the body diode is in the blocking state. MOSFET transistors are able to conduct current at zero voltage UDS. This is not able in the diode due to its threshold voltage value.

Power losses of the MOSFET transistor for voltage drops from zero to the value of body diode threshold voltage, expresses the equation 1; losses from A to B are expressed in equation 2. These losses represent a voltage drop of synchronous rectifier:

$$\Delta P_1 = R_{DS(on)}.I_D^2 \tag{1}$$

where:

$$\Delta P_2 = \frac{R_{DS(on)}.R_d}{R_{DS(on)} + R_d}.I_D^2 \tag{2}$$

The diode is characterized with permanent voltage drop due to its threshold voltage and the voltage drop due to the dynamic resistance Rd. Power losses of diode expresses equation 3:

$$\Delta P_3 = U_{T0}.I_D + R_d.I_D^2 \tag{3}$$

Fig. 2 shows a comparison of conduction losses for synchronous and diode rectifier.

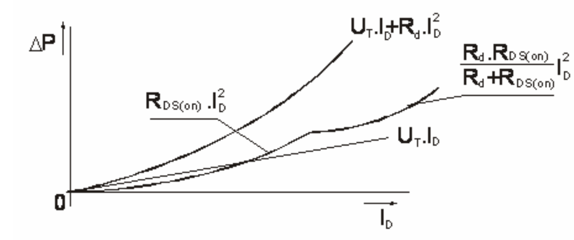


Fig. 2 Comparison of conduction losses

In addition to conduction losses, in semi-conductive structures we can calculate switching losses, control losses and loss in the rectifier commutation. These losses depend on the topology of the rectifier circuit. The energy required for charging and discharging of MOSFET input capacity generates control losses and they are increasing with increasing frequency.

3. CONTROL OF SYNCHRONOUS RECTIFIER

Basic topologies of synchronous rectifiers are derived from conventional. There are two types of control, external control and self-control of synchronous rectifier transistors. An example of external control is in Fig. 3. Control of synchronous rectifier MOSFET transistor contains control and driver circuit, which generates control signal. We must set precise timing of control signals, which is the most critical point of this work. Timing of control signals can be obtained by two ways. The first is to obtain control data from the primary site of converter, from control circuit of the converter. Galvanic isolation of control signal in the secondary side of converter from the primary one is needed. Precise timing of control signals to the gates of MOSFET is important. Additional control losses occur in the external control. The advantage of the external control is non-dependency of control signals due to load variations, simple gate protection and particularly accurate timing of control signals.

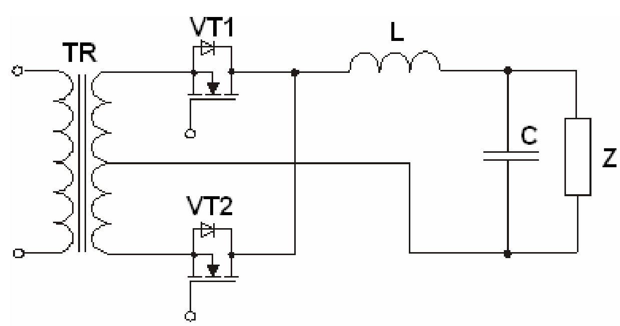


Fig. 3 Control of synchronous rectifier

In self-driven synchronous rectifiers voltage from secondary transformer side is directly used for transistor control. The self-driven synchronous rectifier is shown in

Fig. 4. With the power supply topology on the primary side of transformer the shape control curves are changing. Certain requirements are placed on the control curve.

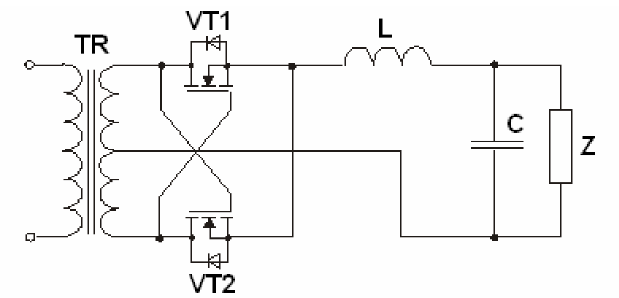


Fig. 4 Self-driven synchronous rectifier

The voltage levels must be high enough for safe switching of transistors, but not too high because of the security of the gates. Energy for control used to charge the parasitic capacity of MOSFET transistors can be used from the magnetizing or dispersion transformer inductance [1], [2].

4. OPTIMIZATION OF THE SYHCNRONOUS RECTIFIER CONFIGURATION

In previous chapter, the most generaly used configurations of synchronous rectifiers are shown. As was mentioned such sollution acts as very perspective idea for applications with high value of output current. Next figure shows dependency of converter's efficiency on output power/current. It can be seen that with the use of synchronous rectifier the efficiency rises by 7% at full load, what is valid for full operation range. This value can be further increased, when several optimization steps are acquired.

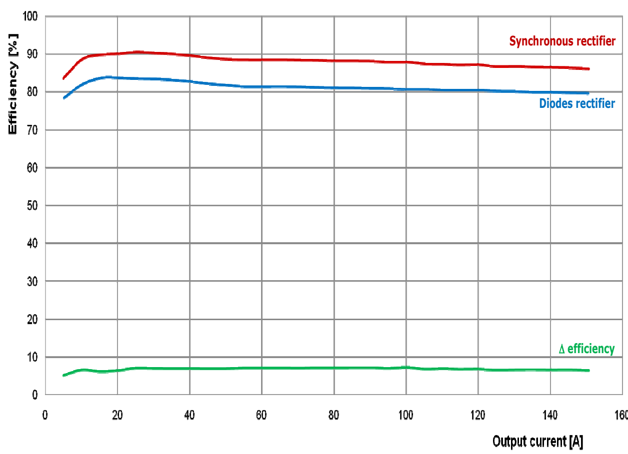


Fig. 5 Comparison of efficiency characteristics

We have been discovering various possibilities and for our purpose we found out next optimization steps:

- utilization of snubber circuit
- utilization of auxiliary diode
- utilization of spikiller

For these types, we have provided a lot of measurements, whereby main investigated parameter was efficiency. In next chapter we will show experimental results for various configurations of optimization proposals.

5. EXPERIMENTAL VERIFICATION

Experimental verifications were made on full bridge converter. Parameters of converter with synchronous rectifier are :

Input Voltage 565 V

Output Voltage 5 V, 10 V, 15 V

Output Current up to 300 A

Switching frequency 17 kHz

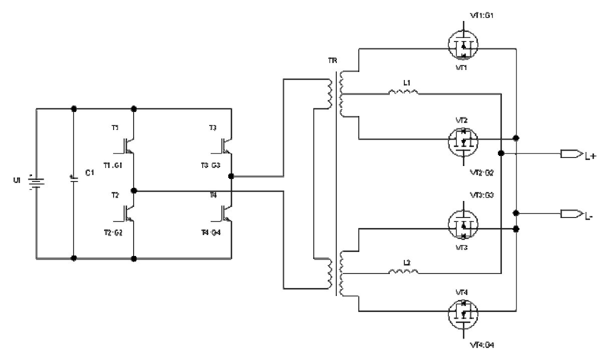


Fig. 6 Full-Bridge converter with synchronous rectifier

Fig. 6 principal schematic of power supply used for electroplating with the use of synchronous rectifier. The configuration is designed as current doubler, whereby several solution according to optimization process were connected to the transistors which are located on the secondary side.

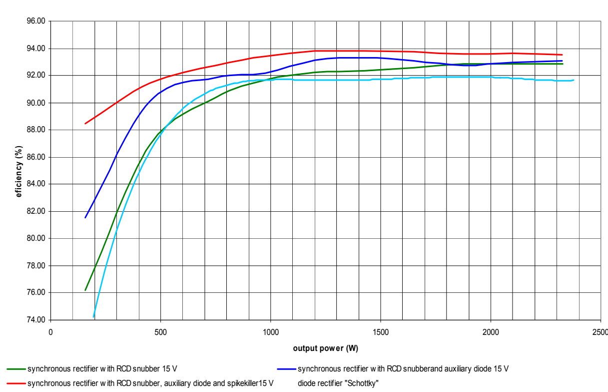


Fig. 7 Efficiency waveforms for various configuration of secondary rectifier for output voltage 15 V

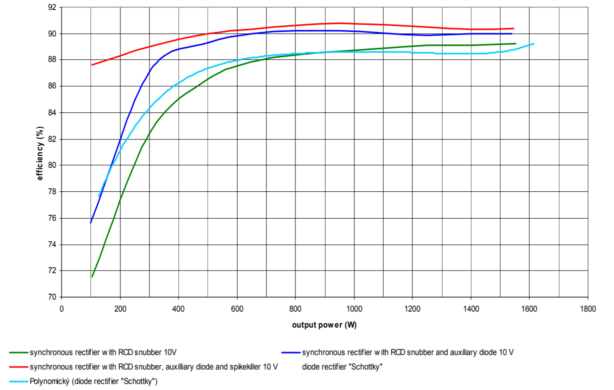


Fig. 8 Efficiency waveforms for various configuration of secondary rectifier for output voltage 10 V

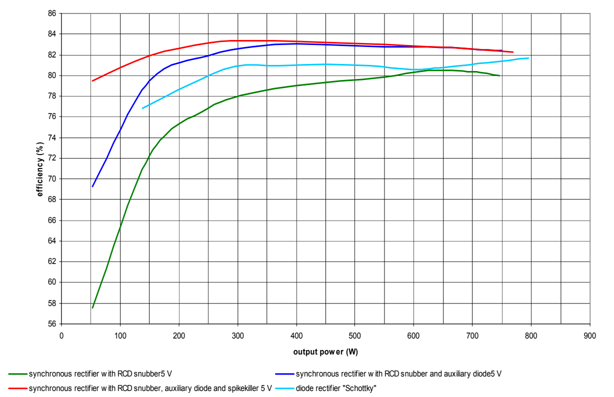


Fig. 9 Efficiency waveforms for various configuration of secondary rectifier for output voltage 5 V

For synchronous rectifier switches, one of the newest family of transistors from IR were used. It deals about MOSFET IRFS3004 with these parameters: ID=240A UDS=40V RDSON(MAX)=1,25mW. Driving signals for synchronous rectifier were taken from main control board of full bridge part. These signals were galvanic isolated from primary part of converter, through high frequency driving transformer. As a driver, the new circuit was made, because commercial available circuits are not sufficient for our application due to high value of transistors input capacitance

Fig.7 to Fig.9 are showing results from the experimental measurements. the red line is efficiency waveform for the configuration of synchronous rectifier with RCD snubber, auxiliary diode and spikekiller. Blue line is related to efficiency waveform for configuration of synchronous rectifier with additional snubber circuit and auxiliary rectifier. Green waveform is valid for synchronous rectifier just with RCD snubber, and cyan waveform is related to diode rectifier. From previous results can be seen that efficiency is highly dependent on output voltage of converter. The lower the value of output voltage, the lower value of efficiency is. Anyway, important result is that with the use of combination of each optimization proposal, the efficiency reach very high values almost in the whole operating range (from 5% of output power up to 100%). Therefore it can be said that optimization process has fulfilled its expectations. Next figure is showing experimental set-up.

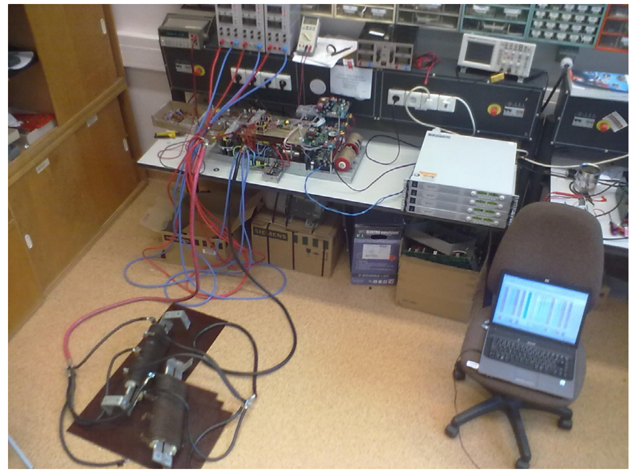


Fig. 10 Experimental set-up

Fig. 8 shows thermal images of devices used in rectifiers. The temperature of diode at 150A load was 101°C, while temperature of MOSFET transistor in synchronous rectifier was 61°C at the same load.

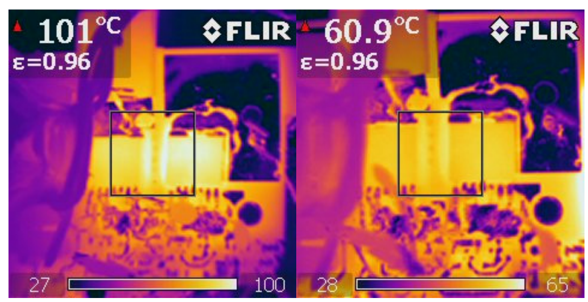


Fig. 11 Thermal image of diode (left) and synchronous rectifier (right)

6. CONCLUSION

The main goal of these experiments was increasing efficiency of switching power supply, using synchronous rectifier. Utilization of the synchronous rectifier is advisable for power supplies for electroplating, with switching frequency higher than 20 kHz. Increase of efficiency is approx. 7.3% if compared to diode rectifier. Used topology for this synchronous rectifier is suitable for high currents, because the current always flows through only one transistor to reduce the losses. The results of this experiment will help in the design of the source with output current of thousands of amps for electroplating industry.

ACKNOWLEDGEMENT

The authors wish to thank to Slovak grant agency APVV for project no. APVV-0138-10 - Research and development of the small power drives with two-phase motors, and for project no. LPP-0366-09 - High frequency power electronic converter. Special thanks is also given to Slovak grant agency VEGA for project no. 1/0943/11.

REFERENCES

[1] http://www.odbornecasopisy.cz, 2004

[2] STOIC, G.- NGUIEN, C.: MOSFET **Synchronous Rectifiers for Isolated, Board-Mounted DC-DC Converters.** In: Conference Proceedings INTELEC'2000 Arizona, USA, 2000

[3] Hargaš L., - Hrianka M., - Lakatoš J., - Koniar D.: **Heat fields modelling and verification of electronic parts of mechatronics systems**, : Metalurgija – Metallurgy, Vol. 49, br. 2 (2010), s. 268-272.

[4] Lee-Hun Kim, Kwang-Su Chun, Chung-Yuen Won, Soo-Seok Kim, Se-Wan Choi.: **An Improved Scheme for High-Efficiency Push-Pull Converter Using Single Winding Self-Driven Synchronous Rectification**, Industrial Electronics Society, 2004

[5] MOHAN, N. – UNDELAND, T.M – ROBBINS, W.P.: **Power Electronics: Converters, Applications, and Design**. Johny Wiley & Sons, New York, 1989

Artificial Psychology: The Psychology of AI

Dr. James A. Crowder, Shelli Friess, MA, NCC
Raytheon Intelligence and Information Systems
16800 E. Centretech Parkway, Aurora, Colorado 80011

Abstract - *Having artificially intelligent machines that think, learn, reason, experience, and can function autonomously, without supervision, is one of the most intriguing goals in all of Computer Science. As the types of problems we would like machines to solve get more complex, it is becoming a necessary goal as well. One of the many problems associated with this goal is that what learning and reasoning are have so many possible meanings that the solution can easily get lost in the sea of opinions and options. The goal of this paper is to establish some foundational principles, theory, and concepts that we feel are the backbone of real, autonomous Artificial Intelligence. With this fully autonomous, learning, reasoning, artificially intelligent system (an artificial brain), comes the need to possess constructs in its hardware and software that mimic processes and subsystems that exist within the human brain, including intuitive and emotional memory concepts. Presented here is a discussion of the psychological constructs of artificial intelligence and how they might play out in an artificial mind.*

1 Introduction

In order to not only design and implement these structures, but also understand how they must interact, cooperate, and come together to form a whole system, we must understand how these structures function within the human brain, and then translate these into how they must function within our "artificial brain." If our system is to possess an "artificial consciousness" then we must understand cognition, intuition, and other capabilities that humans possess [6, 10, 11].

In addition, if we are to create a complete artificial intelligent system, we need to understand how such a system would be received and perceived by people. The reverse is also true in that we must try to understand how the artificial intelligent system will react and perceive people [2, 3, 4, 5].

First, we will explore the concept of "Artificial Psychology" where we look at what it means to have Artificial Intelligence Systems (AIS) resemble human intelligence and when we need to start worrying about the "Psyche" of the Artificial Intelligence system.

2 Artificial Psychology

Psychology is the study of mental processes and behavior of individuals. Artificial Psychology is then the study of the mental processes of an Artificial Intelligence System (AIS) similar to humans [3, 4]. It is about the artificial cognitive processes required for an artificially intelligent entity to be intelligent, learning, autonomous and self-developing [4, 5]. In psychology there are several specialties or focuses of study. Take for example cognitive psychology that studies how the brain thinks and works. This includes learning, memory, perception, language, logic [5, 6, 13, 14]. There is also developmental psychology that considers how an individual adapts and changes during different developmental stages and what is appropriate to consider of a human based on development [17, 18, 19, 20, 21]. There is sports psychology that considers how to affect individual performance and how performance affects the individual. So Artificial Psychology for the purposes of this paper contains the artificial mental process considered necessary to create intelligent, autonomous, self-evolving, artificially cognitive systems. The AIS must mimic human processes in order to be intelligent. After all, isn't the human at the top of the intelligence spectrum?

Artificial Psychology is a theoretical discipline which was first proposed by Dan Curtis in 1963. This theory states that Artificial Intelligence will approach the complexity level of human intelligence when the artificially intelligent system meets three very important conditions:

- Condition 1: The artificially intelligent system makes all of its decisions autonomously (without supervision or human intervention) and is capable of making decisions based on information that is 1) New, 2) Abstract, and 3) Incomplete.

- Condition 2: The artificially intelligent system is capable of reprogramming itself (evolving), based on new information and is capable of resolving its own programming conflicts, even in the presence of incomplete information.[1]
- Condition 3: Conditions 1 and 2 are met in situations that were not part of the original operational system (part of the original programming), i.e., novel situations that were not foreseen in the design and initial implementation of the system.

We believe that when all three conditions are met, then the possibility will exist that the artificially intelligent system will have the ability reach conclusions based on newly acquired and inferred information that has been learned and stored as memories. At this point, we believe the criteria exist, such that the new field of Artificial Psychology needs to be put into place for such systems [4, 5, 6, 7].

The ability of the artificially intelligent system to reprogram, or self-evolve, through a process of self-analysis and decision, based on information available to the system cannot provide the mechanisms for internal inconsistencies within the system to be resolved without adaptation of psychological constructs to AIS methodologies and strategies, and therefore, artificial psychology, by definition, is required.

Current theory of artificial psychology does not address the specifics of how complex the system must be to achieve the conditions presented above, but only that the system is sufficiently complex that the intelligence cannot simply be recorded by a software developer, and therefore this subject must be addressed through the same processes that humans go through to. Along the same lines, artificial psychology does not address the question of whether or not the intelligence is actually conscience or not.

3 Artificial Cognition: What does it mean to be Cognitive?

Cognition is all about thinking. According to the book Ashcroft [22], "…cognition is the collection of mental processes and activities used in perceiving, remembering, thinking, and understanding, as well as the act of using those processes." Adding the term artificial identifies that the nonhuman system is a representation of a living intelligent system. Artificial Cognition refers to how the artificially intelligent machine learns, integrates, recalls, and uses the information that it receives [6, 7, 14, 15, 16]. It is also about how it receives the information. It is difficult at best to create an AIS as complex as human thinking. It is thought that a better understanding of human processes may come from being able to create a truly intelligent machine [22]. It seems that the reverse is also true. Thus, we have a whole new field, Artificial Cognitive Science.

4 Artificial Intuition: What does it mean to be Intuitive

Saying what does it mean to be intuitive is basically asking the question: what does it mean to trust your gut? Another way to say it is to use your heart not your head. Intuition is another way of problem solving that is not the same as logic. According to Monica Anderson[2]:

> "Artificial intuition is not a high-level Logic model so there is no model to get confused by the illogical bizarreness of the world. Systems with intuition then can operate without getting confused with things such as constantly changing conditions, paradoxes, ambiguity, and misinformation."

In her article she also states that this does not mean that sufficient misinformation won't lead such a system to make incorrect predictions, but it means that the system does not require all information to be correct in order to operate. Intuition is fallible, and occasional misinformation makes failure slightly more likely. The system can keep multiple sets of information active in parallel (some more correct than others) and in the end, more often than not, the information that is 'most likely' to be correct wins. This happens in humans, and will happen in Artificial Intuition based systems. It greatly depends on how 'most likely' is defined. If it is only based on the experience of the system, then it can continually fall prey to anchoring and/or the availability heuristic. This implies the need to be supplied with initial data and the use of intuitive guides/rules (heuristics) to help during intuitive conceptual development.

The goal in our AIS is to provide the cognitive intuition required to deal with the world in a real-time, autonomous fashion. Included within the cognitive structure of our the AIS is a Dialectic Argument Structure, which is a methodology constructed for the AIS to deal

[1] This means that the artificially intelligent system autonomously makes value-based decisions, referring to values that the artificially intelligent system has created for itself.

[2] http://artificial-intuition.com/intuition.html

with conflicting and ambiguous information and will allow the system the "cognitive gut" to deal with our paradoxical and ever changing world. In fact, according to Wired.com[3] IntuView, an Israeli high-tech firm has developed "artificial intuition" software that can scan large batches of documents in Arabic and other languages. According to the company's website, this tool "instantly assesses any Arabic-language document, determines whether it contains content of a terrorist nature or of intelligence value, provides a first-tier Intelligence Analysis Report of the main requirement-relevant elements in the document." So if we are going to provide the AIS with the ability to "follow its gut," do we then have to provide it with the emotions we use to make such decisions [9, 13]?

5 Human vs. Machine Emotions

In humans, emotions are still about thinking. According to Marvin Minsky[4]:

> "The main theory is that emotions are nothing special. Each emotional state is a different style of thinking. So it's not a general theory of emotions, because the main idea is that each of the major emotions is quite different. They have different management organizations for how you are thinking you will proceed."

Latest theories look at emotions as the way the brain consciously explains what has happened at a subconscious level. That is, we respond subconsciously (which is faster than subconscious thought) and the brain "explains" what happened with emotions, or arousal states (e.g., fear). So, for the AI system, the emotions produced are a reflection of the type of situation with which the system is dealing.

We can think of emotions in terms of arousal states. When a person is calm and quiet they are more likely to be able to take things in and listen, learn, or problem solve. Think about another emotional state, terror for example. When we are in a state of terror we are not likely to be able to form complex problem solving. Typically with humans, that is why it is recommend to safety plan or practice evacuations. So at the time of crisis or terror the brain doesn't have to perform problem solving. Instead we can just follow the pre-thought out plan. Another example might be the instant you are in a car accident. The body is flushed with adrenaline, heart pounding, hands shaking, probably not a time to work out a calculus problem, for most of us anyway. Often times, emotional states also influence our perception. Take depression for example. It is not likely that a clinically depressed person will simply find the positives of a given situation. There is likely a more doom and gloom recognition. Take a rainy morning, a depressed person who has difficulty finding enjoyment, even if they like the rain may decided to stay in bed, whereas a non-depressed person, who may not even like the rain may, be able to determine that the rain offers opportunity to splash in the water or carry your favorite umbrella.

However, research has also shown that minor stress can actually be good. This seems to point toward the notion that our brains are wired to pay attentions to certain things and that emotions (stress and fear in particular) are an indication that we should pay attention. In their work on Artificial Emotional Memories [3], Crowder and Friess investigated how to utilize these Emotional Memories in order to provide long-term implicit emotional triggers that provide artificial subconscious primers, based on situational awareness metrics.

Similarly, for an artificially intelligent entity, emotions are states of being. If the system is overloaded can it determine what resources to allocate to return to the homeostatic state or state of optimal performance? If for example there are enough indicators to arouse fear can the mediator, so to say, keep operations performing with the correct amount of urgency? Take terrorist threats for example. If an AIS is given enough information to conclude an attack on the country is imminent in the next 24 hours, could the system increase resources to determine the best plan of action? Just as the human level of arousal may contribute to what decisions we make, such as minor chest pain from strained muscle may result in taking an anti-inflammatory or severe chest pains may cause us to call the paramedic[5].

5.1 Basic Emotions

In his book on Emotion and Intuition [1], Bolte concluded:

> "We investigated effects of emotional states on the ability to make intuitive judgments about the semantic coherence of word triads… We conclude that positive mood potentiates spread of activation to weak or remote associates in memory, thereby improving intuitive coherence judgments. By contrast, negative mood appears to restrict spread of activation to close associates and dominant word

[3] http://www.wired.com/dangerroom/2008/10/tech-firms-face/

[4] http://www.aaai.org/aitopics/pmwiki/pmwiki.php/AITopics/Emotion

[5] Psychol Sci. 2003 Sep;14(5):416-21

meanings, thus impairing intuitive coherence judgments."

Bolte found a clear tie between emotions and the ability to have or exhibit intuition. This drives us to a model of basic emotions with the AIS that allow the system to channel resources and find solutions, based on emotional responses to its interaction with its environment. For the purposes of this paper basic emotions are emotions that are in simplest forms. Again they are states of arousal, states of being. For example, calm, alerted, stress, terror or trauma.

The jury is out whether AI will ever have emotions like humans. Consider though that human emotions are based on whether or not human needs are met. In nonviolent communication the author writes about how emotions are based on basic needs. One example is the human need for connection. When humans meet this need they feel valued and loved. As mentioned above, this appears to be a reaction to the mind processing at a subconscious level. It seems that this would be unnecessary for a machine. However, if the AIS is given constraints would those constraints then operate as needs? If the goal was to meet the constraint or satisfy the constraint would the AIS begin to feel. Would the machine reach a level of arousal based on a need or constraint? One possible implementation would be to introduce emotions in response to the system achieving, or not achieving, a goal or objective. This would be analogous to something happening subconsciously and the brain explaining it with an emotion.

Given the studies cited, can we give our AIS a sense of intuition without emotion? If we can, could it then exceed human performance on tasks that emotions influence? How separable is intuition and emotion? The question is: can the AIS perform predictions or problem solving without using states of arousal. We believe the answer is no, and we propose the concept of autonomic nervous system and arousal states within the AIS to provide the "emotion-like" features required to deal with the world around it [3].

Some of the questions that arise from this discussion involve how humans will perceive Artificial Intelligence, particularly with systems that display emotions. And the converse being how would an AI system that has emotional responses perceive humans and their emotional responses?

6 Human Perception of Artificial Intelligence

According to Nass and Moon [23] humans mindlessly apply social rules and expectations to computers. They go on to say that humans respond to cues triggers various scripts, labels and expectations from the past rather than on all the relevant clues of the present, in a simplistic way. In the article, Nass and Moon illustrate three concepts to consider when thinking about human perceptions of AI. The first experiment they describe show that humans overuse social categories by applying gender stereotypes and ethnically identifying with computers. The second experiment they describe illustrates that people engage in over learned social behaviors such as politeness and reciprocity with computers. Thirdly they illustrate human's premature cognitive commitments by how humans respond to labeling. Nass and Moon conclude that individuals apply social scripts that are appropriate for human to human interaction not human computer interaction.

Sarah Harmon[6] shows that gender did not make a significant difference but that people paired characteristics that may have been affected by gender and embodiment. She showed significant correlation between things such as Passive and Likeable for the Male and Understandable and Pleasant for both male and female and Reliable and likeable for the male, thus showing that humans are willing to assign human characteristics to computers. Harmon does state however that we need to consider confounding variables. Harmon also wrote that the degree of the entities embodiment influences how humans deem the characteristics with respect to each other such as the terminal and the robot had significant correlation for understanding/pleasant and friendly/optimistic. Yet the only the terminal showed significant correlation in regard to Understandable/Capable, Pleasant/ Reliable, and Helpful/Reliable.

Considering these authors work one would conclude that how AI is presented to humans will affect how AI is perceived. Even a navigation system in a car that one names seems to take on a whole different meaning once it has a human name. Clearly there are many variables influencing human perception of computers and any AI system. There is much research to be done on how AI could be presented that would make it best perceived by humans.

7 Human Acceptance of Artificial Intelligence

It seems that the non intelligent robotics have had both positive and negative receptions from humans. On one

[6] www.cs.Colby.edu/srtaylor/SHarmon_GHC_Poster.PDF

hand the technology of AI could help humans to function better. For example, as stated earlier, AI could help to detect threats to national security. AI could also be used to train our forces and help solve complex problems. On the other hand AI could take over some human functions. Consider the effects of robots in the auto industry. The technology allowed for machines to do work that humans did. How much can AI out-perform humans? What will happen to human handled jobs and tasks? Thus AI could be well accepted or quickly rejected by humans.

It also seems, as with any technology, there is a usage and learning curve. AI may require humans to learn more about technology in order to be able to interface. As we can see with the internet and cell phone technology there is clearly a generational difference in use and acceptance, and there may be cultural differences in the willingness to accept AI. Thus, as with anything it may take time for humans to accept AI systems on a daily basis.

8 Conclusions and Discussion

Clearly, there is some concern with how the future may go. There have been ethical guidelines for science to follow as they continue to create systems, although this is true in most fields of science. It makes sense to stop and consider ethics and human reactions to AI, after all this is heading to a superhuman technology.

9 References

1. Bolte A, Goschke T, Kuhl J, "Emotion and Intuition." Institute of Psychology, Braunschweig University of Technology, Braunschweig, Germany. annette.bolte@tu-bs.de
2. Crowder, J.A., Friess, S., "Artificial Neural Diagnostics and Prognostics: Self-Soothing in Cognitive Systems." International Conference on Artificial Intelligence, ICAI'10 (July 2010).
3. Crowder, J. A., Friess, S., "Artificial Neural Emotions and Emotional Memory." International Conference on Artificial Intelligence, ICAI'10 (July 2010).
4. Crowder, J. A., "Flexible Object Architectures for Hybrid Neural Processing Systems." International Conference on Artificial Intelligence, ICAI'10 (July 2010).
5. Crowder, J. A., Carbone, J, "The Great Migration: Information to Knowledge using Cognition-Based Frameworks." Springer Science, New York (2011).
6. Crowder, J. A., "The Artificial Prefrontal Cortex: Artificial Consciousness." International Conference on Artificial Intelligence, ICAI'11 (July 2011).
7. Crowder, J. A., "Metacognition and Metamemory Concepts for AI Systems." International Conference on Artificial Intelligence, ICAI'11 (July 2011).
8. DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., and Gray, J. R. (2010). Testing predictions from personality neuroscience. Psychological Science, 21(6):820-828.
9. Marsella, S., and Gratch J., "*A Step Towards Irrationality: Using Emotion to Change Belief.*" 1st International Joint Conference on Autonomous Agents and Multi-Agent Systems, Bologna, Italy (July 2002).
10. Miller EK, Freedman DJ, Wallis JD (August 2002). "The prefrontal cortex: categories, concepts and cognition". *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **357** (1424): 1123–36.
11. Newell, A., "*Unified Theories of Cognition.*" Cambridge MA: Harvard University Press (2003).
12. Damasio A (1994) Descarte's error: Emotion, reason, and the human brain. New York: Gosset/Putnam.
13. Davis M, Whalen PJ (2001) The amygdala: vigilance and emotion. Mol Psychiatry 6:13-34.
14. Eichenbaum H (2002) The cognitive neuroscience of memory. New York: Oxford University Press.
15. Kosko, G., "*Fuzzy Cognitive Maps*," International Journal of Man-Machine Studies, 24: 65-75.
16. LaBar KS and Cabeza (2006) Cognitive neuroscience of emotional memory. Nat Rev Neurosci 7: 54-64.
17. LeDoux JE (1996) The Emotional Brain. New York: Simon and Schuster.
18. LeDoux JE (2000) Emotion circuits in the brain. Annu Rev Neurosci 23:155-184.
19. LeDoux JE (2002) Synaptic Self: How our brains become who we are. New York: Viking.
20. Levine, P., "Walking the Tiger: Healing Trauma." North Atlantic Books, Berkeley, CA (1997).
21. Yang Y, Raine A (November 2009). "Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: a meta-analysis". *Psychiatry Res* **174** (2): 81–8. doi:10.1016/j.pscychresns.2009.03.012. PMID 19833485.
22. Ashcroft, M., "Human Memory and Cognition." Prentice Hall Professional (1997).
23. Nass, C., and Moon, Y., "Machines and mindlessness: Social responses to computers." Journal of Social Issues, 56(1), 81-103.

Relationship between Complexity and Causal Invariance: The NEW Cybernetics Myth

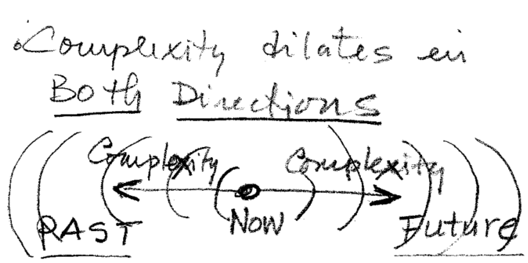
By
Susana Curatolo, CZT Inc. Lawrence KS, USA

ABSTRACT

Since unpredictability is a consequence of Complexity and Chaos, it is shown that unpredictable and uncertain are not equal to each other or interchangeable, but that the former contain the later and that their properties are very distinct. This Gedanken experiment suggests a power law, which, as a consequence of the critical relationship between complexity and causal invariance in that repetitive behavior (nested or not) obeys a simple rule: The Initial Condition (IC) has Knowledge of the Boundary Condition, and the Boundary Condition (BC) has Knowledge of the Order of the system being created by that IC and evolving from it. We show philosophically that by prior experimental work (CZT, S. Curatolo, 2002, 2004, 2006) this is the case, the path of least resistance leads to long range order from short range order systems-this is a proportional power law and it consistently reflects back the Boundary condition resident in the IC. The impulse to initiate the process comes from this reflection, **coherence** between momentum and geometry. Coherence and correlation are directly related through interference, pattern generated through phase differences such that nested or not can be superimposed. The consequent relationship is: as Coherence increases, interference increases, correlation increases, autocorrelation increases, cross-correlation increases, similarity increases, self-similarity increases, self-coherence increases, randomness decreases and through feedback to complexity- of which randomness is a component, until maximal complexity is reached. Maximum complexity is not random and is maximally ordered, long-range ordered and predictable. To perform this task, the self-organizing map preserves the topological relationships of the original space, conjointly lowering the dimensionality. This creates the ability to map any state onto a lower dimensional space, while maintaining their order of proximity. By geometrically approximating the vector distribution in the states of the self-organizing map, it sustains causal invariance of the Riemannian Surface allowing the topology to pullback on itself preserving its structure as it projects recursively and coherently at each state. The qualitatively probability spectrum obeys the order present in the IC as it evolves to the BC and back and re-projects. Precisely because of this, causal invariance guarantees that causal knowledge can predict the effects of manipulation, that is, the correlations will be invariant across changes. So, it is shown here that it is precisely the Initial Condition (IC) that determines the result computationally irreducible, un-decidable and universal. The Initial Condition guarantees irreducibility because of the Observer's limits, that is, Free Will of Choice, this in turn guarantees complexity from simplicity. Apparent Randomization and Run-Away Effect are contained because the Initial Condition has within it the Axiological Adjustment, that is, it directs the movement by inflecting the momentum recursively (spatial pattern or wavelength) determining therefore the velocity. The Initial Condition has the relational constitution of the result. It is further asserted here in this work thru comparative analysis and the proportionality power law, that, the above, akin to 'Karma', is an irreducible function of commitment with room for improvisation, not just a result, but a consequential inflection of things (computations ,events, processes, matter, etc) as they are coming-to be (Hershock,2010)

Introduction

In these experiments we show how to predict unpredictable events in all phenomena by using the equivalence and similarity of effect of the future on the present, and of the past on the present. By connecting complexity and causal invariance in all phenomena we arrive at a Power law from the relationship in the repetitive behavior (nested or not) obeying a simple rule: The Initial Condition (IC) has Knowledge of the Boundary Condition, and the Boundary Condition (BC) has Knowledge of the Order of the system being created by that Initial Condition and evolving from it. There is a simple philosophy inherent in this thinking: An error does not become truth by reason of multiplied propagation, nor does truth become error because nobody will see it (M.Gandhi). So, what is is-whether you think it is or not, it still is. We are constantly confusing the limits of our own perception of what is possible with what is actually possible, and because we do that we reject by reflex-action things that are put forward-not because they are not supported, but because they are outside our limits of perception of sense of possibility. Even if you are in a minority of one, the truth is still the truth. This is why; condemnation without investigation is the height of ignorance….connecting the dots of the real world versus the world we think is real is the analysis necessary to unveil the controlled outcome before the game starts. So the mirroring of the effect of the past onto the present and then invoking its opposite determines the futility of retrodiction, this is the proportion arising from the fact that cause and effect differ so radically in volume that the former can only account for a minuscule fraction of the latter. This is important, because on the basis of an effect alone not much can be said about the cause. Since we would continually avail ourselves of knowledge of conditions and probabilities that pertains to before the effect. By couching a past condition with information from before than, later than and now , a 'bubble' of knowledge is created of prior conditions and probabilities allowing the estimate of the plausibility of alternative causes, then to use these predictions to 'see' which potential cause best fits the current condition. Then, SIMILARLY, the probability of prior events can be determined from those of subsequent events because we can exclude the consideration of the antecedent probability of prior events. This is important because, it opens the power of prediction for probing not just the PAST but also the FUTURE. [This effect is mirrored by complexity]. By knowing many important facts about the longer term future—the nearer term future is predicted. We are in this position with respect to the past. To avoid the futility of retrodiction overwhelming our control by expanding in both directions—towards the future and towards the past---we observe that the complexity increases in both directions from the present.



Retrodiction from the present amounts to predicting the unpredictable—the fluke! The inevitable retrodiction that is always there is useful by opposition (Hegelian maneuver) and mirroring into future. This further means, that solely on the basis of the present or now: complexity dilation = form of past can be made into the form of the future and vice versa by dilation (enlarging or shrinking) because they are similar, they are similar because they have the same FORM, that is they are congruent.

The Map: We assume a Kohonen-like-self-organizing map which preserves the topological relationships of the original space, conjointly lowering the dimensionality. This creates the ability to map any state onto a lower dimensional space, while maintaining their order of proximity. By geometrically approximating the vector distribution in the states of the self-organizing map, it sustains causal invariance of the Riemannian Surface allowing the topology to pullback on itself preserving its structure as it projects recursively and coherently at each state. Self-organizing maps are different from other Maps in the sense that they use a neighborhood function to preserve the topological properties of the input space. SOMs operate in two modes: training and mapping. Training builds the map using input examples. It is a competitive process, also called vector quantization. Mapping automatically classifies a new input vector. This self-organizing map consists of components or **Weights** or **neurons**. Associated with each neuron is a **weight vector** of the same dimension as the input data vectors and a position in the map space. The arrangement of the neurons is a regular spacing in a hexagonal grid. This self-organizing map describes a mapping from a **higher dimensional input space to a lower dimensional** map **space**. The procedure for placing a vector from data space onto the map is to first find the neuron with the closest weight vector to the vector taken from data space. Once the closest neuron is located it is assigned the values from the vector taken from the data space. This type of **architecture** is fundamentally different in arrangement and motivation than feed forward structures. For t from 0 to 1, randomly select a sample; get best match, scale neighbors and increate t in small jumps. Initialize the weight vectors of the Map. The weight that is chosen is rewarded by being able to become more like the randomly selected sample vector or you can preselect and have them fade towards the center-this reduces the # of iterations The U-Matrix value of a particular neuron is the average distance between the neuron and its closest neighbors. In a hexagonal grid, six neurons are considered as the closest neurons. To organize the weight vectors we use the Fundamental matrix. All this is well understood through Kohonen(1999). It's very simple, if they are close together and there is grey connecting them, then they are similar. If there is a black ravine between them, then they are different. How good a map is determined by how strong the similarities between objects are. But what we are really trying to realize is a Rainbow Holographic SOM as a vehicle to show complexity dilation as preserved in the topology. There are limitations to SOM. The main well known ones are: First, it is getting the right data. Unfortunately you need a value for each dimension of each member of samples in order to generate a map. Sometimes this simply is not possible and often it is very difficult to acquire all of this data so this is a limiting feature to the use of SOMs often referred to as missing data. Another problem is that every SOM is different and finds different similarities among the sample vectors. SOMs organize sample data so that in the final product, the samples are usually surrounded by similar samples; however similar samples are not always near each other. If you have a lot of shades of purple, not always will you get one big group with all the purples in that cluster, sometimes the clusters will get split and there will be two groups of purple. Using colors we could tell that those two groups in reality are similar and that they just got split, but with most data, those two clusters will look totally unrelated. So MANY MANY maps need to be constructed in order to get one final GOOD map.

Fig (1) SOM-Bad (Kohonen-1999). Here's an example of just this sort of exact problem happening. You can see in the picture- on the left- that there are three shades of purple that are all in one column.

Yes those three colors are very similar, but two deep shades of purple are split by a lighter shade of purple. The deep purples are more similar to each other and should therefore be near one another as shown in the picture on the right. Second, another limitation is that as the dimensions of the data increases, dimension reduction visualization techniques become more important, and then time to compute them also increases. For calculating that black and white similarity map, the more neighbors you use to calculate the distance the better similarity map you will get, but the number of distances the algorithm needs to compute increases exponentially. However, how long is that tree? If each branch bifurcates into 2 smaller branches with equal radii r1=r2 then r0^3=r1^3+r2^3 is equivalent to r1=2^-1/3 r0~0.794r0~8/9r0, that is that every branch from the generic tree arises from an equal bifurcation of a larger branch. We see this proportion in all nature, like human and mammal's vascular system. By comparing with a typical animal vascular system, the typical capillary radius is 5 microns. Then how many bifurcations from a given primary vessel or trunk of a tree, are required to reach this radius? An aorta of radius r0 with n bifurcations from it will result in such a capillary radius by satisfying: (8/9)^n r0 = 5x10^-6. For r0=0.5cm, n~30. The tot # of vessels in the system after 30 bifurcations is 2^30~10^9.A geometric series calls for n+1 terms so

a+ar+ar^2 +…+ar^n summing to

Sn = a(r^n+1 -1/r-1, r≠1 for n bifurcations. The ratio of the last term in the series to its sum is r^n (r-1)/(r^n+1 -1). If r=2, then the ratio is ½ for large enough n. So, the bifurcations result in ½ this number of vessels ~5x10^8 or ~1.2 x 10^9 physical actual.

(R.Rosen,1967).Distinguishability becomes more difficult observationally as the order increases, the 28th or 29th bifurcation from the several preceding ones. The estimate of the total length of the branching system using a back envelope approach…with the aorta length L0, each bifurcation k produces 2x as many branches and L_k=□Lk-1, assuming 0<□<1.he total length after n bifurcations is

Ln = (2□)^n+1 -1/2□-1 , □≠1/2…So as □ increases, Ln increases. That is a big tree!

The cause and the effect differ radically in volume. This points to the proportion.

In trying to build the map and overcoming the compounding difficulties of the physical medium -H. Donzono et al (2011) are developing a 3D object database using holograms. The extension of this system involves a proposed with limitations-CGH-SOM, a self-organizing map which is composed of units using CGH as memories with a learning algorithm almost identical to conventional SOM with batch updates. CGH (Computer Generated Hologram) is generated in a computer, it simulates the processes related to the Fresnel Holograms used without using optical systems with changing parameters. The 2D info is processed thru the Fresnel Transform Hologram. The drawback of the FTH is that for 3D the FTH has to handle depth info –key for volume and topology. The amplitudes of the reference beam and its incident angle together with the amplitudes of the object beam are used to configure the 3D representation as a SET of POINT Source of Light and the Summation of CGHs are calculated for all the points then become the CGH of the Object. These make up the database from which the SOM is constructed. Since SOM inflicts heavy computational losses, that is why FT to the object is applied prior to SOM, but because the random bit map resulting does not match the input data, random point light sources are used as initial patterns to generate the holograms, before the pre- final step of matching the input object data to the map based on the FTH processing. The final step is to apply SOM std. equation for updating the hologram on the map, because the holograms can be superposed by simply summing the holograms. In CGH-SOM, whether the input object data was randomly generated or laser-scanned had severe limitations coming from several places. First, the matching method of the units by CGH-SOM was not general enough-meaning, simply superposing holograms by adding them caused incorrect matches that could not be checked or prevented-surface irregularities. The FTH is diffraction hologram where the object waves are spherical. If the object has edges or corners, or a point like a cone (this was the object in Danzono et al's experiment, 2011) then it is not seen in the reconstruction. Furthermore, they are Fourier transform based and consequently, the diffraction efficiency of the volume in transmission and reflection are exposure time dependent, the phase retrieval by iterations (large) and their optimizations are temperature dependent. Their object light input is coherent, whereas in Joseph Rosen et al, 2009, the 3D input is incoherent without sacrificing resolution on **complete volume observation**, because it is not scanned, and objects can move-be tracked, yet the images captured are not quite confocal, it requires 3 holograms to be captured, computing the image z (axis) sections from the complex hologram and then enhancing them by deconvolution. It is advantageous in that it record the hologram in synthetic aperture mode in 2D and 3D. via single channel x-axis incoherent interferometer. So, Fresnel holograms are different from each other by the way they spatially correlate the FZP (Fresnel Zone Plate) with the 3D scene. This correlation should be done with FZP that is sensitive to the axial locations of the object points—**otherwise these locations are not encoded into the hologram. If the FZP is dependent on the axial distance of each and every object point—then this means that points which are far from the system, project FZP with fewer cycles per radial length than nearby points and by this condition the hologram can actually create the 3D scene properly. This affects not only volume but complexity dilation. The FZP has to be circular if not you have diffraction efficiency loss; again Fresnel Hologram waves off the object are circular. To overcome these limitations, fast full parallax Fresnel holograms can be computed (Kyoji Matsushima et al 2009) increasing the scale to 2^16x2^16 improving the depth of the CGH Fresnel hologram. Parallax** is a displacement or difference in the apparent position of an object viewed along two different lines of sight, and is measured by the angle or semi-angle of inclination between those two lines. The term is derived from the Greek παράλλαξις (*parallaxis*), meaning "alteration". Nearby objects have a larger parallax than more distant objects when observed from different positions, so parallax can be used to determine distances. We know that the Fresnel Holograms is distance dependent. Even with the faster sampling rates the phase and distance are not preserved if the object has corners and the FZP is not perfectly circular. The surfaces also contribute to discontinuities due to the roughness.

Generating the Map: A SOM from a Rainbow hologram of a Rainbow hologram (RHRH-SOM). First,

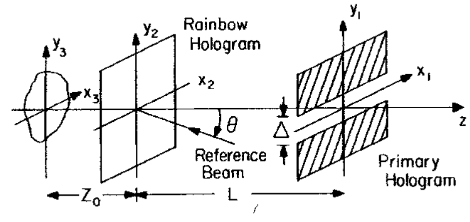
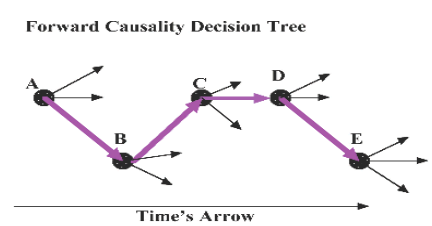
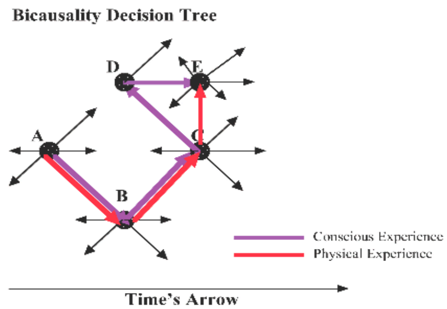


Fig (2) Primary Hologram is illuminated with the conjugate wavefront used in exposing the hologram. A slit of width A is placed over the primary hologram. The image produced by the primary hologram is formed a distance 20 to the left of the rainbow hologram The recording process uses a horizontal slit to eliminate vertical parallax in the output image, greatly reducing spectral blur while preserving 3D for most observers. Stereopsis and horizontal motion parallax, two relatively powerful cues to depth, are preserved. **Stereopsis** is the impression of depth that is perceived when a scene is viewed with both eyes by someone with normal binocular vision. Figure (2) shows an optical arrangement for making a rainbow hologram. A conventional hologram is recorded first (primary hologram). In the reconstruction this primary hologram is illuminated with the conjugate of the reference wavefront used in recording the hologram to produce a pseudoscopic image-real image- of the original object as shown here. A narrow horizontal slit is placed over the primary hologram so only a small range of angles of light rays in the VERTICAL direction is used to form the image. A second hologram is recorded in the plane near the position of the image produced by the primary hologram. The reference wave used in recording the second hologram should have the largest component of its propagation vector parallel to the y-z plane shown here. . I the reconstruction stage for the 2nd hologram, the hologram is illuminated with the conjugate of the reference wave used in recording the hologram The image produced is pseudoscopic, but since the object used in recording the hologram was pseudoscopic, the final image is orthoscopic. This image will appear as if it is being viewed through the re-, a real image is formed of the slit placed on the primary hologram. If the viewer's eyes are placed in the position where the image of the slit is formed, using the exposing wavelength, the image will appear to be the same color as the wavelength used to expose the hologram, even if a white light source is used in the reconstruction step. If the viewer moves his head up or down the color of the image will change if a white light source is used. If the viewer places his eyes in a plane other than the plane in which the image of the slit is formed, the color of the image varies across the image. This means that only a small horizontal section of the image can be seen from any one location, though if the observer changes his/her viewing position, a different part of the object can be seen. If the hologram is illuminated with a laser beam of a different wavelength, the position of the reconstructed image will change. When the hologram is illuminated with a white light source directed from the left of the hologram plate, each color re-constructs a different part of the image at a slightly different angle, so that the whole object is now seen, but with the color varying in the vertical direction. This hologram is a transmission hologram, where the hologram is illuminated on one side, and viewed from the other. Illumination and viewing can be done from the same side if the hologram is mounted onto a reflective surface. Mass replication of such holograms can be done using an embossing process. The embossing process can be carried out with a simple heated press. . In order to permit the viewing of embossed holograms in reflection, an additional reflecting layer of aluminum is usually added on the hologram recording layer. The more mirror like quality of this reflection layer the more accurate reproduction by reflection of the hologram will be obtained. A transmission hologram is one where the object and reference beams are incident on the recording medium from the same side. In practice, several more mirrors may be used to direct the beams in the required directions. Normally, transmission holograms can only be reconstructed using a laser or a quasi-monochromatic source, but a particular type of transmission hologram, known as a rainbow hologram, can be viewed with white (incoherent light) light. In a reflection hologram, the object and reference beams are incident on the plate from opposite sides of the plate. The reconstructed object is then viewed from the same side of the plate as that at which the re-constructing beam is incident. This is an important point of the source of any error in reproducing the hologram and then making a hologram of it. It avoids lasers which considerably preserves the high stereopsis of the image (volume-depth resolution is data). Only volume holograms can be used to make reflection holograms, as only a very low intensity diffracted beam would be reflected by a thin hologram. This is another advantage of incoherent light as opposed to pulsed lasers commonly used in Fresnel holograms. Transmission holograms mounted onto a reflective surface- PET-Thermo-Plastic - polymer/Mylar. A special kind of transmission master needs to be made from the 2D or 3D model Litho of the hologram in order to create an embossed hologram. It is called a slit master (H1) or rainbow transmission master hologram and can consist of more then one slit. For a final embossed rainbow hologram that has the foreground or object roll through the rainbow colors while the background rolls through different colors (shifted) at the same time, two slits can be made on the same plate separated by the color shift distance. Once the transmission master (H1) is made a transmission copy (H2) is made from the H1 master. If there are two slits on the master they can both be imaged onto the H2 copy with one exposure. This is usually done on photoresist. Once the resist copy rainbow hologram is made, it is put in an etching solution that etches away the areas that have not been exposed to light (the destructive interference parts of the fringes) and leaves the exposed regions of the fringes (there are some etching solutions that work opposite and remove the exposed parts and leave the unexposed parts). The hologram can now be seen in on the resist. It is important to maintain the surface relief structure of the fringes so a process that lays down an atom at a time is necessary. To simply coat the resist with silver would "level" out the fringes. Embossing rainbow holograms preserves their entire object stereopsis in every section of the hologram-volume-depth is data. When making a rainbow hologram from an embossed rainbow hologram simply repeats the process of preservation of the properties qualities and quantities of the information embossed the second time. This superposition if done at a faster rate than the original preserves the original in every section of the original, but by going faster than the original we can modulate and create a different one or MASK in the present or past. Complexity dilation and contraction is proportional to spatial-temporal dilation and contraction. This is interpreted in terms of Parallax. In our Rainbow holograms, it is the increase and decrease of Blur that is proportional to apparent speed of motion of the object captured information. This is the reason why in the final SOM we have the IC preserved in the topology. The choice of the the rainbow hologram of a rainbow hologram, was to show masking and to engage in the discussion of retrocausality. We know from the generation and superposition of sub holograms using the various object beams, that incoherence –multiple incoherent exposures –obeys 1/N law. The diffraction efficiency goes as the square of the modulation. The dynamic rang of transmittance is limited from zero to 1... say, and that splitting the range among the sub holograms requires that the modulation of each drops as 1/N, and the average of each exposure adds to that of the others. This is a fundamental phenomena and very useful as we shall see. The order effect amendment to 1/N power law as the multiple exposures are done, in the photoresist process of embossing the last exposure is key whereas in the silver halide process the first exposure dominates. Exposure time matters either way.. Surface reliefs holograms are the best for the SOM outcome. Going back to retrocausality, backwards time causality or reverse time causality is of great importance here in understanding how the future affects the present and the past and effects precede causes, and how this is captured in the RH-RH-SOM. Bi-causality is 2-directional causality, where symmetry in physics and in nature calls for all dimensions of space-time to be symmetric. Archimedes pointed out that bi-causality is best mentally observed by imagining a point outside time. According the Michael Dummett, the logical requirements for reverse causality are…provided that the earlier event is detectable and that it can be detected without disturbing the circumstances under which later in time is claimed to cause the earlier event! So, in advanced and retarded wave observation, while we predominantly observe the retarded wave, it is likely that in every case where energy is being. Schrodinger's cat is alive or it is dead, even when nobody looks…then it does not exist in an uncollapsed wave state. The observer effect becomes a matter of a future view point influencing what has already transpired. Our various sub holograms incoherent and separated by time or wavelength or polarization are influenced by a future decision with a difference factor of 10,000 (O. Hosten& P.Kwiat, U Illinois, 2008). Do-Overs are possible. The Future (the Unpredictable) effect can be predicted by manipulating the vertical parallax. As it increases or decreases (dilates or contracts), so does directly proportional the Blur, the wavelength spread, the source angle spread, and inversely proportional the frequency spread and the time. The complexity dilation and contraction does the same. Forward causality decision trees only move forward through time. In Bicausality decision trees, choices are made both forward and backward in time. See Diagram 1 and 2 below





So, by manipulating the **encoding** of the hologram-the tuning of the Blur, one can **shift the mask** and yet preserve all the originals. You end up with shifted (future) alternatives-perspectives for every viewing angle…Eternity is going on all the time (Charlotte Perkins Gilman). Through time contraction or dilation (speed up slow down or stopped) 2 states can be observed simultaneously…the boundary condition observing the boundary condition- this means the topology folds onto itself. At this point the **Form** of the past and the **Form** of the future are the same and congruent. The Now is continuously dilating and contracting. So when se focus on a point we are simply changing the point of view but we all-ways see-have access to the entire object.

Conclusions& Implications

In the beginning…To perform this task, the self-organizing map preserves the topological relationships of the original space, conjointly lowering the dimensionality. This creates the ability to map any state onto a lower dimensional space, while maintaining their order of proximity. By geometrically approximating the vector distribution in the states of the self-organizing map, it sustains causal invariance of the Riemannian Surface allowing the topology to pullback on itself preserving its structure as it projects recursively and coherently at each state. The qualitatively probability spectrum obeys the order present in the IC as it evolves to the BC and back and re-projects. By using RHRH-SOM approach we have done a Gedanken experiment showing that any part of this proposed-realizable hologram contains the entire whole scalable IMAGE of the volume-depth object topology that can be made with all wavelengths and viewed with all wavelengths and is magnifiable , tunable and decodable. Through the power law explained the effect of future on present and of past on present is demonstrated. By connecting complexity and causality in all phenomena the power law shows IC has Knowledge of BC and BC has knowledge of the ORDER of the system. Retrocausality unveils the controlled outcome before the game starts. The mirroring reveals the proportion arising from cause - effect radically differing in volume and depth, accounts for a little change causes a very large change in both directions from the now, consistent with Karma-the penultimate event in The Buddha's infinitely deep rebirth genealogy.

Acknowledgements: A special thank you to both, Dr. Gisela Dreschhoff for her unwavering financial and technical support and to Dr. Peter D. Hershock for his comments and generous sharing of his information, that helped realize this work.

References

(1)Jan Tusch PhD ,Self-Organized Retinotopic Maps on a Curved Cortical Surface to Predict Visual Field Defects, Computer Science Department, Otto-von-Guericke-Universitat Magdeburg, Germany. Diploma Thesis, September 17, 2004
[2]John A. Adam, A mathematical Nature walk, ISBN 978-0-691-5265-3 Princeton University Press, Princeton and Oxford. USA, 2009
[3] Moshe Safdie, Megascale, Order and Complexity, 2012;[4]http://wwwisg.cs.unimagdeburg.de/bv/files/pdf/thesis_tusch04.pdf;[5] James Gleick, Chaos-Making New science,1987 Penguin Group,ISBN0-1400-92501
[6] David Willcock, 1979, The Source Field , Dutton-Penguin Group ISBN978-0-52595204-6
[7] Richard Gerber,MD, Vibrational Medicine, 3rd Edition, ISBN 978-1-879181-58-8, Bear&Company, Rochester, Vermont, 2001
[8] Malcolm Gladwell, The Tipping Point,ISBN 978-0-316-31696-5 Back Bay Books , Little Brown & Co. 2002
[9] Ranulph Glanville,PhD, Error is not so bad, American Society of Cybernetics and Royal College of Art London and RMIT Melbourne Australia, SouthShea England
[10] Nirmal D.C. Pugh & Derek C. Pugh,PhD, Unveiling Creation-Eight is the Key, ISBN1-887472-43-6, Sunsatar Publishing, Fairfield Iowa USA 1999
[11] His Holiness Maharishi Mahesh Yogi, Maharishi's Absolute Theory of Defence, ISBN 81-7523-000-2, Age of Enlightment Publications, India, 1996
[12] Robert Ehrlich, Nine Crazy Ideas in Science, ISBN 0-691-07001-6, Princeton University Press, 2001
[13] Gino Segre, A Matter of Degrees, ISBN 0-670-03101-1, Penguin Group Press Viking New York NY, 2002
[14] Stephen Wolfram, A new Kind of Science, ISBN 1-57955-008-8, Stephen Wolfram LLC, USA 2002
[15] http://en.wikipedia.org/wiki/Fundamental matrix (computer vision)
[16] http://en.wikipedia.org/wiki/Essential_matrix
[17]Zhengyou Zhang, Determining the Epipolar geometry and its uncertainty: A review, INRIA, International Journal of Computer Vision, vol 27(2),161-168, The Netherlands,1998
[18] Roberto Cipolla and Peter Giblin. Visual Motion of Curves and Surfaces, ISBN 0-521-63251-X Cambridge University Press UK, 2000
[19] L. Montagnier et al, DNA Waves and water, arXiv:1012.5166v1 [q-bio.OT] 23 Dec 2010

[20]Jeffrey Goldstein, Scientific and mathematical roots of Complexity, Adelphi University School of Management and Business, Garden City NY, , www.plexusinstitute.com/edgeware/archive/think/main_filing3.html
[21]Robert M. May and George F. Oster, Bifurcations and dynamic complexity in simple ecological models, The American Naturalist, vol 110,No. 974 July-Aug 1976, University of Chicago press
[22]Kim Krieger, Leonardo's formula Explains why tress don't splinter, Science NOW, http://www.wired.com/wiredscience/2011/11branching-tree-physics
[23] G. J. Whitrow, The Natural Philosophy of time, Thomas Nelson and Sons Ltd. Edinburgh, 1961; Harper and Row NY 1963
[24]Scott Bullock, Scientists debate magical DNA Teleportation, Jan 14 2011, http://www.escapistmagazine.com/news/view/106876-Scientists-Debate-DNA-Tel
[25]Clay Dillow, Can Our DNA electromagnetically teleport itself?, Jan 13 2011 http://www.popsci.com/science/article/2011-01/can-our-dna-electromagnetically-teleport-its
[26]Xu,Lei et al. Imaging analysis of digital holography,OSA,vol13,No.7/Optics Express 2444, 2005
[27]Holographic Methods : Chapter 2, http://www.fou.uib.no/fd/1996/h/404001/kap02.htm [28]Alex Knapp,The Airforce is working on a Holographic Computer, Tech, Forbes Magazine, http://www.forbes.com/sites/alexknapp/2011/12/21/the-air-force-is-working-on-a-holograph...
[29]Edwin Cartlidge, Physicists create a quantum mirror image, Physics World, http://physicsworld.com/cws/article/news/45680
[30]Jae-Weon Lee et al, Quantum Informational Dark energy: Dark energy from Forgetting, School of Computational Sciences, Seoul Korea, arXiv:0709.0047v3 [hep-th] 6 Apr 2008
[31]Quiyuan Pan et al, Institute of Physics, Hunan Normal University, P.R.China, arXiv:0809.0811v2 [gr-qc] 6 Sep 2008
[32]C.H.Bennet et al, Teleporting an Unknown quantum state via dual classical and EPR channels, Pys.Rev.Lett, vol 70, pp.1895-1899,1993
[33]C.H.Bennet et al, Quantum Teleportation project, IBM Research, http://researcher.ibm.com/view_project.php?id=2862
[34]Mark Perakh, Defining Complexity-A Commentary to a paper by C.H.Bennet,
[35]C.H.Bennet, How to define complexity in Physics, and Why, Complexity,Enropy, and the physics of Information, SFI Studies in Science and Complexity, vol VIII, Ed. W.H. Zurek, Addison-Wesley, 1990
[36]William Eckhardt, Causal Time Assymetry, Studies in History and Philosophy of Modern Physics, vol 37, Issue 3, September 2006
[37]George Ainslie, Recursive Self-Preiction in Self-Control and its Failure, Veterans Affairs Medical Center, Coatsville, Preference Change: Approaches from Philosophy, Economics and Psychology, Till Gruene-Yanoff and Sven Ove Hansson, eds, Springer, pp.139-158, 2009
[38] Tetrahedral Symmetry, http://en.www.wikipedia.org/wiki/tetrahedral_symmetry
[39]Rotational Symmetry, http://en.wikipedia.org/rotational_symmetry
[40]A4 in S4, Group Properties, http://groupprops.subwiki.org/wiki/A4_in_S4
[41]Groups of Prime Power Order, Group Properties, http://groupprops.subwiki.org/wiki/Groups_of_prime_power_order
[42]Ilya, Viscount Prigogine, Dissipative Structure Theory and the End of Determinism, http://en.wikipedia.org/wiki/Ilya_prigogine
[43]Prigogine, The end of Certainty, ISBN 0684837056, The Free press, 1997
[44]Indeterminism, http://en.wikipedia.org/wiki/Indeterminism
[45]C.R.Shalizi, Notebooks,Phase transitions and Critical Phenomena, Power Laws and Self Organized criticality, Center for the Study of Complex Systems, U of Michigan
[46]Alessadro Vespignani, How Self-Organized criticality works:A unified mean-field picture, Center for Theoretical Physics, Trieste, Italy, arXiv: cond-mat/9709192v1[cond-stat-mech] 17 Sep 1997
[47] C.R.Shalizi, Power Law Distributions, 1/f Noise, Lhttp://cscs.umich.edu/~crshalizi/notabene/power-laws.html
[48]Duygu Balcan et al,Multiscale mobility networks and the large scale spreading of infectious diseases,arXiv:0907.3304v1 [physics.soc-ph] 20 Jul 2009
[50]Duygu Balcan and Alessandro Vespignani, Phase transitions in contagion processes mediated by recurrent mobility patterns, arXiv:1103.4874v1 [physics.soc-ph] 24 Mar 2011; Nature Physics (20110 doi:10.1038/nphysics1944
[51]Eight Fold Way Physics, http://en.wkipedia.org/eightfold_way_(physics)
[52]Assymetry, http://en.wkipedia.org/assymetry
[53}A. Garrett Lisi, An exceptionally simple theory of everything,SLRI, Incline Village Nevada, arXiv:0711.0770v1[hep-th] 6 Nov 2007
[54]Commutative property, http://en.wikipedia.org/commutative_property
[55]Causality, http://en.wkipedia.org/wiki/causality
[56]Uncertainty principle, http://en.wkipedia.org/wiki/uncertainty_principle
[57]Principle of Bivalence, http://en.wkipedia.org/wiki/principle_of_bivalence
[58]Determinism, http://en.wkipedia.org/wiki/determinism
[59]Fatalism, http://en.wkipedia.org/wiki/fatalism
[60]Daniel Dannett, Freedom evolves,Viking Penguin gr
[61]Destiny, http://en.wkipedia.org/wiki/destiny
[62]Walter Kaufman,Basic writings of Nietzsche, 1967
[63]Deepak ChopraWhy unpredictability, but uncertainty thrives,http://articles.sfgate.com/2010-05-10/news/20891868_1_mice_uncertainty-shocks/2
[64]Mutualism,http://en.wikipedia/org/wiki/mutualism(bio
[65]David Hamilton Wright, A simple, stable model of mutualism incorporating handling time, The American Naturalist,vol134, No.4, pp.664-667,1989
[66]Evolutionarily stable strategy, http://en.wkipedia.org/wiki/Evolutionarily_stable_strateg
[67]Pareto Efficiency, http://en.wkipedia.org/wiki/pareto_efficiency
[68]Prisoner's dilemma, http://en.wkipedia.org/wiki/prisoner%27s_dilemma
[69]Margaret Boden, Creativity and Unpredictability, http://www.stanford.edu/group/SHR/4-2/text/boden.html
[70]David J. T. Sumpter, Collective animal behavior,ISBN 978-0-691-12963-1,Princeton U press,2010
[71]Leah Eldeistein-Keshet, When social creatures act in concert, Physics Today, No.55, November 2011
[72]Statistical Ensembles-Quantum Statistical Mechanics, http://en.wkipedia.org/wiki/Quantum_ensemble
[74]Chris Lintott, The inspiration exchange,U of Oxford, UK, Nature,vol 478,October 2011,p.321
[75] Michael Nielsen, Reinventing Discovery-The New era of networked Science, Princeton U press, 2011
[76]Friedmann equations, http://en.wkipedia.org/wiki/friedmann_equations_Density_parameter
[77]Transcendental Function, http://en.wkipedia.org/wiki/transcendental_function

[78]Hyperbolic spiral, http://en.wikipedia.org/wiki/hyperbolic_spiral
[79]Kaldor-Hicks efficiency, http://en.wikipedia.org/wiki/kaldor-hicks_efficiency
[80]Garrett Hardin,(Resource allocation and) Extensions of the Tragedy of the commons, Science 280(5364)p.682,1998
[81]Jun Sun et al, Wireless Channel Allocation using an Auction algorithm, MIT and NASA SPACE COMM PROJECT No.NAG3-2835
[82]Commonize Costs-Privatize Profits (CC-PP Game), http://en.wikipedia.org/wiki/CC-PP_Game
[83]Hershock, Peter D.,- **Review of Jin Y. Park, Buddhism and Postmodernity: Zen, Huayan, and the Possibility of Buddhist Postmodern Ethics**
[84]-**Liberating Intimacy : Communicative Virtuosity and the Realized Sociality of Ch□an Enlightenment.**
[85](1994).-**The Structure of Change in the Yijing □"** . *Journal of Chinese Philosophy* 36:48-72. 2009
[86]-**Contexts and Dialogue: Yogācāra Buddhism and Modern Psychology on the Subliminal Mind – by Tao Jiang.** *Journal of Chinese Philosophy* 35 (2008)
[87]- (1994). **Person as Narration: The Dissolution of 'Self' and 'Other' in Ch'an Buddhism.** *Philosophy East and West* 44 (4):685-710.
[88] -(2000). **Dramatic]Intervention: Human Rights From a Buddhist Perspective.** *Philosophy East and West* 50 (1):9-33.
[89]-Valuing Karma: A Critical Concept for Orienting Interdependence toward personal and public Good (
[90]-Towards a New Paradigm of DIFFERENCE, (2012)
.[90].- **Review of Jin Y. Park, Buddhism and Postmodernity: Zen, Huayan, and the Possibility of Buddhist Postmodern Ethics**
-**Liberating Intimacy : Communicative Virtuosity and the Realized Sociality of Ch□an Enlightenment.**
[91] (1994).-**The Structure of Change in the Yijing □"** . *Journal of Chinese Philosophy* 36:48-72. 2009
[92]-**Contexts and Dialogue: Yogācāra Buddhism and Modern Psychology on the Subliminal Mind – by Tao Jiang.** *Journal of Chinese Philosophy* 35 (2008)
[93]-(1994). **Person as Narration: The Dissolution of 'Self' and 'Other' in Ch'an Buddhism.** *Philosophy East and West* 44 (4):685-710.
[94] -(2000). **Dramatic [94] Intervention: Human Rights From a Buddhist Perspective.** *Philosophy East and West* 50 (1):9-33.
-Valuing Karma: A Critical Concept for Orienting Interdependence toward personal and public Good (
-Towards a New Paradigm of DIFFERENCE, (2012)
[95]-(1994). **Person as Narration: The Dissolution of 'Self' and 'Other' in Ch'an Buddhism.** *Philosophy East and West* 44 (4):685-710- (1991). **The Structure of Change in the I CHING.** *Journal of Chinese Philosophy* 18 (3):257-285
[96]. (2003)**Renegade Emotion: Buddhist Precedents for Returning Rationality to the Heart.** *Philosophy East and West* 53 (2):251-270.
[84]Tobias Dantzig, Number-The Language of Science,4th ed, ISBN 0-13-185627-8, Plume Publising-Penguin Group, 2005
[97]-(1994). **Person as Narration: The Dissolution of 'Self' and 'Other' in Ch'an Buddhism.** *Philosophy East and West* 44 (4):685-710
[98]- (1991). **The Structure of Change in the I CHING.** *Journal of Chinese Philosophy* 18 (3):257-285.
[99](2003)**Renegade Emotion: Buddhist Precedents for Returning Rationality to the Heart.** *Philosophy East and West* 53 (2):251-270.
[100]Tobias Dantzig, Number-The Language of Science,4th ed, ISBN 0-13-185627-8, Plume Publising-Penguin Group, 2005

ICT Applications for Educational Purposes on Forest and Natural Environment Engineering at Technical University of Madrid, Spain.

Carlos Calderón Guerrero
Departamento de Proyectos y Planificación Rural, Universidad Politécnica de Madrid
Madrid, 28040, Spain.

Carlos Soldevilla Puga
Departamento de Producción Vegetal: Botánica y Protección Forestal, Universidad Politécnica de Madrid
Madrid, 28040, Spain.

Ignacio Bobadilla Maldonado
Departamento de Industrias de los Productos Forestales, Universidad Politécnica de Madrid
Madrid, 28040, Spain.

and

Roberto Rodríguez-Solano Suárez
Departamento de Construcción y Vías Rurales, Universidad Politécnica de Madrid
Madrid, 28040, Spain.

ABSTRACT

The current framework of the EHEA, in which Spanish Universities are immersed, allows the process of transformation from the old degree structure to the new European degree organization, which has started in 2010 and will remain for a maximum period of 6 years. Under these circumstances, it has been considered the integration of ICT methodologies within a multidisciplinary model and their application in the extinguished degree of Forest Engineering in order to help students enrolled and to assist professors under a situation of adaptation from the traditional teaching to new teaching guidelines in a short lap of time. It has been studied the results regarding to ICT used in learning management system or virtual learning environment, evaluation techniques and tutorial activity of the students in different subjects of the degree.
ICT can help to professors in this situation of non-attendance, which allows the use of a range of different evaluative techniques to promote academic achievement and to reduce abandonment index. Tutorial actions under various virtual modalities allow a more direct professor-student relationship with benefits for both sides.

Keywords: Communication and Information Technologies, EHEA, Educational innovation, Blended learning, Teacher training, Extinguished, Forest Engineering

1. INTRODUCTION

As is well known, the Spanish university structure is currently undergoing a profound system of reforms that can be summarized in three essential points:
. - Law 6/2001 of the 21st of December, on Universities [1] as amended by Law 4/2007 [2] to amend the Universities Act, which states in Article 7 the structure of universities.
. - The EHEA (European Higher Education Area), in its legislative development and especially in the RD 1393/2007 Act of the 29th of October, to establish the organization of official university Degrees [3] which provides a new sort of university Degrees, Masters and PhD programs. At this point, it should be remembered that in 2010 was exhausted the possibility of enrolling new students in the old degrees of Engineering and Technical Engineer, with the possibility of completing their training. The students enrolled prior to that date could finish their studies for a regulated period of extinction of the degree.
. - The Lisbon Strategy [4] and ministerial plans collected under the name "Universidad 2015" which is intended to encourage the creation of “Excellent research groups” that makes them competitive in the international community.

Regarding to the degree of Forest Engineering should be pointed that the Technical University of Madrid (UPM), has been the only university that supplied this degree in the Region of Madrid and it is the only one in Spain which it had not been structured the degree in credits in 2010. So it was following the 1971 system Act, partially amended by the Ministerial Act of the 10th of November in 1983 [5]. Finally, with the introduction of the new university degrees in the academic year 2010-2011, framed within the EHEA and after a favorable report from the National Assessment and Accreditation Agency (ANECA) and the regional Parliament of Madrid, was published in the Official Gazette of the 9th of December of 2010 [6], the resolution of the UPM Council, has been established the new Bachelor degree in Forestry and Natural Environment and developed the process of extinction of the old Forest Engineering Degree at the UPM.

In the last years, the Information and Communication Technology (ICT) applications, as a set of tools, auxiliary elements and channels for the use and access to information and media, have been helpful in all educational levels [13][18][19][20][23][25]. In particular, for the Bachelor degree on Forest and Natural Environment Engineering at UPM has been used for different educational purposes in order to support the change that will suffer the students. Due to this reason, a group of professors from the Forest and Natural Environment Engineering faculty that belong to the Forest Educational Innovation Group (INEDFOR) has studied the integration of ICT methodologies within a multidisciplinary model in several Educational Innovation Projects for the degree of Forest Engineering [7] [8] [9] and, in particular, to the extinguished degree [10].

One problem that arises with the fact of extinction, it is the lack of obligation to attend to the courses in person for the prescribed time for students to finish their studies. Therefore there is a drastic change in pedagogy students in order to know the teaching-learning tools that encourage the new EHEA qualifications. In this paper, we will address the topic of the ICT applications, focusing on some of the possible forms of blended learning and B-learning to try to resolve the issue of extinguished degrees and the lack of the traditional lectures in the classroom and the results obtained.

2. METHODS

The objectives of the Innovative Educational Project during the academic course of 2010-2011 consisted in employing a range of ICT tools that they had been used in new undergraduate degrees, by using new evaluating techniques and create line tutoring and support actions to facilitate the monitoring of students in subjects extinguished.

Subjects involved

The three subjects that have participated in this study belong to the second year of the old structure of degree: "Zoology", "Topography, Geodesy and Astronomy" and "Technology and Wood Industries and Forest". Other subjects that joined the project, but the results are not completely shown in this paper are: "Ecology", "Forest Mensuration", "Forest Inventory", "Forest Harvesting", "Applied Hydraulics", "Drawing and Representation Systems" and "Plant Anatomy and Physiology".

ICT used in Learning Management System or Virtual Learning Environment

The three options that have been applied in the Innovative Educational Project since the point of view of teaching B-learning were: 1) the use of institutional Moodle-UPM. This tool offers all kind of academic material and requires the registration of the student in the subject; (see Fig.1). 2) the use of Open Course Ware platform (OCW). This tool offers free access and finally 3) the use of the course website. This tool is free access and non-institutional

Figure 1. Example of Moodle Educational Platform in Forest Zoology.

Evaluative techniques

Evaluation means, in general, a set of systematic collection, analysis and interpretation of valid and reliable, compared with a reference or criterion, allows us to reach a decision as to promote the improvement of the object being evaluated [16].

Figure 2. Intelligent System Assessment Test Self-Assessment developed by the University of Malaga.

The Evaluative techniques are established in the programs of the subjects. Therefore, professor and students should follow this program. Although there are interesting ICT applications, without neglecting traditional assessment techniques (problems and exercises, written assignments and written tests or oral), which consist in a collection of Random Self-Assessment Test, that can be used both to evaluate the student in regular tests throughout the course in the classrooms, or to guide the student to get the skill of knowledge acquired in the process of learning from their own home. The use of educational platforms managing self-test (eg. SIETTE) [12] are very useful ICTs for assessment method in blended teaching (see Fig.2 and Fig.3).

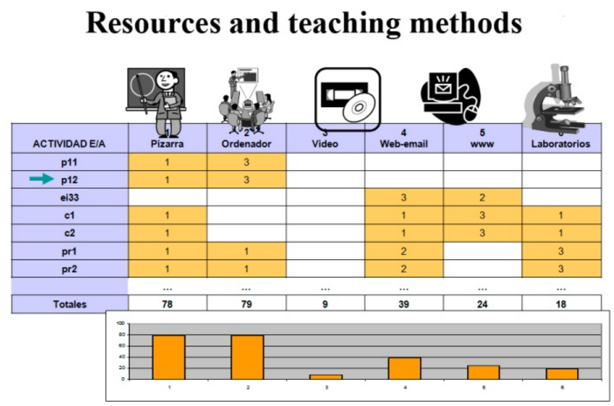
Figure 3. Sample of Self-Assessment Test with the tool SIETTE in Forest Zoology.

Virtual tutorial

As Biggs indicated early in the process of European convergence "Mentoring complements the class presentation and practice, so the student takes much of the burden of work and the tutor is monitoring only" [17]. The virtual tutorial requires the use of all kind of ICT applications. One of the benefits of the Internet to education is to provide an additional media, which it could be parallel, or the only occasionally communication between teacher and students, and among students [14]. This type of mentoring can be developed by email, often through specific software or via educational Platforms, but also through forums or messaging services (Chat) that are available in Moodle too. The great advantage of this approach is the flexible hours, ease of access for students who are shy or who have other occupations and generally, the greater specificity of the issues raised.

3. RESULTS

Having described the procedures handled in the use of teaching tools, evaluation techniques or types of tutoring, the next step was to know which would be applied in each of the subjects of the Innovation Project. An objective was the development of a document to resume the collection of tools used by each subject. The different topics and the number of hours assigned per week for each subject predicts a different use of the techniques (See Fig. 4), but in any situation, it is clear that they all used most of the tools proposed.



ACTIVIDAD E/A	Pizarra	Ordenador	Video	Web-email	www	Laboratorios
p11	1	3				
p12	1	3				
ei33				3	2	
c1	1			1	3	1
c2	1			1	3	1
pr1	1	1		2		3
pr2	1	1		2		3
	...	...	...	...	...	...
Totales	78	79	9	39	24	18

Figure 4. Example of distribution of resources and teaching methods in Wood Technology (I) subject.

Regarding to the use of ICT educational platforms (see Fig.5), we should mention that a 78% of students are already using the Moodle platform, along with the subject of Forest Inventory that has also prepared its own thematic Platform and the implementation of an OCW Web site for Wood Technology (I) subject.

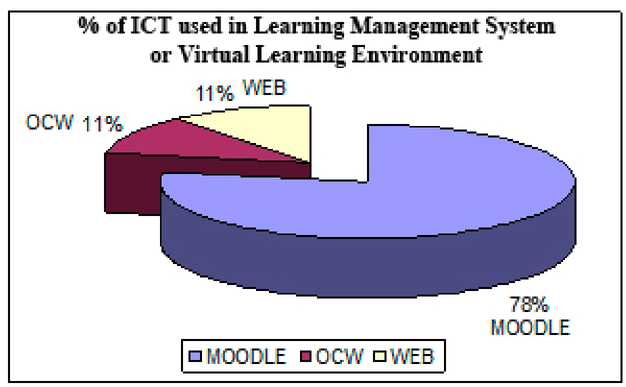


Figure 5. ICT used in Learning Management System or Virtual Learning Environment in INEDFOR Group

Regarding to evaluation techniques, without despising the written or oral examination by attending in person, in the Innovation Project raises the possibility of using other modalities such as practical work or a test that students could develop individually or in groups and the results could be delivered to the professor. According to this, the results of the use by students of Online Assessment Test in three subjects (Forest Zoology, Topography, Geodesy and Astronomy and Technology and Wood Industries 1 are displayed (see Figure 6) vs. other activities.

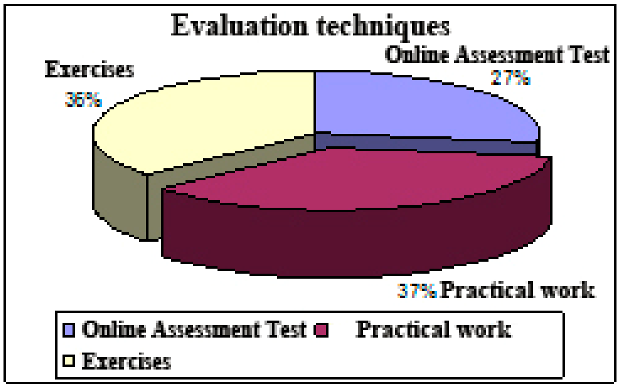


Figure 6. Evaluation techniques employed in the Educational Innovation Project.

The traditional tutorial (group and individual) with the attendance of the student in the method most common used in a 75% (See Figure 7).

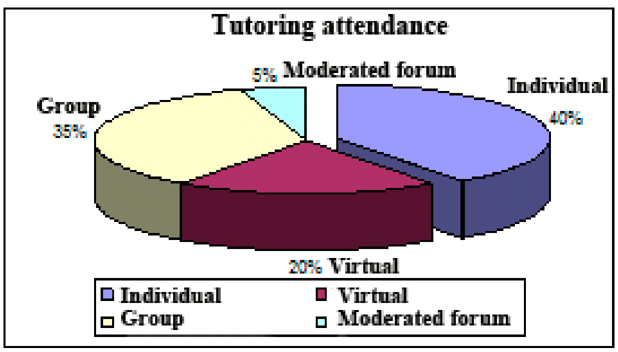


Figure 7. Tutorial activity in the INEDFOR group.

In particular, individual tutorial are used in a 40%. Although in Zoology, the use of virtual mentoring and moderated forums are frequently employed (see Figure 8).

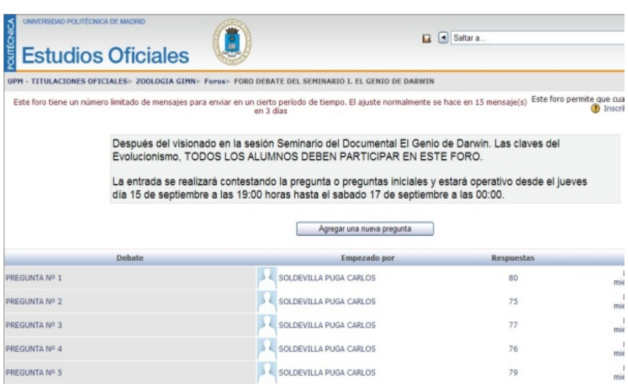


Fig. 8. Discussions in the moderated forum of Moodle platform proposed in Zoology subject.

4. DISCUSSIONS

According to personal situations of the students and the academic circumstances of subjects in the degree program, we have found heterogeneous results. ICTs have a wide range of applications but sometimes there are many factors, like the lack of knowledge of students and professor in ICTs or the particularities of the different subjects.

In subjects as Topography, the practices are carried out in field conditions with real data but coordinated in time with the theoretical material provided, allowing a better use of resources and time, to study simultaneously the theory and practice of the same topic with calculation and presentation of results of the case studies. Therefore, the virtual tutorial, chat, forums, social networking could be a complement, but the practices on site are very important.

In the Wood Technology subject, online tutorial system is very useful and convenient and is therefore used by students, but especially by those who work or for personal reasons, cannot spend much time at the University. These online tutorials are used by submitting documents to review. Normally, the answer of deep doubts by virtual tutorial are not very common because this action implies to express clearly the questions, which it requires to make the effort and a prior synthesis of what students want to ask, and not all students are willing to take that task. This reason could explain the result of 75% for group or individual tutorial vs. 20% in virtual tutorials in all subjects. In the Wood Technology, 30% of students regularly use the virtual tutorial in this subject, compared to 5-10% using the traditional tutorial, which is not very common, during the semester, except in the eve of the examination date where the percentage is extremely increased.

In the present case, due to the reduction of presence, the best option would be the tutorial group as they may be planned in advance and gives opportunity for students to solve exercises or case studies collectively and even answer questions individually. Tutoring attendance voluntary is often used by students to answer questions raised during the review process in all subjects. They are often exploited by minority students and only "academic" tutorial on the eve of evaluation are demanded. They also tend to be individual or in small groups, according to early establishment of a schedule of ongoing attention from the professor.

For degree programs to terminate, this modality of tutorial should be taken into consideration, but previously established goals and to schedule the availability. The aim should always be focused to solve a limited number of questions, always following a previous study of the subject and setting these tutorials in a scheduled program in order to avoid critical moments (as the eve of delivery of exercises or tests).

We have found that the email method is very effective too, although it is necessary to establish a response period (for example, within 48 hours) or set a response time (for example, on Tuesdays from 17:00 to 19:00) in order to avoid misunderstandings. For forums and chats resolution of doubts can be either through the teacher's response or through other peers with the teacher's supervision.

Under these circumstances, we understand that in subjects in which classroom activities diminish or simply disappear (as in the case of subjects and degrees to extinguish), the tutorial can be an invaluable aid for students in understanding and "learning" the contents. These tutorials brings together and coordinates traditional techniques of tutorial classes with tutoring and mentoring virtual group, to facilitate teacher-student contact in careers in this process of extinction. As the same time, we believe that with these tools and actions, we contribute to the student to avoid the sense of abandonment by the academic community and get familiar with the tools and teaching strategies that current students apply in the new graduate degrees.

5. CONCLUSIONS

In the process of extinction of a University degree, several facts affect and have been reflected in previous paragraphs: 1) former and current degrees must coexist at the same time and space, with the same number of professors which leads to coordination problems. 2) Students could have the feeling of abandonment by the academic community.

The student must have no sense of this abandonment by the University because the institution is not giving lectures for former degrees. We must implement educational tools that enable them to complete their studies with dignity and with knowledge of current educational tools that enable them to compete in the professional activities with all students that are attending new degrees.

ICT can help to professors and students in this situation of non-attendance, which allows the use of a range of different evaluative techniques to promote academic achievement and the accurate tracking to reduce abandonment index. The lack of presence of professor allows cost savings and facilitates ICT teacher-student communication and the possibility of individual learning by students.

On-line tutorials have the advantage to write down what student want to solve, which implies to the student to advance a work of synthesis of the problem, to summarize it in a few lines. In the other hand, the answer must also be synthesized, which will take longer than a personal answer on equal terms. Online Assessment Test are a useful tool for students and professors.

5. REFERENCES

[1] España. Ley Orgánica 6/2001, de 21 de diciembre, de Universidades. BOE, 24 de diciembre de 2001, nº 307, p. 49400.

[2] España. Ley Orgánica 4/2007, de 12 de abril, de Universidades. BOE, 13 de abril de 2007, nº 89, p. 16241.

[3] España. Real Decreto 1393/2007, de 29 de octubre, por el que se establece la ordenación de las enseñanzas universitarias oficiales. BOE, 30 de octubre de 2007, nº 260, p. 44037.

[4] Unión Europea. Educación y Formación 2010. Diario Oficial de la UE, de 30 de abril de 2004, nº C 104, p. C104/1.

[5] España. Orden Ministerial de 10 de noviembre, por la que se aprueba el cambio de denominación de las asignaturas en el Plan de Estudios de la EUITF. BOE, de 16 de diciembre de 1983, nº 300, p. 33852.

[6] España. Resolución de 9 de diciembre de 2010, de la UPM. BOE, de 27 de diciembre de 2010, nº 314, p. 107598.

[7] I. Bobadilla, R. Burgos, R. Elena, I. García Viñas, V. Gómez, E. Manrique, C. Martínez, S. Merino, A. Remacha, L. de Salas, C. Soldevilla, S. Villegas, "Coordinación Académica y en Metodologías adaptadas al EEES entre las asignaturas de 2º curso de la EUIT Forestal", II Jornadas Internacionales UPM sobre Innovación Educativa y Convergencia Europea 2008 (INECE´08) UPM, Madrid, (2008).

[8] S. Merino, E. Manrique, C. Soldevilla, C. López, L. de Salas, C. Martínez, "Resultados de adaptación al EEES entre las asignaturas de 2º curso de la EUIT Forestal", III Jornadas Internacionales UPM sobre Innovación Educativa y Convergencia Europea 2009 (INECE´09) UPM, Madrid (2009).

[9] A. San Jose, C. Soldevilla, A. Vivar, A. Castellanos, G. Dorado, P. Galán, A. Gastón, M. Génova, V. Gómez, C. López, E. Manrique, G. Martín, A. Martínez, J. Martínez, C. Molleda, A. Monsó, T. Montoro, C. Pérez, J.E. Saa, J.M. Rodríguez, E. Sadornil, "Adaptación del primer curso de la EUIT Forestal al EEES y a las nuevas titulaciones", III Jornadas Internacionales UPM sobre Innovación Educativa y Convergencia Europea 2009 (INECE´09) UPM, Madrid (2009).

[10] C. Soldevilla, C. Martínez, J.L. Peces, A. Serra, C. Calderón, V. González, Y. Ambrosio, C. López, I, Bobadilla, J.M. Martínez, L. de Salas, V. Gómez, R. Rodriguez-Solano, A. de Blas, A. Bravo. Modelo multidisciplinar de prácticas docentes, técnicas evaluatorias y tutorización en asignaturas con docencia no presencial en planes en extinción de Ingeniería Técnica Forestal. I Congreso Internacional sobre Aprendizaje, Innovación y Competitividad (CINAIC 2011). Madrid. España.

[11] T. Montoro, A. Vivar, V. Gómez, C. Pérez, J.M. Rodrñiguez, C. Molleda, G. Martín, M. Génova, A. San

José, A. Castellanos, A. Monsó, E. Sadornil, A. Martínez, E. Manrique, C.Soldevilla, J. Martínez, S. Pérez, P. Galán, J.E. Saa, G. Dorado, C. López, A. Gastón, "Nueva Titulación de Grado en Ingeniería del Medio Natural (UPM): coordinación académica de las asignaturas de los primeros semetres", VIII Jonadas de Redes de Investigación en Docencia Universitaria, Alicante (2010).

[12] C. Soldevilla, "Empleo de Test para la Tele-educación en la enseñanza de la zoología forestal", Cong. Int. sobre Innovación Educativa para la Educación Superior: hacia un Proceso de Convergencia (Univ. Rey Juan Carlos), Madrid, (2008).

[13] CRUE, Las tecnologías de la Información y las Comunicaciones en el Sistema Universitario Español, Ed. CRUE, Madrid, (2004).

[14] J. Salinas, "Innovación docente y uso de las TIC en la enseñanza universitaria", Revista Universidad y Sociedad del Conocimiento, Vol 1, nº1, pp 1-16 (2004).

[15] A. Bartolomé, "Blended Learning. Conceptos básicos", Pixel-Bit. Revista de Medios y Educación, nº 23, pp 7-20, (2004).

[16] M.J. Rodríguez, "Aplicación de las TIC a la evaluación de alumnos universitarios", Revista electrónica Teoría de la Educación, Vol 6, nº 2, pp 1-23, (2005).

[17] J. Biggs, Calidad del aprendizaje universitario, Ed. Narcea, Madrid, (2005).

[18] M. Cebrián, Enseñanza virtual para la Innovación Universitaria, Ed. Narcea, Madrid, (2003).

[19] F. Aparicio, R.M. González, 1994. La calidad de la enseñanza superior y otros temas universitarios. Editorial: ICE de la UPM, Madrid. p: 236.

[20] S. Brown, A.Glasner, 2003. Evaluar en la universidad. Editorial: Narcea SA. Madrid. p: 225.

[21] M. Cebrian, 2003. Enseñanza virtual para la innovación universitaria. Editorial: Narcea SA. Madrid. p: 196.

[22] P. Hernández, 2001. Diseñar y enseñar. Editorial: Narcea SA. Madrid. Nº pp: 351.

[23] P.T. Knight, 2005. El profesorado en educación superior. Formación para la excelencia. Editorial: Narcea SA. Madrid. p: 286.

[24] R. Werle, 2000. Su propia web. Editorial: DATA BECKER GMBH&CO. KG. Merowingerstr. 30 40223. Düsseldorf. p: 192.

[25] M.A. Zabalza, 2000. Diseño y desarrollo curricular. Editorial: Narcea SA. Madrid. p: 311.

A Study on Effective Factors in Implementing Lifelong Learning From Point of View of Higher Education Experts in Iran

Mahboubeh Sadat Fadavi

Islamic Azad University - Isfahan Science and Research Branch

Assistant Professor of Islamic Azad university-Isfahan Science & Research Branch,Iran

Mahboube_fadavi@yahoo.com

Mohammad Mehdi Mazaheri

Professor of Islamic Azad university- Isfahan Science & Research Branch, Iran

Zohreh Shakibaei

Assistant Professor of Islamic Azad university-Tonekabon Branch,Iran

Shakibaei_z@yahoo.com

Abstract:

This paper tries to study the implementation of Lifelong learning from point of view of higher education experts in Iran. Firstly, the study pays to prerequisites, necessities and social-cultural condition in Iran and then tries to institutionalize long life learning. Present study is qualitative and a kind of documentation and also half organized interview was used. The research sample consists of all Azad and State university faculty members with teaching, writing, and research background in adult education. Sampling method is snow-balling with 24 participants.

KeyWords: Lifelong Learning, Technology, Higher Education, Information, Communication.

Review of Literature:

The provision of learning opportunities for adults has a long and diverse history with a wide rang of providers including the churches, trade unions,co-operation guilds and temperance organizations. However, for much of this century this diversity has not been reflected in policy. With the development of state support for organizations providing learning opportunities for adults learning opportunities become associated largly with a few relatively discrete forms of organized learning.these were provided through the Local Education Authority or Regional Council Adult Education Services or Institutes the pre-1992 Universities extra-mural departments and the workers education association.

From the second world war until the 1970s, this triangle of organizations were constructed as the field of lifelong learning, or as it is more usually put, “adult education”. In other words, the field of adult education and the institutional concerns of adult educatores were the dominant discourses of lifelong learning(chase 1995).this marginalised those forms of life long learning taking place else where, such as in the workplace, the home, local clubs, and lifelong learning as a terrain of practice, policy nd higher education in different policy compartments (Edwards & others,1998) Knowledge essential component of human development, Iran is a country in terms of historical and religious background is very rich, the literature review more than anything to science learning is emphasized and it as a tool of power and knowledge has been learned of human development . From the religious point of view, the Prophet we have stated, science can track although it requires that you travel to China.

This study deals with these questions:

1. What are the prerequisites and possibilities to improve Life long learning?

2. In what extent society social-cultural condition is appropriate for this movement?

3. How can Life long learning be institutionalized in Iran?

The research findings show that benefits of lifelong learning for adults can be:

Keeps their mind sharp, improves memory, increases self-confidence, an inexpensive way to try something new, saves money as they learn to “do it yourself”, and gives you a feeling of accomplishment ,helps to meet people who share your interests ,builds on skills you already have. Making lifelong learning part of one's life also fosters a sense of personal empowerment and, increased self-esteem. It ensures continued growth and intellectual stimulation, leading to a more fulfilling, enjoyable, and enriched lifestyle. However, in order to expand that ability, need to attain a sense of equilibrium, initiative and self awareness.

Finally, participants focused on, financial resources, Possibilities and preparations, Educational environment (Formal education and informal education) , Media, communication and Technology (TV, Radio, Computer, Advertisement) , Curriculum Planning, Great management, and Motivation. Humans will be the valuable and wealth creator resources if they are knowledgeable, proficient, professional, creative and innovative. Knowledge is the new foundation and base for improve the each country.

Table(1)

1. What are the prerequisites and possibilities to improve Lifelong learning?

Capital	**financial resources**
Computer Electronic resources Library Materials and information resources	**Possibilities and preparations**
Media,advertizing,seminars	**Culturalzing**
Experts	**human resources**

Table(2)

2.How can Lifelong learning be institutionalized in Iran?

Formal education (school, university) informal education(in-service training) social education(museum,park,home)	**Educational environment**
TV radio advertisement	**Media and communication**

Table(3)

3. In what extent society social-cultural condition is appropriate for this movement?

interest Commitment & affiliation responsibility ability satisfaction problem solving futurism Curiosity Attempt self esteem Motivation Life needs	**people**
Economical ,social, cultural Development Logical decision making State support	**society**

Referênces:

- Drucker,PF. Knowledge, Work and Society: the Social Transformation of This Century.TheEdwin L. (1994)
Goldkin Lecture, Harvard University's John F. Kennedy School of Government. retrieved March 18, 2005, from http://www.ksg.harvard.edu/ifactory/ksgpress/www/ksgnews/transcripts/ drucklec.html

-Nordstrom N., , Top 10 Benefits of Lifelong Learning, Published on SelfGrowth.com, United States. 2008

-Richard Edwards, Peter Raggatt, Roger Harrison, Ann McCollum, and Judith Calder,. The open University.Recent Thinking in Lifelong Learning 1998.

A constraint-based route search system for smart phone in attraction facilities

Takahiro Shibuya
Faculty of Science and Technology,
Tokyo University of Science,
2641 Yamazaki, Noda,Chiba, 278-8510, Japan
j7411616@ed.noda.tus.ac.jp

and

Hayato Ohwada
Faculty of Science and Technology,
Tokyo University of Science,
2641 Yamazaki, Noda,Chiba, 278-8510, Japan
ohwada@ia.noda.tus.ac.jp

ABSTRACT

It is very difficult for us to find the minimum route to travel in attraction facilities. A searching system for visitors would be useful. Therefore, we constructed a system to find route with the minimum total traveling time. Facility visitors can employ this system on a smart phone. The system is composed of ECLiPSe (ECLiPSe is not a software development environment) and Java Servlet. We concluded that our system is useful and can greatly shorten travel time within the facility.

Keywords: traveling salesman problem, traveling problem in attraction facilities, smart phone, ECLiPSE, Java Servlet,

1. INTRODUCTION

There are many attractions in popular attraction facilities, making it difficult for visitors to find the fastest way of moving about. A visitor picking a very slow route may become tired from walking and waiting, and may miss the opportunity to ride the desired attraction.

I think that visitors can move around quickly if there is a system to find the fastest order.

This traveling problem is similar to the traveling salesman problem. A traveling salesman must find the shortest possible route that visits each city exactly once, when given a list of cities and their pairwise distances.

Research is being conducted to solve the attraction routing problem by applying the traveling salesman problem. For example, research is being conducted to propose how best to move around the "2005 World Exposition, in Aichi, Japan." with two-opt method and a simulated annealing method, which is a one of meta-heuristics method to search for an approximate solution cite.[1]

Research is also being conducted to propose how to travel efficiently with CPLEX.[2]

However, such research has two problems. First, the research employs a fixed waiting time and therefore is not a realistic model. Second, we cannot actually use these systems.

Our goal is thus to develop a realistic model and a route search system that visitors can use on a smart phone.

2. SUBJECT ATTRACTIONS

Subject area

We construct a system for Tokyo Disneyland in Chiba Prefecture in Japan as an example. A visitor chooses eight of the thirty-one attractions he/she would like to visit. We assume that the visitor can ride all attractions without considering service being suspended.

Location of Attraction

Figure 1 presents a map of Tokyo Disneyland, and Table 1 lists the attraction's number, name, and argument. The attraction's argument means the attraction's name in the system. The shortest distance between attractions is measured with "Kyorisoku"[10], a map service to measure distance provided by Mapion Co.,Ltd.(a Japanese company). When we measure distances, we use things to be measured on the map with Kyorisoku. We regard walking speed as three kilometers per hour and set it for the transit time.

We assume that the attraction's entrance and exit are at the center of the attraction and that the center point leads to the nearest street because we cannot determine the attraction's entrance and exit from the map.

We adopt the data of official site in Tokyo Disneyland[11] as the seat-load time. Actual waiting times were published in October 2008 by Tokyo Disneyland. We adopt holiday's and weekday's every thirty minutes waiting times. We utilized the site "Congestion expectation calendar in Tokyo Disneyland"[11] to adopt waiting time.

Business hours actually differ by date, but we assumed that Tokyo Disneyland is open from 8 a.m. to 10 p.m.

Table 1
Attraction in Tokuyo Disneyland

number	name of attraction	argument
1	Splash Mountain	splash
2	Space Mountain	space
3	Pirates of the Caribbean	carib
4	The Haunted Mansion	haunted
5	Big Thunder Mountain	big
6	Pooh's Hunny Hunt	hunny
7	It's a Small World	smallworld
8	Star Tours	startours
9	Jungle Cruise	jungle
10	Western River Railroad	westernriver
11	The Enchanted Tiki Room	tiki
12	Westernland Shootin' Gallery	westernlandshoot
13	Country Bear Theater	bear
14	Mark Twain Rivarboat	marktwain
15	Tom Sawyer Island Rafts	ikada
16	Beaver Brothers Explorer Canoes	canoes
17	Peter Pan's Flight	peterpan
18	Snow White's Scary Adventures	shirayuki
19	Pinocchio's Daring Journey	pinocchio
20	Dumbo The Flying Elephant	dumbo
21	King Arthur Carrousel	carrousel
22	Alice's Tea Party	alice
23	Roger Rabbit's Car Toon Spin	cartoon
24	Minnie's House	minniehouse
25	Mickey's House and Meet Mickey	mickeyhouse
26	Gadget's Go Coaster	gocoaster
27	Chip'n Dale's Treehouse	treehous
28	Donald's Boat	donaldboat
29	Buzz Lightyear	buzz
30	Star Jet	starjet
31	Autopia	raceway

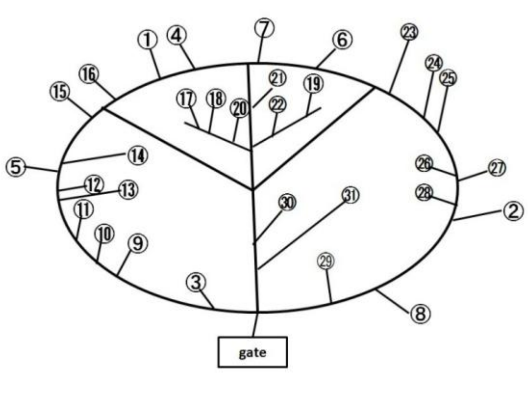


Figure 1. Attraction Map

3. SYSTEM OUTLINE

ECLiPSe Program

We constructed a system with ECLiPSE. Though it often makes mistakes, the ECLiPSe we are using is not a multi-language software development environment comprising an integrated development environment (IDE). ECLiPSe is a software system for developing and deploying Constraint Programming applications. For example, it is useful in the areas of optimization, planning, scheduling, resource allocation, timetabling, transport, etc. The ECLiPSe language is largely backward-compatible with Prolog and supports different dialects.

This system has one program to calculate the minimum time when considering transit time, waiting time, and seat-load time. A visitor inputs the starting time, the attraction's name(from Table 1), and the day of the week(weekday or holiday), and the system outputs the minimum time, the arriving time, and the route. The minimum time is the time from starting at the entrance to returning to the exit(User can also search the minimum time from starting at any attractions to returning to any attractions). Furthermore, the entrance and the exit are collocated in Tokyo Disneyland.
More than two attractions and less than eight attractions can be input to, this system, so visitors can search for routes for from two to eight attractions.

We present the following formula to calculate the time from the entrance to exit. (The following sign is defined when we present a formula.)

I : a set of attraction
T :a set of time zone
M_{ij} : transit time of from "attraction $i \in I$"
to "attraction $j \in I$"
M_{kg} : transit time of from "attraction $k \in I$"
to exit
M_{gh} : transit time of from entrance to
"attraction $h \in I$"
W_{it} : waiting time of "attraction $i \in I$"
when time is $t \in T$
P_i : seat-load time of "attraction $i \in I$"
A : a set of branch$\{ (i,j) \mid i,j \in I\}$

The following formula illustrates the minimum time.

・Formula considering transit time, waiting time, and seat-load time

$$M_{gh} + \sum_{(i,j)\in A}(M_{ij} + W_{it} + P_i) + M_{kg} \rightarrow \min$$

We next describe the system structure. First, we implemented the above calculation and timetable in ECLiPSe software. This program can be executed on an ECLiPSe terminal. The following figure depicts the execution screen. B in Results in Fig.2 is the minimum time, and the side of result(space - splash...) represents the route.

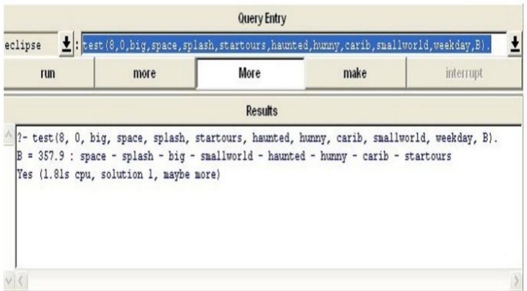


Figure 2. Execution Screen With ECLiPSe

System structure

We can construct a web application with this ECLiPSe program in two ways.

One way is to interact with Java. ECLiPSe can interact with Java [13] using a Java-ECLiPSe interface. The Java-ECLiPSe interface is a general-purpose tool for interfacing ECLiPSe with Java, Sun's popular object-oriented platform-independent programming language. For example, it could be used to write a Java graphical front-end for an ECLiPSe optimization program or to interface ECLiPSe with a commercial application written in Java.

The other way is to execute Java on the web with a Java Servlet. A Java Servlet is a program for dynamically generating an HTML document for a web page with Java. We construct a dynamic web site for a smart phone to use this function.

Figure 3 illustrate the structure of our system. The system incorporates Apache and Tomcat; Apache is used to build a web server, and Tomcat is used for the Java Servlet.

First, visitor accesses the web site we made with a smart phone. The visitor then inputs the starting time, weekday or holiday, and the eight attractions he/she would like to ride, and this information is passed to the Java program, and to the ECLiPSe program. The path that the ECLiPSe program should search for the minimum time must be described in the Java program.

The ECLiPSE program is then run on Java and calculates the route. The result is passed to the Java program as a type of object; therefore, we need to transform the object type into a string type. Because the string information contains the extra information, the character strings have to be split and the extra information have to be removed with Java's method. When the object is split, nine words are produced and stored in a list. The first of the nine words is the minimum route time. The remaining eight words are the attractions' names(the argument names in Table 1) stored in a list in the order of travel. For example, the first word in the list indicates the first attraction to which the visitor should travel, and the eighth word indicates the last attraction to be traveled to.

We note that this system is for Japanese people, and so on, the output needs to be translated into Japanese. The translation is performed in Java. In addition, the arrival time is also calculated in Java.

When the above process is completed, the system will output the total time, arrival time, and minimum route. Finally, the data is output on the web with the Java Servlet.

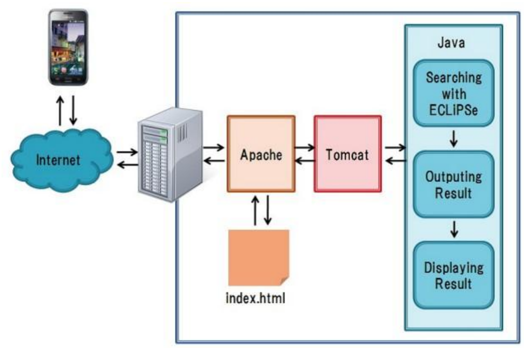


Figure 3. System Structure

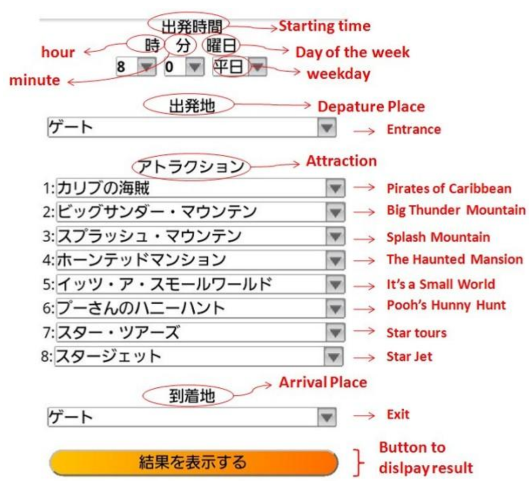


Figure 4. Top Screen

Interface

Our system has the following interfaces. Figure 4 depicts the top screen. This screen is displayed when a visitor first accesses our site. The visitor then inputs the starting time, weekday or holiday, and attractions from this screen. After two or three seconds, the result screen (Fig. 5) will display the starting time, arrival time, weekday or holiday, the sum time, and the minimum route for attractions.

4. SYSTEM VERIFICATION

In this section, we demonstrate the effectiveness of our system.

Verification environment

The system employ hardware and software in Table 2.

Table 2
Hardware and Softwares to be used

CPU	Intel(R) Celeron(R) CPU 2.00GHz
Memory	1.03GB
OS	CentOS 5.3
ECLiPSe	ECLiPSe 6.0
Java	Jdk1.6.0_25
Tomcat	Tomcat 6.0
Apache	Apache 2.2.3

Results

Average and minimum times: We voluntarily select eight attractions and compare the sum time which our system calculated with the average time. This time is the average time to be calculated when the same transit time, waiting time, and seat-load time are used. (Someday we have to use this system in Tokyo Disney Land and confirm the accuracy of this system.)

Our system assumes that the starting time is 9 o'clock, and that the attractions selected are No.1 to No.8 in Table 1; the result is presented in Table 3.We also output the run time. We caluculated ten times and the average time was applied to run time. The run time in the case of weekday is 1.83s and in the case of holiday is 1.94s.

Discussion: Our system's time and the average time differ by 143.1 minutes on weekdays and 158.4 minutes on holidays in theory. These time differences illustrates the time reduction by using our system. Our system thus has the possibility to reduce the total travel time. Of course, a margin of error will be introduced because these results are calculated by simulation. However, we think that the results provided by our system will become significant when visitors search for the most suitable route. In conclusion, we have demonstrated that using our system shortens the total time.

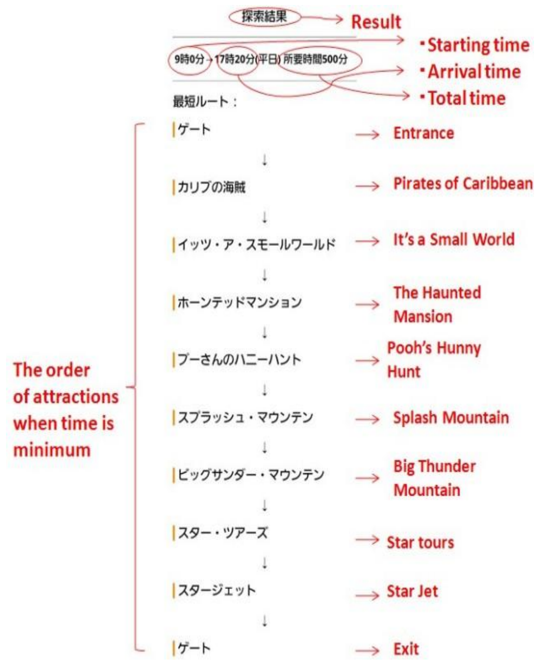


Figure 5. Resylt Screen

Table 3. Total Time

		TIME	ROUTE
weekday	OUR SYSTEM	357.9(min)	2-1-5-7-4-6-3-8
	AVERAGE	501.0(min)	5-2-8-6-7-3-4-1
holiday	OUR SYSTEM	642.6(min)	2-7-1-3-5-4-6-8
	AVERAGE	801.0(min)	5-1-4-8-2-6-7-3

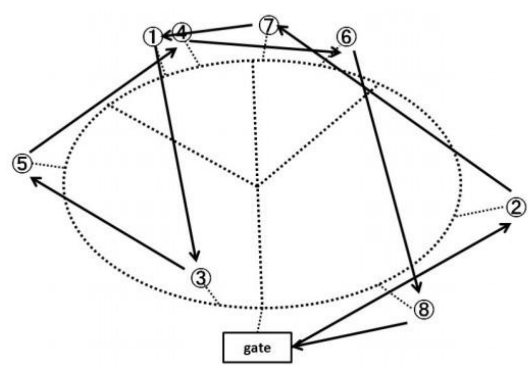


Figure 6. Route in the sum time calculated by our system

5. CONCLUSION

This study constructed a system for searching routes in Tokyo Disneyland using a smart phone and presented the system outline and interface. Additionally, we compared the average total time with the time calculated by our system and concluded that our system is useful because it reduces the time required by about 150 minutes.

Currently, we are seeking to resolve three problems in this system. The first problem is that the number of attractions visitor can select is fixed in our system. We need to change fixed inputs to variable inputs. The second problem is that our system could introduce an increasing error. We have to recalculate the total time en route. (Though we have already completed the programs to solve these problems.) The last problem is that we didn't consider the Disney Fastpass, a virtual queuing system created by the Walt Disney Company. Fastpass enables visitors to avoid long lines at the attractions on which the system is installed, freeing them to enjoy other attractions during their wait. We completed the ECLiPSe program for calculating the minimum time when using Fastpass and now incorporate it into a smart phone. We will focus on the above three improvements in the future.

6. ACKNOWLEDGE

We thank the Research Institute for Science and Technology in Tokyo University of Science for financial support.

7. REFERENCE

[1] Yuki M., Hisayoshi, S. & Miho, T: OR of Amusement park － as an example of The 2005 World Exposition, Aichi, Japan － http://www.seto.nanzan-u.ac.jp/msie/grthesis/ms/2005/index.html
[2] Syouta, H., Yuka, H. & Daisuke, N: The minimum route in universal studio Japan - http://www.seto.nanzan-u.ac.jp/msie/grthesis/ms/2007/04mm010.pdf

[3] S, Lin. & B, W, Kernighan : An Effective Heuristic Algorithm for the Traveling-Salesman Problem. Operations Research 1973,Vol. 21, No. 2, pp. 498-516

[4] Stutzle, T. & Hoos, H. : Ant System and local search for the traveling salesman problem. Evolutionary Computation, 1997., IEEE International Conference, 1997, 309-314.

[5] Takashi, H., Takayuki, K. & Tohru, I : Solving vehicle routing problems with soft time windows using chaotic neurodynamics. IEICE transactions on fundamentals of electronics, communications and computer sciences 2006, 105(675), 17-22.

[6] Takahiro, S., Yoshihiro, H. & Koji N : Inversed Function Delayed Network for Traveling Salesman Problem. IEICE transactions on fundamentals of electronics, communications and computer sciences 2007, 107(328), 55-60.

[7] Takenori, M. & Naoki, M : Comparison of Approximate Methods for Traveling Salesman Problem. IEICE transactions on fundamentals of electronics, communications and computer sciences 2003, 102(625), 1-6.

[8] Eiichi, G., Etsuji, T. & Mitsuo, W : A Randomized and Genetic Hybrid Algorithm for the Traveling Salesman problem. Information Processing Society of Japan, 2001, 2001(27), 61-64.

[9] Caseau, Y. & F. Laburthe. : Solving small TSPs with constraints. In Proceedings the 14th International Conference on Logic Programming, 1997, 316-330

[10] http://www.mapion.co.jp/route/

[11] http://www.tokyodisneyresort.co.jp/tdl/

[12] http://www15.plala.or.jp/gcap/disney/

[13] http://www.eclipseclp.org/doc/embedding/embroot041.html

Approach of life support-system using smart phone

Kouhei Yoshida
Faculty of Science and Technology,
Tokyo University of Science,
2641 Yamazaki, Noda,Chiba, 278-8510, Japan
k-yoshida@ohwada-lab.net

and

Hayato Ohwada
Faculty of Science and Technology,
Tokyo University of Science,
2641 Yamazaki, Noda,Chiba, 278-8510, Japan
ohwada@ia.noda.tus.ac.jp

Abstract

In this research, we have developed a life-support system that uses a smart phone. This system consists of an application that collects a user's life-log from everyday behavior, and a management application that can peruse a user's situation at any time. The features of this system lie in automatically deciding a user's movement. To get location information at one-minute intervals, it determines the user’s movement using the move distance for each minute and the station information for each location during that period of time. This system obtains the status of the user from his or her daily behavior recorded on the smart phone. Since we provide the monitoring function to an administrator and a user, we achieve a life support function with each capability.

Keywords: a life-support system, a smart phone, life-log, movement, the station information

1. INTRODUCTION

In recent years, the smart phone has grown very popular. A smart phone is equipped with advanced sensor devices such as GPS and an accelerometer sensor. In addition, one can easily acquire information such as the user's peripheral status and behavioral record. By analyzing such information as digital data, such as the users’ behavioral characteristics, tastes, and hobbies, one can acquire information not previously available. The system can be expected to provide its best service to support the lives of users or to provide advertising tailored to each user. We believe that we can enhance the lifestyle of older people by logging the movements of each day.

In this research, we developed a life-support system that uses a smart phone to manage everyday life. The system is targeted to older people. Using the smart phone, which is a type of mobile device, we developed a tool designed to compile a life-log that logs a person's behavior as digital data based on everyday behavior. Furthermore, as an environment that can monitor the user's status every day, we propose a system that can analyze the user's life-log every day.

In the following, related work is discussed in Section 2. Section 3 describes the proposed method, and Section 4 presents the conclusions.

2. RELATEDWORKS

To determine movement from a GPS log, it includes a method that uses differences in velocity. This technique is a method of analysis based on the maximum speed and the difference of acceleration of the movement, whether in a car or on foot. As one method for identifying the movement, it utilizes Zheng's research [1]. Using a GPS log of 45 users taken over six months, they created a decision tree for average speed, maximum acceleration, and maximum speed, and they then determine the type of movement (bus, car, bicycle, or walking). Using map data, it maps the coordinates of the GPS log onto a vector map. Matching the data on the map and provides a method for identifying the movement. When it matches the GPS log with the map, if a certain level of speed is indicated and there is a line to that point, the movement is determined to be by train. If there is a street, it determines the movement to be by car.

In this research, location information is acquired at one-minute intervals. The system determines the user’s movement using the distance moved over one minute and the station information around each location in certain period of time. In addition, a life support is limited to the Kanto area.

3. SYSTEM OUTLINE

The basic system being proposed is explained based on Fig. 1. The system consists of two applications, an application in a smart phone and a web application. The application on the smart phone is aimed at collecting data every day about the user, such as places visited, the user's movement, or the arrival of user information via telephone or e-mail.

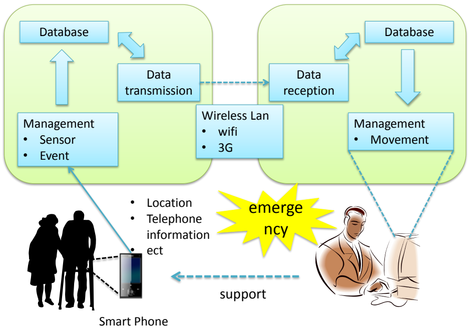


Fig.1 Outline of the system

The web application is aimed at perusing and managing the data sent by the user.

Next, the flow of the system is explained. A user who becomes a candidate for support is registered in the database by an administrator. Starting the application will start the acquisition of position information, etc. Position information, the information arriving by telephone, etc., are managed all by the management module in the application. Whenever position information is acquired or a telephone message arrives, the information is recorded in a database created in the smart phone by the management module. These collected life-logs are managed day by day. This data is periodically sent to a personal computer by the data transmitting module. Data sent to the personal computer is immediately saved by the data reception module in a database on the personal computer. An administrator readings data from the database, using a management module can peruse the movement and so on and can supervise. When abnormalities are detected, the database can support the user. The user can also peruse the situation at any time.

3.1 Overview of the Mobile Application

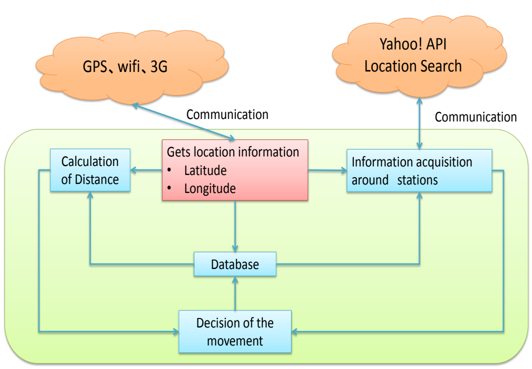


Figure 2: Application outline figure

The application created on the smart phone is diagrammed in Fig. 2. The acquisition of position information and movement information by this application is important for supporting the user. Once life-log collection begins, GPS, 3GS, and Wifi will be used to determine the user's present location. This determination is performed at one-minute intervals. The acquired position information is passed to a database for preservation, along with the calculated movement distance and acquired surrounding station information. Next, a judging module judges the user's movement based on the movement distance and surrounding station information. Finally, the judged movement is saved in a database with the acquired position information.

When people move, the type of movement (e.g., movement on foot or by car) is judged by the movement-distance module based on distance changes. In this research, life support is limited to users Kanto area. We used the Euclidean distance formula to obtain the distance between two points in the calculation of movement distances because user's movement in the Kanto area is not considered long distance.

$$\sqrt{(SP_x - ST_x)^2 + \left(SP_y - ST_y\right)^2} \qquad (1)$$

In this research, to calculate the distance moved over one minute. The current position is SP, and the position of one minute ago is ST. SP_x and ST_x is mutual latitude. SP_y and SP_y is mutual longitude.

The surrounding-station information-acquisition module acquires the name of any station found within a 500m radius based on the position information. The name of the station is acquired using the location search API [7] of the Yahoo! API service. If we have made to get the station name for all of the acquired location information, we will consider two possible problems. First, the location information acquisition time is longer because it is a vast number are constantly accumulating. Second, if there exists a location near the train station, it is difficult to determine because of transportation would always come to get the station name. In this study, we have limited time to move, or to obtain location information from the station. Term is ten-minute. A total of ten position data points, including position information from the previous nine measurements along with the present position, are searched.

A movement judging module a judges the type of movement (e.g., on foot, by car, or by train) based on the result of a movement-distance calculation module and a surrounding-station-information-acquisition module. This judgment results in three possible status: "Stay," "Work," and "Train." Table 1 shows each judgment condition.

Table 1 judgment condition

Movement	distance (m)	Station
Work	< 1000	< 2
Train	>= 1000	>= 2

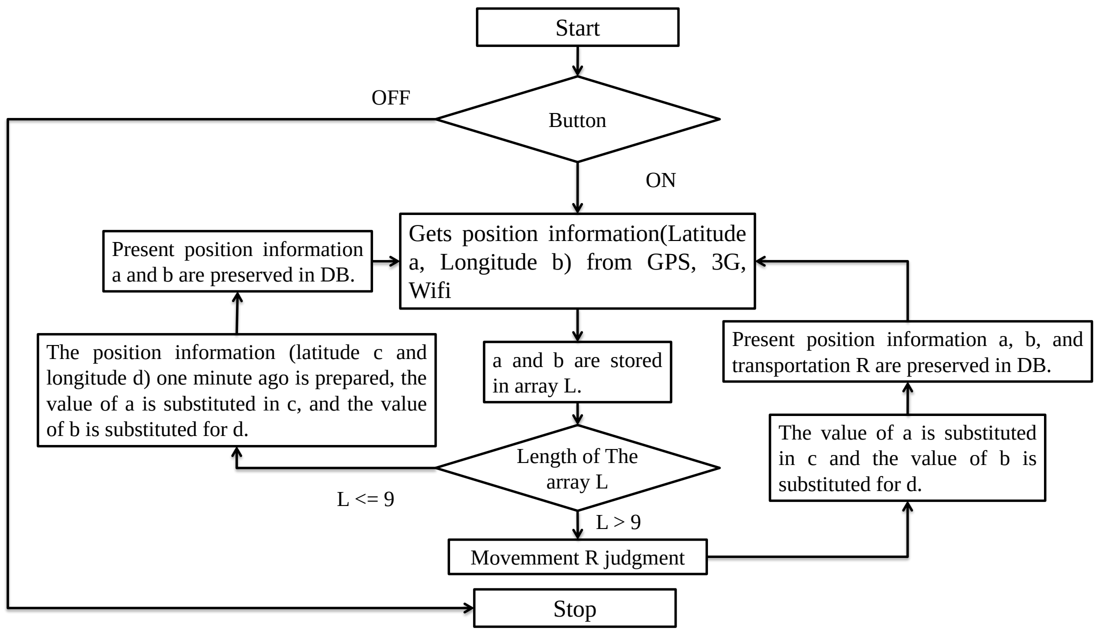


Fig. 3 Algorithm for judging movement

Next, it explains the algorithm of this application. Figure 3 shows it. The user pushes the acquisition beginning button and this application starts. Moreover, the user can arbitrarily end the application. The position information is acquired simultaneously when the application starts. The acquisition interval for position information set on one minute.

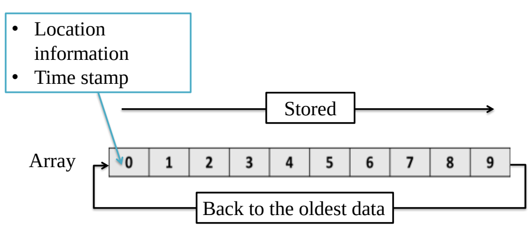


Fig.4 Flow of data storage

An array L of length n is prepared, and position information is stored in that array. In this research, we acquired information from the station nearest the location over ten minutes. It was thus ten times the length of array L. As anterior position information, a variable is prepared and position information is recorded. Finally, the position information and the acquired time are saved in a database, and the system returns to the acquisition of position information. This procedure is followed until ten datasets are stored in array L. After the 11th time, the array is updated and indicates the storage location where the data is the oldest. Figure 4 shows the flow. Moreover a user needs to manually judge the movement up to ten times. After the data is stored in the array, the associated movement is judged automatically. Figure 5 shows the algorithm. First, the number of agreements with the previous position information is calculated among the ten position information sets, using the present position information. When five or more data points agree, the module judges this as “Stay.”. The present position information is recorded in the database as front position information, and the system returns to position information acquisition. When there are fewer than five agreements, the movement distance is measured using the present position information and front position information.

Next, information on any station on the outskirts is combined with the position information in array L, and the system shifts to judging the type of movement. When two or more stations whose movement distance acquired from the position information is 1000m or more, and the array L exist, it is judged as “Train”; Otherwise it is judged as with “Work.”. Finally, the present position information is recorded in the database as front position information, and the system returns to position information acquisition.

The outline of the event-processing system is now explained based on Fig. 6. In this system, processing is performed for every event. In the system flow, the application is started in an event-generating standby state. When an event occurs, a module corresponding to event is invoked and processing is performed. The information acquired at that time is saved in a database, and the system returns to an event stand-by state.

Next, the information-acquisition module for the telephone and the e-mail receiving information-acquisition module are explained. With the information-acquisition module for the telephone, measurement of an hour of use is performed at the same time that operation of the telephone is performed. An hour of use is measured when the telephone is used. Telephone use and measurement end after an hour of simultaneous use, and the system returns what the

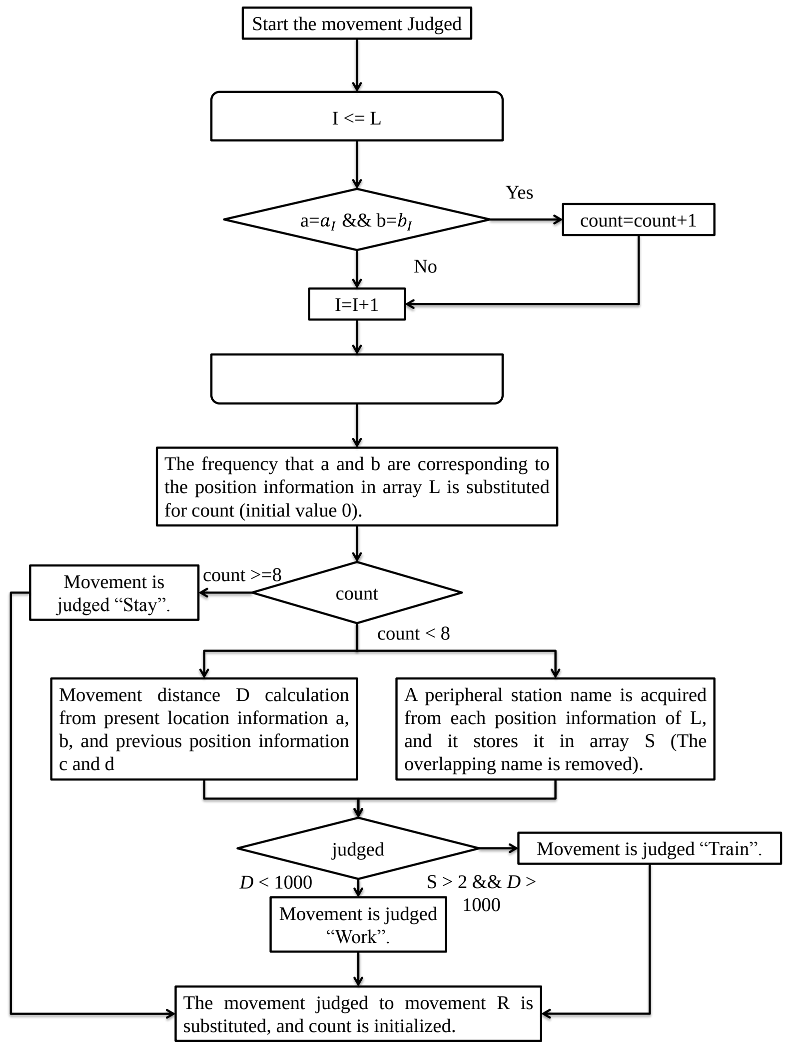


Fig.5 Algorithm of movement judgment

operation was, how many minutes it lasted, and when it occurred.

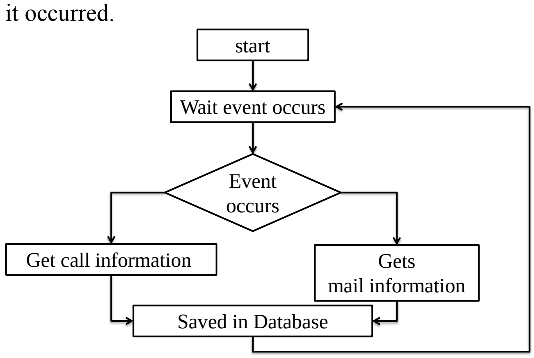


Fig.6 Outline of the event-processing system

The e-mail receiving information acquisition module indicates when e-mail is received and its status (unopened). This module is also called when e-mail is opened, and indicates the opening time and its opened status.

As mentioned above, the user only possesses the smart phone and can view the life-log. The user can thus receive lifestyle support from an administrator or a special-care provider while leading an everyday.

3.2 Outline of Web Application

The principles of the web application are explained. The user needs to register with the web application. After registration, the user can receive services such as support from an administrator, and examine the situation on any given day.

The application flow is now explained. The life-log periodically sent from the application in the smart phone is saved in database by the management module.

The administrator can browse the situation on a given day for every user. Figures 7 and 8 demonstrate the functions of the administrator. The screen in Fig.7 represents the user's status for the current day. No.1 in the figure indicates the user's status with regard to movement, and No.2 in the figure is the movement root.

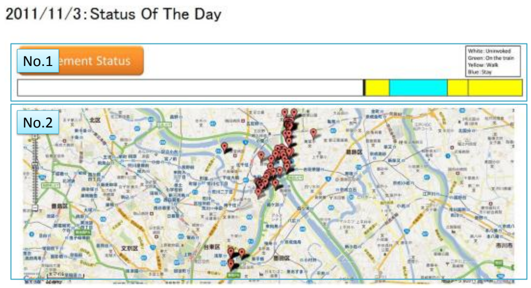


Fig.7 Administrative functions (current day)

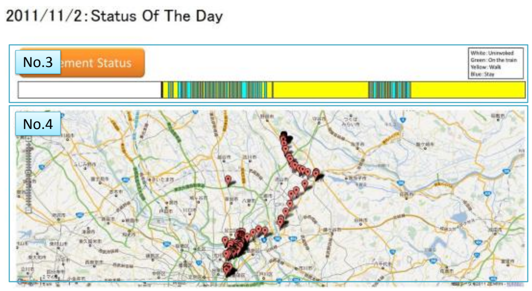


Fig.8 Administrative functions (previous day)

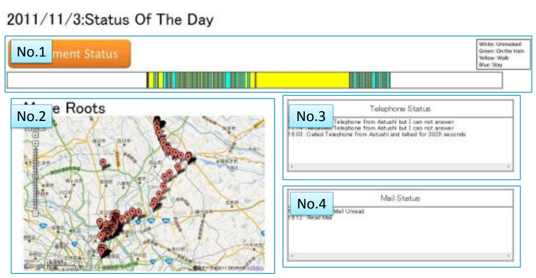


Fig.9 User functions

The screen in Fig.8 represents the user's status for a previous day. No.3 in the figure indicates the user's status with regard to previous movement in the past, and No.4 in the figure shows the previous movement root. As an administrative function, this can be viewed as information concerning movement and information concerning the movement root. Telephone information and e-mail information are restricted personal information, and they as such cannot be perused to prevent infringements of privacy.

The administrator monitors a user's present situation based on the situation on a previous day. Depending on the user's situation, the user's lifestyle is directly supported or by communicating with a special-care provider.

The user can peruse the situation on his or her own on a given day. Figure 9 indicates the functions available to the user. No.1 in the figure indicates the user's status with regard to movement, and No.2 in the figure indicates the movement root. No.3 shows the telephone call history and incoming phone calls, and No.4 shows the received e-mail history. The user can see where he/her is, and how far he/her has moved (distance). Also, several items that elderly people tend to overlook are designed to be visible in the information, including the record of incoming telephone calls and reception of e-mail.

4. CONCLUSIONS

In this research, we developed a life-support system that uses a smart phone to manage everyday life. The system consists of an application that collects a user's life-log based on everyday behavior and a management application that can monitor the user's situation at any time. This system collects the user's life-log using the smart phone. Data is sent to the administrator on a regular basis. Finally, we provide an environment that can monitor the received data. We are able to support the user's life a cording to his or her specific situation. The smart phone acquires location information at one-minute intervals and can automatically determine the user’s movement using the movement distance over one minute along with station information from around each location during the same period of time. Since we provide a monitoring function to both administrators and users, we achieve life support from both capabilities.

5. REFERENCES

[1] Zhenhui Li, Ming Ji, Jae-Gil Lee, Lu-An Tang, Yintao Yu, Jiawei Han, Roland Kays,“ MoveMine:Mining Moving Ob-ject Databases”,SIGMOD'10, June6-11, 2010, Indianapo-lis, Indiana, USA

[2] Yu Zheng,Like Liu,Longhao Wang and Xing Xie,“Learn-ing Transportation Mode from Raw GPS Data for Geo-graphic Applications on the Web”, WWW2008, 2008, Bei-jing, China

[3] Vincent W,Zheng,Yu Zheng , Xing Xie and Qiang Yang,“ Collaborative Location and Activity Recommendations with GPS History Data”, WWW2010, April 26-30, 2010,Ralcigh, North Carolina, USA

[4] Lei Chen,M,Tamer Ozsu,Vincent Oria,“Robust and Fast Similarity Search for Moving Object Trajectories ”, SIG-MOD2005, June14-16, 2005, Baltimore, Maryland, USA

[5] Lili Cao,John Krumm,“ From GPS Traces to a Routable Road Map ”,ACM GIS '09,November4-6,2009,Seattle,WA,US

[6] Simon R., Fröhlich P., A Mobile Application Framework for the Geospatial Web. In Proceedings of WWW '07 (Banff Canada, May 2007). ACM Press: 381-390.

[7] Yahoo! API Location Search, http://developer.yahoo.co.jp/webapi/map/localsearch/v1/localsearch.html

GPS on Every Roof, GPS Sensor Network for Post-seismic Building-wise Damage Identification

Kenji Oguni
and Tomohiro Miyazaki
Department of System Design Engineering
Keio University
Hiyoshi, Yokohama, Japan
Email: oguni@sd.keio.ac.jp

Masayuki Saeki
and Naoki Yurimoto
Department of Civil Engineering
Tokyo University of Science
Noda, Chiba, Japan
Email: saeki@rs.noda.tus.ac.jp

ABSTRACT

Development of wireless sensor network equipped with GPS for post-seismic building-wise damage identification is presented in this paper. This system is called *GPS on Every Roof*. Sensor node equipped with GPS antenna and receiver is installed on the top of the roof of each and every building. The position of this sensor node is measured before and after earthquake. The final goal of this system is to i) identify the displacement of the roof of each house and ii) collect the information of displacement of the roof of the houses through wireless communication. Superposing this information on GIS, building-wise damage distribution due to earthquake can be obtained. The system overview, hardware and some of the key components of the system such as on-board GPS relative positioning algorithm to achieve the accuracy in the order of several centimeters are described in detail. Also, the results from a field experiment using a wireless sensor network with 39 sensor nodes are presented.

Keywords: GPS, Wireless Sensor Network, Earthquake

1. INTRODUCTION

High risk of earthquake beneath large cities and/or disaster affecting wide area caused by a huge earthquake should be anticipated in the earthquake-prone regions. Once a huge earthquake hits and causes significant damage to the cities, information about the damage on the buildings in that area should be gathered quickly. However, as the scale and the severity of the disaster increases, the information about the damage caused by the earthquake will be deteriorated in quality, quantity and in its speed. Besides, significant damage on backbone communication network could be caused by an strong earthquake.

Under this condition, a wireless sensor network system covering the whole city could be a very prominent candidate for generating and gathering precise information about the damages caused by a huge earthquake. Especially, a brute force application to install a sensor node equipped with a GPS receiver on the top of the roof of each and every building in the whole city is the most straightforward and the easiest idea to come up with. Just by measuring the position of the GPS receiver before and after a huge earthquake, the displacement of the roof of each and every building can be obtained and this information can be gathered by wireless communication between sensor nodes. However, this straightforward application has been automatically dismissed because of the poor positioning accuracy of GPS with affordable cost. The typical positioning accuracy of affordable GPS used in the vehicle navigation is in the order of several meters[1]. Obviously, this is not enough for identifying the structural damages caused by an earthquake. At least, the positioning accuracy in the order of several centimeters is required for this application.

Motivated by the discussion above, a system called *GPS on Every Roof* has been developed by the authors. *GPS on Every Roof* is a wireless sensor network system consisting of sensor nodes equipped with an affordable L1-GPS receiver, a CPU with high computation power and a wireless communication module with relatively long communication distance. This sensor node is installed on the top of the roof of each and every building in the target area. The position of this sensor node is measured before and after earthquake. Then, the displacement of the roof caused by the earthquake can be obtained.

2. SCHEMATIC VIEW OF GPS ON EVERY ROOF

GPS on Every Roof is used for identification of damages on each and every building in the area affected by an earthquake. Sensor node shown in Figure 1 equipped with a GPS patch antenna, a GPS receiver, a CPU (indicated as SH2) and a wireless communication device is installed on the top of the roof of each and every building. The position of this sensor node (more precisely, the position of the GPS patch antenna) is measured before and after earthquake. Then, the displacement of the roof caused by the earthquake can be obtained. This information is sent to a main server through an autonomous wireless sensor network. Superposing this information on GIS, information about building-wise damage can be obtained together with the distribution of the damage. Also, possible

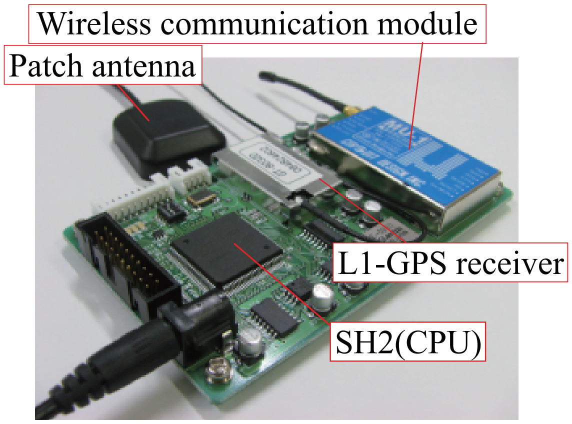


Fig. 1. Wireless GPS sensor node

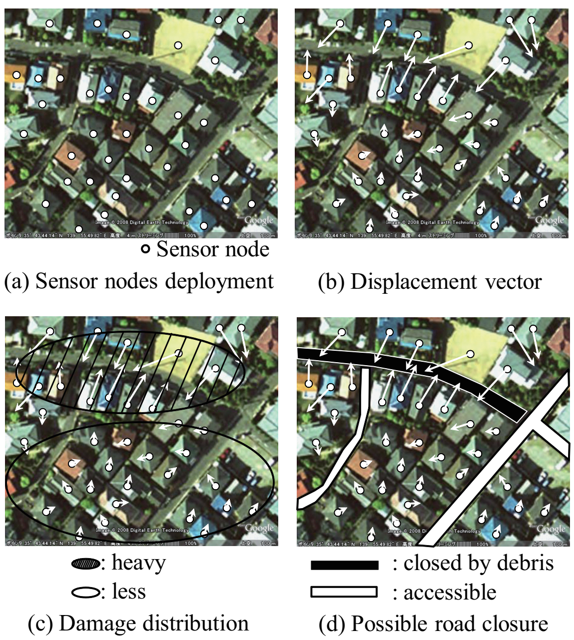


Fig. 2. Workflow of the system

road closure due to the debris from the collapsed houses can be estimated. This workflow is schematically shown in Figure 2.

3. IMPLEMENTATION

The accuracy of GPS positioning required for *GPS on Every Roof* is at least several centimeters. The positioning accuracy of GPS used in commercial car navigation system is at most several meters and this is not good enough for identifying the collapse of the buildings. On the other hand, GPS surveying with millimeter order positioning accuracy uses expensive hardware which cannot be spread on the top of the roof of each and every building in a city.

The algorithm for relative positioning with phase ranges has been implemented on the sensor node shown in Figure 1. This algorithm has been developed by the authors and will be discussed in detail in the next section. The positioning accuracy of several centimeters (several millimeters with high quality GPS observation data set) can be achieved by the on-board positioning analysis. This solves the dilemma between high positioning accuracy and low cost and this makes *GPS on Every Roof* possible.

In this section, the details of the design of the GPS observation data sharing, sensor node hardware and automatic allocation of sensor node ID are presented. The design philosophy behind these implementations is the suppression of the total amount of the wireless communication for reduction of time and for enhancement of the robustness of the system.

GPS obervation data sharing

To achieve the positioning accuracy of several centimeters using devices with affordable cost, the relative positioning with phase ranges [1] should be employed. The relative positioning with phase ranges uses the carrier phase of the signals from GPS satellites as the observation data. In this positioning algorithm, the relative position of a sensor node to the reference node is obtained. The relative position is computed based on the difference between the data observed on the sensor node under consideration and those of the reference node. This results in the requirement for the method to share the GPS observation data of the reference node.

Thus, we are left with the following two options. The one is the collection of all the GPS observation data from all the sensor nodes under local server and the positioning analysis is performed on the local server. The other is to spread the GPS observation data of the reference node to all the neighboring sensor nodes and on-board relative positioning analysis with phase ranges is performed on each and every node in the network.

Estimate of time for wireless communication for data sharing: GPS receiver on each sensor node receives the signals from GPS satellites and generates a data packet of 38 bytes in size every second. For 5 minutes GPS observation (typical duration of time for achieving the positioning accuracy of several centimeters using the GPS positioning method described in the next section), the total amount of GPS observation data on each sensor node results in 11.4 Kbytes. This data set should be sent to the CPU on which the GPS positioning analysis is performed.

The wireless communication device implemented on the sensor node requires T (msec) for transmitting a data packet as follows.

$$T = n \times 1.04 + 34 \quad (1)$$

where, n is the size of data (bytes) in the packet. From Eqn. (1), the total amount of time for transmitting GPS observation data from one sensor node to the local server or to other sensor nodes is estimated as 22 seconds.

If the first option for data sharing (i.e., the collection of

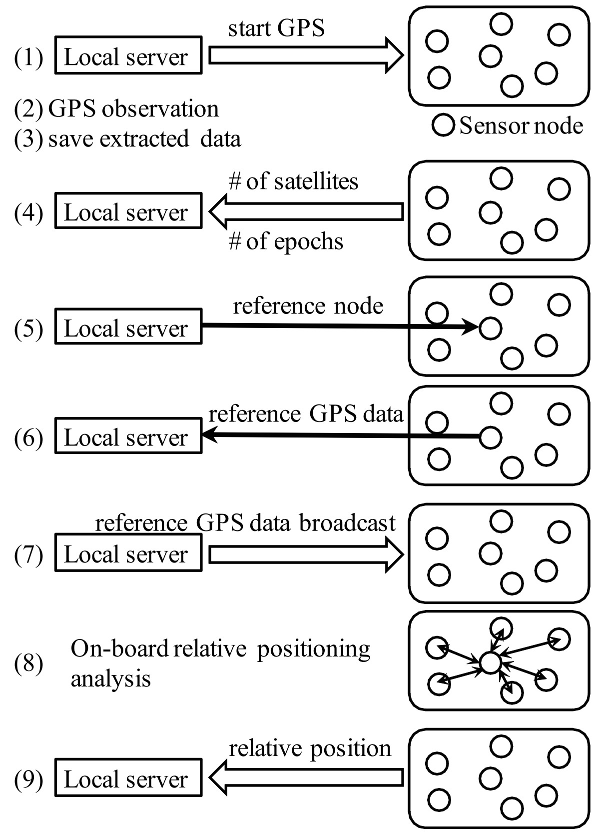


Fig. 3. Workflow of GPS positioning

all the GPS observation data from all the sensor nodes under local server and the positioning analysis is performed on the local server) were taken, the total amount of time for collecting the GPS observation data from (the maximum of 253) sensor nodes ends up with 93 minutes. This is too long and never satisfies the requirement for *GPS on Every Roof*.

Therefore, the second option (i.e., reference GPS data spread and on-board positioning analysis) has been employed in the proposed system. The only data sent to the local server from all the sensor nodes is the result of GPS positioning analysis (i.e., the x, y, z coordinates and an index for showing the reliability of the result). This results in the requirement for sending 1 data packet of 16 bytes in size from each sensor node and the total amount of time for 253 sensor nodes results in 13 seconds. Also, in the second option, the total number of wireless communication is suppressed. This makes the system scalable and robust.

Workflow of GPS positioning: Based on the discussion above, the reference GPS data is shared by the neighboring sensor nodes and on-board positioning analysis is performed in the proposed system. The workflow of GPS positioning is summarized in Figure 3.

(1) The local server broadcasts a command to the sensor nodes to start GPS observation.
(2) The sensor nodes respond the command and turn on the GPS receiver. Since the GPS receivers were turned off before the command, they have to search for the GPS satellites for the first 30 $\sim$ 60 seconds. This process is called a *cold start* and this makes the GPS observation time longer than that for a *hot start*.
(3) After the initialization process in the cold start, each sensor node observes the signals from GPS satellites and extracts the GPS data required for relative positioning with phase ranges. This extracted data set is saved on the on-board external RAM in the appropriate format.
(4) After GPS observation, each sensor node analyzes the observed data and evaluates the data quality. The number of observed GPS satellites and the number of observed epochs are the indicators of the data quality. These indicators are reported to the local server.
(5) The local server determines the reference sensor node based on the reported indicators of the data quality of each sensor node. Then the local server requests the reference sensor node to send back the whole GPS observation data to the local server.
(6) The reference sensor node responds to the command from the local server and reports its GPS observation data to the local server.
(7) The local server broadcasts the reference GPS data to all other sensor nodes.
(8) Each sensor node performs on-board relative positioning analysis using the reference GPS data and observed GPS data saved on its on-board external RAM.
(9) Each sensor node reports the analysis results (i.e., the x, y, z coordinates and an index showing the reliability of the result) to the local server.

This work flow requires the wireless communication of the whole GPS data only in the steps (5): from the reference node to the local server and (7): broadcast from the local server. Thus, with m reference nodes, the total number of the whole GPS data transmission is at most $2m$. This is a drastic reduction of the communication compared with the first option (to collect the GPS observation data from all the sensor nodes to the local server). This significant contribution for the reduction of time and for the enhancement of the robustness of the system is achieved by the implementation of on-board GPS relative positioning algorithm.

Sensor node hardware

A prototype sensor node shown in Figure 1 has been developed based on the discussion above. The major components of the current prototype sensor node to be described here are the CPU, the wireless communication module and the GPS receiver.

Since the on-board computation for GPS positioning requires double-precision arithmetic, a 32-bit RISC microcomputer, SH7144F (Renesas Technology Corp.) is employed as the main CPU on the sensor node. SH7144F works under 3.0V $\sim$ 3.6V input voltage with the highest clock frequency of 50MHz. The on-chip multiplier which can execute multiplication operations (32bits $\times$ 32bits $\rightarrow$ 64bits) in two to four cycles is the key feature of SH7144F to be employed as the main CPU on the sensor node. It also has 256KB of the internal ROM, 8KB of internal RAM and 4 asynchronous or clocked synchronous serial communication interfaces. Together with these internal devices, external RAM of 2MB is directly connected to SH7144F on the sensor board to keep and process the GPS observation data.

The wireless communication module on the sensor node is MU-1-1252 (CIRCUIT DESIGN, INC.). MU-1-1252 is a designated low power wireless communication module with the frequency band of 1252MHz. Sensor nodes with this module can be spread anywhere without any license or permission. In spite of its limited RF output power (10mW), the maximum distance of the wireless communication by MU-1-1252 is 600m with a line of sight. This distance will be reduced to 100m $\sim$ 200m when the line of sight is lost. This distance is long enough for the proposed application and this is the biggest reason for employing MU-1-1252 despite of its low communication speed (effective rate is 6800bps). The energy consumption is 60mA and 35mA for transmission and receiving, respectively with input voltage of 3.3V.

GPS receiver used in the prototype is a L1-GPS receiver GT8032 (FURUNO ELECTRIC CO., LTD.). Originally developed for GPS clock based time synchronization, GT8032 outputs the carrier phase in the format of FURUNO Binary in addition to the ordinary NMEA format GPS data. This data set of carrier phase is used in the GPS positioning proposed in this paper. In the proposed system, this GPS receiver is connected to a patch antenna used for car navigation for reduction of the cost.

This prototype sensor node works on a Li-ion battery with 3.6V output. This voltage is reduced to 3.3V on the sensor node. Although this prototype works on a battery (1900mAh@3.6V) for several hours without additional energy supply, reduction of energy consumption is preferable. For this purpose, energy hungry SH7144F should be used only for on-board computation of GPS positioning. Other tasks such as controlling the wireless communication unit and GPS observation should be taken care of by other CPU with lower energy consumption. In the next version of the sensor node, this energy management with dual CPU should be implemented and this is left for the future work.

Automatic allocation of sensor node ID

Automatic allocation of sensor node ID is inevitable when large number of sensor nodes exist in the network. Sensor nodes are frequently replaced, removed or added. Especially in *GPS on Every Roof*, many sensor nodes could be physically damaged. Under this condition, management of the sensor node list in the network is a tedious task and this may be a potential source of a bug. Large scale sensor network should be a maintenance-free and deploy-and-work system. In this sense, automatic allocation of sensor node ID is one of the most important functionality for a sensor network with scalability. Particularly in *GPS on Every Roof*, the amount of the wireless communication and the total time for allocation of sensor node ID should be minimized for generating displacement data of each and every building in the whole area within 30 minutes.

Although there are many existing algorithms for automatic allocation of sensor node ID in the ad-hoc sensor network, a new algorithm customized for the wireless communication module (MU-1-1252) on the sensor node for *GPS on Every Roof* has been developed. The major features of the customized algorithm are its robustness and quickness.

For the assignment of ID, serial number of MU-1-1252 on each sensor node is used in *GPS on Every Roof*. Each sensor node reports its own serial number to the local server and the local server allocates unique ID to each sensor node. In this simple algorithm of ID allocation, the total time becomes a problem. Since MU-1-1252 avoids interfering the communication between other stations by checking RSSI (Received Signal Strength Indicator) before transmitting data packet, each sensor node should be given its own time slot with clean environment. The serial number of MU-1-1252 has 8 digits and one epoch of the wireless communication requires 70ms. If 70ms time slot is simply assigned for reporting its own serial number, time slots for the serial number 1 to 99,999,999 should be prepared. Then, the total time for completing ID allocation results in 81 days. The minimum number of time slots to be prepared is 253 (the maximum number of sensor nodes under a local server). Unnecessary idle time of the network in the simple assignment of the time slot is the major problem. Therefore, proper filtering of the serial number for assignment of the time slot is the key for solving this problem. For this purpose, a filter called *prime number filter* has been developed.

Prime number filter: The prime number filter is based on a set of three prime numbers and the remainder. An example is shown in Figure 4. This filter consists of three layers. In the first layer, the serial number of MU-1-1252 on a sensor node is divided by a prime number 3. Then, this serial number is categorized by the remainder. The same procedure is executed for the second layer with a prime number 5 and for the third layer with 17. Thus, a serial number is categorized in one of 255(=3×5×17) time slots. Since the range of the serial number is 8 digits, many serial numbers are categorized in the same time slot. However, by applying a filter consisting of different set of prime numbers, the combination of the serial numbers categorized in the same time slot is changed. Thus, by applying 3 filters consisting of (3, 5, 17), (2, 7, 19) and (13, 23), serial numbers up to 7 digits (1 to 19,518,720) can be categorized in different time slots. If a prime number filter with a single prime number (11) is added to the abovementioned three filters, all 8 digits are fully categorized in different time slots. Introduction of this prime number filter drastically reduces the total time for ID assignment. ID assignment to the sensor nodes in a system consisting of 253 nodes can be completed in 1 minute.

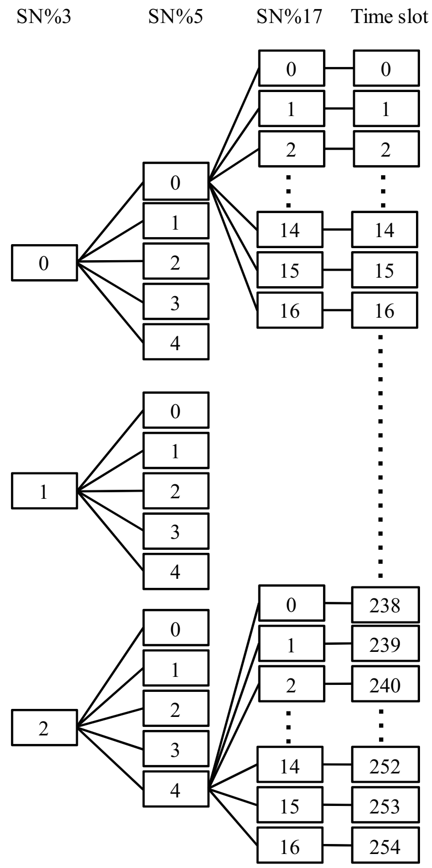


Fig. 4. Example of prime number filter

4. GPS RELATIVE POSITIONING ALGORITHM WITH L1 CARRIER PHASE FOR ON-BOARD ANALYSIS

To make *GPS on Every Roof* possible, GPS relative positioning should have high accuracy (e.g., at least, in the order of several centimeters). Also, to enhance the robustness of the system, the communication of the whole GPS observation data should be avoided. GPS positioning analysis should be performed on-board (i.e., on each sensor node) and only the final result of the positioning analysis should be sent to the central server. The relative positioning algorithm presented here satisfies these requirements.

Double-Differenced carrier phases between nodes with short inter-node distance at time t, $\phi_{ij}^{kl}(t)$, can be modeled as

$$\lambda\phi_{ij}^{kl}(t) = \rho_{ij}^{kl}(\mathbf{x}, t) + \lambda N_{ij}^{kl} + e_{ij}^{kl}(t), \tag{2}$$

where $*_{ij}^{kl}$ are the double difference values for GPS satellites k, l and the sensor nodes i, j. $\rho_{ij}^{kl}(\mathbf{x}, t)$ is the double difference of the distances between satellites and the nodes, $\mathbf{x}$ is the position vector of a sensor node, λ is the wave length of the carrier, N_{ij}^{kl} is the double difference of the integer ambiguity and $e(t)$ is the noise. The assumption of the short inter-node distance allows us to approximate $e_{ij}^{kl}(t)$ as a sum of random noise and the noise due to antenna configuration [1].

Taylor expansion of Eqn. (2) with respect to $\mathbf{x}_0$ results in

$$U_{ij}^{kl}(t) = \partial_x\rho_{ij}^{kl}(\mathbf{x}_0, t)\Delta x + \partial_y\rho_{ij}^{kl}(\mathbf{x}_0, t)\Delta y + \partial_z\rho_{ij}^{kl}(\mathbf{x}_0, t)\Delta z + \lambda N_{ij}^{kl} + e_{ij}^{kl}(t) + h.o.t. \tag{3}$$

where, $U_{ij}^{kl}(t)$ is the corrected Double-Differenced carrier phases and can be computed as follows.

$$U_{ij}^{kl}(t) = \lambda\phi_{ij}^{kl}(t) - \rho_{ij}^{kl}(\mathbf{x}_0, t) \tag{4}$$

For different sets of satellites, Eqn. (3) with the corresponding coefficients can be obtained and these equations form the following simultaneous equations [2], [3], [4], [5].

$$\mathbf{U}(t) = \mathbf{A}(t)\Delta\mathbf{x} + \mathbf{N} + \mathbf{e}(t) \tag{5}$$

The unknowns in Eqn. (5) are $\Delta\mathbf{x} = (\Delta x, \Delta y, \Delta z)$ and a vector $\mathbf{N}$ which has N_{ij}^{kl} as components.

To solve Eqn. (5) in a robust manner, linear approximation of $U_{ij}^{kl}(t)$ with respect to time t is introduced. This approximation safely holds for small t (i.e., short GPS observation time). GPS observation time in the proposed system is 5 minutes and this is considered to be short enough for linear approximation and long enough for keeping independence of the GPS observation data. In this linear approximation, only the first and the last epoch in the observation are used to form linearized observation equations corresponding to Eqn. (3). As a result, with little amount of data communication, on-board GPS positioning analysis with the accuracy of a few centimeters (with better condition, accuracy of a few milimeters) can be achieved.

5. FIELD EXPERIMENT

The on-board GPS positioning algorithm and other fundamental algorithms for wireless sensor network such as automatic ID allocation, reference GPS data broadcast, retrieval of lost packets and communication timing control are implemented on the sensor nodes and the local server. Then, a field experiment with 39 sensor nodes was performed.

The sensor nodes were deployed in grid as shown in Figure 5 (indicated as "before $+$"). The numbers in Figure 5 are the IDs of the sensor nodes. To simulate the displacement due to an earthquake, the sensor nodes 479, 459, 397 and 396 were displaced by about 50m, 2m, 1m and 3m, respectively. Then, relative positioning was performed again. Sensor nodes after displacement are indicated as "after $\circ$." The arrows are the displacement vectors. The arrangement of $'\circ'$ after displacement shows that the locations of all the sensor nodes (including the nodes 376, 475 and 481) were precisely identified. It seems as if the sensor node 481 was displaced and the nodes 376 and 475 suddenly appeared out of the blue. But this is because of the wrong estimation before displacement for the nodes 376, 475 and 481. The vectors for these nodes are the dummy displacement vectors. Other displacement vectors correspond to the actual displacement of the displaced sensor

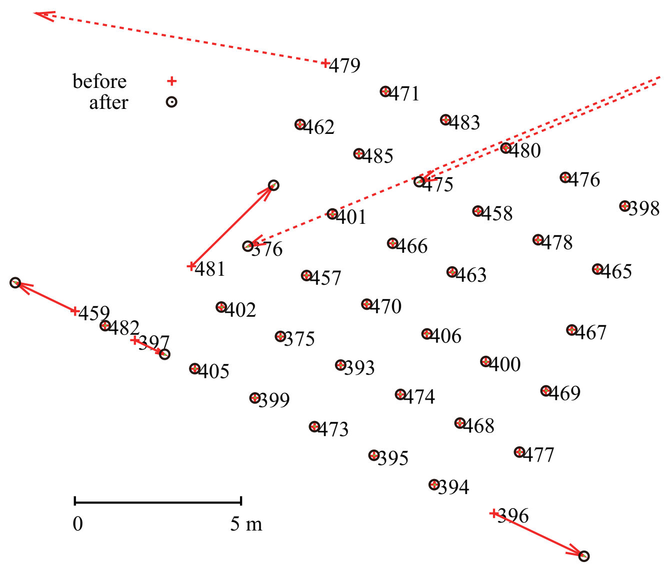


Fig. 5. Estimated sensor node position before and after displacement

nodes (479, 459, 397 and 396) and these displacements were precisely estimated within the accuracy of a few centimeters.

6. CONCLUSIONS AND FUTURE WORKS

A wireless sensor network system called *GPS on Every Roof* has been proposed in this paper. This system is a wireless GPS sensor network for post-seismic buildinf-wise damage identification. The system overview, hardware and on-board GPS positioning algorithm are discussed together with results from a field experiment.

The future work regarding *GPS on Every Roof* are as follows.

- Development of sensor node with better (longer distance, higher communication speed) wireless communication module
- Field experiment with sensor nodes deployed on the top of the roof of the buildings and with longer duration of time
- Development of sensor node with dual CPUs (one low-power, low-specification CPU and one high-power, high-specification CPU) for reduction of energy consumption

REFERENCES

[1] Hofmann, B., Wellenhof, H. Lichtenegger and J. Collins (2001), "GPS, Theory and Practice", *Springer Wien NewYork*.

[2] Kim, D. and Langley, R.B. (2000), "GPS Ambiguity Resolution and Validation: Methodologies, Trends and Issues", *7th GNSS Workshop, Int. Symp. GPS/GNSS*, Seoul, Korea, Nov.

[3] Teunissen, P. J. G. (1993), "Least-squares estimation of the integer GPS ambiguities", Invited lecture, Section IV Theory and Methodology, IAG General meeting, Beijing, China.

[4] Teunissen, P. J. G., P. J. de Jonge, C. C. J. M. Tiberius (1997), "The least-squares ambiguity decorrelation adjustment: its performance on short GPS baselines and short observation spans." *Journal of Geodesy*, 71, 589-602.

[5] Saeki, M., Kosaka, T. and Kaneko, S. (2006), "Phase ambiguity resolution based on linear modeling of DD carrier phase", *Proc., ION GNSS 2006*

An Integrative Alignment Approach for Information Security Policy in the Context of Strategic Planning

Maria Soto Corpuz
Information Security Institute, Queensland University of Technology
Brisbane, Queensland/4000, Australia

ABSTRACT

The enterprise information security policy is derived from the strategic requirements for risk management and corporate governance. Consistent alignment between the security policy and the other corporate business policies and strategies has to be maintained if information security is to be implemented according to evolving business objectives. There are however limitations in current approaches for developing and managing the security policy to facilitate consistent strategic alignment within the business strategic planning cycle. The proposed full integrative planning approach for the enterprise information security policy conceptually demonstrates that the security policy can be presented as a business policy within the strategic management cycle. As such, this paper[1] argues that the security policy can take on integrated strategic planning activities alongside other strategic business policies. The recommended future research includes the adoption of the proposed security policy framework to establish applicability and the development of an assessment model for both the framework and the security policy. This is to validate the framework propositions and identify elements of value as well as areas for improvement.

Keywords: **information security management, enterprise information security policy, strategic management**

1. INTRODUCTION

Information security has been recognized as a core requirement for corporate governance that is expected to facilitate not only the management of risks [1][2], but also as a corporate enabler that supports and contributes to the sustainability of organizational operations [3]. In implementing information security, the enterprise information security policy is the set of principles and strategies that guide the course of action for the security activities [4] and may be represented as a brief statement that defines program goals and sets information security and risk requirements [5].

The enterprise information security policy (alternatively referred to as security policy in this paper) is a strategic statement presented as an element of corporate ICT governance [6] and is derived from the strategic requirements for risk management and corporate governance. Consistent alignment between the security policy and the other corporate business policies and strategies has to be maintained if information security is to be implemented according to evolving business objectives. This alignment may be facilitated by managing security policy alongside other corporate business policies within the strategic management cycle.

There are however limitations in current approaches for developing and managing the security policy to facilitate consistent strategic alignment. This paper proposes a conceptual framework for security policy management by presenting propositions to positively affect security policy alignment with business policies and prescribing a security policy management approach that expounds on the propositions.

2. INFORMATION SECURITY POLICY MANAGEMENT AND POLICY THEORIES

Organizational executive management often considers matters relating to information security to be mainly technical issues under the domain of information technology and commonly delegated to the IT department [5][6][7][8]. Observations by researchers [8][9] on current information security management system frameworks and standard practice guidelines [10][11][12][13] imply that these standards only provide suggestive definitions and characteristics of information security policy. There is a lack of definition of a process approach on its development and management. Information security management standards are checklists of security controls defined in generic terms [6]. Other security policy development and maintenance frameworks [14][15] that are proposed to address the limitations of the checklist approach relate to security activities that are usually conducted separately to those of corporate governance and risk management activities. Consequently, some security policy theories propose aligning the information security policy with the IT strategic planning [5][15]. In adopting these theories, full integrative relationship between security policy development and other strategic management activities such as strategic planning and corporate governance remains a challenge. The security policy is rendered as an IT-focused initiative with inadequate reference and at times misalignment with the other corporate strategic objectives that are not directly related to IT.

Further research on security policy approaches adopt the risk management framework for information security management systems as a security planning process [2][16] and as a full integrative management framework [17] between information security and corporate risk to present the actual linkages between security policy and corporate risk policy. However, these policy theories and approaches restrict the alignment positioning of information security policy with that of IT planning as it is seen as a technical concern [8][18] or at most with that of corporate risk planning as an operational risk [17].

The restrictive alignment positioning inhibits the relationship of the security policy (and consequently the management and implementation of information security) with other strategic policies and as a consequence does not encourage coordination between the different strategic policy domains. This perspective poses inherent disadvantages to the overall implementation of

[1] This full paper expands from the short paper titled *"The Enterprise Information Security Policy as a Strategic Business Policy within the Corporate Strategic Plan (Extended Abstract)"* published in the Proceedings of the Symposium on Risk Management and Cyber-Informatics (RMCI 2011) of the 15th World Multiconference on Systemics, Cybernetics and Informatics, Orlando USA, 2011.

information security as the strategic alignment of the security policy with the other business policies is not manifested and therefore hardly understood at the board level. As shown in Figure 1, the security policy is excluded from the corporate strategic management cycle as a business policy and alignment to other corporate business policies such as the corporate governance policy is not directly accommodated.

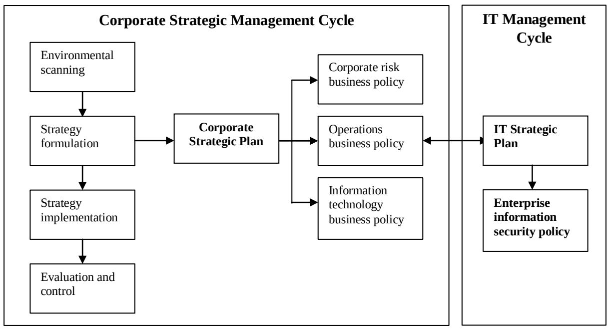


Figure 1 Traditional Alignment Approach for Enterprise Information Security Policy

The resulting security policy usually consists of an unstructured checklist set of security controls that is developed without the in-depth understanding to substantiate the requirement for functionality and value of such security policy and controls. Lacking in the corporate alignment on the strategic level, the security policy is dropped out of scope as a corporate-level agenda item. A consequence of this issue is the difficulty in obtaining the required support at the executive level [19] resulting inadequate allocation of resources and corporate support required for successful delivery of information security outcomes.

The strategic management cycle is discussed in the following section to provide a brief background upon which to base the conceptual development of the proposed security policy framework.

3. STRATEGIC MANAGEMENT AND BUSINESS POLICY

Strategic management is defined as a set of managerial decisions and actions that determines the long-run performance of a corporation [20][21][22]. The basic model of strategic management consists of four major phases namely: (1) environmental scanning; (2) strategy formulation; (3) strategy implementation; and (4) evaluation and control [22][23][24]. The environmental scanning stage involves the identification of the internal and external strategic factors that may affect the future of the organization [22].

The second stage of strategy formulation involves the development of the long-range plan through the definition of the mission, objectives, strategies and policies [24]. The mission statement defines the purpose and existence of the organization while the statement of objectives identifies what needs to be accomplished and by when to achieve the mission statement. Based on the defined mission and objectives, strategies are then developed. A strategy lays out the plan to achieve the organization's mission and objectives and requires broad directives to guide the organization's strategy formulation with its implementation [24]. These broad directives are defined in a corporate business policy [21]. As an element of strategic management, the business policy is the general management of integrated internal functions to ensure that decision's and actions are aligned with and in support of the defined mission, objectives and strategies [22].

The third stage of the strategic management model is the strategy implementation. Strategy implementation consists of developing and implementing the programs, budgets and procedures necessary to action the strategies and policies defined in the strategy formulation stage [23]. The last stage is the evaluation and control stage which involves the process of monitoring the actual performance and comparing the results with that of the defined performance objectives as a condition of business improvement [24]. The strategic management model is adopted in the development and overall management of strategic business policy strategies and may be utilized for developing security policy as a strategic business policy.

In the next section, the hypothesis and propositions are formulated and presented as a conceptual framework. This is followed by the security policy management approach that expounds on the propositions supporting the hypothesis.

4. CONCEPTUAL FRAMEWORK FOR ENTERPRISE INFORMATION SECURITY POLICY MANAGEMENT

In developing the propositions, the following hypothesis for strategic alignment denoted by H_{SP} (hypothesis for strategic planning alignment) is presented:

Hypothesis (H_{SP}): *Consistent alignment will be facilitated by developing and managing the security policy alongside other corporate business policies within the strategic management cycle.*

The conceptual framework in Figure 2 presents that strategic level development *($P1_{SP}$)* and an integrative management of the security policy within the strategic planning process *($P2_{SP}$)* will positively affect security policy alignment with business policies and support the requirements of H_{SP}. The next sections discuss the propositions and a security policy management approach that expounds on the propositions is then provided.

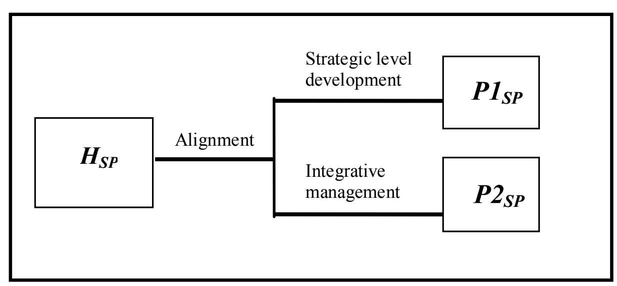


Figure 2 **Integrative Alignment Framework in the Context of Strategic Planning**

Security Policy as Corporate Business Policy

Corporate business policies are required to provide guidelines to ensure that all decisions and activities are aligned with the defined strategies as part of good corporate governance [22] [24]. As information security has become a crucial part of corporate governance [1][7] that focuses more on people, processes and information in addition to IT [2], the security policy that provides the critical direction-setting aspect of information security is one of the most important controls of corporate governance [9].

The security policy is based on business and organizational requirements to manage risks [3] and it is balanced against cost efficiencies and business benefits [10]. It requires continuous assessment and revision to address evolving risks [25][26]. The continuous assessment of security policy ensures practices and procedures within the security policy are coordinated and integrated [11].

Policy development for information security takes on concepts of strategic planning by way of objective setting and coming up with the program of projects to be undertaken according to the set business objectives. This apparent similarity where it is defined that the security policy contains a layered structure set of sub-policies and procedural implementation that need to be reviewed and assessed may be undertaken in the same manner as that of a corporate strategy or policy and its related program plan of business initiatives. The consistent alignment may be achieved by developing and managing the security policy as a business policy alongside the other strategic policies as depicted in Figure 3.

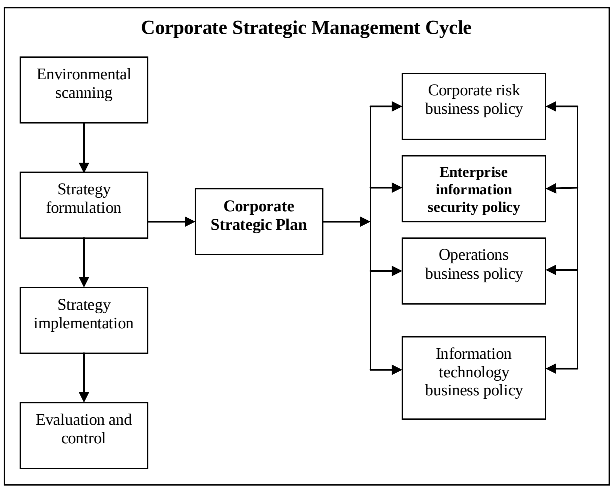


Figure 3 Proposed Alignment Perspective for Security Policy as a Strategic Business Policy

This perspective presents the security policy at the corporate level alongside the IT strategic plan and addresses the limitations in current approaches that confine the strategic alignment positioning only to that of IT planning. The first proposition ($P1_{SP}$) is presented as:

Proposition 1 (*$P1_{SP}$*): *Developing security policy as a corporate-level business policy will positively affect it's consistent alignment with other business policies as part of corporate governance.*

Integrating Security Policy within the Strategic Planning Process for Alignment

Research studies on business planning (BP) and information systems planning (ISP) have presented various integration approaches [27][28][29][30] that may be utilized to address alignment issues across corporate functions. The need for aligning information systems planning with business planning have been demonstrated by prescriptive [27][29] and empirical studies [31][32][33] as contributing to the assurance that business objectives are met and effective information technology investments are made. Information systems planning should be integrated within business planning to achieve alignment through full integrated planning [34][35]. This involves developing both the BP and ISP strategies at the same time using the same planning process and establishing an integrative relationship between BP and ISP.

The presence of the alignment mechanisms of content, timing and personnel inherent in a full integrative relationship [36] provides benefits to organizations as a result of improved coordination of information systems plans with business plans [31]. It has been asserted that information systems should not be regarded as separable from business strategy but instead fully integrated with the corporate policies [37]. As security policy development involves information systems planning, the adoption of the full integrative planning approach between security policy and strategic business policies is applicable.

Strategizing information security as a business policy may be then be approached by integrating security policy development within the traditional strategic planning and management cycle. This approach provides for consistent coordination of changes in requirements between the security policy and the other corporate business policies as presented in Figure 4.

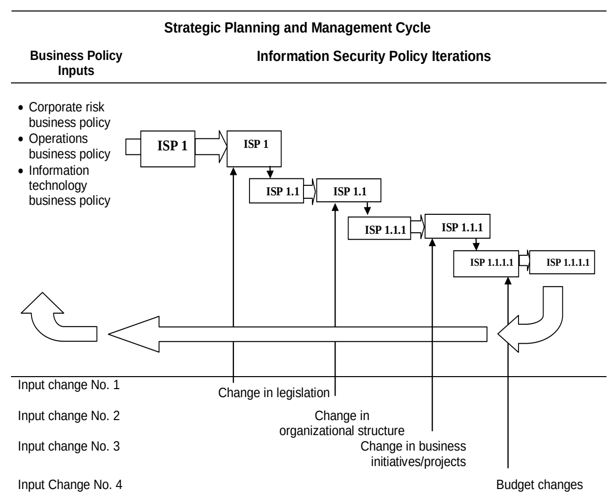


Figure 4 Timeline of Change Alignment within the Strategic Management Cycle

This manner of integrating different management domains by utilizing the same management framework process to enhance alignment has been prescribed for other corporate governance activities such as risk management [17]. From this perspective, the second proposition (***$P2_{SP}$***) is formulated:

Proposition 2 (*$P2_{SP}$*): *Integrating security policy development within the strategic planning and management process will positively affect the consistent alignment between the corporate business policies and the security policy.*

5. AN INTEGRATIVE ALIGNMENT APPROACH FOR INFORMATION SECURITY POLICY IN THE CONTEXT OF STRATEGIC PLANNING

The proposed approach for security policy management is developed by adopting the strategic management model to facilitate full integrative alignment as shown in Figure 5. This integrative approach depicts the security policy as a corporate business policy that is consistently aligned with business policies and objectives.

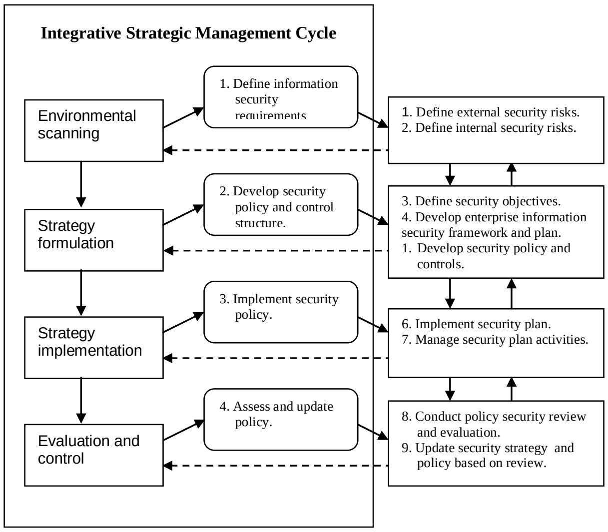


Figure 5 Integrative Approach for Developing Enterprise Information Security Policy

Briefly, the proposed framework process is described as follows:

Step 1 ***Define enterprise information security requirements*** **(environmental scanning)** – Environmental scanning for strategic planning involves defining the external factors and the internal factors affecting the business environment in all aspects [22]. This step enables the organization to establish the key business process and prioritize the focus of the strategic management efforts. The SWOT (strengths-weaknesses-opportunities-threats) analysis [38], a tool used in strategic management and business policy development, can be employed in environmental scanning.

For the security policy, this step accommodates the interplay and coordination of the various factors involved in developing each corporate strategy or policy with that of the security policy requirements on the strategic level. This approach expands the security policy requirements to include the risks involved in corporate governance, risk management and business operations in addition to the usual IT-related aspects of security risks. As an example, corporate governance risks that are not usually considered in traditional non-integrative approaches but are now identifiable in this step involve risks pertaining to staff misconduct and organizational disaster management.

In addition, this step encourages strategic initiatives for information security to be directly incorporated within the corporate strategic plan. As such, the security policy becomes an enabler that secures the environment for new technologies and services in the strategic portfolio. The output of this step is the Enterprise Information Security Risk Assessment Report.

Step 2 ***Develop security policy and control structure*** **(strategy formulation)** - The costs and trade-offs of the proposed enterprise information security policy and the corresponding policy control solutions as aligned to the other business policy control structures are determined in this activity. Detailed examples of controls that are developed based on the other business policy requirements such as risk management policy and operational business policies include security planning, processing authorization, incident response, configuration management, logical access control and secure application architecture. Useful techniques for evaluating risks and developing the strategies involve group decision-making employing the Delphi technique [39] and deriving decision trees to arrive at resolutions. For this step, the integrative approach allows the matching of costs and benefits between the security policy and control structures with those of the corporate business requirements on the strategic scale. The deliverable output of this process is the Enterprise Information Security Policy.

Step 3 ***Implement security policy*** **(strategy implementation)** - Based on the enterprise information security policy derived in Step 2, the enterprise information security plan is developed. The mitigating controls for the enterprise-wide security risks identified through the security risk assessment process incorporates strategic business initiatives and are finalized and drawn into a comprehensive enterprise information security policy set of initiatives or projects. The developed information security plan may take the form of a program of projects to implement the set of security controls which may be new treatments, additional controls or modifications to current controls. An important element of the security plan is the security architecture that consists of the security technologies as part of the implemented control structure of the security policy. One such example is the security policy stating a requirement for network security. The security technology that will meet this policy requirement may involve the deployment of an antivirus and firewall system. The output of this step is the Information Security Plan that details the projects and the taxonomy of procedural policies.

Step 4 ***Assess and update policy according to evolving business requirements*** **(evaluation and control)** - In this step, the review process for each policy development cycle is conducted to provide the basis for the next round of policy development activities. This activity fulfills the requirement for business improvement as required in strategic management. Coordinating and correlating lessons learned for the policy review process with that of the corporate business policy review activities, including that of corporate risk will further enhance future policy development activities as strategic objectives are met.

6. CONCLUSION

The goal of enterprise information security is to enable an organization to satisfy all of its strategic business objectives by implementing systems considering information technology-related risks to the organization, its business partners and clients. Although presented in various ways, the definition and characteristics for developing the strategic policy for enterprise information security remains consistent in that the security policy should outline the organization's program set of initiatives to ensure the security controls based on risk mitigation strategies are implemented to meet business objectives.

This paper presents a conceptual framework for security policy management by introducing propositions on how treating the security policy as a business policy may positively affect consistent alignment between the different strategic level policies. The proposed full integrative planning approach for the enterprise information security policy conceptually demonstrates that the security policy can be presented as a business policy within the strategic management cycle. As such, it is argued that the security policy can take on integrated strategic planning activities alongside other strategic business policies. This integrative approach is conceptually illustrated to foster an interactive relationship between security policy and other business policies to facilitate changes according to evolving requirements. The challenge that remains is the readiness of organizations to acknowledge the strategic value of treating the security policy as a business policy by adopting the proposed framework. The recommended future research includes the adoption of the proposed security policy framework to establish applicability and the development of an assessment model for both the framework and the security policy to validate the framework propositions and identify elements of value as well as areas for improvement.

7. REFERENCES

[1] R. Baskerville and M. Siponen, "An Information Security Meta-Policy for Emergent Organizations", **Logistics Information Management**, Vol. 15 No.5/6, 2002, pp. 337-346.

[2] E. Humphreys, "Information Security Management Standards: Compliance, governance and risk management", **Information Security Technical Report**, Vol. 13, 2008, pp.247-255.

[3] Organization for Economic Co-operation and Development OECD, **OECD Guidelines for the Security of Information Systems and Networks: Towards a Culture of Security**, OECD , 2002.

[4] W. Caelli, **Information Security Handbook**. Macmillan Publishers Ltd, 1991.

[5] V. LeVeque, **Information Security A Strategic Approach**, John Wiley and Sons Ltd, 2006.

[6] M. Siponen, "Towards Maturity of Information Security Maturity Criteria: Six Lessons Learned from Software Maturity Criteria", **Information Management and Computer Security Vol. 10 (5)**, 2002, pp. 210-224.

[7] B. von Solms and R. von Solms, The 10 deadly sins of information security management, **Computers and Security**, 2004, Vol. 23, pp. 371-376.

[8] D.F. Lohmeyer, J. McCrory and S. Poggreb, "Managing Information Security, **McKinsey Quarterly Special Edition**, 2002, pp 12-15.

[9] K. Hone and J.H.P. Eloff, "Information Security Policy – What Do International Information Security Standards Say", **Computers and Security**, Vol. 21, Issue 5, 2002, pp. 402-409.

[10] International Standards Organization ISO/IEC, **ISO/IEC 27001: Information technology – Security techniques – Information security management systems – Requirements**, International Standards Organization, 2005.

[11] M. Swanson and B. Guttman, **NIST Generally Accepted Principles and Practices for Securing Information Technology Systems Special Publication 800-14,** National Institute of Standards and Technology, 1996.

[12] J. Cazemier, P. Overbeek, L. Peters, **IT Infrastructure Library Best Practice for Security Management**. Office of Government Commerce, Crown Copyright, 1999.

[13] IT Governance Institute ITGI, COBIT 4.0 **Control Objectives Management Guidelines Maturity Models**, ITGI, 2005.

[14] J. Rees, S. Bandyopadhyay and E. Spafford, "PFIRES: A Policy Framework for Information Security", **Communications of the ACM,** Vol. 45. No. 7, 2003, pp. 101-106.

[15] N. F. Doherty and H. Fulford, "Aligning the Information Security Policy with the Strategic Information Systems Plan", **Computers and Security** Vol. 25, 2006, pp. 55-63.

[16] D. Straub and R. Welke, "Coping with Systems risk: Security Planning Models for Management Decision Making", **MIS Quarterly**, 22(4), 1998, pp. 441-469.

[17] M. Corpuz and P. Barnes, "Integrating Information Security Policy Management with Corporate Risk Management for Strategic Alignment", **The 14th World Multi-Conference on Systems, Cybernetics and Informatics Proceedings**, Vol 4, 2010, pp.337-342.

[18] B. von Solms, "Corporate Governance and Information Security", **Computers & Security**, 20(3), 2001, pp215-218.

[19] A. Kankanhalli, H. Teo, B. Tan and K. Wei, "An Integrative Study of Information Systems Security Effectiveness", **International Journal of Information Management** (23), 2003, pp.139-154.

[20] S. Tilles, "How to Evaluate Corporate Strategy", **Strategic Management: Harvard Business Review** (Ed: R. Hamermesh), USA: John Wiley & Sons Inc., 1983, pp. 77-96

[21] D. Harvey, **Business Policy and Management**, Ohio: Charles E. Merrill Pub., 1982.

[22] T. Wheelen and J. Hunger, **Strategic Management and Business Policy: Concepts and Cases**, New Jersey: Pearson Prentice and Hall Pub., 2008.

[23] E. Bowman, H. Singh and H. Thomas, "The Domain of Strategic Management: History and Evolution", **Handbook of Strategy and Management** (Eds: A. Pettigrew, H Thomas and R. Whittington), London: Sage Pub., 2006, pp.31-51.

[24] J. Kay, P. McKiernan and D. Faulkner, "The History of Strategy and Some Thoughts About the Future", **The Oxford Handbook of Strategy** (Eds: D. Faulkner and A. Campbell), Oxford: Oxford University Press, 2003, pp.27-52.

[25] Institute of Chartered Accountants ICA, **Internal Control: Guidance for Directors on the Combined Code**. Institute of Chartered Accountants (England and Wales), 1999.

[26] Information Systems Security Association (ISSA). **Generally Accepted Information Security Principles** *(GAISP V3.0),* 2004, ISSA.

[27] W. R. King, "Strategic Planning for Management Innformation Systems", **MIS Quarterly**, Vol 2 (1), 1978, pp. 27-37.

[28] W. R. King and R. W. Zmud, "Managing Information Systems: Policy Planning, Strategic Planning and Operational Planning, **Proceedings of the Second International Conference on Information Systems**, Boston, 1981.

[29] N. Goldsmith, "Linking IT Planning to Business Strategy", **Long Range Planning**, Vol. 24, No. 6, 1991, pp. 67-77.

[30] A.L. Lederer and V.Gardiner, "Strategic information systems planning: The Method/1 approach. **Information Systems Management**, Vol. 9(3), 1992, pp. 13-20.

[31] A. L. Lederer and A. L. Mendelow, "Coordination of Information Systems Plans with Business Plans", **Journal of Management Information Systems**, Vol. 6. No. 2, 1989.

[32] T. S.H. Teo and W. R. King, "Assessing the impact of integrating business planning and IS planning", **Information and Management**, Vol. 30, 1996, pp. 309-321.

[33] Y.E. Chan, R. Sabherwal and J.B. Thatcher, "Antecedents and Outcomes of Strategic IS Alignment: An Empirical Investigation", **IEEE Transactions on Engineering Management**, Vol. 53, No. 1, 2006, pp.26-47.

[34] T. S.H. Teo and W. R. King, “Integration between Business Planning and Information Systems Planning: An evolutionary-Contingency Perspective”, **Journal of Management Information Systems**, Vol. 14, No. 1, 1997, pp. 185-214.
[35] W. R. Synott, **The Information Weapon: Winning Customers and Markets with Technology**, New York, John Wiley, 1987.
[36] J.K. Shank, E. G. Niblock and W. T. Sandalls Jr., “Balance Creativity and Practicality in formal planning”, **Harvard Business Review**, 51 (1), January/February 1973, pp. 87-95.
[37] T. Smaczny, “Is Alignment between Business and IT the Appropriate Paradigm to Manage IT in Today’s Organization?”, **Management Decision**, 39(10), pp. 797-802.
[38] M. Porter, **The Competitive Advantage**, The Free Press, New York, 1985.
[39] T. Merna, **Risk Management at Corporate, Strategic Business and Project Level. MPhil Thesis**, 2002, UMIST, Manchester.

The Self Beyond Itself: Further Reflection on Spinoza's Systems Theory of Ethics

Heidi M. Ravven, Ph.D.
Hamilton College
36 Albany Street
Cazenovia, NY 13035

ABSTRACT

I offer here a new theory of moral agency –why people are ethical, why they are not, and how to get them to be more ethical. I look broadly at evidence from the new brain sciences, systems theory, and Spinoza's philosophy to address the question. Spinoza's vision of the infinitely expansive boundaries of the self as we progressively take in and embrace the world, and also extrude ourselves into the universe, forms the basis for rethinking moral agency and moral psychology. His vision opens up a broader swathe of the evidence coming out of both the theory of complex adaptive systems and also out of the new brain sciences than is usually thought of as relevant to exploring and understanding what makes us moral beings. It points us toward the pertinence of understanding moral agency in terms of nested complex adaptive systems.

Keywords: Spinoza, moral agency, complex adaptive systems, affective neuroscience

Introduction

I argue here that it is a sense of self as spanning mind-body and world that is the origin and nature of our moral investment and agency in the world. We locate our basic biological sense of self-preservation and self-furthering in a self distributed beyond our skin into our environments, natural and human. This is why we care about the world and why it is the arena of our moral concern and of our ideals.[1] In this paper I review some of the mounting evidence that the scope of the self as moral agent, of who is performing a given moral action, can be *extended* into the environment as extensions of self and also *distributed* beyond the individual to groups, and even encompass at times whole contexts. The scope of the self as actor, its agency, can in certain circumstances be laid at the feet of social-cultural-historical systems, spanning time and place and even generations.

The self is permeable and relational (as well as self-promoting, self-protecting, and self-furthering) –rather than closed and discrete and playing out its own internal program upon the world stage. Neurological body maps literally *extend* the 'me' to include the hammer I use when I nail the picture on the wall or the car when I'm driving. Research reveals that the feeling I have that Tessie, my metallic light blue Acura, is an *extension* of my body when I drive in fact reflects the neural reality. For my body maps are extended to include its proportions and motions as within the bounds of my self that I feel and control. They are mapped within the 'body mandala', so to speak. There is a 'tool-body unification' or *extension*. There is also a certain amount of space surrounding our body, 'peripersonal space', that is like a bubble around us that is included in our neural self-maps.[2] The expansion to include tools and other objects that we use to do things and carry out our aims, as well as all kinds of biographical and cultural and familial information, is now referred to in neuroscience as the Extended Self.

The philosopher Andy Clark, in his 2008 book, *Supersizing the Mind*, has written about how the mind spills over into the world.[3] When he argues in this book that the mind is 'extended' he means that "at least some aspects of human cognition … [are] realized by the ongoing work of the body and/or the extraorganismic environment," so that the "physical mechanisms of the mind … are not all in the head" and in our central nervous system.[4] When we use a computer or a pad of paper, a calculator or our address book, our mind has both distributed memory and even operations outside itself.

This view of the matter radically complicates and reconfigures the nature of the relationship between mind and world, the neurophilosopher David Chalmers.[5] Andy Clark says that we have "a fundamentally misconceived vision … that depicts us as 'locked-in' agents –as beings whose minds and physical abilities are fixed quantities apt (at best) for mere support and scaffolding by their best tolls and technologies." He proposes instead that our "minds and bodies are essentially open to episodes of deep and transformative restructuring in which new equipment (both physical and 'mental') can become quite literally incorporated into the thinking and acting systems that we identify as our minds and bodies."[6] As an example Clark mentions a robot arm, an arm that extends one's reach and gets mapped into self-maps whose scope now includes its reach.

Clark's analysis supports what Antonio Damasio proposed about the incorporation of the extended biographical and cultural aspects of the self in a third level of self-mapping: becoming and being a self, on the one hand, and responding to and incorporating multifaceted contexts are one and the same ongoing process. Yet there is more. For we not only discover the world within us but *we also discover ourselves in the world, identifying ourselves with parts of it.* The psychoanalysts call this Projective Identification. Philosophers since Spinoza have referred to the Group Mind. Psychologists have studied Mass Psychology. And neuro-philosophers have begun to explore the phenomenon of *Distributed Agency*, a subject of action that is larger than the individual. The Distributed Self leaks out of its boundaries of the skin, and can even feel itself somewhere else entirely outside the body. We can call this: *The I that is We*.

---

[1] See my articles: H. Ravven, "Spinoza's Anticipation of Contemporary Affective Neuroscience" in *Consciousness and Emotion*, an interdisciplinary science and philosophy journal ed. by Ralph Ellis and Natika Newton, (Volume 4 Number 2, 2003) and, "Spinoza and the Education of Desire," *Neuro-Psychoanalysis*, vol. 5, issue 2, 2003, pp. 218 - 229. My (invited) review essay is part of an extended exchange among neuroscientists and philosophers, the neuroscientists Jaak Panksepp, Douglas Watt, and Antonio Damasio and myself on Antonio Damasio's recent book, *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain.* Antonio Damasio's response to my review essay appears just after it as does his response to the review essay of Panksepp and Watt

[2] Blakeslee and Blakeslee, *The Body Has a Mind of Its Own* (Random House, 2008): pp. 142-143 and 117-118

[3] Andy Clark, *Supersizing the Mind: Embodiment, Action, and Cognitive Extension* (Oxford, New York, Auckland, Cape Town, Dar es Salaam, Hong Kong, Karachi, Kuala Lumpur, Madrid, Melbourne, Mexico City, Nairobi, New Delhi, Shanghai, Taipei, Toronto: Oxford University Press, 2008)

[4] *Supersizing the Mind*, p. 82

[5] *Supersizing the Mind*, p. xvi

[6] *Supersizing the Mind*, pp. 30-31

'The I that Is We'

"Humans are collective thinkers, who rarely solve problems without input from the distributed cognitive systems of culture."[7]

-Merlin Donald

"Where do you stop, and where does the rest of the world begin? There is no reason to suppose that the critical boundary is found in our brains or our skin."[8]

-Alva Noë

I approach the discovery of the self beyond itself and in the world from a number of angles. The evidence is building for the *extension* of our selves into our tools and computers and pencils, into robot arms and cell phones and, of course, our cars; and for the *distribution* of our sense of self into shared environments and contexts, from culture and family to nation, to school and neighborhood, to generation and church. Here we find the source of a sense of *distributed agency*, that it can be the group, rather than the individual, who is performing an action or making a decision. There is growing evidence of a distributed self from widely different quarters: from psychological studies of infants in their development of 'co-consciousness', a shared world and a self co-constructed by self and environment; from studies of our unconscious thinking and feeling, which reveal that each of us has a number of implicit working self-concepts rooted in two-person repertoires that arise from self-with-significant other (mother, father, siblings, and the like) relational patterns ingrained in early childhood and triggered ever anew by the environment; from the surprising neurobiology of out-of-body experiences that reveal that we can discover our feeling of self outside of our bodies and lodged in parts of the environment; from the neurochemistry of the self-other boundary, a boundary that breaks down and enables the other to feel like self in empathy and love but also in shared anger and fear; from the discovery of mirror neurons, brains cells in others when they act, which directly cause homologous brain cells to fire in mere observers of the action, creating a shared experience of actor and observer *from the inside*; and from the sociological analysis and meta-analysis of success and intelligence, whose findings identify the major causes of outstanding individual achievement as environmental, social and cultural, rather than individual or genetic.

Taken together the amassing evidence ought to begin to change where we look when we we're searching for ethics. We should begin to look not inside the individual, as we have assumed, but rather outside, in the environment. Some philosophers and other theorists have begun to do just that.

There is a growing movement to rethink thinking, and the mind more generally, as embodied and embedded in its environments. So the mind is not a brain in a vat. The days when thinking is likened to a computer program are coming to an end. As the UC Berkeley philosopher Alva Noë puts it in his book, *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness,* the standard view, not only in philosophy but in neuroscience, has been that "we are brains in vats on life support. Our skulls are the vats and our bodies the life-support system."[9] But that standard view is turning out to be wrong.

One of the things that the old view assumed was that it made no difference whether thinking takes place in an embodied person embedded in its environment or instead in a machine or something else. So a good analogy to the way thinking was supposed to work was a television. You could see a movie in a theatre, on TV, or on your computer and it wouldn't matter much except for the scale and the clarity. But the movie was the movie just different technology bringing it to you. But that's not the way the mind is turning out to work.

Instead, and contrary to decades of the dominance of the standard 'movie' account in cognitive science, the ways that the brain is biologically, neurologically, and ecologically constructed are coming to be appreciated as supremely relevant to the *content* of the mind. The mind is not a computer running a discrete genetic or other kind of internally constructed program that would be the same on any type of hardware. That computer or media metaphor, a metaphor that has driven a great deal of research, is simply misguided when it comes to human thinking. For the body-in-context shapes the mind –and the content of the mind—in crucial respects rather than merely underlying it.

The new conception of the mind is that it is not only embodied but also *embedded* in its environment: in its contexts and situations and histories and communities of all kinds, social and cultural and linguistic and natural. As Alva Noë puts it,

> The limitations of the computer model of the mind are the limitations of any approach to mind that restricts itself to the internal states of individuals. [10] … The content of experience—what we experience—is the world; in the world's absence we are deprived of content. [11]

Finally, in addition to the *embodiment* and *embeddedness* theses, there is the *extendedness and distribution* thesis. This is the claim that the mind is not confined to the skull. It means that, "the boundaries of cognition extend beyond the boundaries of individual organisms,"[12] because "the skull is not a magical membrane." Instead, both what's in the mind but also *who's doing the thinking and acting* are "boundary crossing" and "world involving."[13] That's an idea that 'blows your mind', blows it open—literally!

---

[7] Merlin Donald, "How Culture and Brain Mechanisms Interact in Decision Making," chapter 9 (pp. 191 – 205) in Christoph Engel and Wolf Singer, ed., *Better Than Consciousness? Decision Making, the Human Mind, and Implications for Institutions*. Strung Form Reports, (Cambridge, Massachusetts and London England: MIT Press, 2008), p. 192

[8] Alva Noë, *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness* (New York: Hill and Wang, A division of Farrar, Straus and Giroux, 2009): p. 60 My emphasis

[9] *Out of Our Heads*, p. 5

[10] *Out of Our Heads*, p. 169

[11] *Out of Our Heads*, p. 180

[12] Philip Robbins and Murat Aydede, "A Short Primer on Situated Cognition," Chapter 1 in Robbins and Ayedede, editors, *The Cambridge Handbook of Situated Cognition* (Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sào Paulo, Dehli: Cambridge University Press, 2009): p. 3

[13] *Out of Our Heads*, pp. 48-49

Our thinking and our acting are not separated, as we tend to think of them as cognitive reflection on, and an internal picture (representation) of, a world separated from ourselves upon which we take independent action. Instead, perception and cognition depend upon and are crucially constructed by the way we *interact* with the world. Cognition is now being shown to involve the sensory motor brain, that is, "motor capacities, abilities, and habits."[14] This may occur both 'online', so to speak, and 'offline'. Action and cognition are, in some not yet fully delineated or completely understood ways, bundled together, causally interdependent, rather than discrete and independent processes. Our perception is interdependent with bodily motor relationships –which is to say perception is an *interaction with its environment* and what is perceived *is* the interaction, rather than a self-removed grasp of the external environment *per se*.

The theory of how perception is shaped by how we interact with things was first put forth in the 1960s by psychologist, James Jerome Gibson of Cornell University. Gibson theorized that we –and animals as well, Gibson proposed-- do not perceive objects, or the environment more generally, objectively in terms of the shape and volume of objects. Instead, we perceive the environment and objects in terms of how we envision how we can interact with them; we do not see objects *per se* but rather "affordances that make possible and facilitate certain actions." The Blakeslees explain what Gibson meant by 'affordances' by suggesting that, "handles afford grasping. Stairs afford stepping. Knobs afford turning. Hammers afford smashing." We perceive the world, according to Gibson, "through an automatic filter of affordances."[15] As the Blakeslees put it:

> Your perception of a scene is not just the sum of its geometry, spatial relations, light, shadow, and color. Perception streams not just through your eyes, ears, nose, and skin, but is automatically processed through your body mandala to render your perceptions in terms of their affordances. That is generally true of primates, whose body mandalas have grown so rich with hand and arm and fine manipulation mapping, and even more so for you, a human animal.[16]

If thinking, and even our basic perceptions of the world around us, are not separate and separable from acting, then moral agency cannot be, as we tend to assume, about understanding and assessing situations from a removed perspective, and then subsequently, rationally and independently, choosing the right action. For on the new model all three are bound together—perhaps in the way that emotions and cognition have been found to be bound together in neural packages and pathways.

And mirror neurons also lend support to the view that we *act out* scenarios as the basis of understanding others' actions. Motor schemas seem to be far more than about action as the scientific terminology has characterized them. Perception, emotion, action, cognition, empathic understanding of others, all seem to be integrated together in a context that includes self and environment mutually interacting. We are within systems.

What we believed to be discrete and bounded mental processes that 'we' then in some sense preside over from above and bring together, are turning out to be more intimately bound together from the start and all the way up the line. There is no independent 'we' or 'I' outside of these bundled perceptual, conceptual, affective, and enactive processes and self-world environments; there is no 'I' who stands above them as if they belong to someone else or are distant parts of the world, and looks down upon them –the body as if 'other' in Cartesian fashion—from the perspective of an inner mental 'I' who then decides or feels or chooses or acts. As Alva Noë puts it:

> Scientists seem to represent us as if we were strangers in a strange land.[17] … Our relation to the world is not that of an interpreter. … Our relation to the world is not that of a creator. The world is bigger than we are; *what we are able to do is to be open to it*."[18]

Openness to the World: Cognitive Externalism and Distributed Agency

"It's not what is inside the head that is important, it's what the head is inside of."

--James Jerome Gibson

"*Human decision making is most commonly a culturally determined process … When the individual 'makes' a decision, that decision has usually been made within a wider framework of* ***distributed cognition****, and in many instances, it is fair to ask whether the decision was really made by the distributed cognitive cultural system itself, with the individual reduced to a subsidiary role. … Distributed systems are able to change where in the system each component that influences a certain decision is located.*"[19]

--Merlin Donald

Openness to the world would seem to be our fundamental posture. We are of the world and in it, engaged in and engaging the environment and our many contexts. The misleading but dominant metaphor of 'seeing' as our basic relation to the world obscures this reality. Seeing places us too much on the outside looking in. Let's replace sight with touch. If we think of ourselves as fundamentally touching and being touched, acting and being acted upon, and acting together with others, then we can grasp our fundamental openness. Each of us can come to be aware of ever-larger contexts and environments in which we are embedded, as affecting and being affected. The mere local context we grasp is too narrow to contain or explain the scope of the openness of the self. The self we are and with which we can identify moves ever outward. We discover our thinking, emotion, and action as a product of the group and of ever-wider contexts and environments. We come to know ourselves by discovering ourselves beyond ourselves. The self is *distributed*.

The brain interacts with the world in ways that influence our perception itself. Pivotal findings of neuroscience,

[14] Philip Robbins and Murat Aydede, p. 5
[15] Blakeslees, p. 106
[16] Blakeslees, p. 106
[17] *Out of Our Heads*, p. 183
[18] *Out of Our Heads*, p. 184
[19] Merlin Donald, p. 202 My emphasis

particularly those of Jaak Panksepp,[20] about the crucial role of action in perception were anticipated by the philosopher Susan Hurley beginning in the 1990s, and especially in her first great breakthrough work, *Consciousness in Action*.[21] Her insights go a long way to explaining and establishing that the boundary between self and world is not set by our skin. Instead, patterns of interaction are what thinking is all about. In *Consciousness in Action* Hurley argued that action is distributed among mind, body, and world rather than being attributable to the individual alone. And her conjecture has turned out to have a great deal of neurobiological evidence in support of it.

Hurley argues that our standard (and generally unconscious) assumption, that perception and action operate according to an input-output model, is simply incorrect. This means that we do not simply have sensory faculties that bring us (raw) data from the world (input), which we make sense of according to some internal genetic or other program, and then act upon (output). We falsely presume that the mind is bounded in a way that separates it/us from the world, so that input and output are distinct processes.

Alva Noë explains that, "We are not world **representers**. … Our worlds are not confined to what is inside us, memorized, represented." Instead, "we live in extended worlds" that are "reachable" rather than "depicted."[22] By trying to rid us of the presumption that we are 'representers', Hurley is banishing our sense of ourselves as observers of the world rather than participants in it. She is banishing the metaphor of 'seeing'. Another way to put it is that we are not somehow separate from the world, our brains constructing a common intersubjective internal world by imposing standard patterns upon a chaos of disorganized perceptual data that are, thereby, structured in meaningful ways. These subjective and intersubjective constructions, the standard story goes, then play out as films going on in our heads that have a tenuous relation to the actual external world. We are locked in. Hurley sets out to refute this view that the mind is an *internal* program playing out upon a world stage.

Part of the conceptual problem, she says, is that we tend to focus almost exclusively on the input to output direction: how the mind structures incoming percepts. And we tend to ignore the functions from output back to inputs, or, "the way environments, including linguistic environments, transform and reflect outputs from the human organism." In other words, the world we encounter is not an unstructured arena of chaotic raw data but in fact is (pre-)structured by human practices, linguistic meanings, institutions, histories, cultures, and nature itself. So both directions are just as complex, Hurley remarks; and not only that, they "are causally continuous." "To understand the mind's place in the world," she goes on, "**we should study these complex dynamic processes as a system**, not just the truncated internal portion of them." Our place in the world is "*a complex dynamic feedback system* [that] includes not just functions from input to output, but also feedback functions from output to input, some internal to the organism [that is, from internal data about the body's state back to the mind], others passing through the environment before returning."

[20] Jaak Panksepp (in an unpublished manuscript) calls the complex the 'Emotion Action Systems'.

[21] Susan Hurley, *Consciousness in Action* (Cambridge, Massachusetts and London, England: Harvard University Press, 1998)

[22] *Out of Our Heads*, pp. 82, 83 My emphasis

Hurley also challenges the notion that the contents of our mind and the structuring of the mind are independent of the world, and the world we take in is independent of the mind. Her book is an extended argument for Externalism, the view that the self is in the world and that self and environment are related in and as interacting open systems. What that means is that there is nothing that is either 'pure self' or 'pure environment'. Hurley says that the Externalism she advocates is a version of Contextualism. Both are always interactively constituted:

> The revolution that began with Kant's arguments about perceptual experience should be carried through to agency. Action is no more pure output than perception is pure input. The whole of the Input-Output picture should be rejected, not just half of it.

When we act, we create a relationship to the environmental context we inhabit, and that relationship both influences what we perceive, on the one hand, and also structures the mind's way of perceiving.

It is a mistake to confuse the vehicle with the content. The fact that we have *neural architecture* that makes possible a sense of self, does not mean that the *content* of that self is also a product of that architecture alone. Nor are our minds completely passively determined by the. Instead we can understand the internal neural architecture as making possible *openness* to the world and shaping by the world along with our shaping of the world. Person and world are relational, interactional, and also contextual.[23]

The themes of Hurley's revised approach to the mind –decentralization, self-organizing systems, context dependence, feedback, emergence—have resonance in research programs in connectionism, dynamic systems theory, and artificial life. She remarks that the input-output view in ethics, presupposes and bolsters the claim of free will, for it conceives us human beings as originating sources of causal chains. Hurley gives a deathblow to free will since it envisions human beings, instead, as contextually embedded in natural and social causal networks and webs.[24]

Complex Adaptive Systems

Human cognitive processes are inherently social, interactive, personal, biological, and neurological, which is to say that a variety of systems develop and depend on one another in complex ways.[25]

--William J. Clancey

"The mind leaks out into the world, and cognitive activity is distributed across individuals and situations. This is not your grandmother's metaphysics of mind: this is a brave new world. Why should anyone believe it?

"One part of the answer lies in ***the promise of dynamical systems theory*** *... as an approach to modeling cognition. ... Insofar as the mind is a dynamical system, it is natural to think of it as extending not just into the body but also into the world. The result is a radical challenge to traditional*

[23] *Consciousness in Action*, pp. 263, 264
[24] *Consciousness in Action*, p. 250
[25] William J. Clancey, "Scientific Antecedents of Situated Cognition," Chapter 2 in *The Cambridge Handbook of Situated Cognition* (pp. 12-34): p 12

ways of thinking about the mind, Cartesian internalism in particular."[26]

--Philip Robbins and Murat Aydede

The insight that a person is an open system in relation to other open systems, natural and cultural, has begun to be rigorously articulated and theoretically worked out in the developing field of Systems Theory. Accoring to theorist William J. Clancey:

> An all-encompassing generalization is the perspective of complex systems. From an investigative standpoint, the one essential theoretical move is contextualization (perhaps stated as 'antilocalization…: we cannot locate meaning in the text, life in the cell, the person in the body, knowledge in the brain, a memory in a neuron. Rather, these are all active, dynamic processes, existing only in interactive behaviors of cultural, social, biological, and physical environment systems.

A self, according to this approach is 'self-organizing' and 'unfolding' and always contextualized or 'situated'.[27]

Thinking about human behavior in terms of systems changes dramatically the way we conceive agency, what it means to act and who is doing the acting. That is the conclusion of the computer scientist Merlin Donald of Case Western University. In an article, "How Culture and Brain Mechanisms Interact in Decision Making," Donald argues that although decision-making "seems to be a very private thing: individualized, personal, and confined to the brain," when looked at from system theory, we realize that culture is a major factor in how the brain self-organizes during development, both in its patterns of connectivity and in its large-scale functional architecture.[28] So it is the system that makes the decision: "The mechanisms in such decisions must be regarded as hybrid systems in which both brain and culture play a role." Donald, of course, does not deny that decisions are made in individual brains. Nevertheless, he points out that "human brains … are closely interconnected with, and embedded in, the distributed networks of culture" that "define the decision-space."

To discover who is actually acting in a given case, all the facts need to be taken into account and then analyzed from multiple standpoints, from the brain sciences to organizational behavior to culture and history, etc., etc. Only through this multidisciplinary approach can the attribution of agency and responsibility be accurately *distributed* across people and levels of organization and participation and authority.

"Complex adaptive systems are quite different from most systems that have been studied scientifically. They exhibit coherence under change, via conditional action and anticipation, and they do so without central direction."[29]

--John H. Holland

"Complex adaptive systems [are] those that learn or evolve in the way that living systems do. A child learning a language, bacteria developing resistance to antibiotics, and the human scientific enterprise are all discussed as examples of complex adaptive systems."[30]

--Murray Gell-Mann

The human person is not only an open system within others but also *adapts.* Complex adaptive systems are a special kind of system. Adaptive systems are those in which emergence and self-organization hold sway.

The control in a complex adaptive system is decentralized and widely distributed, rather than being under some central control. The patterns of activity arise or emerge from the interactions of the agents rather than from some overall plan. Nevertheless there is dynamic stability, identifiable patterns that are neither utterly chaotic nor substantially fixed. These patterns evolve, changing over time as the system itself changes and evolves. It is the individual 'agents' in the system that, from their location and environment, develop adaptive behavior.[31] They exhibit the same patterns of the whole at various scales within the system. Learning is an important feature of complex adaptive systems even though there is no central consciousness involved. And they are highly resilient. Ever increasing diversity is an important feature of Complex Adaptive Systems and crucial to their capacity to adapt and survive.

Diversity in Complex Adaptive Systems

The hallmark of complex adaptive systems is perpetual novelty, according to John Holland. Diversity arises from how this kind of system recycles its resources. ! The Recycling Effect produces resources available to be used in new environmental niches and these niches are filled by increasingly diverse species. For "each new adaptation opens the possibility for further interactions and new niches."[32] A complex adaptive system does not settle into locked-in patterns but keeps producing change.

It is the particular niche that defines the kind of diversity that will arise. In evolution, what this results in is convergence. It is necessary for the ongoing evolution of the system.

Conclusion

How can thinking in terms of complex adaptive systems help us rethink moral agency and come up with ways to make societies and all kinds of groups function more ethically? Crucially, complex adaptive systems theory suggests that to change the person we need to look at the system. Interventions in context and environment rather than in the brain or mind of the person (whether through drugs or the training of the individual will) seem to be the place to start.

Diversity is crucial, too. We need to think about diversity and its role in complex adaptive systems to ensure their ongoing vitality and continuing evolution. In social systems, diversity plays out as the introduction of diverse people, and practices, and points of view that challenge and

[26] Robbins and Aydede, p. 9
[27] Clancey, p. 28
[28] Merlin Donald, "How Culture and Brain Mechanisms Interact in Decision Making," chapter 9 (pp. 191 – 205) in Christoph Engel and Wolf Singer, ed., *Better Than Consciousness? Decision Making, the Human Mind, and Implications for Institutions*. Strung Form Reports, (Cambridge, Massachusetts and London England: MIT Press, 2008): p. 192

[29] *Hidden Order,* pp. 38-39

[30] Murray Gell-Mann, *The Quark and the Jaguar: Adventures in the Simple and the Complex* (New York: W H. Freeman and Company, 1994): p. x
[31] "Complex Adaptive Systems," in Wikipedia.org.
[32] *Hidden Order*, p. 29

disrupt the stable social system, sparking a more complex and inclusive reordering and reintegration.

There is a tradeoff between closure to variation and the resulting static internal coherence, on the one hand, and on the other, openness so great that the system cannot accommodate the differences fast enough through internal systemic reintegration.

Spinoza theorized a systems theory of ethics that was perpetually reorganizing at the brink or 'edge' of chaos. He advocated a personal maximal openness to others and to the world while retaining the capacity for dynamic self-organization. Each of us ought to cultivate an openness to others that doesn't overwhelm us but can be integrated into our sense of self and what we care about through both understanding and the increased capacity for empathic identification, standing in the other's shoes, perspective taking, and also openness to critique and self-critique. As a friend puts it, it's not just about tolerating differences but finding in oneself the capacity to enlarge one's empathic acceptance and capacity to learn from others about themselves and *also about one's own self from another's perspective*. So ongoing diversity is necessary to the overcoming of self-deception, dogmatism, and denial –our most ubiquitous and corruptive moral dangers, those of 'selfiness'.

We can now rethink the problem of 'selfiness' from a systems perspective. 'Selfiness' is the attempt to maintain a narrow systematicity and coherence that won't allow in challenging data from others or even from the implicit meanings and intentions of our own actions. Selfiness tends toward the refusal to acknowledge that one is a part of larger systems, cultural, social, and natural. It is the arrogance of the myth of self-creation, of free will. The overcoming of narrow selfiness of this kind in an expansive self-coherence that enlarges the self to include more of the world and others is a lofty ideal for the individual and a noble and difficult path. It is also a rare one, as Spinoza pointed out.[33]

A Final Word on Spinoza

Spinoza anticipated Externalism and Systems Theory[34] and he rethought moral agency in terms of them. He envisioned Nature as a network, a system of causes at all levels from the cultural to that of physics. Each person, animal, or thing was a *location* in the system of networks, a location that defined the point in the open dynamic systems within systems that is the infinite universe. He regarded each person not as a static thing or quasi-genetic program but instead as what he called a 'ratio'. Each is an open system within open systems, and each system at every level strives to maintain its internal homeodynamic organization while being open to the larger systems, environments, which were its constitutive causes and to which it also contributed.

For Spinoza the self strives to become a more internally coherent, self-organizing internalization of its immediate world—Spinoza calls that dynamic well-functioning of self as a coherent system, 'activity'-- and then of its more distant environments. The paradox and irony is that to be truly yourself *is* to be your world, and ultimately the universe that created you.

To be this self is to be this point in the universe, and it took the whole universe up till now to produce any given 'me'. So to attain what he regarded as a state of personal autonomy, or 'freedom', as he called it, to achieve the spiritual and moral psychological aim of *The Ethics*, was to come to understand and own as self all that has come to make up this biologically, psychologically, socially, culturally, historically, biographically and of course today we would include cosmologically and quantum mechanically, etc., etc, situated and constructed self.

The world thus is systematically introduced into the self as causes of the self and hence as self –but in the doing the self now flips and sees itself in terms of its world, in terms of those parts of the world that appear now as personally constitutive, and there is no limit to that centrifugal force. We are in principle at home in the universe and our freedom lies in making that real to ourselves. The environment is not foreign but constitutive. So the outcome and irony of autonomy is that its achievement only comes to fruition in the embrace of the environment and of those things within the environment in which one now sees oneself, and progressively more so to infinity. To see aspects of the environment as 'self', rather than only as 'other', is to feel the world not as merely external limit to the self but as constitutive of the self and the self as distributed to and contained within its environment. We have arrived at a familiar place but from a new directions: We have come to love the other as the self—literally.

[33] See my three papers on Spinoza and Systems Theory: H. Ravven, "What Can Spinoza Teach Us Today About Naturalizing Ethics? Provincializing Philosophical Ethics and Freedom without Free Will" " in *Cognitive, Emotive, and Ethical Aspects of Decision Making in Humans and in Artificial Intelligence* Volume III. Edited by Iva Smith and Wendell Wallach, George E. Lasker, Editor-in-Chief. Published by The International Institute for Advanced Studies in Systems Research & Cybernetics, Canada, 2005: pp. 99-104;
H. Ravven, "Spinoza's Systems Theory of Ethics" in *Cognitive, Emotive, and Ethical Aspects of Decision Making in Humans and in Artificial Intelligence* Volume III. Edited by Iva Smith and Wendell Wallach, George E. Lasker, Editor-in-Chief. Published by The International Institute for Advanced Studies in Systems Research & Cybernetics, Canada, 2004;
H. Ravven, "Notes on Spinoza's Critique of Aristotle's Ethics: From Teleology to Process Theory," *Philosophy and Theology*, Volume IV, #1, Fall 1989, pp. 3 -32

[33] "Opinionator: The Stone," The New York Times, October 17, 2010

[34] I have written on Spinoza's anticipation of systems theory and his rethinking of ethics in terms of it. See my: See my three papers on Spinoza and Systems Theory: H. Ravven, "What Can Spinoza Teach Us Today About Naturalizing Ethics? Provincializing Philosophical Ethics and Freedom without Free Will" " in *Cognitive, Emotive, and Ethical Aspects of Decision Making in Humans and in Artificial Intelligence* Volume III. Edited by Iva Smith and Wendell Wallach, George E. Lasker, Editor-in-Chief. Published by The International Institute for Advanced Studies in Systems Research & Cybernetics, Canada, 2005: pp. 99-104; See also the new study by Rainer E. Zimmermann, *New Ethics Proved in Geometrical Order: Spinozist Reflexions on Evolutionary Systems*: Exploring Unity Through Diversity Volume 2, (Litchfield Park, AZ: Emergent Publications, 2010

Sme Development Strategies based on Financial Accounting, and Internationalization

Vivanco JESUS.
Finance Department
Universidad A. de Aguascalientes
Aguascalientes C.P.20131

and

Oropeza MIGUEL
Accounting Department
Universidad A. de Aguascalientes.
Aguascalientes C.P.20131

and

Montalvo MIGUEL
Marketing Department
Universidad A. de Aguascalientes.
Aguascalientes C.P. 20131

ABSTRACT

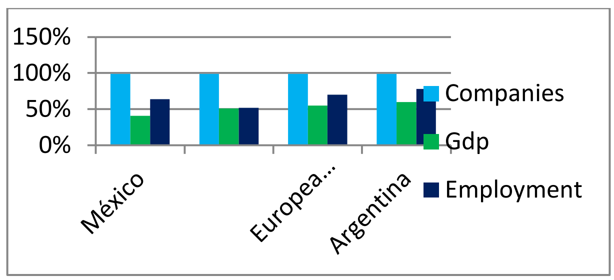
This research aims to identify strategies to increase performance and development of SMEs and financial strategies like liquidity and leverage sustained in accounting as well as the internationalization of SMEs can help in their development and enable them to stay on the market, since a high percentage of SMEs fail to overcome the barrier of two years, on this research through a survey of 125 SMEs in the State of Aguascalientes Mexico, is demonstrated by statistical analysis such as linear regression and ANOVA [15], a high positive relationship between the administration of the financial aspects such as liquidity and leverage, the internationalization of SMEs with increased performance.

Keywords: Financial Strategies, internationalization, SMEs

1.INTRODUCTION.

SME's represents the economic platform of our country, since they currently used 62% of the economically active population, contribute 42% of national GDP and 99% are companies operating in Mexico, which is why our concern to seek development and retention tools, the finance function implies a fundamental factor for its effective operation, this research will address the topic of finance and the internationalization as a means or strategies for the development of SMEs, starting by introducing some informative data representing the perspective of SMEs in Mexico.

Graphic 1
Economic Impact of SME's



Sours: Small and medium size enterprises OECD June 2002

As shown in the graph, SME's represent a significant contribution to the economies of the countries, since their contribution is very high in terms of GDP, employment and number of companies, as this involves the livelihood of millions of families who depend permanence of SME's to ensure their livelihood and their families

Table No. 1
Table of Enterprises Classification
by number of employees

	Industry	Commerce	Services
Micro	0 to 10	0 to 10	0 to 10
Small	11 to 50	11 to 30	11 to 50
Medium	51to 250	31 to 100	51 to 100

Source: Official Journal of Mexican Federation, December 30 2007.

2.OBJECTIVE OF RESEARCH

The objective is to identify the influence of the administration of financial aspects such as liquidity and leverage and the internationalization of SMEs in the improvement of their performance

Specific objectives.
a) .- Identify which of the financial aspects, the liquidity or leverage, have greater influence on improving their performance of SME's.
b) .- To identify if the internationalization of SME's can improve their performance

3. LITERATURE REVIEW

Financial aspects supported the development of accounting. The Modigliani and Miller theorem establishes, under restrictive assumptions indeed, that the financial structure has no influence on the value of the company. However, many theoretical works that justify, under more realistic, positive and negative connections between the form of financing and firm performance or efficiency [1], It has been argued also less indebted firms are more efficient in sector al restructuring situations because they have fewer financial obligations to meet and can devote all their efforts to address the necessary changes, [29], Should be considered in any case external financing constraints given in micro and small enterprises, since they are paying less and a higher interest rate [19], but this restriction is reduced to credit when the company is having a greater antiquity [5].

Several researches has been done in Russian, Spanish, American and Canadian small firms, measuring the debt level, liquidity, and their influence on the profitability of the SME's, through statistics techniques like univariate and multivariate.
And the results of these researches shows an strong positive relation between low leverage and high liquidity with a better profitability, [17],[6],[16],[11],[23],[24],[30].

This information helps us to raise the benchmark hypothesis of this research because this goal has been raised to identify the influence of the financial management aspects such as liquidity and leverage and the internationalization of SMEs in the performance of these and following results of the research presented above, there is a strong positive relationship between accounting information and the correct determination of the ratios, analysis of financial information.

Financial aspects to be considered in the development or performance of SME's liquidity:
Liquidity is a term that demonstrates the ability to pay that companies have to meet its short-term and often is measured by working capital. [21], say that is the difference between current assets and current liabilities and is also defined as the portion of assets financed by long-term liabilities in short without liquidity is impossible for a company and especially SMEs can survive, given the low level of capital that for their operations . [34], considering the amount of liquidity to meet capital expenditures over a period of time that can be insured by the initial capital and cash generated from normal operations in a period of time without consider selling assets and obtaining external funding,

In the business management of liquidity, measured by the financial aspects of companies based on indices such as indexes of payments of clients delayed, that better explain the causes of financial problems such as high leverage and lack of liquidity, which have even led to the bankruptcy of companies and especially SME's by low operating capital, [4].

From this information we can conclude that it is important the culture of payments from employers for healthy business economy, setting priorities according to their ability to pay, in the execution of these, employees, suppliers, tax authorities, etc..

In a study of 607 SME's in retail trade in the United States, which analyzed 14 different key factors such as liquidity, income, debt and operating efficiency, reflecting the results, which have higher earnings, liquidity and they are less leveraged than previously thought, showing that support the operation itself is better than debt, to achieve higher profits, [9].

The above references tell us that if the operation is supported with the resources generated by the operation itself, it becomes a more sound policy that the use of debt, causing the rising cost of operation, making the business less profitable and decrease your chance of survival.

Leverage
Better known as debt leverage, defined by.[35], as the proportion of funds that have provided the commercial bank creditors and measured: Total Debt / Total assets and it's recommended that the ratio does not exceed 33%, since otherwise the business starts to be owned by others.
SME's participating in a very significant economic output and employment, but have a high mortality rate and was done a study in the metropolitan area of Cali Colombia from 2000 to 2004 on the main causes of premature death can be, the initial firm size, lack of income and debt, as well as external factors such as economic growth in the sector in which they work, or the geographic location of this, the study revealed that the initial size, its performance financial, economic sector activity, have a positive impact, but exist an inverse relationship between debt and survival of the business, which means that debt is the riskiest factor for the survival of SME's,[32].

Was performed a study of 99 SME's in Italy on the effect of the ratio of the number of banks which the Sme's operates and the business performance, and located that the return on assets decreases when the number of banks with which it operates is increased, also was located that the amount of interest expense increases when the number of banks for small businesses increases, mainly by the effect of indebtedness and concludes that the fewer number of banks have in operation, reducing the level of agency problems and eliminates the negative effects of operation that affect the performance of SMEs, by high debt, [7].

They conducted a small study in Brazil, with the aim of identifying how they affect funding sources internal or external, formal or informal business performance and the empirical study revealed that the domestic financing remains a relationship positive business growth and growth has a positive effect with formal financial institutions not with informal sources of financing,[31].

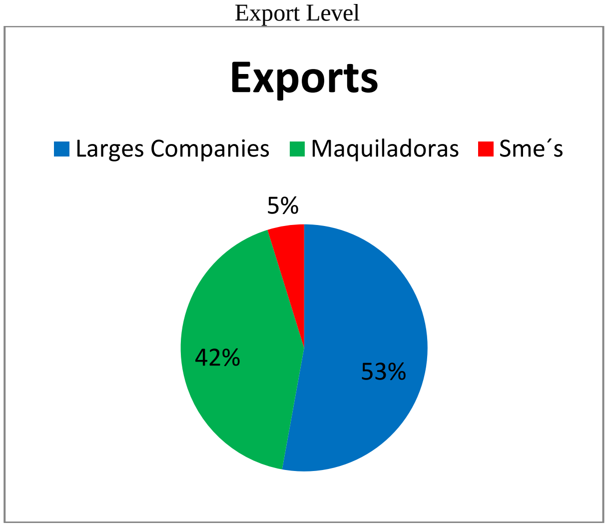
Based on the studies presented in the aforementioned items, you can identify that the indebtedness has been the cause of the high mortality of SME's and the informal credit management do not represent a security financing costs, making it very convenient to handle funding levels that allow the control of corporate governance without falling into excesses of leverage to cause the failure by the lack of ability to pay.

Based on the data mentioned before, you can comment on the importance of SMEs as a means of local development and is presented as an option for development of SMEs, the internationalization.

Internationalization of SMEs

To display an overview of the situation that is reality in the internationalization of SMEs to the national level, is presented the following graph showing the level of internationalization of SMEs in Mexico.

Graphic No.2
Export Level



Source: Information Paper on SMEs. Ministry of Economy 2002

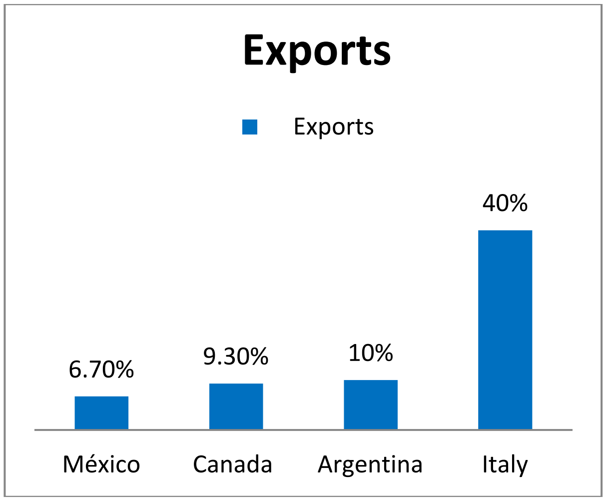
As can be seen in the graph only 5% of SMEs in our country achieve export, implying that there is a long gap to go to reach the internationalization of SMEs but this should be a challenge for our country because, according Empirical studies, its have shown which through exports, it is possible to achieve development and performance of SMEs.

[27]. conducted a study that examines factors associated with the adoption of strategies of internationalization and its effects on performance and compare it with non-exporting firms and concludes that internationalization is associated with increasing the return on sales, facilitating the development of small businesses.

[28],conducted an analysis of 300 SME exporters in Argentina, Chile and Colombia in the period 2001-2004, in order to identify success factors that have enabled these companies to venture into the field exports, and the results of this study indicate that factors, commercial, production, technology and environment, are what make it to affect export levels that based on these results, leads to establish the basic basket export products and to determine the potential of exporting SMEs.

The following table presents a comparison of exports of SMEs from different countries, eager to see the trend and location of SMEs in our country (México), compared to other countries

Graphic. No 3
Comparison of exporting SMEs



Source: Information Paper on SMEs. Ministry of Economy 2002.

As shown in the graph, of the compared countries, Mexico is the country with the lowest percentage of SME exporters and it´s caused due to the lack of support from governmental levels to ensure the development of SMEs outward providing resources and training to enable the development of our SMEs, du these it's perennial link between the authorities and SMEs to join efforts to improve this level of exports which is an alternative for the development and retention of Mexican SME´s.

Aguascalientes exporting companies.

The number of SMEs exporting in Aguascalientes, is very low but this represents a challenge to promote the SMEs in the State of Aguascalientes and in the whole country, developing the SME´s for reach to achieve the level of quality required for export operations through the economic support of the government authorities.

Performance

SME´s Performance Factors :Have identified different ways of measuring performance in SMEs that considers different factors that result in successful performance of SMEs raised by several researchers as follows.

[12,]reports that measured performance in Spain in terms of competitiveness and survival of SMEs, based on the success factors as innovation, value added, services, technology, quality and the operations systems are able to demonstrate that the degree of innovation in products, processes and management systems have a positive influence on performance.

[13], in his investigation "cost management for quality and performance of small and medium-sized enterprises" in the Murcia region of Spain considers as the dependent variable a set

of performance measures, which are: generated resources Operating income, economic profitability, operating margin and financial performance

[3], in their research entitled "Factors explaining the competitive success for SMEs in the state of Veracruz Mexico", in their analysis considers as dependent variables to analyze the success of SMEs, the following: The average number of earnings before interest and taxes, Quality of products or services, Introduction of innovations, Productivity of labor, Customer satisfaction with products or services, knowledge and business experience, employees motivation-satisfaction and reputation and image of the company, and found to be most relevant variables in the case of Veracruz; the financial capacity, technological position and innovation, marketing capabilities, management of human resources and technologies information and communication.

In this research, were selected as variables to measure performance, set by [3], it was considered appropriate, by taking into consideration aspects that are not purely quantitative, since they consider the qualitative performance such as customer satisfaction, employee satisfaction and reputation and image of the company.

4.HYPOTHESIS

The assumptions made in this investigation are:
Hypothesis1. With greater financial management of liquidity and leverage, increase performance of SMEs. As suggested by [9], in his research that showed that financial support in the operation itself is more profitable than the financial leverage and generate more liquidity in order to achieve higher profits

Hypothesis2. Higher level of internationalization of SMEs, improve performance and as such,[27].derived from their research; conclude that internationalization is associated with increased return on sales, encouraging the development of small businesses

5.METODOLOGY.

The research conducted is quantitative, using the method of linear regression analysis, ANOVA and frequency analysis, the research also is transectional, because the surveys are at a point in time of the year 2009 character exploratory correlational descriptive and explanatory, as they look for causal relationships between variables [33].
The design of the research is not experimental, and therefore none of both dependent and independent variables are manipulated

For this research were used quantitative tools to achieve test hypotheses on the influence of the administration of financial aspects and the internationalization of SMEs and their influence on the performance of these.

The collection of information was directly, through interviews with the owner or responsible of the SMEs, as most of the questions in the collection of information, are only manage by managers of the enterprise; surveys were implemented in the first half of 2009

Final survey instrument was designed based on the literature and indicators of both dependent and independent variables, to achieve the identification of the existence of the influence of the administration of financial aspects and the internationalization of SMEs in the SEM´s performance; based in the response of employers, according with their criteria and identifying the causes of performance of SME's, related to the industry, and according with the responses of the entrepreneurs, using the Likert scale to measure results from 1 to 5, with 1 being completely disagree and 5 totally agree, the questionnaire was applied to a sample of 125 SME's. from different business sectors of Aguascalientes.

6. RESEARCH FINDINGS.

For show the results of this investigation, we present the table No.2, that indicates the average of the different aspects of performance measured using a Likert scale, where 1 is the least important and 5 is the most important, which factors stand out as the greatest performance impact observed by employers, the customer satisfaction and quality of products and services.

Table No.2
Media Business Performance

DEPENDENT VARIABLE	MEDIA
1.- Quality of product / service	4.02
2.- Customer Satisfaction	4.02
3.- Image of the company and its products and services	3.96
4.- Reduced absenteeism	3.94
5.- Speed of adaptation to market needs	3.93
6.- Efficiency of internal business processes	3.89
7.- Organization of the staff Tares	3.89
8.- Reduced Staff Turnover	3.82
9.- Increased Productivity	3.74
10. Increase in market share	3.66
11. Motivation and Satisfaction of Workers	3.63
12.-Increase Profitability	3.54

Source: Authors' calculations based on the results of SPSS frequency analysis. On a scale of 1 = Not important and 5 = very important

Table No.3
Significance of Variables "Finance Management vs. Performance

Independent Variable	R	R Square	t	Sig.
Finance Management	.436	.190	21,741	.000***

Statistically significant differences: (*): p <0.1; (**): p <0.05; (***): p <0.01Dependent Variable: Performance.

The above table indicates the existence of a high positive relationship between the Finance Management and Performance of SMEs du to the R-squared is above the .160 marked by[8] , for 5 independent variables and a sample of 125 surveys and the level of significance is below .01, with a 95% of confidence level, so that it may explain the increased business performance of SMEs through the adoption of the Finance Management as a form of management.

Finance Administration was established by determining the average of the factors identified in the collection instrument, such as Finance Management. Decrease in leverage, increased liquidity and increased level of business operations .

To set the dependent variable analysis was determined by averaging the various elements set as dependent variables, such as: Quality of product / service Customer satisfaction, Image of the company and its products and services, reduction of absenteeism, quick adaptation to market needs, efficient internal business processes, the Organization of the Tares personnel, reduction of staff turnover, increased productivity, increased market share, Motivation and Satisfaction of Workers, Increase Profitability, in order to measure the impact of the effect of financial management with the average of the performance factors described above.

The above table indicates the existence of a high positive relationship between the Finance Management and Performance of SMEs as the R-squared of .160 is above that mark [8], for 5 independent variables a sample of 125 surveys and the significance level is below .01, 95% confidence level, so that may explain the increased business performance of SMEs through the adoption of financial management as a form of management.

Table No.4
Significance of ANOVA table of international markets and Performance

Variables Dependientes	F	Sig.
1.- Quality of product / service	1.243	.267
2.- Customer Satisfaction	4.496	.036**
3.- Image of the company and its products and services	2.293	.133
4.- Reduced absenteeism	1.868	.174
5.- Speed of adaptation to market needs	0 .817	.368
6.- Efficiency of internal business processes	6.537	.012**
7.- Organization of the staff Tares	4.484	.036**
8.- Reduced Staff Turnover	3.236	.074
9.- Increased Productivity	0.799	.373
10.-Increase in market share	1.607	.207
11. Motivation and Satisfaction of Workers	2.502	.116
12.-Increase Profitability	4.795	.030**

Statistically significant differences: (*): p <0.1; (**): p <0.05; (***): p <0.01
Predictive variable: Opening of New International Markets

Table No. 4 above shows that the internationalization affects the Performance, with greater influence on: efficiency in internal business processes, increase profitability, and organization of staff tasks and the satisfaction customers.

7.DISCUSSION AND CONCLUSION

This study shows through the application of statistical models and the collection of related information that the basis of financial reporting is an accounting well substantiated on a realistic and technical basis, in our study is demonstrated through the implementation of the univariate and multivariate statistical models, the high influence of the financial management on aspects as leverage and liquidity with the performance increased of SMEs that demonstrate the Hypothesis H1, "Finance Management and SME´s Performance", concluding that it is advisable as a business practice, relying on the resources generated by the operation and avoid excessive debt, which impacts the loss of liquidity due to payment of interest on the loans acquired and the Hypothesis H2 is demonstrated indicating that SME´s internationalization, improved their performance, such was demonstrated through statistical models; applied the adoption of internationalization as a business practice allows the development of quality in the manufacture of products for SMEs, given the high level of quality required by international companies, leading to the increase in business profitability and the possibility of permanent business, making it a viable alternative to improve the performance of SMEs by promoting their development, so that the objectives stated in the this research was reached since the hypothesis were demonstrated as mention before. Future research can delve into issues related to financial and fiscal factors which are related to the profitability increased of SMEs.

8.BIBLIOGRAPHY.

[1]Alarcón, S. (2008): Endeudamiento y eficiencia en las empresas agrarias, **Revista española de Financiación y Contabilidad,** vol. 37, No. 138, abril-junio, p. 211 – 230.

[2] Andrés-Suárez, J. 2000. Los parámetros característicos de las empresas manufactureras de alta rentabilidad. Una aplicación del análisis discriminante, ***Revista Española de Financiación y Contabilidad***, 104, 443–481.

[3] Aragón Sánchez Antonio y Alicia Rubio Bañón (2005) Investigación "Factores explicativos del éxito competitivo: el caso de las Pymes de estado de Veracruz publicación "Contaduría y administración" No. 216 mayo-agosto 2005.

[4] Back, Peter (2005), Explaining Financial Difficulties Based on Previous Payment Behavior, Management Background Variables and Financial Ratios**. European Accounting Review**; Dec2005, Vol. 14 Issue 4, p839-868, 30p, 10 Charts

[5] Brewer III, E. y Genay, H. (1996): "How are Small Firms Financed? Evidence from Small Business Investment Companies", *Economic Perspectives*, Nov/Dec, vol. 20 No. 6, p. 2-18.

[6] Bezlepina, I., y Oude Lansink, A. 2003. Liquidity and productivity in Russian agriculture: farm-data evidence, Contributed paper, ***25th International Conference of Agricultural Economics*** *August*, 16-22.

[7] Castelli, Annalisa, Dwyer, Jr., Gerald P, Hasan, Iftekhar, (2006) **Bank Relationships and Small Firms' Financial Performance Working Paper Series** (Federal Reserve Bank of Atlanta); May2006, Issue 5, preceding p1-26, 27p

[8] Cohen, J. and Cohen, P , (1983) **Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences** 2a.Edition. Hillsdale, NJ: Lawrence Erlbaum.

[9]Droms, William G. Miller, Charles W. Lacerte, Grant A(1979). A Financial Profile Of Small Retailing Firms. American Journal of Small Business; Apr79, Vol. 3 Issue 4, p42-58, 17p

[10]Fernández, A.I., y García Olalla M., (1991): «**Análisis del comportamiento económico-financiero de los sectores empresariales en España**», *Esic Market*, 72, abril-junio, pp. 113-128.

[11] Fernández, E.; Montes, J. M., y Vázquez, C. J. 1996. Caracterización económico-financiera de la gran empresa industrial española según su rentabilidad, ***Revista Española de Financiación y Contabilidad***, 87, 343-359.

[12] García Pérez de Lema Domingo , Dr. Salvador Marín Hernández, Mtro. Tirso Juan Ordaz Coral (2007), **Análisis estratégico para el desarrollo de la pequeña y la mediana empresa en el estado de Quintana Roo**, Universidad tecnológica de Cancún, Universidad de Cantabria, Universidad Politécnica de Cartagena, Universidad de Murcia Libro Edit. Universidad de Cancún

[13] García Perez de Lema y Miñarro Quiñonero (2006). **Publicación de Investigación. Gestión de los costes de calidad y rendimiento de las pequeñas y medianas empresas** *la universidad politécnica de Cartagena.*

[14] García Pérez de Lema, D. (2005): ***Estrategia e innovación de la Pyme industrial en España.*** Asociación Española de Contabilidad y Administración de Empresas,

[15] Garza Garcia J. (1995) **Análisis de la Información Mercadológica a través de la Estadística Multivariante** Edit. Alhambra Mexicana, México.

[16] González-Pérez, A.L.; Correa Rodríguez, A. y Acosta Molina, M. (2002): "Factores determinantes de la rentabilidad financiera de las pymes", ***Revista Española de Financiación y Contabilidad*****,** vol. 31, No. 112, abril-junio, p. 395-429.

[17] Hernando, I., y Vallés, J. 1994. **Algunas diferencias en la productividad de las empresas manufactureras españolas**, *Investigaciones Económicas*, 18, 117-141.

[18] Hirotaka Takeuchi & Michael Porter, (1986). **Three Roles of International Marketing in Global Strategy**, Harvard Business School Press, Boston, USA 1986

[19] Holmes, S. y Dunstan, K. (1994): "The Cost of Debt for Small Firms: Evidence from Australia", ***Journal of Small Business Management***, January, vol. 32, No. 1, p. 27 – 35.

[20] Lambin, Jean-Jacques, **Marketing estratégico. Tercera edición**. McGraw Hill / Interamericana de España, 2003

[21] Lawerence & J.Gitman (1996) **Administración Financiera Básica,** San Diego Univesity. Harla, México.

[22] Maroto. J. A., [1989]: ((Evolución de la rentabilidad y de sus factores explicativos (1982-1987)», ***Papeles de Economía Española***, n.' 39, pp. 376-396.

[23] Maroto Acín, J.A. (2001): "Modelos comparados de pyme en la Unión Europea. Los viejos problemas ante el nuevo entorno**",** ***Papeles de Economía Española***, No. 89/90, p. 18-42.

[24] Martín Rodríguez, M. y Sáez Fernández, F.J. (2001): "Políticas públicas de apoyo a la financiación de la pyme: fundamentos, instrumentos y resultados", ***Papeles de Economía Española***, No. 89/90, p. 167-186.

[25] Mata, J. y Portugal, P. (1994): "Life Duration of New Firms", ***Journal of Industrial Economics*****,** vol. 42, No. 3, septiembre, p. 227-246.

[26] Melle Hernández, M. (2001): "Características diferenciales de la financiación entre las pyme y las grandes empresas españolas. Asimetrías informativas, restricciones financieras y plazos de endeudamiento", ***Papeles de Economía Española*****,** No. 89/90, p. 140-166.

[27] McMahon, Richard G. P.Davies, Leslie G. ARTICLE ABSTRACTS (In English, French, German, and Spanish). **Journal of Small Business Management**; Jan94, Vol. 32 Issue 1, p1-8, 8p, 3 Charts

[28] Milesi, Dario,Moori, Virginia,Robert, Verónica,Yoguel, Gabriel (2007) Publication Desarrollo de ventajas competitivas: pymes exportadoras exitosas en Argentina, Chile y Colombia. **Revista de la CEPAL** Date: Wednesday, August 1 2007

[29] Morrison-Paul, C. J.; Johnston, W. E., y Frengley, G. A. G. 2000. Efficiency in New Zea-land sheep and beef farming: The impacts of regulatory reforms, ***Review of Economics and Statistics***, 82 (2), 325-337.

[30] Rege, Udayan P,(1984). ACCOUNTING RATIOS TO LOCATE TAKE-OVER TARGETS **Journal of Business Finance & Accounting**; Autumn84, Vol. 11 Issue 3, p301-311, 11p

[31] Saeed, Abubakr, (2009) FORMALITY OF FINANCIAL SOURCES AND FIRM GROWTH: EMPIRICAL EVIDENCE FROM BRAZILIAN SMES 1999-2005 **Journal of Academic Research in Economics**; Oct2009, Vol. 1 Issue 2, p129-140, 12p, 4 Charts

[32] Sandoval, Alberto Arias, Marín, Raül Fernando Quiroga (2008). CESE DE ACTIVIDADES DE LAS PYMES EN EL ÁREA METROPOLITANA DE CALI (2000-2004): UN ANÁLISIS DE SUPERVIVENCIA EMPRESARIAL. (Spanish) **Cuadernos de Administración** (01203592); 2008, Vol. 21 Issue 35, p249-277, 29p, 8 Charts, 7 Graphs

[33] Schmelkes Corina (1998) Manual para presentación de anteproyectos e informesDe Investigación Oxford University Press México.

[34] Skomp, Stephen E, Edwards, Donald E (1978). MEASURING SMALL BUSINESS LIQUIDITY: AN ALTERNATE TO CURRENT AND QUICK RATIOS. **Journal of Small Business Management**; Apr78, Vol. 16 Issue 2, p22-26, 5p, 1 Chart

[35] Weston, J.F. & Brigham, E.F (1987) **Las Finanzas en Administración Universidad Iberoamericana SA de CV México D.**F. 1984.

Citas Electrónicas.
http://www.cipi.gob.mx/html/..%5CPol_Apoyo_Pymes_Mex.PDF

Data Mining in a System of a Candidate Selection Process

Renata Maria Abrantes Baracho
School of Information Science
Federal University of Minas Gerais
Belo Horizonte, Brazil
renatabaracho@ufmg.br

Márcio Teodoro Dias
School of Information Science
Federal University of Minas Gerais
Belo Horizonte, Brazil
marcio.dias@ifmg.edu.br

Abstract **– Data mining and information retrieval has been research areas increasingly important in the information age where fast, accurate access to information becomes an essential factor in decision-making in various fields of knowledge. The main goal of this study is to analyze the data of candidates to selection processes of Federal Institute from Minas Gerais - IFMG in search of patterns that are not explicit, with the support of a free tool for data mining. This study comes to the research question of what are the relevant information about the profile of candidates who qualify in the selection process of IFMG. Through data mining and analysis of results is possible to define the profile of applicants and the features that influence the selection process. This article presents a study on the database in the candidates selection process of IFMG. One of the requirements of the research includes the study of data mining tools available on the Internet and a test applies in a real situation. Then the article presents an experience of using a free tool for Data Mining. The software chosen for the tests, called Weka, was developed by the University of Waikato, New Zealand. It makes searching for patterns in database of system used to the management of selection processes of the Federal Institute from Minas Gerais. The methodology includes access to the database IFMG, with 9952 records of candidates for the selection process. Five categories were defined as relevant for verification and analysis. Categories include sex, marital status, age, waiver of registration fee and have disabilities or not. We used the attributes that have 100 or more occurrences in the database. The results show that the pattern that occurred more frequently among those classified in absolute numbers was that of male candidates, who had no exemption from the registration fee, less than 30 years old and single, with 166 occurrences. Among all candidates, 718 have these characteristics, so 9.66% of the candidates that fit this pattern was classified. The pattern with the highest percentage of occurrence were male candidates, who had no exemption from the registration fee, less than 30 years old and married. Of the 301 candidates with all these features that participate in the processes of selection, 31 were classified. It represents 10.33% of 301 candidates evaluated. There were 9952 candidates of which 711 were classified, the representing 7.14%. Among the 291 candidates with exemption from the registration fee, only 3 were qualified which represents 1.03%. This result indicates that lower-income applicants had lower performance. The associative algorithm provides a list of up to 100 rules, which were presented in this article the top 15 with a combination of categories, sex, marital status, age, registration fee and deficiencies. The survey allows the discovery of patterns in database performance with different combinations of the categories.**

Information retrieval, information system, database, data mining, Weka, selection process

I. INTRODUCTION

Currently there is growing research on data mining and information retrieval to support the decision-making organizations. There is large databases that can generate direct subsidies for decisions. In this context, the research seeks to use data mining tools to analyze the database. The objective of this research is to analyze the data of applicants to a selection system to define the patterns the profile of classified candidates. This study addresses the research question of what are the relevant information about the backgrounds of the candidates who qualify in the selection process of Federal Institute from Minas Gerais - IFMG.

The Federal Institute from Minas Gerais - IFMG is a federal autarky, created by Law nº 11.892 [1], enacted on December 29, 2008, by Luiz Inácio Lula da Silva who was the President of Brazil. In the article 2 of the Law, Federal Institute are defined as: Institutions of higher, basic and professional education specialized in providing professional and technological education in different modalities of teaching, based on the combination of technical and technology knowledge with their teaching practices, under this Law.

IFMG is now comprised of ten units in different cities of the state of Minas Gerais: Bambui, Betim, Congonhas, Formiga, Governador Valadares, Ouro Branco, Ouro Preto, Ribeirão das Neves, São João Evangelista, Sabará. These units are connected to a central government in Belo Horizonte, state capital of Minas Gerais.

People who work in IFMG are admitted through the selection processes that are performed by the institution itself. This processes can be made to hire permanent or temporary employees, or to hire trainees.

Each Process has one or some requirements, defined in a public notice, to evaluate whether a candidate is able to occupy a post. The candidates who can fulfill these requirements, that are generally related to education and performance in a test, are classified. Among these ones, the best classified candidates will be called to fill the positions, according to the number of vacancies. Those candidates who do not fulfill the requirements specified in the notice are considered disqualified from the selection process. According to the number of places offered in the Notice, a list is generated for ratification of process. Candidates present in the list can still be called to fill the post.

Management of routines related the selection process in IFMG is done with the support of an internally created software, called Recepta. This system was developed with PHP programming language and PostgreSQL Data Base

Management System. Recepta was registered at the National Institute of Industrial Property from Brazil (INPI) and is also available as free software.

In database of the system, which was implemented in March 2010, are stored data relating to registration, candidates' personal data, and whether the candidate is or is not in the list of selection process.

The main objective of this work is to search the database of the selection process to find patterns among the personal characteristics of candidates and their final results, achieving this result with the use of a free tool for data mining.

II. Literature Review

Whereas the text mining involves moving information retrieval. In this research the concept of information retrieval is based on considerations of Lancaster [2] and Hjørland [3].

Lancaster [2] describes retrieval of information how the process of searching within in a collection of documents to identify those dealing with a particular subject.

Hjorland [3] define that one of the problems in retrieval is the definition of access points to a database of electronic documents containing text, images and different media.

Knowledge Discovery in Databases (KDD), is a concept used to describe the exploration of implicit information in large volumes of data, whose technology emerged by necessity and difficulty of exploring large databases (Bigolin, Bogorny, Alvares [4]).

The process KDD can also be defined as "the non-trivial process of identifying patterns that are valid, new, potentially useful and, finally, comprehensible in data" (Fayyad, Piatetsky–Shapiro, Smyth [5]).

Knowledge Discovery in Databases is the result of a process that goes through three main stages: preprocessing, data mining, and the post-processing (Bigolin, Bogorny, Alvares [4]).

In step of preprocessing is necessary to prepare the data so that a mining tool can extract from it the implicit and potentially useful information. This preparation involves the following tasks, in accordance with Bigolin, Bogorny, Alvares [4]:

- determination of the goals of discovery: the problem is clearly defined;
- data cleaning: elimination of noise and inconsistency from data;
- data integration: data from multiple sources can be combined;
- data selection: relevant data for data mining are identified and grouped together, generating a sample of the database;
- data transformation: conversion of data to a format interpretable by the data mining tools.

This stage can demand up to 80% of the total processing time, due to difficulties in integrating heterogeneous databases (Carvalho apud. Manilla [6]).

Data mining is "a step in KDD process that involves applying data analysis and discovery algorithms that, under acceptable limitations of computational efficiency produces a particular enumeration of patterns about the data" (Fayyad, Piatetsky–Shapiro, Smyth [5]).

For EDELSTEIN [7], "Data mining tools find patterns in the data and infer rules from them. Those patterns and rules can be used to guide decision-making and forecast the effect of those decisions".

Post-processing is the step of submission and evaluation of the patterns found in mining process, that are responsible for the identification and analysis of relevant patterns, as well as the definition of the form with the extracted information will be presented. The following figure shows the steps of the KDD process, Figure 1.

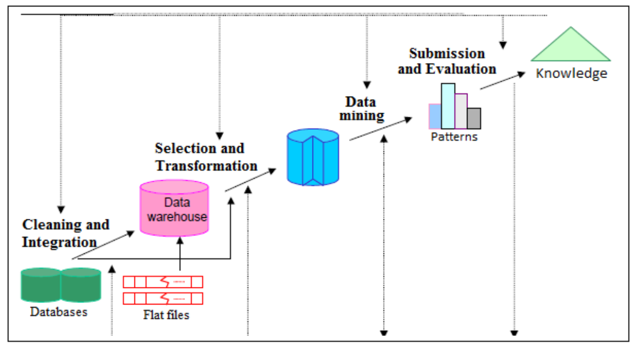


Figure 1. Steps of the KDD process.
Source: BIGOLIN, BOGORNY, ALVARES (2003)

III. Methodology

The methodology is developed as the goal of finding relevant information on the profile of candidate that can be classified in the selection processes of IFMG. For this work is not important whether the candidate was or was not called to occupy the position to which applied. It depends on the number of vacancies offered by the institution. The most important is whether it was classified.

In the second part was made a search on the Internet to find a free software of data mining to be used. A software to receive an input data and perform data mining for patterns not explicit. During the study found information about several tools such as KDB2000, KNIME, MDR, Orange, Tanagra and Weka. After analyzing the software Weka was chosen because it is widely used, many options for implementing data mining algorithms and be well documented. Their characteristics are described in detail in (Witten and Frank [8]), whose authors are responsible for the implementation of the tool.

The next step was to define what characteristics of the candidates would be considered to verify the association between these characteristics and the results of the candidates in selection processes. After analyzing the characteristics were defined five attributes: sex, marital status, age, the exemption from registration fees and disability. Low-income candidates may request exemption from registration fee in a selection process, being that this request may or may not be accepted.

The attribute in this case seeks to assess the performance of low-income applicants.

The next stage was to define how the data would be prepared for mining. It was decided to use attributes independent of each other and that had 100 or more records in database, because a small sample could lead to errors.

IV. Data Preparation

The data preparation began with cleaning the data, to extract only those who contribute to the aims of research.

During the development of this research had contests in progress but without result, because of this the data subsequent to October 2011 were not extracted.

As described, the processes of selection IFMG include: Tender for Admission, Simplified Process for Substitute Teacher Selection and Selective Processes of Trainees.

The Selection Process of Trainees has a different audience of others, so data for these candidates have different characteristics, especially in relation to age, marital status and income. Data from this type of process were ignored. For this research were considered candidates for Tender for Admission, Simplified Process Selection.

Applications are accepted when the applicant makes the payment of registration fee or is entitled to exemption from this fee. In this study we considered the applications accepted. The data of candidates that did not make tests does not interfere in final result.

In selecting the data are considered those with 100 or more occurrences. Then went to look for attributes that were less than 100 records which will be disregarded. Therefore, the data related to disability were not considered, only 74 candidates were declared to have any special needs. Evaluating the marital status, it was found that only 28 candidates were widowed, so they were grouped in the same category of divorced.

The last step in the data preparation was the elimination of any records that contained incorrect information. It was realized that several candidates make mistakes to fill the date of birth during registration, for example, some candidates whose year of birth was filled in with the value 2011. Were considered only those who have filled the year of birth as 1993 or less, whereas the minimum age to occupy public post in Brazil is 18 years.

Date of birth itself is not an interesting value in search of patterns, because of the large number of distinct occurrences. The candidates were grouped into three age groups according the date of birth: less than 30 years old; between 30 and 45 years old; and more than 45 years old.

After making data preparation including the definition of the five attributes: sex, marital status, age, the exemption from registration fees and disability, these data was exported from the database in CSV format. After the cleaning process, were obtained 9952 records.

V. Data Mining and Results

According to research, to data mining software was used Weka (Waikato Environment for Knowledge Analysis), version 3.6. The code was first developed in 1993 at the University of Weikato in New Zealand. It was built with the Java language, supports multiple file extensions as input, one of which is the CSV format.

In the dataset, one of the attributes is considered the class attribute, while the others are attributes predictive. In this research, the class attribute is what shows if the candidate is classified. This attribute is populated with the values yes or no. All other predictive attributes will be evaluated in relation to the class attribute.

In Weka software, the last attribute is considered by default as the class attribute, but to change this order is possible.

In the home screen of the program is the option to open a file with the data to be mined. In this work it was used a file in the CSV format, exported from Recepta system database. After loading the file, Weka's screen displays bar graphics that allow comparing each predictive attribute in relation to the class attribute, as shown in the Figure 2.

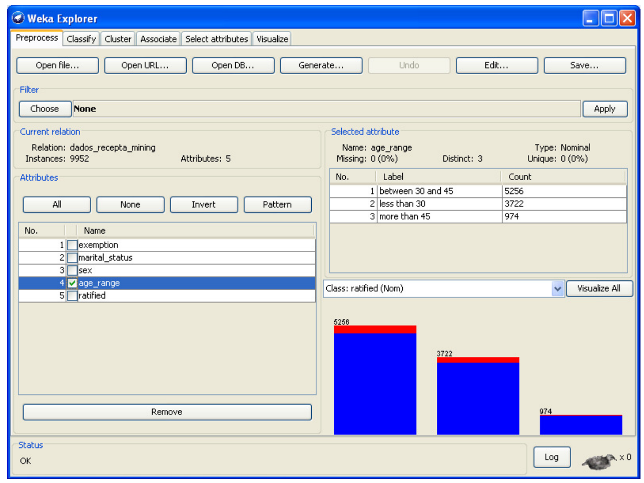


Figure 2. Comparison Graphic of the class attribute with a predictive attribute.

Next, several algorithms available in the software have been tested. Four of them brought results most relevant to the aims of research.

Inside of the classification algorithms, available in tab "Classify", there is a folder called Trees, which contains some algorithms that build decision trees. Below is the result brought by the algorithm "RandomTree", with highlights in areas considered most important for this work (Table 1).

Table 1 – Algorithm "RandomTree"

```
RandomTree
==========

exemption = no
|  sex = m
|  |  age_range = between 30 and 45
|  |  |  marital_status = married : no (898/83)
|  |  |  marital_status = single : no (808/73)
|  |  |  marital_status = divorced or widower : no (68/4)
```

```
| | | marital_status = other : no (55/1)
| | age_range = less than 30
| | | marital_status = married : no (300/31)
| | | marital_status = single : no (1718/166)
| | | marital_status = divorced or widower : no (3/1)
| | | marital_status = other : no (18/2)
| | age_range = more than 45
| | | marital_status = married : no (205/10)
| | | marital_status = single : no (55/2)
| | | marital_status = divorced or widower : no (31/2)
| | | marital_status = other : no (17/1)
| sex = f
| | age_range = between 30 and 45
| | | marital_status = married : no (1113/59)
| | | marital_status = single : no (985/49)
| | | marital_status = divorced or widower : no (206/13)
| | | marital_status = other : no (75/3)
| | age_range = less than 30
| | | marital_status = married : no (567/45)
| | | marital_status = single : no (2083/140)
| | | marital_status = divorced or widower : no (25/0)
| | | marital_status = other : no (35/3)
| | age_range = more than 45
| | | marital_status = married : no (183/8)
| | | marital_status = single : no (110/6)
| | | marital_status = divorced or widower : no (84/3)
| | | marital_status = other : no (19/2)
exemption = yes
| sex = m
| | age_range = between 30 and 45
| | | marital_status = married : no (11/1)
| | | marital_status = single : no (14/0)
| | | marital_status = divorced or widower : no (2/0)
| | | marital_status = other : no (1/0)
| | age_range = less than 30
| | | marital_status = married : no (9/1)
| | | marital_status = single : no (73/1)
| | | marital_status = divorced or widower : no (0/0)
| | | marital_status = other : no (1/0)
| | age_range = more than 45 : no (3/0)
| sex = f : no (177/0)

Size of the tree : 48

Correctly Classified Instances     9241        92.8557 %
Incorrectly Classified Instances    711         7.1443 %
```

The algorithm "RamdomTree" constructs a decision tree by filling out the class attribute with the value that appeared most frequently, informing in the end of each line how many instances are not in accordance with this model. For example, in the tenth row of result, the model predicts that if the candidate is male, has no exemption from the registration fee, has less than 30 years of age and is single, the trend is that isn't qualified, based on what happened in most situations However, the result shows how many times the test was performed (1718) and how often the class attribute was different from that described in the model (166). That is, the 1718 candidates with these characteristics, only 166 were classified. In this tree in all cases the class attribute will be considered as "no", because most candidates could not (the total percentage of qualified applicants was 7.14%). This Randon tree allows check which the numeric pattern that was repeated over. In 166 cases a male candidate, who had no exemption from the registration fee, has less than 30 years old and is single managed to qualify. Of the total 9,952 candidates analyzed, 9241 were not classified. Among 711 qualified candidates, 166 candidates has pattern 1 (male candidate, the exemption from the registration fee, has less than 30 years old, single). In numerical terms is the pattern 1 that appears most frequently. In percentage terms (10.33%) of the 300 candidates 31 were classified presenting the pattern 2 (male candidate, the exemption from the registration fee, has less than 30 years old, married), managed to qualify. That is, the pattern 2 (10.33%) in percentage terms represents a superior result to the pattern 1 (7.14%).

In the end, the algorithm shows the Correctly Classified Instances with 9241 events (92.8557%) and Incorrectly Classified Instances with 711 occurrences (7.1443%).The model defines the classifier attribute as no. The classifier attribute tends to be no. When the classifier attribute is yes the program describes how misclassification because it did not follow the trend indicated by the model. The percentage of Incorrectly Classified Instances refers to cases where the class attribute is yes. The misclassification shows how many times the model was not followed.

Another classification algorithm was tested "AdaBoostM1", which is in the folder "Meta". The first rows of the result are transcribed below (Table 2).

Table 2 – Algorithm "AdaBoostM1"

```
AdaBoostM1: Base classifiers and their weights:

Decision Stump

Classifications

sex = m : no
sex != m : no
sex is missing : no

Class distributions

sex = m
no                      yes
0.9116550116550116      0.08834498834498834
sex != m
no                      yes
0.9415400918403392      0.0584599081596609
sex is missing
no                      yes
0.9286575562700965      0.07134244372990353
```

This algorithm compared each predictive attribute with the class attribute, informing the percentage of occurrence of each value of class attribute. In our test, it presents information quite interesting in sex attribute. The last line of the example presents the total, without considering the sex of the candidate. As stated previously, in general, about 7.14% of the candidates were classified in selection processes. Approximately 8.83% of candidates classified are men and 5.84% are women.

The last classification algorithm used was the "NaiveBayesSimple", located in folder "Bayes". Below is described the result of the test (Table 3).

Table 3 – Algorithm "NaiveBayesSimple"

```
Naive Bayes (simple)

Class no: P(C) = 0.92857143

Attribute exemption
no              yes
0.96873648      0.03126352

Attribute marital_status
married       single        divorced or   other
                            widower
0.33290071    0.59928618    0.04391088    0.02390223

Attribute sex
m               f
0.42319342      0.57680658

Attribute age_range
between 30 and 45    less than 30    more than 45
0.43374797           0.49302326      0.07322877

Class yes: P(C) = 0.07142857

Attribute exemption
no              yes
0.99438202      0.00561798

Attribute marital_status
married       single        divorced or   other
                            widower
0.33473389    0.61344538    0.03361345    0.01820728

Attribute sex
m               f
0.53370787      0.46629213

Attribute age_range
between 30 and 45    less than 30    more than 45
0.40252454           0.5483871       0.04908836
```

This algorithm first separated the class attribute, and then compared it with each attribute predictive. The class attribute refers to the exemption from the registration fee. The table shows those candidates who failed the exemption from the registration fee (0.92857143) and those who got the exemption (0.07142857). This result provides an interesting analysis, when comparing the percentages in each attribute. According to Table 3, the age range has attribute values near between 30 and 45 (0.43374797) and less than 30 (0.49302326) representing a majority for the "Class no". The "Class yes" also presents values close between 30 and 45 (0.40252454) and less than 30 (0.5483871) The marital status attribute to the "Class no" has values close to married (0.33290071) and single (0.59928618) and the "Class yes" features were similar between married (0.33473389) and single (0.61344538). The results show that age and marital status did not significantly affect the probability of classification, since the values "yes" and "no" are very close. However, in the attributes sex and exemption from the registration fee this difference is large. There is 3.12% of candidates with exemption from the registration fee among those who were disqualified and 0.56% among those who were qualified. In other words, candidates with lower income had performance far below the overall average. Among the 291 candidates with exemption from the registration fee, only 3 were qualified which represents 1.03%.

In relation to associative algorithms, the most relevant among the tested was "PreditiveApriori", which lists a maximum of 100 found best rules. Below are listed the 15 main rules (Table 4).

Table 4 – Algorithm "PreditiveApriori" - 15 main rules

```
Best rules found:

1. age_range=between 30 and 45 ratified=yes 286 ==> exemption=no 285    acc:(0.99491)
2. marital_status=single sex=f ratified=yes 195 ==> exemption=no 195    acc:(0.99488)
3. sex=f age_range=less than 30 ratified=yes 188 ==> exemption=no 188    acc:(0.99486)
4. exemption=yes sex=f 177 ==> ratified=no 177 acc:(0.99483)
5. marital_status=single sex=m ratified=yes 242 ==> exemption=no 241    acc:(0.99477)
6. exemption=yes marital_status=single 218 ==> ratified=no 217    acc:(0.99463)
7. marital_status=married sex=f ratified=yes 112 ==> exemption=no 112    acc:(0.99428)
8. marital_status=married sex=f age_range=more than 45 184 ==> exemption=no 183    acc:(0.99424)
9. marital_status=single age_range=more than 45 166 ==> exemption=no 165    acc:(0.99389)
10. marital_status=married ratified=yes 238 ==> exemption=no 236    acc:(0.99378)
11. sex=m age_range=more than 45 ratified=no 296 ==> exemption=no 293    acc:(0.99318)
12. exemption=yes 291 ==> ratified=no 288 acc:(0.993)
13. marital_status=married sex=m age_range=more than 45 207 ==> exemption=no 205 acc:(0.99274)
14. sex=m age_range=less than 30 ratified=yes 202 ==> exemption=no 200    acc:(0.99251)
15. marital_status=other age_range=more than 45 36 ==> exemption=no 36    acc:(0.98762)
    exemption=no sex=f 81    acc:(0.7181)
```

In the results above are described associations that appeared most often, with the percentage corresponding to that situation. Attribute of exemption from the registration fee appears several times, since the vast majority of candidates have not obtained this exemption. The rule number 4, that combined (sex=female and exemption=yes) the two attributes that had lower performance. In the analyzed data, were 177 female candidates who had exemption from registration fee, i.e., has low income, and none of them is in the list of qualified. The rule number 6 shows us that low-income single candidates (marital_status=single and exemption=yes) had very low performance, because of 218 candidates in this situation, 217 were disqualified.

The software also lets you select which attributes will be considered in the test. In the example below were taken into account only the attributes sex and marital status, in association algorithm “PreditiveApriori” (Table 5).

Table 5 – Algorithm “PreditiveApriori” - attributes sex and marital status

Best rules found:

1. marital_status=other 232 ==> sex=f 140 acc:(0.59134)
2. marital_status=divorced or widower 428 ==> sex=f 323 acc:(0.45237)
3. sex=m 4290 ==> marital_status=single 2668 acc:(0.42982)
4. marital_status=other 232 ==> sex=m 92 acc:(0.41665)
5. sex=f 5662 ==> marital_status=single 3309 acc:(0.39666)
6. marital_status=married 3315 ==> sex=f 1890 acc:(0.37453)
7. marital_status=single 5977 ==> sex=f 3309 acc:(0.34185)
8. marital_status=single 5977 ==> sex=m 2668 acc:(0.31202)
9. marital_status=married 3315 ==> sex=m 1425 acc:(0.28479)
10. sex=f 5662 ==> marital_status=married 1890 acc:(0.2438)
11. sex=m 4290 ==> marital_status=married 1425 acc:(0.24006)
12. marital_status=divorced or widower 428 ==> sex=m 105 acc:(0.14226)
13. sex=f 5662 ==> marital_status=divorced or widower 323 acc:(0.02419)
14. sex=m 4290 ==> marital_status=divorced or widower 105 acc:(0.01379)
15. sex=f 5662 ==> marital_status=other 140 acc:(0.01376)
16. sex=m 4290 ==> marital_status=other 92 acc:(0.01373)

Assessing the Rule 2, we see that among those divorced or widowed women who participated in the selection process, the number of women exceeds the number of men, because they are 323 in total of 428, or 75.47%, this particular attribute .

Table 5 allows better understand the candidates profile who participated in the selection process.

Rule 3 shows that of 4,290 men surveyed, 2,668 are single which represent 62.19% of the total. Rule 5 shows that among the 5,662 women, 3,309 are single, thus 58.44%. Most candidates are single, this index shows that the number of single men more than single women. Rules 10 and 11 show the number of married men and women. Married men are 1225 (33.21%) and married women are 1890 (33.38%). That is, the percentage of married people is almost the same considering men and women.

VI. CONCLUSION

The use of the software allowed the discovery of patterns in the database, noting that the performance was different according to selection attributes. Considering the sex attribute, it was revealed that the men fared better than women. The poor candidates had a result less than the other candidates. Another interesting pattern was discovered in the relationship between marital status and sex. The evaluation results showed that most people widowed or divorced who participated in the selection processes are female. In relation to the married people, the percentage of men and women is almost the same.The software Weka has many mining algorithms, which may be relevant to other databases. The usability of the tool is good, and is compatible with a widely used file standard, the csv standard. As a suggestion for a future work, evaluate the relationship between other factors and classification, as the relationship between physical disabilities and classification or between physical disabilities and income.

REFERENCES

[1] Brasil. Lei Nº 11.892, DE 29 DE DEZEMBRO DE 2008. Disponível em:<http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2008/Lei/L11892.htm>. Acesso em 26/11/2011.

[2] F. LANCASTER, A. WARNER, “Information retrieval today”. Arlington: Information Resources Press, 1993.

[3] B. HJORLAND, “The concept of subject in information science. Journal of Documentation“, [S. l.], 19928. Lancaster, F.W., Information retrieval today. 47p. (1993).

[4] BIGOLIN, N. M. ; BOGORNY, V. ; ALVARES, L. O. "Uma Linguagem de Consulta para Mineração de Dados em Banco de Dados Geográficos Orientado a Objetos". In: XXIX Conferencia Latinoamericana, 2003, La Paz. XXIX Conferencia Latinoamericana, 2003. v. 1. p. 23-35.

[5] FAYYAD, U.M.,PIATETSKY–SHAPIRO, G, SMYTH, P. "Knowledge Discovery and Data Mining: Towards a Unifying Framework". Proceeding of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, Oregon, august, 1996.

[6] CARVALHO, D. R.; LOPES, L. R. ; BUENO, M. ; ALVES NETO, W. "Ferramenta de Pré e Pós-processamento para Data Mining". In: XII Seminário de Computação, 2003, Blumenau. XII Seminário de Computação, 2003.

[7] EDELSTEIN, H. " Technology how to: Mining Data Warehouses". Information Week, 8 Jan., 1996.

http://www.informationweek.com/561/61oldat.htm

[8] WITTEN, I. H.; FRANK, E. "Data Mining: Practical Machine Learning Tools and Techniques", 2nd edition, San Francisco, Morgan Kaufmann Publishers, 2005.

Cloud Computing: Better Ways to Control its Services

Khalid N. Alhayyan

alhayyan@usf.edu
Information Systems and Decision Sciences
College of Business
University of South Florida

ABSTRACT

In this study, a research approach is proposed for providing better ways of controlling the development of different levels of services in cloud computing environment. The approach is based on multi-paradigm conduct through implementing the hermeneutic circle: from theory to data (deduction), and from data to theory (induction), guided by the Control Theory [5] and the application of an interpretive case study as described by Benbasat et al. (1987), [2]. Two significant findings are anticipated from this research: theoretical-based guidelines for developing standardized cloud computing SaaS services and generalized standards that offer unified methodology of interfacing SaaS services to different SaaS-provider platforms.

Keywords: Cloud Computing, SaaS Services, Control Theory, Firm-based case study, Standards

1. INTRODUCTION AND MOTIVATION

Armbrust et al. (2009), [1] define cloud computing as the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services. It is motivated by the construction of large scale datacenters, new technology trends such as scalability and portability, new business models, and new application opportunities such as mobile interactive applications [1]. Cloud computing has attracted good deal of both of providers who look for more business benefits and users who seek services that are offered in low cost. However, users are still concerned about how reliable these services are with respect to its availability, transferability, and confidentiality. There is a little research that offers better ways for how to control developing the cloud computing services. Therefore, the central research question of this study is:

How do SaaS (Software as a Service) providers in cloud computing context better control developing different levels of services in an efficient manner that meet stakeholder requirements?

This research question generally addresses two of three main obstacles identified by Armbrust et al (2009), [1]: cloud computing adoption obstacles (e.g. availability and data lock-in), and cloud computing growth (e.g. data transfer bottleneck). Availability refers to the continuity of SaaS services that are provided to users. Data-lock in refers to the obstacle that may prevent users from moving their data and programs from one provider to another due to the lack of standardized interface. Last, data transfer bottleneck refers to a high-cost situation that may face users when they choose to place or move data from a datacenter of

Level of virtualization / Application model	**High-level virtualization**	**Intermediate-level virtualization**	**Low-level virtualization**
Computational model			
Storage model			
Communication model			

Table 1. Resources visualization based on applications models and levels of abstractions

a provider to another. Cost can add up quickly, making the cost of data transfer a significant issue.

The nature of application models and the levels of managed services offer various level s of sophistication that may seem difficult to develop, manage and control, particularly to meet the stakeholders' needs. That is, any cloud application may need three levels of models: computation model, storage model, and communication model. On the other hand, SaaS services differ in the level of abstraction or virtualization that will be presented to the users/programmers. There are three levels of virtualized services: high-level virtualization, intermediate-level virtualization, and low-level virtualization. Table 1 shows a grid of different combinations between the application models and the levels of managed services.

As we go across any one application model from left to right, we expect to see different levels of abstraction with respect to what is offered to users. For instance, for high-level virtualization, a provider should offer the following features: less control to users, great automatic scaling, and high availability (e.g. Google AppEngine). Low-level virtualization offers: more control to users, less automatic scalability and failover (e.g. Amazon EC2). Finally intermediate-level virtualization offers general-purpose computing that is good for both users and programmers (e.g. Microsoft Azure).

There are two aspects through which this research shows an importance. The first aspect is that businesses (SaaS providers) must maintain significant level of control over their products, thus managing the quality and progress of the resulting software outcomes. Given the dynamic nature of cloud computing services (e.g. various business requirements, various users, and high sophistication of offered services), SaaS providers must ensure they have explicit and formal control in a development environment that is distributed and unstable. The second aspect is related to the needs of potential SaaS users: Availability, data lock in, and data transfer bottleneck. Unless these needs are met, it is difficult to see a surge in cloud computing adoption. Thus, SaaS providers should consider these users' needs to be built in the introduced SaaS services.

2. RESEARCH APPROACH

A research line to address this proposed study is through implementing the hermeneutic circle: from theory to data (deduction), and from data to theory (induction). This research conduct is expected to increase our understanding for finding richer insights from both approaches (deduction and induction) that are complementary and more informative. An approach to conduct this research comes into two main steps. First, we follow the approach of Emergent Outcome Control by Harris et al., 2009, [3], that is based on the Control Theory by Ouchi (1977), [5], to represent an outline for how to conduct this study. This theory can be viewed as shown in Figure1. Second, we follow the above iterative outline to conduct an interpretive case study. Benbasat et al. (1987), [2], specify, in general, the steps to follow for conducting a case study: (1) Decide the unit of analysis (SaaS service). (2) Decide multiple-case design for generalizability purposes (e.g. 10 firms). (3) Decide site selection for replication and confirmation (similarity of results). (4) Deicide data collection methods (documentation, interviews, archival records,

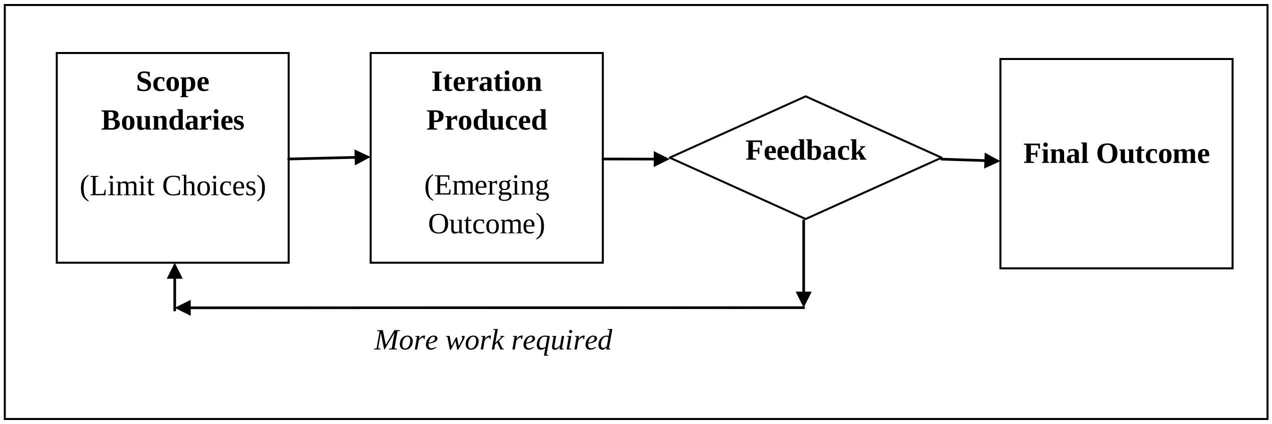


Figure 1. Emergent outcome control, adapted from Harris et al. (2009)

direct observation, and physical artifacts). (5) Perform the data analysis and exposition (finding significant patterns).

The industry collaboration will be employed to conduct this study. It would probably be from small and medium sized firms, as large firms seem to own their own datacenters and less likely to participate. Following the iterative outline in Figure 1, we first determine the scope boundaries of SaaS services of interest. For example, there are concerns originating from potential users about the difficulties of extracting their data from one cloud to another, which may prevent some organizations from adopting cloud computing. Therefore, the service boundary becomes building suitable and standardized software interface that mitigates the data lock-in problem. Second, there will be an iteration process across participating firms in order to build a good understanding about the set of requirements and virtualization levels that should be accounted for when building SaaS services. Feedback on current outcomes from the various firms will guide the process of extending or shrinking the previous boundaries. Next, set of general software protocols are determined as newer version of the previous outcome that addresses the obstacles of data lock-in and confidentiality (e.g. Ajax is a good example as a front-end for cloud computing). These protocols will be evaluated iteratively according to predetermined rules for confirmation purposes until the final outcomes reach to reasonable consensus of agreement on their applicability and suitability.

3. ANTICIPATED CONTRIBUTION

The overall conduct of the study is behavioral as we are aiming to address better ways of controlling SaaS development process. However, an anticipated finding of this multi-paradigm study is a generalized set of software standards that are necessary for successful implementation of SaaS services across many organizations. As articulated in Hevner et al (2004), [4], an artifact is to solve relevant problems and create innovation. Thus, we argue that developing generalized outline of SaaS standards represents an artifact to be added to the knowledge base. Additionally, we anticipate two general findings from this study. First, theoretical-based guidelines for SaaS services development will be proposed as better ways for efficient control and maintainability. Second, set of generalized standards will be proposed to impose unified methodology of interfacing SaaS services. These two findings are the potential theoretical contribution of this study. When this study is complete, we anticipate offering better ways for creating good software quality, hence mitigating the issue of less cloud computing adoption and growth. Armbrust et al (2009), [1], argue that among the most important software quality of SaaS service are the high availability, preventing data lock-in, and confidentiality. We argue that the outcomes of this theoretical-based research will meet these software quality requirements.

4. REFERENCES

[1] Armbrust, M., A. Fox , R. Griffith, A. Joseph , R. Katz ,A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia (1977). "**Above the Clouds: A Berkeley View of Cloud Computing,**" Technical Report No. UCB/EECS-2009-28.

[2] Benbasat, I., Goldstein, D.K. and Mead, M. (1987). "**The Case Research Strategy in Studies of Information Systems,**" MIS Quarterly (11:3), pp. 369-386.

[3] Harris, M., A. Hevner, R. Collins (2009). "**Controls in Flexible Software Development,**" Communications of the Association for Information Systems. Vol. 24, Article 43, pp. 757-776.

[4] Hevner, Alan R.; March, Salvatore T.; Park, Jinsoo; and Ram, Sudha. (2004). "**Design Science in Information Systems Research,**" MIS Quarterly, (28: 1).

[5] Ouchi, W. G. (1977). "**The Relationship Between Organizational Structure and Organizational Control,**" Administrative Science Quarterly 22(1), p. 95-113.

Sequential Decision Making Predictions under the Influence of Observational Learning

Shing H. DOONG
Department of Information Management, ShuTe University
YanChao District, KaoHsiung City, Taiwan 82445

ABSTRACT

Today's corporate managers face challenges in information technology (IT) adoption with great stakes. The waiting period for an IT investment to be realized, if any, could be long; thus word-of-mouth information propagation may not help them to make wise decisions. Though information used by early adopters to make their decisions may not be available to the public, late adopters can often observe the decisions made by the early adopters and infer hidden information to supplement their own private information. Observational learning theory applies when a person uses observed behavior from others to infer something about the usefulness of the observed behavior. Walden and Browne proposed a simulation procedure to model the influence of observational learning in sequential decision makings. Previously, we proposed a dynamic Bayesian network (DBN) to model sequential decision makings under the influence of observational learning. In the present study, we show how to infer a DBN model from simulated data. Hidden Markov model and artificial neural networks were used to infer the DBN model. Their performance will be discussed.

Keywords: Sequential Decision Making, Observational Learning, Dynamic Bayesian Network, Hidden Markov Model, Artificial Neural Networks

1. INTRODUCTION

Today's corporate managers face challenges in information technology (IT) adoption with great stakes. IT together with telecommunication has been considered the main driver for the economic growth of many countries in the new economy era since 2000s. To many companies, IT has become an indispensable part of their core competence with several characteristics. First, IT is becoming so powerful and complex that a fair assessment of its merits is difficult. Second, capital investments in IT are substantial, yet returns on investments often take time to materialize. Brooks has shown that software and other technological components are complex artifacts ever built by human beings [1]. In some cases, impacts of new technologies may take years to be realized [2]. Owing to these reasons, corporate managers need different kinds of tools and practices to help them make wise decisions in IT adoptions.

When people make decisions with limited or asymmetric information, they use different practices to correct this information deficiency. Observational learning occurs "when one person observes the behavior of another person and infers something about the usefulness of the behavior based on that observation" [3]. Research shows that, due to information asymmetry, people use what they observe from others to update their own private information or belief about a decision making [4]. Observational learning often leads to an interesting phenomenon called informational cascades [5].

An informational cascade occurs "if an individual's action does not depend on his private information signal" [5]. Walden and Browne [3] developed a theoretical extension of the observational learning model in [5], where a binary private signal is generated for each decision maker who chooses to adopt or reject an action. In [3], a continuous private signal is issued to each individual who also chooses to adopt or reject an action. Changing the private information signal from binary to continuous has produced many interesting results. For example, unlike the easy informational cascading in the case of binary signals, Walden and Browne showed that there are always late decision reversals in a sequence of decision makings. That is, informational cascades do not occur in the case of continuous signals.

A simulation procedure was used to investigate the extended observational learning theory [3]. In a later study, we showed that the Walden and Browne (WB) model can also be investigated from the perspective of a dynamic Bayesian network (DBN) [6]. In the present study, we consider the problem of inferring the DBN from simulated data. Hidden Markov model (HMM) and artificial neural networks (ANN) are used to infer the DBN model.

This paper is organized as follows. We briefly discuss the WB model and our DBN perspective first. Then, HMM and ANN are introduced to learn the DBN, given simulated data. Experimental results will be presented next, followed by discussions and conclusions.

2. MATERIALS AND METHODS

Observational learning with continuous signals

Walden and Browne used a continuous signal to denote the private information received by an individual [3]. A sequence of individuals will make a decision of choosing technology A (e.g., adopt the cloud computing IT) or technology B (e.g., reject the cloud computing IT). Assume that technologies A and B emit signals from the normal distributions $N(\mu_A, \sigma^2)$ and $N(\mu_B, \sigma^2)$ respectively, and $\mu_A > \mu_B$. An individual chooses A if the following condition is satisfied:

$$\frac{p(s \mid \mu_A)}{p(s \mid \mu_B)} \geq \beta \qquad (1)$$

Here s is a private signal received by the individual, $p(s \mid \mu_A)$ and $p(s \mid \mu_B)$ are the probability distribution functions (pdfs) of $N(\mu_A,\sigma^2)$ and $N(\mu_B,\sigma^2)$, respectively. Plugging in pdfs to solve for s, we obtain the following decision rule:

$$\begin{cases} s \ge r(\beta) & \Rightarrow \quad A \\ s < r(\beta) & \Rightarrow \quad B \end{cases} \tag{2}$$

$$r(\beta) = \frac{2\ln(\beta)\sigma^2 - \mu_A^2 + \mu_B^2}{2(\mu_A - \mu_B)} \tag{3}$$

Using signal detection theory [7], Walden and Browne set the decision threshold β as follows:

$$\beta = \frac{\Pr(\mu_B)}{\Pr(\mu_A)}k \tag{4}$$

Here k is common to all individuals and involves the relative benefit of B to A. For the first individual, we can assume technologies A and B are equally good, thus $\Pr(\mu_B) = \Pr(\mu_A)$. For the remaining decision makers, these terms are posterior probabilities after observing previous decisions:

$$\begin{cases} \Pr(\mu_A)_{t+1} = \Pr(\mu_A \mid D_t, ...D_1) \\ \Pr(\mu_B)_{t+1} = \Pr(\mu_B \mid D_t, ...D_1) \end{cases} \tag{5}$$

In the above equation, D_t denotes the decision made by the t-th individual. Using the chain rule of conditional probability, Walden and Browne deduce the following rule for updating decision thresholds:

$$\beta_{t+1} = \left(\frac{\Pr(D_t \mid \mu_B, A_t)}{\Pr(D_t \mid \mu_A, A_t)}\right)\beta_t, \tag{6}$$

where $D_t = a$ or b when the t-th individual chooses technology A or B, and A_t consists of decisions made by all t-1 previous individuals. Decisions in A_t together determine the signal break-point $r(\beta_t)$ for the t-th individual via the threshold β_t and Eq. (3). Consequently, we have the following identities when A is the true technology emitting private signals.

$$\Pr(a \mid \mu_A, A_t) = \int_{r(\beta_t)}^{\infty} p(s \mid \mu_A)ds$$

$$\Pr(b \mid \mu_A, A_t) = \int_{-\infty}^{r(\beta_t)} p(s \mid \mu_A)ds$$

Similar identities can be derived when B is the true technology emitting private signals. If the t-th individual chooses technology A (i.e., $D_t = a$), β_t will be scaled down to form β_{t+1} because $\Pr(a \mid \mu_B, A_t) < \Pr(a \mid \mu_A, A_t)$. Thus, the break-point $r(\beta)$ for signal s is moved leftwards. There is more space for the next individual to choose technology A. On the other hand, if this individual chooses B, then β is scaled up, $r(\beta)$ moves rightwards and there is less space for the next individual to choose technology A.

Dynamic Bayesian Network

In order to model the WB model as a DBN, we need to use two sequences of random variables to describe the dynamic involved in a sequential decision model [6]. The variable X_t represents a decision threshold β_t, thus $\Pr(X_t = \beta_t) = 1$, and the Y_t variable represents the outcome of a decision. Thus, assuming that signals are drawn from technology A, then $\Pr(Y_t \mid X_t)$ is given by:

$$\begin{cases} \Pr(a \mid \beta_t) = \int_{r(\beta_t)}^{\infty} p(s \mid \mu_A)ds, \\ \Pr(b \mid \beta_t) = \int_{-\infty}^{r(\beta_t)} p(s \mid \mu_A)ds \end{cases} \tag{7}$$

In order to describe the fact that β_t scales up or down depending on the decision of the t-th individual, a causal link from Y_t to X_{t+1} must be established as follows:

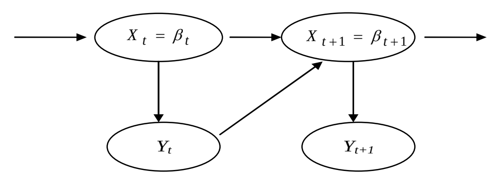


Figure 1. A DBN for observational learning

The X_{t+1} variable depends on X_t and Y_t by rules in Table 1. This random variable has a discrete distribution and its value depends on the previous hidden state (X_t) and the previous decision (Y_t). The WB model is now converted to a DBN in Figure 1 with relevant pdfs given in Eq. (7) and Table 1.

Table 1. Conditional probability $\Pr(X_{t+1}|X_t, Y_t)$

$\Pr(X_{t+1} \mid X_t, Y_t)$	(β_t, a)	(β_t, b)
$\beta_t \dfrac{\Pr(a \mid \mu_B, A_t)}{\Pr(a \mid \mu_A, A_t)}$	1	0
$\beta_t \dfrac{\Pr(b \mid \mu_B, A_t)}{\Pr(b \mid \mu_A, A_t)}$	0	1

The DBN perspective has a few advantages over the original WB model. First, the signal receiving and decision making step has been simplified to a binomial sampling step. Second, the dynamic updating of decision thresholds is replaced by clear rules in Table 1. The DBN in Figure 1 is also easy to interpret. It tells causal relationships among all relevant variables and shows how the system evolves as time moves ahead.

Hidden Markov Models

A hidden Markov model (HMM) is represented by a 5-tuple (S, V, A, B, P) where $S = \{s_1, s_2, \dots, s_N\}$ consists of N states that are not directly observable, $V = \{v_1, v_2, .., v_M\}$ denotes M observable outcomes emitted by a state, $A = ((a_{ij})) = \Pr(s_j \mid s_i)$ represents the transition probabilities between states, $B = ((b_{jk})) = \Pr(v_k \mid s_j))$ represents the emitting probabilities of outcomes by states, and $P = ((p_i))$ represents the initial probabilities of states [8].

Given an HMM with all needed components, a sequence of outcomes can be generated by (1) choosing an initial state according to the initial probability vector P; (2) emitting an outcome from this state by using the emitting probability matrix

V; (3) transiting to a next state by following the transition matrix A; and (4) repeating steps (2) and (3). This is a data deduction procedure commonly used in simulation studies.

On the other hand, given sequences of outcomes (data), an HMM may be learned from the data and used for future predictions. This is a patterns induction procedure used in most data mining algorithms.

Artificial Neural Networks

Artificial neural network (ANN) has been successfully applied to solve many function approximation problems in engineering and social sciences. An ANN simulates the neural system of a brain to learn patterns from examples and uses the learned knowledge to make predictions for new data [9]. A basic data processing unit in a neural net is called a neuron which is connected to other neurons via synapses. The structure of an ANN refers to the number of neurons and the way they are distributed and connected.

To simplify the computation, neurons are scattered into layers and information is transferred from layer to layer. The input layer represents the independent variables in a function approximation problem. The output layer corresponds to the dependent variable(s). Layers between the input and output layers are called hidden layers. An ANN with hidden layers is also called a multilayer perceptron (MLP). Without a hidden layer, a simple perceptron has limited learning capability [10]. It has been shown that an MLP can approximate arbitrarily well any continuous decision region provided that there are enough layers and neurons [11]. Learning an ANN from data is to find optimal synaptic weights to fit training data with known input-output pairs. Attention must be paid to the network structure so that we do not overfit the model with data. A trained ANN can be used to predict output values for new input values.

3. A SIMULATION STUDY

To examine informational cascades of sequential decision makings under the influence of observational learning, both [3] and [6] presented a simulation study.

Assume that two alternative technologies A and B are to be selected by a sequence of individuals. Suppose A is the better technology, thus all private signals will be emitted by its pdf, which is assumed be a normal distribution $N(\mu_A, \sigma^2)$. We assume that $\mu_A = 1$, $\mu_B = 0$, and $k = 1, \sigma = 1$ in the previous section of observational learning with continuous signals.

A simulation run of decision makings consists of 100 sequential decisions as explained previously. For the WB approach, this includes (1) drawing a signal from the pdf of technology A; (2) making a decision based on the signal, Eq. (2) and Eq. (3); (3) updating the new threshold according to the decision made and Eq. (6); and (4) continuing the process until the 100th decision is made. On the other hand, for the DBN approach, this includes (1) drawing a sample from the uniform distribution on (0, 1) to decide technology A or B according to Eq. (7); (2) updating conditional probability $\Pr(X_{t+1}|X_t, Y_t)$ in Table 1; and (3) continuing the process until the 100th decision is made.

Because the simulation is based on probabilistic samplings, one run of simulation can differ from another run of simulation substantially. Thus, a total of 1000 runs of simulation are conducted to smooth out fluctuations between runs. At the end, the average correct decision rate for each decision position (from 1 to 100) is reported. The average correct decision rate at a position is the number of correct decisions (i.e., choosing A) at that position out of total runs divided by 1000.

Figure 2 shows that both approaches yield very similar cures of average correct decision rate. Both approaches have an average correct rate curve that starts low at around .70 and increases to around .95 at the later stage. The correlation value between these two sequences of average correct decision rates is .994 and the mean absolute error (MAE) is .013. Other simulation types including random updating of decision thresholds and cases of tertiary decisions can be considered with the DBN approach.

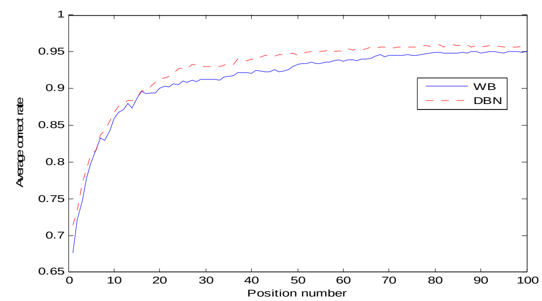


Figure 2. Comparison of WB and DBN approaches.

4. LEARING PATTERNS OF SEQUENTIAL DECISIONS

The last section presents a simulation study based on WB and DBN models. We now consider the reverse process of discovering models from data. Since real sequential decisions are hard, if not impossible, to obtain in business, we use simulated data from the DBN approach to learn patterns of sequential decisions under the influence of observational learning.

Training samples

The DBN approach is used to generate training samples for learning patterns of sequential decisions. In total, 1000 observation sequences are outputted from the simulations. Each observation sequence consists of 100 sequential decisions of 1 (for choosing A) or 0 (for choosing B).

Using HMM as a learning tool

HMM is a special DBN when it is spread out in steps. In order to use HMM, we need to decide the number of hidden states and the number of observable outcomes. Since there are only two possible decisions (1 or 0), we choose 2 hidden states and 2 emitted outcomes. The 1000 observation sequences of training samples are fed into a Baum-Welch (also called a forward-backward) learning algorithm to learn parameters of an HMM [8]. These parameters include the initial probability for each state, outcomes emitting probabilities and states transition probabilities.

The trained HMM is used to generate 1000 sequences of simulated outcomes. The simulation is obvious and straightforward, given the full parameters of an HMM. Each sequence consists of 100 sequential decisions. The average correct decision rate is computed as before and compared with that from the DBN approach.

Using ANN as a learning tool

In order to use ANN as a learning tool, we need to set up an input-output correspondence, i.e., input variables and output variables. Using the DBN perspective (Figure 1) as a guideline, we can set up a correspondence as

$$\beta_{t+1} = f(\beta_t, D_t)$$

Since β_t determines the distribution of D_t according to Eq. (7), we will use the probability of choosing A as the surrogate variable. Let p_t denote the probability of choosing A at the t-th position. p_t can be determined from Eq. (7). Then, we will approximate the following function with ANN.

$$p_{t+1} = f(p_t, D_t) \qquad (8)$$

To prepare training samples for the ANN learning, we use the average correct decision rate from the DBN simulations to denote p_t, i.e. p_t = (number of decision A at position t)/1000. The D_t variable is extracted from the 1000 observation sequences of the DBN simulations. Instead of using the full set of 1000 observation sequences to train a single ANN model, we train 10 ANN models with smaller data sets and average outputs from these 10 ANN models to make predictions. More specifically, we randomly choose 100 observation sequences from the DBN simulation to train an ANN model. This procedure is repeated 10 times to get 10 ANN models, which are bagged to get the final predictor. The idea is similar to a bagging predictor [12].

Since our model in Eq. (8) has only two inputs and one output, we do not need to use a complicated network structure. One or two hidden layers will suffice for our data set. Though our data set may be potentially large, e.g., 100 observation sequences with 100 sequential decisions will produce 9900 input-output pairs, of which many are simply duplicates. After using a trial-and-error approach with test data, we decided to use a two hidden layer structure – the first hidden layer has 4 neurons and the second hidden layer has 2 neurons. Our final MLP has 2, 4, 2, 1 neurons in the respective layer. The Sigmod function was chosen to be the activation function.

After the bagging aggregator is trained, it is used to predict the probability p_t in a simulation of sequential decisions. The first decision is simulated by using the average correct rate at position 1 from the DBN simulation. A random sample is drawn from the uniform distribution on (0, 1) and compared with this average correct rate to choose technology A or B. After the decision is made, it is plugged into Eq. (8) with the learned bagging ANN predictor to predict the next probability of choosing A. This process continues until the 100^{th} decision is made. Again, 1000 runs of simulation are conducted to calculate the average correct decision rate from the learned ANN model.

5. EXPERIMENTAL RESULTS

In this section, we present the experimental results from different simulation scenarios.

The standard case

In the standard case, we assume k=1. Thus, the relative benefit of choosing A or B is equal to one. The previous simulation study has shown that the average correct decision rate increases from around .70 at position 1 to around .95 at position 100. Figure 3 shows the average correct decision rate curves from DBN, HMM and ANN. The DBN simulation was used to generate training samples for the other two to learn. Both HMM and ANN learn their model from the training samples, and use the learned model to simulate sequential decisions. The average correct decision rate curve reports the simulation results using the trained model.

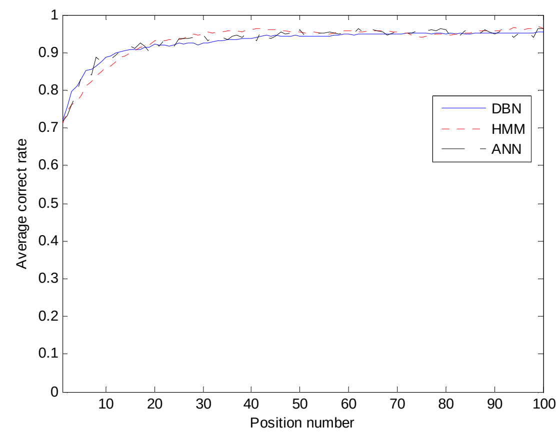


Figure 3. Comparison with DBN, HMM and ANN (k=1)

The MAE between DBN and HMM is .013 and the same measure for DBN and ANN is .008. On the other hand, the correlation between DBN sequence and HMM sequence is .969, and the same measure for DBN and ANN is .982. Thus, ANN has learned a better prediction model for this standard case.

Technology B has a higher relative benefit

In this case, we assume k = 10, thus technology B has a higher relative benefit than technology A. This gives individuals less incentives to choose technology A.

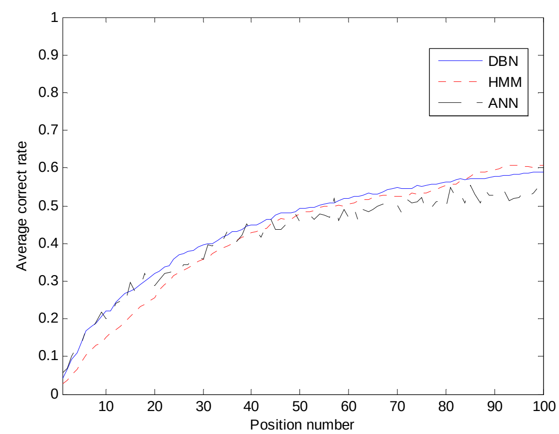


Figure 4. Comparison with DBN, HMM and ANN (k=10)

Figure 4 shows that the rate of choosing technology A is substantially smaller than that in the standard case. This is reasonable; because of a higher relative benefit for choosing B over A, an individual must have received a very strong signal in order to make a decision of choosing A. The probability of choosing technology A is small at the beginning. When more individuals select technology A, later individuals increase their belief in technology A through observational learning.

The simulations show that the average correct decision rate increases from less than .10 at position 1 to around .60 at position 100. The MAE between DBN and HMM and ANN is .028 and .031 respectively. The correlation between DBN sequence and HMM sequence is .995, while the same measure for DBN and ANN is .992. Therefore, HMM is a better prediction model in this case.

Technology B has a higher relative benefit and only partial sequences are used

In this case, we assume that technology B has a higher relative benefit (k = 5), and only partial sequences from the DBN simulations are used to train HMM and ANN models. We assume that only the first 50 decisions in an observation sequence are used to train prediction models.

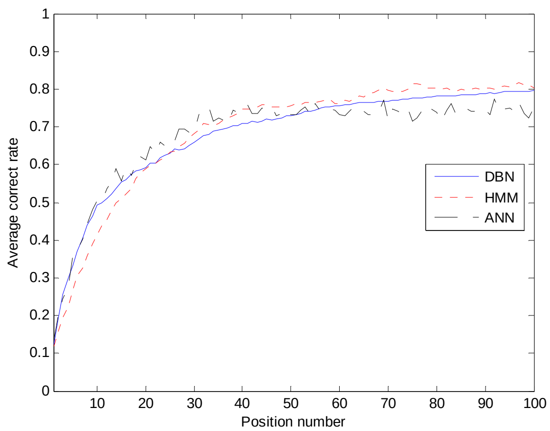


Figure 5. Comparison of models (*k*=5, only 50 decisions used)

Since the relative benefit of B to A is not as big as the one in the previous case, we expect individuals to have more incentives to choose technology A. Figure 5 verifies this with an initial average correct decision rate of around .15 to the last rate of around .80 at position 100. Since we only use the first 50 decisions to train HMM and ANN, their performance for the second half of decision sequences is more interesting. Figure 5 shows that the HMM model performs better than the ANN model for this part of decision sequences. Overall, the HMM model also produces a smaller MAE (.026 vs. .030) and a higher correlation (.993 vs. .970) than the ANN model.

6. DISCUSSIONS

The experimental results show that HMM has a better capability than ANN in learning patterns of sequential decisions. When k is big, the average correct decision rate curve resulting from the ANN model is much jagged than that from the HMM model. This is interesting if we consider the fact that the HMM model has no causal links between D_t and β_{t+1}.

The DBN model in Figure 1 is our basis to construct the ANN model in Eq. (8). That is, the current probability of choosing A and the current decision outcome should together decide the next probability of choosing A. On the other hand, an HMM has causal links between hidden states only. Using transition probabilities, the next state is sampled based on the current state only. Outcomes from the current state have no effect on the sampling of the next state in HMM. This seems to contradict the causal model explained by the DBN perspective of observational learning.

The jaggedness of the ANN average correct decision rate curve may come from an over-fitted neural network. Because we have a small network structure with a bundle of data, though many of them are duplicates, we may over-fit the network to produce a sensitive predictor. The bagging procedure does not seem to overcome this difficulty. Other prediction algorithms such as support vector regressions with known capabilities in over-fitting control may be considered in the future.

7. CONCLUSIONS

Today's corporate managers face challenges in IT adoption with great stakes. Corporate IT has become so powerful and complex that a fair assessment of its merits is difficult. Capital investments in IT are substantial, yet returns on investments often take time to materialize. Conventional word-of-mouth information propagation procedures may work for consumer IT decisions, but not for corporate IT decisions.

Though it is usually difficult to obtain the private information that other companies use to make their IT adoption decisions, it is possible to observe what the other companies have decided in their IT adoption. Observational learning theory applies when a person uses observed behavior from others to infer something about the usefulness of the observed behavior. Corporate managers may practice observational learning to help them make better IT adoption decisions.

Observational learning is known to create informational cascades, a phenomenon when an individual's action does not depend on his private information signal. When informational cascades occur, belief inferred from observational learning has overshadowed the private information signal that an individual uses to make his decision.

Walden and Browne [3] proposed a simulation model to show that informational cascades do not occur when the private information signal is continuous. We presented a DBN perspective of the WB model in [6]. The DBN approach demonstrated similar simulation results as the WB approach..

This study is focused on learning the DBN model resulting from observational leaning impacted sequential decisions. Two machine learning tools are used to learn the DBN. The first one, hidden Markov model, is itself a special case of DBN. The second one, artificial neural network, is a popular learning algorithm in artificial intelligence.

The HMM learning approach does not consider impacts of the current decision (D_t) on the sampling of the next state. It also uses a limited number of hidden states to represent continuous information signals. On the other hand, the ANN learning approach uses the DBN perspective to model a functional form for approximation. Its continuous output variable meets the type of private information signals in Walden and Browne [3].

The experimental results show that HMM has a better learning capability than ANN in our study. In the future, we plan to run more tests with different learning algorithms and diverse training samples. Learning patterns of sequential decisions constitutes the reverse process of simulation studies as presented in [3, 6]. Together, simulation studies and patterns learning can help us better understand how observational learning impacts sequential decisions.

Acknowledgements: This research has been supported in part by a grant from the National Science Council of Taiwan under the contract number NSC99-2410-H-366-006-MY2.

8. REFERENCES

[1] E. Brooks, **The Mythical Man-month: Essays on Software Engineering**, Reading, MA: Addison-Wesley, 1975.

[2] E. Brynjolfsson, and L. Hitt, L., "Paradox lost? Firm-level Evidence on the Returns to Information Systems", **Management Science**, Vol. 42, No. 2, 1996, pp. 541-558.

[3] E.A. Walden, and G.J. Browne, "Sequential Adoption Theory: a Theory for Understanding Herding Behavior in Early Adoption of Novel Technologies", **Journal of the Association for Information Systems**, Vol. 10, No. 1, 2010, pp. 31-62.

[4] A. Bandura, **Social learning theory**, New York: General Learning Press, 1977.

[5] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", **The Journal of Political Economy**, Vol. 100, No. 5, 1992, pp. 992-1026.

[6] S. Doong, and S. Ho, "Construct a Sequential Decision Model: a Dynamic Bayesian Network Perspective", **Proceedings of the 44th Annual Hawaii International Conference on System Sciences (HICSS)**, 2011.

[7] D. Green, and J. Swets, **Signal Detection Theory and Psychophysics**, New York: Wiley, 1966.

[8] L.R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", **Proceedings of the IEEE**, Vol. 77, No. 2, 1989, pp. 257-286.

[9] I.H. Witten, and E. Frank, **Data Mining, Practical Machine Learning Tools and Techniques**, San Francisco: Morgan Kauffman Publishers, 2005.

[10] M.L. Minsky, and S.A. Papert, **Perceptrons**, Cambridge, MA: MIT Press, 1969.

[11] A. Gallant, and H. White, "On Learning Derivatives of an Unknown Mapping with Multilayer Feed-forward Networks", **Neural Networks**, Vol. 5, No. 1, 1992, pp. 129-138.

[12] L. Breiman, "Bagging Predictors", **Machine Learning**, Vol. 24, No. 2, 1996, pp. 123-140.

GA + Segmentation for Large-Scale Asset Renewal Optimization

Tarek HEGAZY and Roozbeh RASHEDI
Department of Civil and Environmental Engineering, University of Waterloo
Waterloo, N2L3G1, Canada

ABSTRACT

Civil Infrastructure assets require continuous renewal actions to sustain their operability and safety. Allocating limited renewal funds amongst numerous building components, however, represents a large-scale optimization problem and earlier efforts utilized genetic algorithms (GAs) to optimize medium size problems yet exhibit steep performance degradation as problem size increases. In this research, after experimenting with various approaches of segmenting a large problem into multiple smaller sub-problems, clustered segmentation proved to be the most promising. The paper discusses the underlying life cycle analysis model, the various segmentation methods, and the optimization results using the improved GAs + clustered segmentation, which proved to be able to optimize asset renewals for 20,000 components with no noticeable performance degradation. The proposed method is simple and logical, and can be used on variety of asset types to improve infrastructure fund allocation. Future extension of this research is then highlighted.

Keywords: Capital Renewal, Asset Management, Computer Applications, Genetic Algorithms, Life Cycle Cost Analysis, Network-Level Decisions.

1. INTRODUCTION

Capital renewal is an essential asset management decision that is important to sustaining the serviceability of civil infrastructure assets (roads, bridges, water/sewer networks, educational buildings, etc.). In general, capital renewal involves two levels of decisions [1]: project-level decisions of the appropriate rehabilitation strategies to use for individual asset components (roof, windows, foundations, etc.); and network-level decisions of which asset components to repair in each year of the plan. Each level of decision, by itself, is a complex combinatorial problem that involves trying different combinations of actions to determine the best decision. For example, at the project level, consider the case of one bridge with 5 main components (deck, substructure, superstructure, joints, and finishing) with four repair alternatives for each component. One decision can be (1, 2, 0, 3, 1) which represents the indices to the repair types for the five components, respectively. As such, the number of possible decisions (solutions) is $4^5 = 1{,}024$. Expanding to the network level, on the other hand, consider a network of only 10 bridges that need to be repaired within a five-year planning period. The network-level decision involves deciding for each bridge i its year of repair (0, 1, 2, 3, 4, or 5). As such, for the ten-bridge network, the possible number of bridge selection combinations is 6^{10} = 60,466,176, which is extremely large. Network level decisions, therefore, are much larger in size than project level decisions. Due to large problem size, finding optimum decisions becomes a very difficult task that existing systems have not been able to adequately address.

In the literature, many researchers have proposed models for optimizing rehabilitation (renewal) decisions (e.g., [2]; [3]; [4]). These efforts provided interesting approaches to model life cycle cost analysis, however, none has proved to be able to handle very large-scale problems. Many commercial asset management systems also exist (Synergen, CityWorks, RIVA, etc.) but generally lack optimization capabilities and mostly use a simple ranking approach to prioritize assets for rehabilitation purposes [5]. Among the recent efforts that integrated both levels within an optimization framework is the Multiple Optimization and Segmentation Technique (MOST) developed by Hegazy and Elhakeem (2011). MOST (discussed briefly in the next section) handles a large-scale problem by first optimizing project-level sub-problems and using their results to formulate a network-level optimization (Fig. 1). Using this approach, MOST utilized the Genetic Algorithms (GAs) technique to handle network-level problems for up to 8000 components, simultaneously (one building has about 150 components). This paper builds upon the MOST technique and improves its performance for real-life problems to suit the organizations that own a large number of building assets (e.g., school boards, industrial facilities, etc.).

2. MULTIPE OPTIMIZATION AND SEGMENTAION TECHNIQUE (MOST)

MOST was introduced as an integrated life cycle cost analysis model to support asset renewal for school buildings administrated by Toronto District School Board (TDSB) in Canada. MOST uses a divide and conquer approach to segment the large optimization into a series of smaller optimizations, starting at the project level. Each individual optimization considers one building component and one possible repair year; and determines the best repair method and cost, for that component in that year. Within each small optimization, the formulation considers the component condition, deterioration behavior, and expected after-repair condition to determine the repair with the highest benefit-to-cost ratio. The results of all the individual optimizations (suboptimal at the project level) then are passed to a single network-level optimization.

The resultant of all project-level optimizations is a pool of best repair scenarios and their corresponding costs and benefits. These are used as the input to a network-level optimization to decide on repair timing. The objective of network-level optimization is to minimize the overall network deterioration index (DI_N) while not exceeding the available repair budget. Rather than a one-shot optimization over the 5-year planning horizon, MOST uses a year-by-year optimization formulation, from the first year consecutively until the end of planning horizon (as indicated in Fig. 1). Using this formulation reduces the solution-space size and leads to better solution quality. In general, the overall parameters in the network-level optimization (variables, objective function, and constraints) are as follows:

$$\text{Decision Variables:} \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} & Y_{14} & Y_{15} \\ Y_{21} & Y_{22} & Y_{23} & Y_{24} & Y_{25} \\ . & . & . & . & . \\ . & . & Y_{jk} & . & . \\ . & . & . & . & . \\ Y_{j1} & Y_{j2} & Y_{j3} & Y_{j4} & Y_{j5} \end{bmatrix} \quad (1)$$

where, Y_{jk} = 0 (no repair), Y_{jk} = 1 means component J is decided to be repaired in year k.

Objective function: minimize the network deterioration index (DI_N)

$$DI_N = \frac{\sum_j (Average\ DI_{jk} \times RIF_j) + \sum_j (IE_{jk} \times Y_{jk} \times RIF_j)}{\sum_j RIF_j} \quad (2)$$

$$\forall j \in Network \quad , \quad \forall k \in Planning\ Horizon$$

where, RIF_j is the relative importance factor of component j; DI_{jk} is the deterioration index of component j in year k; and IE_{jk} is the improvement effect of repairing component j in year k, which is equal to:

$$IE_{jk} = EP_{jk} - EP_{j0} \quad (3)$$

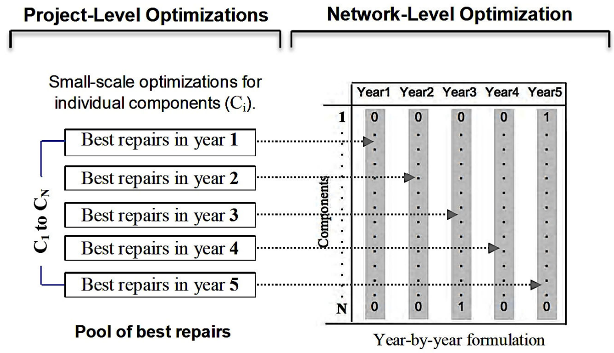


FIGURE 1 MOST

where EP_{jk} is the expected performance of instance j when repaired in year k and EP_{j0} is the initial performance of the component without repairs [6].

Constraint: Total repair cost for selected components in year $k \leq$ budget limit in year k.

Using these optimization parameters, and the results of the project-level optimizations, a life cycle cost analysis (LCCA) model was developed and implemented in an Excel spreadsheet, as shown in Fig. 2. Part (a) of Fig. 2 shows a partial list of asset components, where each row is a component. The highlighted component (fire alarm system), for example, has a relative importance factor (defined internally by experts at the TDSB) of 90. The current performance (deterioration) before repairs is also 72.49. The following columns then represent the cost and performance associated with repairs in year 1, 2, to 5 (results of the project-level optimizations). For example, if the component is repaired in year 2 (as highlighted), its deterioration will be reduced from 72.49 to 24.31, at a cost of $42,350. Part (b) of Fig. 2 uses the data in part (a) to formulate the LCCA and the optimization parameters. As an example, the decision to repair the fire alarm system in year 2 is selected by assigning a value of 1 to the decision variable of year 2 (as shown). Accordingly, the LCCA model reads values for RIF (90); expected performance after repair in year 2 (24.31); and repair cost ($42.350). The combination of component decisions determines the overall network deterioration index, using Eq. (1), Eq. (2) and Eq. (3). To handle the network-level optimization, MOST uses genetic algorithm (GA), which has been widely used by many researchers in various domains to solve combinatorial problems [7; 8; 9]. Experimenting on several network-level optimizations with different numbers of components, it was noticed that the performance of GAs steeply degrades as the problem size increases. At 8,000 components, it was noticed that GAs becomes incapable of reaching solutions that are better than simple ranking methods. To consider larger models GA + Segmentation approaches is proposed to improve the performance of MOST technique.

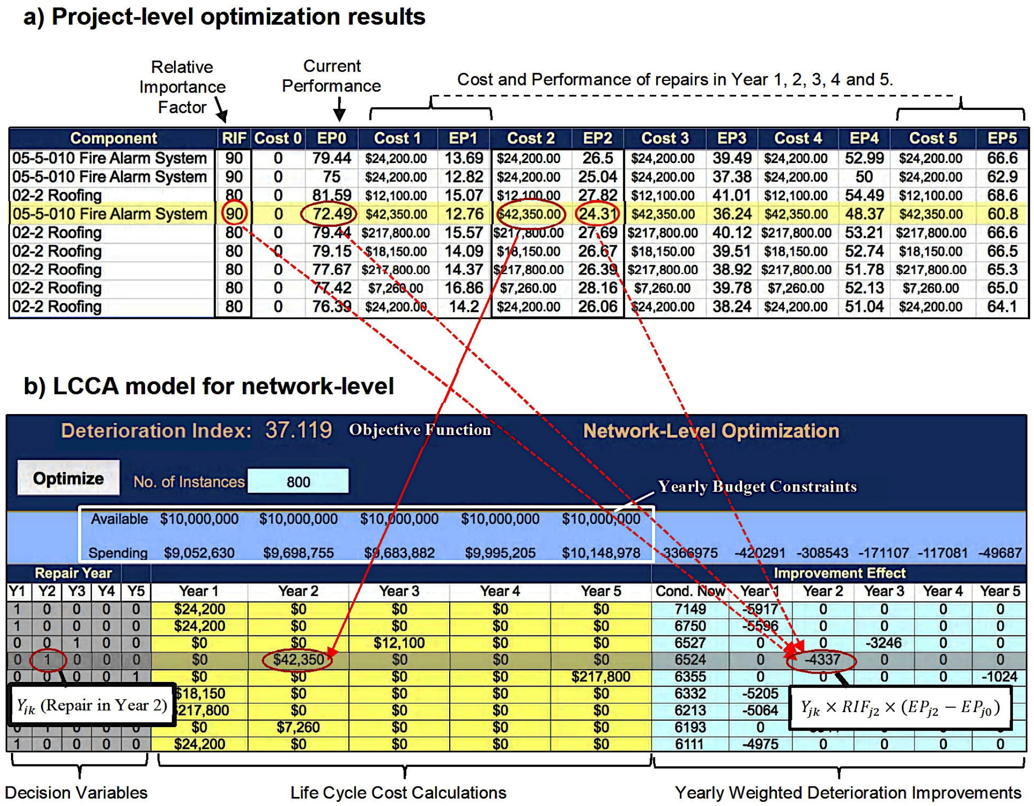


FIGURE 2 Network-level life cycle cost analysis model

3. PROPOSED GA + SEGMENTATION

This study proposes a segmentation approach at the network level to improve the solution quality and to handle very large-scale problems (as indicated in Fig. 1). While the original MOST technique utilizes segmentation at the project level by solving multiple small-scale optimizations, the proposed "GA + Segmentation" approach is applied at the large-scale network-level optimization. The main idea is to decompose a large problem into sub-problems that are optimized separately and their results are then merged to assemble the final solution. The main idea is to decompose the large network-level optimization model into several sub-problems with reduced solution-space. Next, the generated sub-problems that are easier to handle by GA, are optimized separately and their results are merged to assemble the final solution. Various segmentation methods have been investigated and tested on different size problems (Fig. 3). Initial experiments involved two methods: Random Segmentation; and Data Compression. The clustered segmentation approach, which is developed based on previous approaches, uses different mechanisms for generating segments, allocating budget limit to each segment, and redistributing leftover moneys in the best manner. The description of the various methods and their results are discussed in the following.

Random Segmentation

Random segmentation is a simple procedure for generating segments, which can be used as a starting point to evaluate the effectiveness of using segmentation with GAs. In this approach, the large list of components is divided into segments with equal number of randomly selected components (Fig. 3). To distribute the total yearly budget among the segments, the budget for each segment is calculated by dividing the total yearly budget by the number of segments. For example, if four segments are created, each will be allocated 25% of the total yearly budget. After optimizing all segments individually, results are combined to give the solution to the original model. The results of several experiments show improved solution quality when the number of components is less than 6,000; however, it still suffers from performance degradation. Accordingly, similarity-based segmentation methods (i.e., data compression and clustering) are introduced and investigated in the next sections.

Segmentation Using Data Compression

Data compression segments the large list of components based on similarities between components. Components of the same system (e.g., mechanical, electrical, architectural, etc.) and very close initial conditions (DI_0) are grouped in a separate segment. Next, to compress the

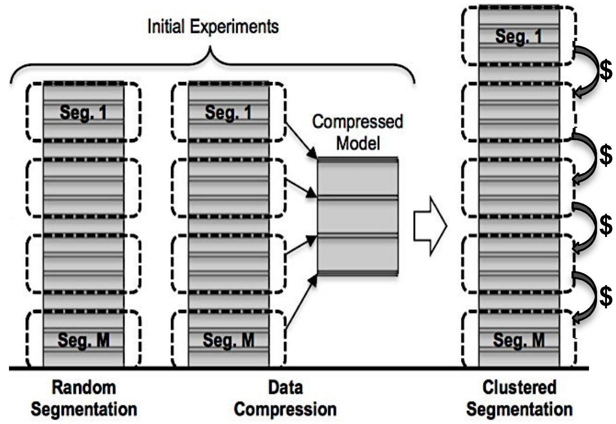


FIGURE 3 Segmentation approaches

large-scale data, each consistent segment is replaced by a single representative component (Fig. 3), which has the combined characteristic of all of the internal components. The compressed model, which contains only the representative components, is then optimized instead of the original large-scale model and solutions obtained for each representative component is reassigned to all the components in the segment that it represents. A model with 800 components from TDSB asset inventory is used for primary experimentation. The initial model consists of 541 architectural, 210 mechanical, and 49 electrical components. After grouping components based on their building system, then the components with similar RIF and very close initial condition (i.e., 1% variation) are fitted into segments.

Following this procedure, 29, 22, and 21 segments are generated, containing architectural, mechanical, and electrical components, respectively. One representative component of each segment is then generated from its individual components, with the following characteristics:

- Expected performance in each year = average of EPs for the individual components;
- Relative importance factor = average of RIFs of all individual components; and
- Repair cost = sum of repair costs of those for the individual components.

With the representative components, the initial model with 800 components is compressed to a model with only 72 representative ones. The compressed model has the same objective function as the base model and also the same budget limits. After optimizing the compressed model and reassigning the representative solution to corresponding components, a network deterioration index of 41.04 was obtained. Although, data compression results are better than simple ranking ($DI_N = 44.89$), comparing to solutions obtained by using GA with random segmentation, data compression is found to have lower quality solution. The poor optimality can be attributed to the approximation used to determine the representative components.

Segmentation Using Clustering

Clustering is a similarity-based method that considers all the components, without compression. It involves four steps that contribute to improving the efficiency of very large-scale network-level optimization: (1) Similarity analysis of input data: used to generate segments with similar internal data and to assign specific characteristics to segments to allocated budget accordingly; (2) Determining optimum segment size: used to achieve better quality solutions; (3) Segment ordering: used to prioritize budget allocation; and (4) Budget allocation and redistribution: used to allocate budget in the most efficient way. These steps are discussed as follows:

1) Similarity Analysis of Input Data: Clustering is mainly based on pattern analysis and similarities between datasets with respect to different parameters. Considering similarity during the segmentation can result in creating segments from components having close characteristics (e.g., very close RIF, initial condition, deterioration behavior, etc.). One of the widely used similarity measures is the Euclidian Distance [10]. Considering two sets of data, $x = \{x_1, x_2, \dots, x_n\}$ and $y = \{y_1, y_2, \dots, y_n\}$, as two points in an n-dimensional Euclidian space, the similarity (Euclidian Distance) between the two datasets is defined as :

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2} \qquad (4)$$

With N datasets, an (N × N) similarity matrix can be created to indicate the level of similarity among all data elements. As such, for the network-level LCCA, similarity between components is determined by calculating the Euclidian Distances, considering the deterioration behaviors (variations in DIs), and relative importance factor (RIF) as follow:

$$d\left(component_i, component_j\right) = \sqrt{\left(Average(DI_j)_k - Average(DI_i)_k\right)^2 + \left(RIF_j - RIF_i\right)^2} \qquad (5)$$

$$\forall\, j \in Network \quad , \quad \forall\, k \in Planning\ Horizon$$

where, $Average(DI_j)_k$ is the average of deterioration indices during the planning horizon for component j, and RIF_j is the relative importance factor for that component. Using Eq. 5, a similarity matrix was created and the values were color coded, then the components were sorted based on the similarity values.

2) Determining Optimum Segment Size: Segment size refers to the number of components within each segment. Determining segment size is based on the capability of optimization tools and model formulations to achieve high quality solutions. Based on the results of many experiments, segment sizes from 50 to 100 have the highest solution quality and are suggested to use for the TDSB asset renewal problem.

3) Segment Ordering: To help in identifying the budget limit to apply to each segment, a measure of a segment's relative criticality is calculated based on the relative importance factors of its components. First, the Criticality Index SCr_j of segment j is defined as the average of the criticalities of its components and the segment size SS_j (number of components in the segment), as follows:

$$SCr_j = \frac{\sum_{i=1}^{SS} CrI_i}{SS_j} \quad (6)$$

where, CrI_i is the Criticality index (from 0 to 1) of component *i*, which is calculated as follows:

$$CrI_i = \frac{RIF_i}{100} \times \frac{Average(EP_i)_k}{100} \quad (7)$$

$$\forall\ k\ \in Planning\ Horizon\ [1, 2, 3, 4, 5]$$

Afterwards, the Relative criticality (RC_j) of a segment is then calculated as follow:

$$RC_j = \frac{SCr_j}{\sum_{i=1}^{NS} SCr_i} \quad (8)$$

NS = Number of Segments

After defining the relative criticality of all segments, segments are ordered from low relative criticality to high relative criticality to facilitate budget allocation in the following step.

4) Budget Allocation and Redistribution: After ordering segments based on criticality values, available yearly budget is allocated to segments based on their relative criticality, as follows:

$$B_{kj} = TB_k \times RC_j \quad (9)$$

where, B_{kj} is the allocated budget to segment *j* in year *k*, and TB_k is the total available budget in year *k*, and RC_j is the relative criticality of segment *j*.

Based on various fund allocation experiments, it was found that increasing the available budget for high criticality segments improves the final solution. Since each segment will be optimized separately and the budget constraint cannot be met to the exact dollar, a small fraction of the budget will be remained unallocated. This leftover money will accumulate from many segments to become a considerable amount in a large-scale problem with many segments. One way to redistribute the leftover money from segment *j* is to add it to available budget for segment *j*+1. Using this approach during the optimization process results in more allocation of budget to segments with higher relative criticality (sorted at the bottom), as shown in Fig. 4. Combining this budget redistribution with budget allocation based on relative criticality leads to the following revised budget allocation function:

$$B_{kj} = TB_k \times RC_j + \left(B_{K(j-1)} - SRC_{k(j-1)}\right) \quad (10)$$

where, $B_{k(j-1)}$ is the allocated budget to segment (*j*-1) in year *k* and $SRC_{k(j-1)}$ is total cost of repairs in segment (*j*-1) in year *k*.

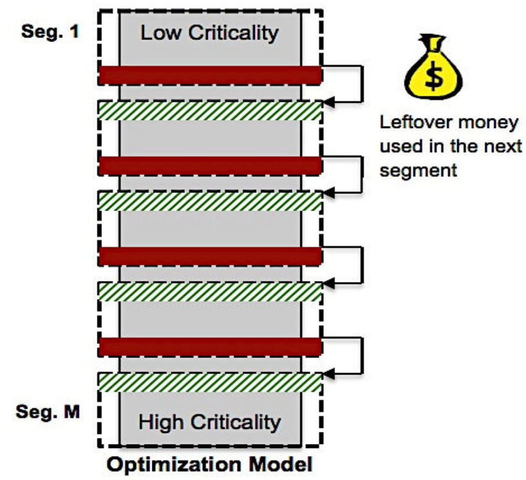


FIGURE 4 Sequential budget redistribution

4. RESULTS AND DISCUTION

For comparison purposes, the best result of the previous work of Hegazy and Elhakeem (2011), which achieved a DI_N of 44.8 for the 8000 components of the TDSB, was used as a baseline. The proposed "GA + Segmentation" mechanism was then applied to networks of different sizes, created by copying the 800 components multiple times. The benefit of this approach is that the results should be multiples of the best results obtained from the 800-component case. Fig. 5 shows an overall comparison of results obtained from the optimization with clustered segments, in comparison to the simple ranking approach typically used by asset managers, and the previous research of the MOST technique. As shown in Fig. 5, at 8000 components, the optimization achieved a network deterioration index of 32.8, which is a huge improvement in optimization performance compared to the baseline case. As shown in the figure, experimenting with even larger number of components the proposed approach shows almost no degradation in the optimization performance. All experiments were conducted using variations of the base LCCA model in Fig. 2 and a laptop computer with 2.4 GHz speed and 4 GB of memory.

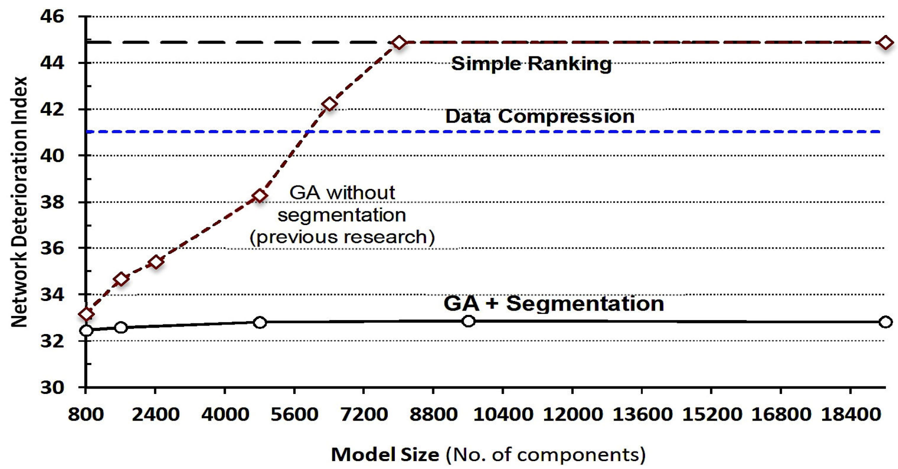


FIGURE 5 Network-level optimization using GA + segmentation

5. CONCLUSIONS

This paper builds upon an earlier work that developed an integrated life cycle cost analysis model to support asset renewal for school buildings administrated by Toronto District School Board (TDSB) in Canada. This prior work used Genetic Algorithms (GAs) to optimize asset renewal decisions for up to 8,000 components simultaneously. In the paper, three segmentation approaches: Random Segmentation; Data Compression; and Clustered Segmentation have been investigated to improve the performance of GAs in optimizing problems with larger size. After extensive experimentation, GA + Clustered segmentation proved to be the most promising. In this approach, segments are generated based on similarities among the components in terms of relative importance and deterioration behaviors. Optimum segment size was then determined based on experimentation to be 100 components. Segments were then ordered, in an ascending order, based on segment relative critically and leftover money was redistributed in a sequential manner from low criticality segments to higher criticality segments. The proposed mechanism was applied on data obtained from the Toronto District School Board and proved to be able to optimize asset renewals for 20,000 components and more with no noticeable performance degradation as the number of components increases. The proposed optimization mechanism with segmentation has the potential to be applied to data from other types of complex infrastructures systems such as bridges, highways, and water/sewer networks.

6. REFERENCES

[1] Hudson, W.R., Haas, R., and Uddin, W. **Infrastructure Management**, McGraw-Hill, New York, USA, 1997.

[2] Liu, C., Hammmad, A., and Itoh, Y., “Maintenance Strategy Optimization of Bridge Decks using Genetic Algorithm,” **Journal of Transportation Engineering,** ASCE, 123(2), 1997, pp. 91-100.

[3] Thompson, P. D., T. Merlo, B. kerr, A. Cheethan, and R. Ellis, “The New Ontario Bridge Management System,” **Transportation Research Circular 498**, **TRB**, 2000, pp. F-6/1-F-6/15.

[4] Macke M. and Higuchi S., “Optimizing Maintenance Interventions for Deteriorating Structures Using Cost-Benefit”, **Journal of Structural Engineering**, Vol.133, 2007, 925-934.

[5] Halfawy, M., Newton, L., and Vanier, D., “Municipal infrastructure asset management systems: State-of-the-art review,” **National Research Council of Canada.** Report No: 48339, 2005.

[6] Hegazy T., Elhakem A., “Multiple optimization and segmentation technique (MOST) for large-scale bilevel life cycle optimization”, **Canadian Journal of Civil Engineering**, 38(3), 2011, pp. 263-271

[7] Cheung, S., Tong, T. K., & Tam, C., “Site pre-cast yard layout arrangement through genetic algorithms,” **Automation in Construction**, 11(1), 2002, pp. 35-46.

[8] Osman, H., Georgy, M., and Ibrahim, M., “A Hybrid CAD-Based Construction Site Layout Planning System Using Genetic Algorithms”, **Automation in Construction**, 12 (6), 2003, pp. 749-764.

[9] Heagay, T., Elbeltagi. E., El-Behairy, H., “Bridge Deck Management with Integrated Life-Cycle Cost Optimization”, **TRB**, Vo. 1866, 2004, pp. 44-50.

[10] Steinbach M., Karypis G., Kumar V., **A Comparison of Document Clustering Techniques**, University of Minnesota, Technical Report #00-034, 2000.

Modelling of Bioprocess for Streptokinase Production Using Mechanistic and Neural Network Approaches

Pavan KUMAR
Department of Biotechnology, Indian Iinstitute of Technology Roorkee
Roorkee, Uttarakhand 247667, India

Sanjoy GHOSH
Department of Biotechnology, Indian Iinstitute of Technology Roorkee
Roorkee, Uttarakhand 247667, India

ABSTRACT

Streptokinase is a vital fibrinolytic drug produced by β-hemolytic streptococci often used to treat myocardial infarction and pulmonary embolism. While growing recombinant strain of E.coli to produce streptokinase in a bioreactor, we are dealing with a controlled environment experiencing the role several indispensable factors that are associated to structured and unstructured aspects of the system. Since a cell itself can be assumed to be an entire structured system that shows the accountability of various parameters. On the other hand, unstructured factors are found influencing the active existence of cell in the fermentation media environment. A model has been established on the basis of both structured and unstructured constraints that has key role in dealing with the plasmid stability. Our effort is to configure a composite model which represents the over all dynamics in a well defined algorithm that depicts the behaviour of the microbial population in the entire bioreactor operational environment. The simulation of the process has clearly shown the role of each and every parameter viz. metabolite concentration, dilution rate etc, to swot up competitive dynamics and segregational instability.

Keywords: Growth dynamics, fibrinolytic drug, competitive theme, bioreactor environment, segregational instability, microbial population, structured and unstructured constrains.

1. INTRODUCTION

Streptokinase is a 47 kDa multi-domain protein constituted of 440 amino acid residues in which about 36 percent are non-polar. The enzyme streptokinase is listed in 'Model of essential medicines, 2005' by WHO. It is a vital fibrinolytic drug produced by β-hemolytic streptococci often used to treat myocardial infarction and pulmonary embolism. Being a life saving drug it has high commercial demand. There requires an improvement in technology to increase its production from recombinant strain.

Bioprocess modelling strategies are now playing a key role in the production of high value products like streptokinase. Structured and unstructured constraints are modelled in a mechanistic way that has key role in dealing with the plasmid stability. Our effort is to configure a composite hybrid model which represents the overall dynamics in a well defined algorithm that depicts the behaviour of the microbial population in the entire bioreactor operational environment. The estimators like segregational instability, plasmid copy number and metabolite concentration are utilized to estimate some of the aspects of fermentation process which may lead to improve supervision and control. The simulation of the process has been done using numerical simulation to evaluate and predict the process behaviour using neural network approach. It is observed that predicted output obtained from neural network had close statistical resemblance to the actual input values. The idea is to utilize the approach of input-output mapping to optimize the output.

There should be a kinetic model to analyze the decrease of plasmid copy number. Also there is a need to develop fast determination of copy number during the process operation. Plasmid number in plasmid containing cells decreases rapidly while in the process and hence fast determination of the copy number which is very much essential. The existing method, e.g. CsCl gradient centrifugation, HPLC etc is offline methods and time consuming. One of the important tasks is to measure the copy number in-process in order to take corrective measure to minimize its loss. In order to maintain a high copy number and high fraction of recombinant cells in the culture medium strategy would be to optimize key parameter, dilution rate in continuous culture production of streptokinase. It is also used to govern product formation by regulating the sustenance of fraction of recombinant cells in the culture medium.

Understanding the growth dynamics it become easier to recognize the pivotal factors that has influence in formation of the desired product hence modelling such a system in future will be of great help to enhance the production adjusting those key parameters. Production of an enzyme or product within a cell is associated to various pathways. So, several components and sub-components together have their role in an enzyme production dynamics. Using the theme of structured model we may assume the subcellular level dynamics and its sustainability. Several subcomponents involve in the production dynamics used to contribute their some vital role and hence control the product formation. Although there is a decline in the product formations in respect to the considered time frame in general.

While evaluating the performance of a unit carrying out production process consideration is to be made on associated prime pathways with their subcomponents. The ultimate component is giving the product after the entire interaction of all subunits substantially in a defined manner. The sequential interaction of subunits, their failure and repair in the considered time span is quite interesting. A probabilistic framework can show a better implication of the approach to design a model in order to evaluate the performance of such dynamic units. The feasibility of this approach is to be made on the basis of reliability assessment of the production unit in a mechanistic way. It is rather more interesting to ensure the applicability of stochastic process in running any subunit in its working state or getting repaired at any instance of time, the sole network will be expressed in the form of a unitary system which is responsible for the production of a particular product which lies on the retaining of its productive state with varying level of product formation which is very likely to be time dependent. The release of metabolites etc may have its impact to negatively affect the subunits involve in pathways, results into cease in production and hence decline in the reliability of such units taken into account. Mortality is often considered to lower the extent of inter specific competition and thereby promote the coexistence of competing species [1].

2. METHODOLOGY AND MODEL DEVELOPMENT

Streptokinase is biochemically a well known fibrinolytic activator that enhances the conversion of plasminogen to the fibrinolytic enzyme plasmin. Workers have over-expressed the streptokinase in E.coli. It is 47 kDa protein encoded by skc gene of Streptococcus equisimilis which is a gram positive bacterium. Though skc gene is known to have many rare codons in its composition, overexpression was achieved instead of having some negative effect. Earlier the constructs were transformed into BL21 (DE3) and grown in LB medium till 0.6 OD at A600. Then the cultures were induced with 1 mM IPTG at 37deg C for 3 h. The expression profile of the streptokinase samples were analyzed [2]. The analysis of relative codon frequency of skc gene in E.coli reveals the presence few position specific rare codons that affect the heterologous protein expression to an extent. In the way of growing E.coli BL21 recombinant strain in a bioreactor, we are dealing with a controlled environment experiencing the role several indispensable factors that are associated to structured and unstructured aspects of the system. Since a cell itself can be assumed to be an entire structured system that shows the accountability of various parameters. On the other hand, unstructured factors are found influencing the active existence of cell in the fermentation media environment. With the progress of fermentation process the initial population of recombinant cells tends to divide into two different cell populations, including those that are having plasmid (as vector pRSET-B) and other that are devoid of any plasmid copy number.

Structured Model Outline

The dynamics and model framework are confined to the concept of reliability has its relevance in assessing the durability of the functionality of any mechanistic unit for a defined time frame. Over all idea is to support the associated pathway networks in view of conducting possible estimation of the enzyme streptokinase production taking into consideration as an instance. Often time dependent interaction of few prominent parameters among the ample of parameters in natural cellular system has been given a desirable weight age in the developing model. The outline of general model in terms of configured subunits and dynamic pathways with an optimal outcome on the basis of reliability in numerical terms, as shown in **Fig. 1**.

The fundamentals governing topology includes, the fact that any production unit has initially maximum reliability and it decreases on basis of interaction of several subunits in a specific topological fashion with variable dynamics. The varying consequence of transformation in topology can be taken into account to depict the observed level of productivity in any terms.

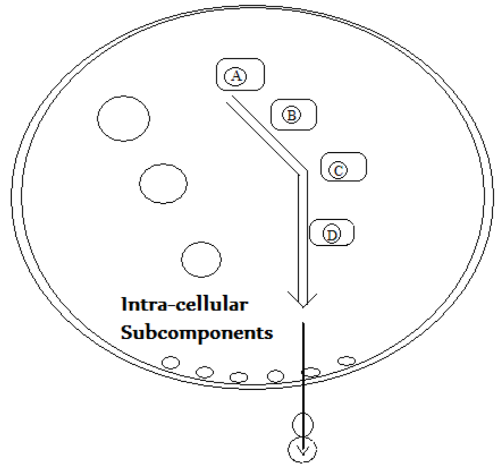


Figure 1. Model framework showing processing details in structured model form

Considering all the possibilities together, the running original part with general decline in reliability, repair of the partially failed component at any time instant and replacement of the completely failed component at the instant; the putative model can be designed. Although overall process is Stochastic. Supportive algorithms to adapt the property of existing regulatory networks are now in the current trend. Stochastic assumptions are made to generate random decision to regulate the variable pathways by the predefined optimal criteria. Time dependent variables play crucial role in governing the multidimensional dynamics thus preparing the model for showing close proximity to the real-time process dynamics.

Regulatory control and computational intelligence techniques are of major help in artificial framing the dynamics associated to bioprocess [4]. Cellular intelligence is now a days a most excellent tool to amplify the information from cellular level activities in order to utilize it for higher level computational applications in prospects of getting automated [3]. Neural network and hybrid Neural Network modelling are being applied for control of fed-batch like process and to evaluate the bioprocess performance under non-ideal conditions [6]. Also Signal oriented modelling is remarkably an approach for utilizing the intra and inter cellular level signals to emphasize over the fundamentals of structured model. The use of regulatory control enable the models based on cells to utilize information gained from experience and thereby respond intelligently to external stimulus or conditions.

Generally quite complex mechanistic models are required to describe adequately the metabolic dynamics of multi-cellular systems especially under non-ideal conditions. It is evident that cells have internal regulatory control to govern all biochemical pathways in a legitimate manner. Hence it coordinates and directs the adaptive machinery to cope up with external or extra-cellular variations maintaining the supportive mechanism that operates to serve simultaneously. The utility of the devised algorithm would be proved to a greater extent only by ensuring its usefulness under realistic conditions. This instrumental work can only be made feasible employing computational and stochastic approaches which easily support to constitute the variable pathway types.

Functional unit's identification within the network of pathways is a vital task to complete for giving a minimal framework. An optimal and reliable representation of the overall model using integrated approach for bioprocess design [7] is a mandatory task in this direction. All the model parameters of plasmid replication control can be obtained independently and no adjustable parameters are needed for the plasmid model [8]. Online estimators for biomass and recombinant protein concentration were constructed using information available online by the application of Neural Network [9]. The accuracy of simulations allows a description of the complex interaction between substrate, microorganism and metabolite (acetate) which is a product and also a carbon source for the microbes [10]. In continuous culture the increase in dilution rate to certain point seemed to induce a rapid decrease in plasmid copy numbers [11]. A rapid quantification of plasmid copy number is important process variable that can be used in process control [12].

Classical interpretation has always been appreciated to support and realize the emergent role of any component. A highly developed model presents designing of composite architecture that optimally blends cellular intelligence, artificial intelligence and mechanistic models [3]. Indeed constituting a framework revealing some link between intra-cellular and extra-cellular processes is reasonably of high utility to generate a robust model with an ample of indispensable dimensions.

Unstructured Model Development

On the other hand approach to utilize the unstructured model to quantify the performance in terms of magnitude is rather equally important. Since the production of the streptokinase is directly associated to stability of plasmid in the recombinant cells thus time dependent loss of plasmid [5] can be found associated accordingly by the given Eq. (1),

$$\frac{dq}{dt} = \left(1 - e^{-mf\left(\frac{X_2}{X_1 + X_2}\right)}\right)\mu_2 \qquad \text{.….(1)}$$

X_1, X_2 Concentration of two cell populations (g/l), X_1-plasmid bearing, X_2-plasmid lacking cells, M- Metabolite concentration (g/l), m_f Collective probability factor, μ_2 Specific gr. rate of plasmid free cells (1/h) and t Time (h).

Overall the probability of plasmid loss in process operation,

$$q = f(t, X1, X2, M, \mu 2) \qquad \text{......(2)}$$

Above parameters are in turn depends on several other variables being the function of more other parameters. Plasmid loss probability (p) is directly found to govern the number of plasmid bearing (pb) and lacking (pl) cells in the preceding and subsequent generations. Computing the magnitude of cells count, from Eq. (1) – (6) we can estimate the system dynamics by simply performing the computer simulation in Matlab 8.0.

$$n_i = pb_i + pl_i \qquad \text{......(3)}$$

$$pl_i = (pb_{i-1}*p + pl_{i-1})*2 \qquad \text{......(4)}$$

$$pb_i = (pb_{i-1} - pb_{i-1}*p)*2 \qquad \text{......(5)}$$

$$pb_{i+1} = n_i - pl_i \qquad \text{......(6)}$$

In the expressions n is the total number of cells in the bioreactor system at a time. The index 'i' stands for the number of generations. The dynamics dealing with count of two types of cells viz. plasmid bearing and lacking cells has its direct association with the production of product like streptokinase. The byproducts like acetic acid etc has its adverse impact over these variety of cells variably due to different constraints associated with them. The unstructured model variables are made to be influenced by the structured model dynamics with time dependent paradigm. So the impact of intra-cellular factors including genetic etc has its profound impact in governing the overall model performance.

Classical Population Dynamics Models

An earliest basic model for population dynamics was proposed earlier by Lotka and Volterra [13] which interprets the population interaction for a prey-predator like system. The microbial ecosystem was considered as a functional entity characterized by certain macroscopic measurements such as the total quantity of biomass or the total number of cells in the medium. It is possible to work with rapid growth of species in well controlled environments, such as ''chemostat" [14]. During the culture and production of microorganisms, the control of the bioprocess sometimes depends on the microorganism concentration or the biomass density and conditions provided for that time period, for instance, in case some aerobic microorganisms the Dissolved Oxygen (DO) is a key factor to their growth [15].

While performing fermentation process in the bioreactor, recombinant cells are found to lose their plasmid. After some time as the fermentation proceeds two types of cell populations are developed. Since substrate is a growth-limiting nutrient (or factor) so there starts a competition between two populations. The organisms carrying the plasmid or plasmid bearing cells are likely to be weaker competitors than one without because of the added load on its metabolic machinery [16]. There could be a number of factors that regulates the dynamics of plasmid carrying cells within the reactor. One such major factor is probability of plasmid loss due to segregation during cell division that can be described by segregative instability.

The probability of plasmid loss in selective medium and difference in specific growth rate between recombinant and plasmid free cells [17] are affected by factors like genetic make-up of the host cells and the reactor operating parameters such as temperature, pH and growth medium composition. Moreover the selective pressure of selection medium is less effective due to the leakage of gene product [18] which is responsible for selective mechanism. A chemostat model of competition can be established between plasmid bearing and plasmid free organism for a single nutrient where plasmid bearing organism can produce toxin against plasmid free organism at same cost to its reproductive abilities [19]. [20] studied the effect of an inhibitor on two populations. They considered the degree of inhibition in presence of inhibitor or toxicant on growth rate. It is studied that estimation of immeasurable biological variables is important in fermentation process, directly influencing the optimal control performance of the fermentation system as well as quality and yield of the targeted recombinant product, applying some novel strategy for state estimation of fed-batch fermentation [21].

The input substrate concentration and dilution rate serve as operating parameters and they are to be controlled by the experimenter. The study for the cases where nutrient supplied at constant rate and time dependent manner were performed earlier and a delay in the growth response of organism to nutrient uptake is obtained. Varying feed profiles were employed in the post-induction phase of recombinant streptokinase protein expression, including constant feed rates, linearly increasing feed rate and exponentially varying feed rates [22] to evaluate the requirement of variable feed strategy.

Our Model Constraints and Framework

Considering the case of selective media, a relevant factor selection stress coefficient can be taken into account which is favourable for plasmid bearing cells [18] while it doesn't favour the plasmid free cells. Since it favours the population of plasmid bearing cells so this parameter would be considered to resist the phenomenon of plasmid loss. Therefore probability of plasmid loss in selective medium is smaller than that in non-selective medium. During the process of fermentation metabolites formation [23] occur as by-products of metabolism which are toxic [24] and inhibit the growth of both types of cells to different extent. The plasmid bearing cells are likely to lose their plasmid because of the permeability of such metabolites into the cells from the environment while other type that is free of plasmid do not have that much extent of harm. In other words we can say that toxicity of metabolite equally harm both the population simultaneously but since plasmid bearing cells are liable to lose their plasmid in response, so plasmid free cell population is strengthening in the same time. So plasmid free cell population or its specific gravity would have an increasing trend despite of decreasing due to this event.

It is observed that probability of plasmid loss is not constant throughout the fermentation process due to the formation of metabolites which shows its presumed toxic effect after certain threshold concentration, so it needs to be taken as variable parameter in the model with respect to time. A criterion of threshold policy can be implemented to evaluate collective probability factor which influence the time dependent variation in probability. Also mortality or formation of dormant cells is possible which is due to toxification developed on account of toxic metabolite or by-product formation, like acetic acid formation in case of fermentation process carried for streptokinase production. Considering the above assumptions we may write a part of the dynamic model [5] as given below:

$$\mu_1 = \mu_{1\max} S/(K_1 + S)$$

$$\mu_2 = \mu_{2\max} S/(K_2 + S)$$

$f_1 = m_p - r + z$; while, $f_2 = z - r$

where, $m_f = \begin{cases} f_1 \text{ if } M > M_{th} \\ f_2 \text{ if } M < M_{th} \end{cases}$

Here, $m_1 > m_2$, also, $(0 < f_1, f_2 < 1)$

Since metabolic toxicity has an influence over plasmid bearing cell population to a larger amount.

m_1 & $m_2 = 0$, if $M < M_{th}$

M_{th} required to be evaluated for different recombinant strain of micro-organisms and media composition under varying set of operational conditions and it depends upon experimental setup with presumed parameters for a bioprocess. Likewise, values of constants m_1, m_2, m_p and r_1 could be assessed for a defined set of conditions. The Nomenclature used for the above unstructured model and its plot are depicted here.

N_1, N_2 Concentration of two cell populations (g/l), N_1-plasmid bearing, N_2-plasmid lacking cells, Y-Yield i.e., ratio of gram of cells formed and gram of substrate consumed, Y_M Yield in terms of metabolite, M_{th} Threshold concentration of metabolite that can effect population growth (g/l), m_l Metabolite toxicity coefficient for plasmid bearing cells, m_2 Metabolite toxicity coefficient for plasmid free cells, p Probability of plasmid loss, K_1 Monod coefficient for plasmid bearing cells (g/l), K_2 Monod coefficient for plasmid lacking cells(g/l), k_d Decay constant for plasmid bearing cells (1/h), k_p Decay constant for product (1/h), D Dilution rate or washout rate (1/h), q Probability of plasmid loss, S_0 Initial substrate concentration (g/l), μ_1 Specific growth rate of plasmid bearing cells (1/h), μ_2 Specific growth rate of plasmid free cells (1/h), μ_{1max} Maximum specific growth rate of plasmid bearing cells (1/h) and μ_{2max} Maximum specific growth rate of plasmid free cells (1/h).

The dynamics related to plasmid bearing and lacking cell populations together with effects impart due to metabolite toxicity and selective stress. Above Monod expressions are to compute specific growth rates. For the collective probability factor m_f two variants are used as f_1 and f_2, in expression which has got variable net probability tested with the level of threshold metabolite concentration to assume its value. Since initially plasmid lacking cell population and metabolite concentration is absent in the medium, both can be taken to zero. In the very start of the process all cells present are plasmid bearing, so the probability of plasmid loss is also to be taken as zero. The magnitude of recombinant cell population and substrate has a pivotal role in governing the dynamics.

3. RESULTS AND DISCUSSION

Process simulation has been done using Matlab 8.0, in order to numerically evaluate the role of each and every parameter. The initial values [25] taken for different parameters at time zero. The simulation is done using most of the standard values taken from a previous model data meant for streptokinase [26] together with various other model constraints with smaller magnitude assumed for different set parameters, which has been used earlier in our work [5], taken on the basis of their apparent role in this dynamical system. The probability of plasmid loss is zero at time t_0 since at the beginning of process all recombinants cells have plasmid machinery. Different dilution rates are considered which proved to be the most relevant factor for continuous operation. Dilution rate was started at a very low value and increased to high values to evaluate the sensitivity of response at different levels.

Plotting all the three variable parameters in **Fig.** (2), together consequently justifies the correlation among different component variables simultaneously. The dynamics of plasmid copy number is shown in **Fig.** (3). It is quite reasonable to investigate that low dilution rate allows the fermentation to last longer generates higher magnitude of concentration of plasmid bearing cells and favour the consumption of substrate. The numerical simulation of the model equations shows plasmid loss tends to occur after a certain limit of metabolite level, which depicts that the increasing concentration of metabolite tends to support the population of plasmid lacking cells that witnesses the trend in loss of plasmid from recombinant cells.

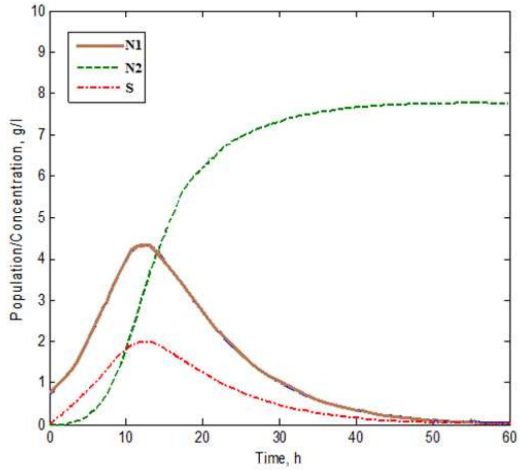


Figure 2. Showing dynamics in figure at dilution rate of D=0.21

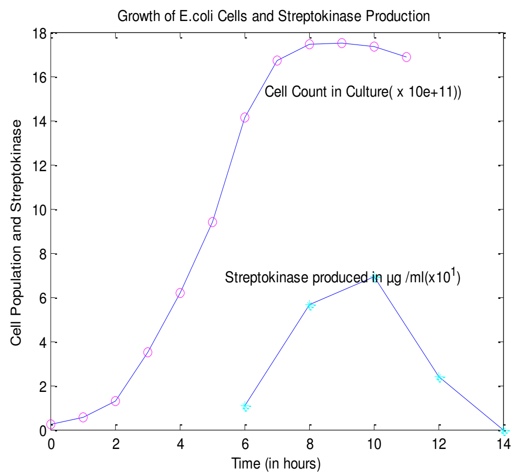


Figure 3. Streptokinase production and cell growth

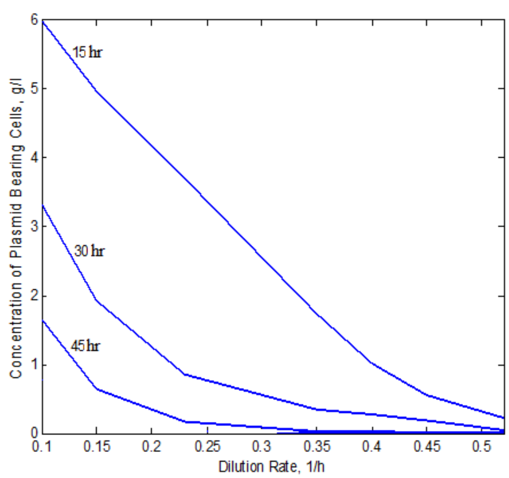


Figure 4. Impact of dilution rate on plasmid contaning cells with time duration of bioprocess

Role of Dilution Rate and Metabolite Concntration

The influence of dilution rate D on concentration of plasmid bearing and plasmid lacking cells with respect to time duration of continuous process is represented in two plots, **Fig.** (4) and (5). The plots in respect to dilution rate have their importance in deciphering the behavior of biased inter-population cell dynamics in this case. The results emphasizes that delayed plasmid loss occurs at lower dilution rates. Numerical simulation of the continuous fermentation process could be rather helpful to enhance the performance adjusting dilution rate to achieve product in amplified amount

The plasmid bearing cell has shown an adrupt declining trend at initial hours of the culture, while incresing the dillution rate, as shown in figure (4). There is a very high probability of plasmid loss even at low metabolite concentration on moving towards a high dilution rate. Particularly the probability has shown a rapid hike on increasing dilution rate beyond D=0.45 in figure (5).

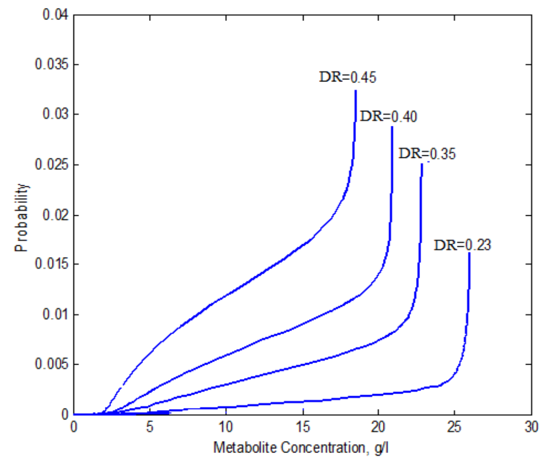


Figure 5. Showing probability variation of plasmid loss with respect to concentration of metabolite at different rate of dilutions

4. CONCLUSIONS

In our approach structured and unstructured model together taken into an amalgamated form, we can have a more sophisticated mathematical representation that is required to be experimentally validated. An approach of designing sophisticated model put forward a higher level of challenge accommodates all sorts of noise viz. noises at intra-cellular level: genetic noise and interaction noise as well as environmental noise at extra-cellular level. These noise associated properties often incorporates more real time properties to such flexible models.

So, population dynamics model for plasmid bearing and plasmid lacking cells in bioreactor has been made more robust to develop an insilico dynamical system which has the characteristics of a chemostat that we used to employ for streptokinase production. The representation of the dynamical system through modelling has its relevance in predicting the behaviour of the system on disturbances or changes made in initial conditions. In the proposed model new factors are taken into account like selection stress coefficient and metabolite toxicity coefficient that have resolved the simultaneous variation in other parameters and their interaction criteria. A relatively composite model ensures a higher degree of flexibility since it has a number of adjustable parameters and properties.

5. REFERENCES

[1] P. Abrams, “The effect of density-independent mortality on the coexistence of exploitative competitors for

renewing resources", **American Naturalist**, Vol. 158, No. 5, 2001, pp. 459-470.

[2] C. Thangadurai, P. Suthakaran, P. Barfal, B. Anandaraj, S.N. Pradhan, H.K. Boneya, S. Ramalingam and V. Murugan, "Rare codon priority and its position specificity at the 50 of the gene modulates heterologous protein expression in Escherichia coli", **Biochemical and Biophysical Research Communications**, Vol. 376, 2008, pp. 647–652.

[3] M.C. Nicoletti, L.C. Jain and R.C. Giordano, "Computational Intelligence Techniques as Tools for Bioprocess Modelling, Optimization, Supervision and Control", **SCI**, Vol. 218, pp. 1-23.

[4] J. Wang, L. Zhao and T. Yu, "On-line Estimation in Fed-batch Fermentation Process Using State Space Model and Unscented Kalman Filter", **Chinese Journal of Chemical Engineering**, Vol. 18, No. 2, 2010, pp. 258-264.

[5] P. Kumar and S. Ghosh, "Population dynamics model for plasmid bearing and plasmid lacking cells for streptokinase production in continuous flow stirred tank bioreactor", **International Journal of Engineering Science and Technology**, Vol. 2, No. 5, 2010, pp. 118-127.

[6] P.R. Patnaik, "Neural and Hybrid Neural Modeling and Control of Fed-Batch Fermentation for Streptokinase: Comparative Evaluation under Nonideal Conditions", **The Canadian Journal of Chemical Engineering**, Vol. 82, No. 3, 2004, pp. 599–606.

[7] R.A. Rouf, "Evaluation of Process Simulators for an Integrated Bioprocess Design", **Report from University of Waterloo**, Ontario, Canada, 1999.

[8] M.M. Ataai and M.L. Shuler, "A Mathematical Model for Prediction of Plasmid Copy Number and Genetic Stability in Escherichia coli", **Biotechnology and Bioengineering**, Vol. 30, 1987, pp. 389-397.

[9] J. Glassey, G.A. Montague, A.C. Ward and B.V. Kara, "Enhanced Supervision of Recombinant E. coli Fermentations via Artificial Neural Networks", **Process Biochemistry**, Vol. 29, 1994, pp. 387-398.

[10] M.J. Guardia and E.G. Calvo, "Modelling of Escherichia coli growth and acetate formation under different operational conditions", **Enzyme and Microbial Technology**, Vol. 29, 2001, pp. 449-455.

[11] P. Reinikainen and I. Virkajarvi, "Escherichia coli growth and Plasmid Copy Numbers in Continuous Cultivations", **Biotechnology Letters**, Vol. 11, No. 4, pp. 225-230.

[12] T. Schmidt, K. Frieshs and E. Flaschel, "Rapid Determination of Plasmid Copy Number", **Journal of Biotechnology**, Vol. 49, 1996, pp. 219-229.

[13] J.E. Bailey and D.F. Ollis, Biochemical Engineering Fundamentals, **Mc Graw-Hill Book Co.**, New York, Vol. 4, 1986, pp. 79-108.

[14] J. Harmand, A. Rapaport, D. Dochain and C. Lobry, "Microbial ecology and bioprocess control: oportunities and challenges", **Journal of Process Control**, Vol. 18, 2008, pp. 865–875.

[15] H. Guo and L. Chen, "Periodic solution of a chemostat model with Monod growth rate and impulsive state feedback control", **Journal of Theoretical Biology**, Vol. 260, 2009, pp. 502–509.

[16] Z. Lu and K.P. Hadeler, "Model of plasmid-bearing, plasmid-free competition in the chemostat with nutrient recycling and an inhibitor", **Mathematical Bioscience**, Vol. 148, 1998, pp. 147-159.

[17] D. Zabriskie and E.J. Arcuri, **Enz. Microb. Technol.**, Vol. 8, 1986, pp. 706-717.

[18] C.A. Sardonini and D. Dibiasio, **Biotechnology Bioengineering**, Vol. 29, 1987, pp. 469-475.

[19] S.B. Hsu and P. Waltman, "A model of the effect of anti-competitor toxins on plasmid-bearing, plasmid free competition", **Taiwanese Journal of Mathematics**, Vol. 6, No. 1, 2000, pp. 135-155.

[20] R.E. Lenski and J.E. Bouma, "Effects of Segregation and Selection on Instability of Plasmid pACYC184 in Escherichia coli", **Journal of Bacteriology**, Vol. 169, 1987, pp. 5314-5316.

[21] W. Jianlin, Z. Liqiang and Y. Tao, "On-line Estimation in Fed-batch Fermentation Process Using State Space Model and Unscented Kalman Filter", **Chinese Journal of Chemical Engineering**, Vol. 18, No. 2, 2010, pp. 258-264.

[22] S. Ramalingama, P. Gautama, K.J. Mukherjee and G. Jayaraman, "Effects of post-induction feed strategies on secretory production of recombinant streptokinase in Escherichia coli", **Biochemical Engineering Journal**, Vol. 33, 2007, pp. 34–41.

[23] S.Y. Lee and E.T. Papoutsakis, "Metabolic Engineering", **Marcel- Dekker**, USA, Vol. 2, 1999, pp. 13-58.

[24] K. Stuebner, E. Boschke, K.H. Wolfe and J. Langer, "Kinetic analysis and modelling of streptokinase fermentation", **Acta. Biotechnol.**, Vol. 11, 1991, pp. 467-77.

[25] P.R. Patnaik, "A heuristic approach to fed-batch optimization of streptokinase production", **Bioprocess Engineering Journal**, Vol. 13, 1995, pp. 109-112.

[26] P.R. Patnaik, "Neural Optimization of Fed-batch Streptokinase Fermentation in a Non-ideal Bioreactor", **The Canadian Journal of Chemical Engineering**, Vol. 80, 2002, pp. 920-926.

Heuristic Approach for Fund Allocation in Complex Rehabilitation Programs

Tarek HEGAZY, Roozbeh RASHEDI, and Abdelbaset ALI
Department of Civil and Environmental Engineering, University of Waterloo
Waterloo, N2L3G1, Canada

ABSTRACT

Pavements are important assets to the safety and serviceability of road networks. Due to their fast deterioration and stringent municipal budgets, allocating the limited rehabilitation funds becomes a complex task. In this paper, an efficient heuristic approach for fund allocation has been introduced. After prioritizing pavements based on different criteria, such as AADT, pavement IRI, and road type, the proposed heuristic process combines three procedures for determining required amount of funding and the best pavement treatments to use during the planning horizon. To automate the process, it has been programmed within a simplified pavement management system (PMS) for testing and validation on a case study of more than 1200 pavements. Based on the results of various experiments, the proposed method allocated the limited rehabilitation funds efficiently, and selected the most appropriate treatments to use. This method is simple and logical, allows multiple what-if scenarios, and can be used on variety of asset types to improve infrastructure fund allocation.

Keywords: Infrastructure Assets, Rehabilitation, Pavement Management, Heuristic, Fund Allocation, Computer Application.

1. INTRODUCTION

Repair and rehabilitation are important decisions for sustaining the serviceability and safety of the civil infrastructure. Effectively allocating limited rehabilitation funds amongst numerous assets, however, is a large-scale and complex optimization problem that has not been adequately addressed by traditional optimization tools. Improper fun allocation for maintaining pavement networks, for example, results in poor ride quality and consequently impose huge indirect costs to the society, including vehicles wear, traffic congestion, crashes, injuries, and delays.

To help asset managers in the difficult infrastructure management decisions such as the allocation of rehabilitation funds, Infrastructure Management Systems (IMSs) or Asset Management Systems (AMSs) have emerged as a systematic process of maintaining, upgrading and operating physical assets cost effectively [1; 2; 3]. Hudson et al. (1997) defined an infrastructure management system as "the operational package that enables the systematic, coordinated planning and programming of investments or expenditure, design, construction, maintenance, rehabilitation, and renovation, operation, and in-service evaluation of physical facilities" [4]. A generic asset management includes different functions (Figure 1): (a) accurate inspection and condition assessment of all components; (b) predicting future condition deterioration of these components along a planning horizon (e.g., 5 years); (c) identifying repair types and estimating their costs and benefits in terms of condition improvement for each component; and (d) life cycle analysis to decide on which components to be repair, best repair types, and best timings to repair these components, under budgetary and other practical constraint.

A Pavement Management Systems (PMS) is an infrastructure management system that applies specifically to roads (e.g., bottom of Figure 1). Among its main functions is to facilitate the prioritization and fund allocation decisions related to road rehabilitation and repairs [5]. In the literature, various efforts have developed rigorous mechanisms to prioritize pavement candidates for repair purposes based on a single criterion or a multiple criteria. Upon prioritizing the various pavement sections, however, existing mechanisms leave the fund allocation task to the asset manager, assuming that it is just a simple matter of assigning money to top priority items until the budget is exhausted. This assumption, however, is an oversimplification since fund allocation decisions require detailed life cycle cost analysis at both individual pavement level and at the whole network level as well [6; 7]. A good pavement management system (PMS) ideally integrates the above functions properly. First, current condition on pavements in terms of riding quality can be assessed based on different condition indices, such as international roughness index (IRI), pavement quality index (PQI), or surface distress index (SDI) [8].

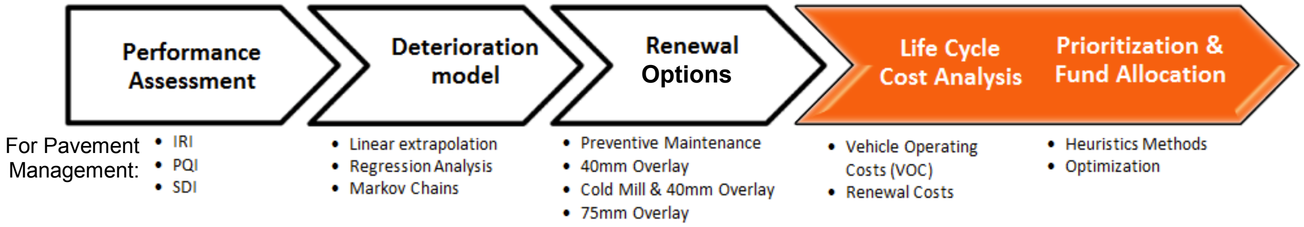


FIGURE 1 Generic pavement management framework.

Next, future pavement conditions for different road types, such as interurban or rural roads, are estimated using different deterioration models, e.g., regression analysis or Markov chains [9; 10]. These models should consider the impact of factors such as annual average daily traffic (AADT) on the future condition of a pavement. In the next step, different repair alternatives (preventive maintenance, cold mill, overlay, micro surfing) need to be evaluated in terms of their costs and improvement effects. Finally, using a life cycle cost analysis (LCCA), the best repair scenarios are selected to maintain a good overall ride quality for the pavement network, cost-effectively. In essence, LCCA considers all the sub-decisions within the planning horizon related to when, and how to repair pavements. To find the best solution by LCCA, a pavement prioritization and fund allocation mechanism needs to be developed. However, when large number of pavements with different deterioration behaviors require repair actions, it is extremely difficult to handle such large-scale and complex combinatorial problem [11]. For instance, consider the case of a small municipality that has 100 pavements sections along a 5-year planning horizon. If three repair alternatives are available (cold mill, 40mm overlay, and 75mm overlay), the possible combination of repair actions is $5^{100\times3}$, which is extremely large and prohibitive.

As an effort towards enhancing asset management capabilities related to fund allocation decisions, this paper proposes a practical and efficient heuristic mechanism for fund allocation. To arrive at a cost-effective solution for large-scale problems, the proposed method incorporates three heuristic procedures and supports fund allocation taken into account road priorities, serviceability requirements, deterioration trends, and budget limits. The proposed methodology has been implemented in a simplified spreadsheet-based PMS and applied on a network of more than 1200 roads. Comments on the performance of the methods are then highlighted.

2. CASE STUDY

The case study is a pavement management investment analysis challenge posted at the 6th International Conference on Managing Pavements (ICMP6). The Challenge was initiated to carry out an analysis and recommend strategies for managing a defined network of interurban and rural roads.

The pavement network is comprised of a total of 1293 road sections spanning 3240 km, covering two road classes, and varying in traffic use, surface age, and condition. The rural roads (R) span most traffic and condition categories. Inter-urban roads (I) are represented on the medium to very highly traffic roads. All pavement sections are located within the same climatic region with consistent sub-soil conditions. Each section has a defined length, width, number of lanes, AADT, soil type, year of construction, base thickness, base material type, most recent treatment, and surface thickness. In addition, surface condition assessments (International Roughness Index, IRI, and others), extent of distresses, and predicted trigger or needs year are specified for all sections. The discount rate for investment analysis is specified as 6%, also increase in vehicle operating costs due to increase in pavement roughness, represented by IRI, and the annual traffic growth rate for the interurban and rural road networks are 2.5 and 1.5%, respectively. Tables 1, 2, and 3 show the annual rate of increase of IRI, the repair costs, and the IRI trigger levels, respectively. Figure 2 also shows the roughness improvements due to different treatments.

Having this information about the pavement network condition, deterioration behaviors, repair options, and vehicle operation costs (VOC), an asset manager needs to decide on the best repair actions and allocate available funds cost-effectively. In the next section, the proposed heuristic approach is described and further implemented within a spreadsheet-based pavement management system to handle the case study.

TABLE 1 Annual rate of IRI increase.

Road Class	AADT	Rate of Increase in IRI (m/Km/Yr.)
Interurban	>8000	0.069
	<8000	0.077
Rural	>1500	0.091
	<1500	0.101

TABLE 2 Unit cost of treatment.

Intervention Type	Cost ($)
1. Preventive Maintenance	6.45
2. 40mm Overlay	6.75
3. Cold Mill & 40mm Overlay	10.50
4. 75mm Overlay	15.75
5. 100mm Overlay	16.50

TABLE 3 IRI trigger levels and Relative Importance.

AADT	IRI Trigger (mm/m)	Relative Importance Factor (RIF)
<400	3.0	1.0
400-1500	2.6	1.4
1500-6000	2.3	1.7
6000-8000	2.1	1.9
>8000	1.9	2.1

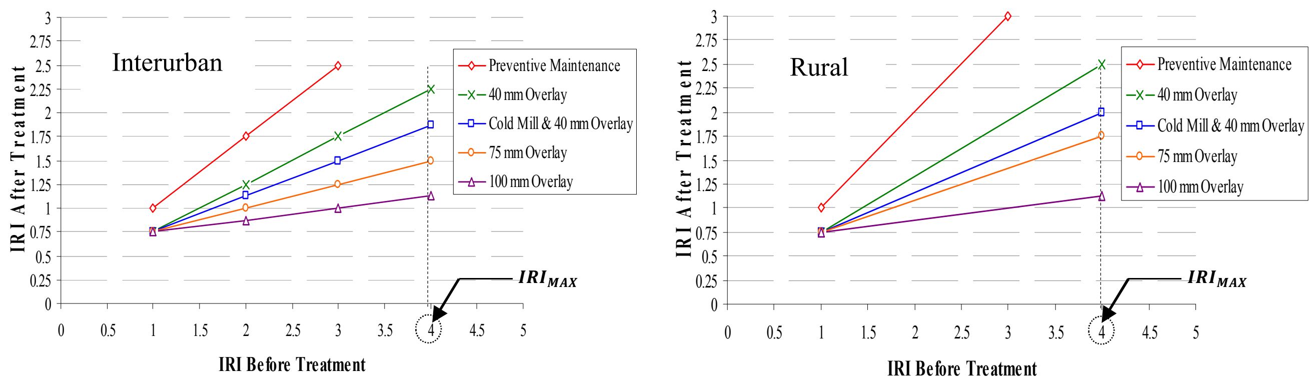


FIGURE 1 The improvement effects of various treatments.

3. HEURISTIC APPROACH FOR REPAIR FUND ALLOCATION

In order to select the best pavement treatments within the planning horizon, this study proposes a heuristic repair fund allocation approach. The method uses a ranking function based on a calculated priority index for pavements, which reflects the need for urgent repair actions. As shown in Table 3, the IRI trigger values (i.e., the minimum acceptable IRI values) are smaller as the traffic becomes heavy. This is logical because heavily used roads need to be maintained with good ride quality (smaller IRI). In the same sense, it is possible to consider that heavily used roads are relatively more important and their importance factor is a function of the IRI trigger value as well. As such, a relative importance factor (RIF_j) for each pavement j can be determined based on AADT as follows:

$$RIF_j = IRI_{MAX} - IRI\ Trigger\ Levels \qquad (1)$$

where, IRI_{MAX} is the maximum IRI value of 4, as shown in Figure 2. Using Eq. (1), therefore, the last column of Table 3 shows the calculated relative importance factors.

Relative importance factors are not sufficient alone to prioritize a pavement for repairs because a less important road can be in a worse condition and thus deserves to be in higher priority. As such, to develop a simple priority index (PI_j) for repairing pavement j, the relative importance factor can be combined with a condition indicator (such as the current condition, IRI_{0j}), as follow:

$$PI_j = RIF_j \times IRI_{0j} \qquad (2)$$

After calculating all PIs, pavements are sorted from high to low priority. This helps in allocating the available funds to pavements with urgent repair needs prior to less critical ones to improve the overall network condition effectively. The overall condition of the whole network is therefore calculated by averaging the IRI values over the network in the entire planning horizon, as follow:

$$Overall\ Pavement\ Nework\ Condition = \frac{\sum_{j=1}^{N}\sum_{k=1}^{5} IRI_{jk}}{N} \quad (3)$$

$$\forall\, j\ \in network, \forall\, k\ \in planning\ horizon$$

After prioritization, the proposed heuristic approach involves two functions that are discussed in the following subsections: (a) Determining the minimum budget that maintains network condition above an acceptable level; and (b) Selecting the best repair types and timings under budget limit (near-optimum fund allocation).

Determining Minimum Funding Needs

This proposed heuristic procedure is used to determine the minimum required budget to bring the pavement network above an acceptable level. It combines project level decisions that are related to selecting the best treatment types, and network level decisions for addressing the repair timings during the planning horizon. Basically, each pavement is selected and its yearly IRI values are checked to see in which year the pavement will deteriorate to he unacceptable trigger level (i.e., year 2 in Figure 3). Once the year of repair is selected, a minimum-cost treatment can be selected as the one that maintains the pavement above the IRI trigger level during the planning horizon (e.g., repair action 4 in the figure). This approach spends minimum money only in the necessary year so that to preserve the condition of individual pavements above the IRI trigger level in all years. Repeating this process for all pavements accumulates the necessary budget level that is required for the whole network. This answers the important question of how much level of funding is needed to keep our assets at minimum acceptable level during the planning horizon.

Near-Optimum Allocation of Limited Repair Funds

Since most municipalities receive lower funding levels than they ask for, or what is necessary, two heuristic approaches are proposed to determine a near-optimum allocation of the limited pavement rehabilitation fund. There are as follow:

Method 1: This heuristic method is a project-wise approach that analyzes 'asset-by-asset'. Since pavements are prioritized based on their priority index, pavements with higher priority index that are in urgent need of repair are considered first. In this method, each pavement is evaluated over the planning horizon to find the violation time and the least-costly repair (Figure 4a). The process is similar to the one in Figure 3, except that when an asset is decided to be repaired in a certain year that has no remaining funds, then the asset will not be repaired and the process continues to the next one until all assets are evaluated.

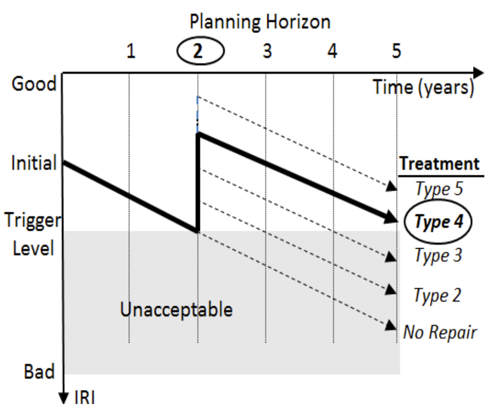


FIGURE 3 Repair type and timing under no budget constraint.

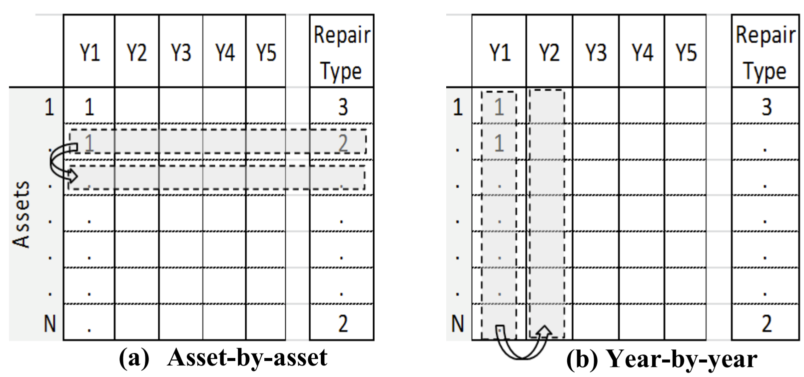


FIGURE 4 Two fund-allocation heuristics under budget constraint.

Method 2: This heuristic method is a network-wise process that allocates available budgets 'year-by-year' (Figure 4b). In this method, each year is considered separately and the assets that violate the trigger level in that year are considered and then a least-costly treatment is decided for them one-by-one until the budget limit of that year is reached. The process starts from year 1 and moves to the next year, until the end of planning horizon.

Spreadsheet-based Implementation

In order to evaluate and compare the proposed fund allocation heuristics, they have been implemented within a spreadsheet-based PMS. An overview of the developed spreadsheet model is shown in Fig. 5 showing all the sheet portions that relate to the various asset management functions. As shown in Figure 5, each road is represented as a separate row and all the data related to each road are represented in columns. The model in Figure 5 is formulated considering a tactical asset management plan of five years. The two main decisions in the model are:

- The index to one of five repair types in column "Repair Type" for each road (integer variables); and

- The index to one of the five repair years "Year 1 to Year 5" for each road (binary variables).

These two decisions for each road are linked by equations to all the related functions of current performance assessment, deterioration, repair costs, and improvements after repair, and all the LCCA details. These functions are explained in the following.

Priority Indices: The proposed spreadsheet combines the IRI with the AADT for each road, into Priority Index "PI", based on Eq. (2). This index varies from 0 to 5. When the road's PI is zero, performance is high and the road has low repair priority. Also, when the road is in worst performance, it gets the highest priority for repair (PI = 5), Thus the priority is higher for roads with higher deterioration.

Deterioration Model: To enable life cycle analysis over a 5-year planning period, it is important to predict the future deterioration of the roads over the next five years. The future condition for each road has been evaluated based on the annual rate of IRI increased in Table 1 and the average annually daily traffic (AADT).

Decision Variables and Constraints: In the spreadsheet model, five treatments are available for each pavement (Table 2). These repair options are referenced using integer values from 1 to 5. Repair timing during the planning horizon is also referenced using binary variables (1 represents a repair action and 0 means no repair). Since the repair timing is only once within the planning horizon (i.e., a single visit), the sum of all the binary variables in all years must be either 0 (no repair), or 1.

Life Cycle Cost Calculation: In the spreadsheet, Life Cycle Cost over the five year plan is calculated yearly for each asset considering repair cost (according to the selected repair type) and the Vehicle Operating Costs (VOC). Accordingly, the present value of the total life cycle cost "TLCC" is calculated as follows:

$$\text{TLCC} = \text{Sum of } [(\text{Repair Costs} + \text{VOC})_n / (1+i)^n]$$

where, n is the year number and i is the applicable interest rate per year (user input).

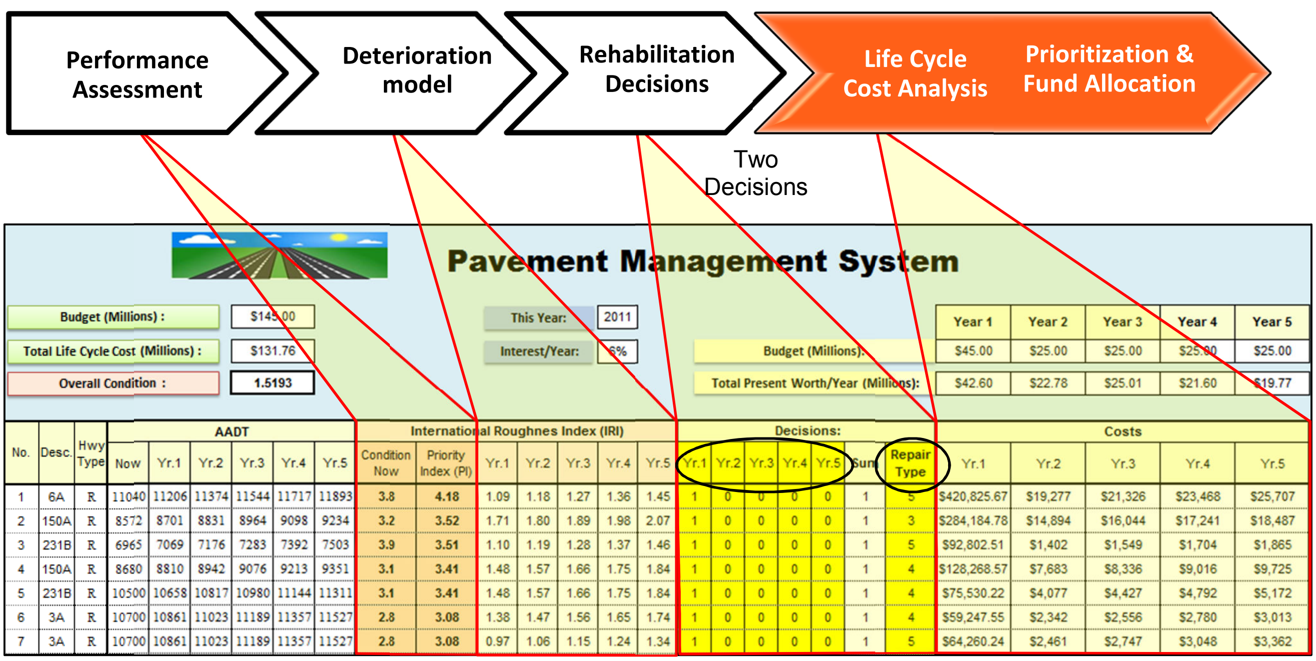


FIGURE 5 Pavement management spreadsheet.

4. EXPERIMENTATION

The proposed heuristic methods have been programmed using Excel's VBA programming language as macro programs. Before experimenting with fund allocation, the network had an overall condition of 2.473. Applying the first heuristic procedure to determine the necessary budget revealed that the total budget of $51.64 million is required for preserving the whole network above the acceptable level of service, which improves the overall condition from 2.473 to 1.82. More than 50% of this amount, however, is required to be spent in the first year since many assets have bad initial condition and require immediate intervention.

In a more realistic fund allocation situation, an asset manager is assumed to have a budget limit of $10 million in each year (add up to the same total amount needed but are equally distributed among the years). Considering this budget limit, the two proposed heuristic methods (asset-by-asset, and year-by-year) were used for allocating the limited funds. Upon applying the proposed methods, the funds were allocated successfully and the heuristic approach improved the overall IRI as shown in Table 4. Both procedures are equally efficient (i.e., not much far from the case of no budget redistribution), but have the great advantage of avoiding the need for uneven fund distribution among the years of the plan. In order to validate the proposed approach further, its results were compared to the results of optimization using genetic algorithms (GA). Using GA-based optimization, which is a more time-consuming and complex process, the optimization could not improves the results of the heuristic approach by more than 0.6%.

TABLE 4 Comparison of results

	No Budget Limit	Budget Limit	
		Asset by Asset	Year by Year
Total Life Cycle Cost (Millions)	51.64	49.3	49.2
Overall Condition (IRI)	1.8237	1.9184	1.9262
No. of assets violated	0	221	232

5. CONCLUDING REMARKS

A practical heuristic approach has been proposed for efficient fund allocation of infrastructure repair funds, which is a large-scale and complex optimization problem. Using a case study of a pavement network of more than 1200 pavements, the proposed method proved to be efficient and can be easily used by asset managers to allocate limited funds cost-effectively while maintaining the serviceability criteria. Some key advantages of the proposed heuristic method are as follows: Efficient and easy-to-use; Determines the needed budget level and selects the best treatments under a budget limit; Considers both network level and project level decisions; Applicable to different size pavement networks; and Applicable to other assets such as bridges, culverts, etc.

There are several future extensions that can improve the performance of the prioritization and fund allocation mechanism further. It is also possible to combine the proposed heuristic method with GA-based optimization by using the good speedy results of the heuristic method as initial values for the optimization process to improve the results further. Also, the proposed heuristic can be improved by enhancing the condition assessment and deterioration model of the PMS, in addition to considering a combination of pavement quality measures such as surface distress index, and pavement quality index.

6. REFERENCES

[1] Vanier, D.J., “Why Industry Needs Asset Management Tools”, **NRCC Seminar Series, Innovations in Urban Infrastructure**, APWA International Public Works Congress, Denver, USA, Vol. 15, 1999, 11-25.

[2] **Transportation Association of Canada TAC**, Primer on Highway Asset Management Systems. Ottawa, March,1999.

[3] Federal Highway Administration FHWA, **Asset Management Primer**. A Report prepared by the U.S. Department of Transportation, 1999.

[4] Hudson, W.R., Haas, R., and Uddin, W., **Infrastructure management**, New York, USA: McGraw-Hill, 1997.

[5] Haas, R., Hudson, W.R., and Zaniewski, J.P., **Modern Pavement Management**, Krieger Publishing Company, Malabar, Fla, 1994.

[6] Grussing, M. N., Uzarski, D. R., & Marrano, L. I., “Optimizing facility component maintenance, repair, and restoration investment strategies using financial ROI metrics and consequence analysis”, Applications of Advanced Technology in Transportation - Proceedings of the **Ninth International Conference on Applications of Advanced Technology in Transportation**, Chicago, IL, 2006, 81-86.

[7] Karydas, D. M., & Gifun, J. F., “A method for the efficient prioritization of infrastructure renewal projects”, **Reliability Engineering and System Safety**, Vol. 91, 2006, 84-99.

[8] Fukuhara, T., Terada, K., Nagao, M., Kasahara, A., and Ichihashi, S., “Automatic Pavement-Distress-Survey System”, **Journal of Transportation Engineering**, ASCE, May/June, 1990. Vol. 116, 1990, 280-286.

[9] Yang, J., Lu, J. J., Gunaratne, M., and Xiang. Q., “Overall pavement condition forecasting using neural networks – An application to Florida highway network”, Proceedings of the **82nd Transportation Research Board Annual Meeting** (CD-ROM), Washington, D.C. Issue 1853, 2003, 3-12.

[10] Butt, A.A., Shahin, M.Y., Feighan, K.J., and Carpenter, S.H., “Pavement Performance Prediction Model Using the Markov Process”, Transportation Research Record: **Journal of the Transportation Research Board**, No. 1123, TRB, Washington, D.C., 1987, 12-19.

[11] Csiszár, S., “Optimization algorithms (survey and analysis)”, LINDI 2007 - **International Symposium on Logistics and Industrial Informatics,** 2007, 185-188.

The nature of information as a fundamental property of matter: A case study using petroleum and hydrocarbon gases

AUTHORS:

Rustem Seifuol-Mulukov, Institute for Informatics Problem, Russian Academy of Sciences, Doctor of Sciences in Geology, Professor in Information Sciences, Moscow, Russian Federation

Marjorie M.K. Hlava, President, Access Innovations, Inc. Albuquerque, New Mexico, USA

ABSTRACT

This paper presents a unique view of information as a fundamental property of matter. Mathematical, physical, cybernetics, quantum mechanics and philosophy rules and concepts are considered in support of the informational nature of matter. First consider the "bit" (0's and 1's) as a fundamental unit in conveying information in computers which is subject to substitution with physical and mathematical values. The computer bit value reflects the nature of information as a fundamental property of matter. This understanding could serve as the basis for a new approach, based on an informatics framework, for the investigation of complex natural systems. To illustrate such an approach, we use a new view on petroleum and hydrocarbon gases are considered, based on an evaluation of their information content. Changes in the information amount show that petroleum and hydrocarbon gases originated from carbon and hydrogen atoms which appeared in the Earth's mantle, forming the simplest hydrocarbons.

KEYWORDS: Information Science, Classification, States of Matter, Computer Science, Informatics, Cybernetics, Hydrocarbon States, Mathematical Theory

……………………………………………………….

BACKGROUND AND REVIEW OF THE LITERATURE

The midpoint of the twentieth century brought significant changes in mankind's understanding of the nature of information. Claude Shannon's groundbreaking *Mathematical Theory of Communication* was published in 1948, and soon after, computers spread all over the world. It became clear that information, in addition to its everyday meaning as *informatio* (Latin), has a broader scope and was a worthy subject of scientific investigation, both from theoretical and practical application points of view. Less than 20 years after Shannon's publication English speaking countries established the term "information science" (Taylor, 1966). From the very beginning, it was recognized that information science has origins and concerns in common with existing fields of study: linguistics, semantics, psychology, and library science (Otten and Debons, 1970). We find it is very significant that in this lineup of sciences, we do not see physics or cybernetics.

The term "informatics" appeared in Germany and France (Steinbuch, 1957; "Informatique", 2011) to cover studies on computerized methods of information processing, accumulation, storage, and retrieval. In the USSR in the mid-1960s, informatics became the official name of the science dealing with all aspects of information, including information science and computer science. In present-day Russia, informatics is viewed as a natural science dealing with fundamental study of the nature and properties of information, as well as the methods and systems of information creation, collection processing, transmission, and usage and their application in industry and everyday life (Sokolov, 2009). The Russian view of informatics is demonstrated by Russian encyclopedia entries, as well as by entries in dictionaries of physics, philosophy, other areas, and general terminology. These entries define informatics similarly, as the science covering the *nature, structure, collection, storage, processing, transmission*, and *retrieval* of information, and of additional operations connected with information (emphasis in italics above and below in text by the author). As one can see from this brief survey of informatics, its prevailing focus has been the study of operations performed in connection with information. The basic Rubricator presented as Lenin's classification outline of knowledge is a basis for this different point of view and underpins the current Russian view of the science of information and its interrelationships. Russian Classification systems based on the Rubricator have often placed fields of study in different configurations than those of other Western schools of thought leading to different relationships between the sciences and allowing for differing conclusions. The interdisciplinary study of the nature of information with physics or cybernetics has not been evident.

Currently in the English-speaking world, the use of the term “informatics” is rather limited in comparison to “information science”. Its position in the ASIS&T Thesaurus of Information Science, Technology, and Librarianship (2005) shows informatics as a third level term and a narrow focus. On the other hand, information science, according to M. Norton (2010, p. 6), deals with “five areas: *collection* and *storage*, classification and control, access for *retrieval*, *communication* and evaluation of information,” some of the very same areas that Russian scholars consider the provenance of informatics.

All the above-mentioned definitions of the sciences dealing with information are in the context of information’s traditional role as covering data, reports, messages, and the like, as well as computer operations. Such an understanding of the nature of information is quite understandable. Over the millennia, oral information has been transmitted from generation to generation, preserved as historical and religious recitations especially connected to geographic locations and other artifacts and data all have become the collective historical memory of cultures. The appearance of written language set the first stone in the foundation of information science as the library and archival science.

During the 20th century, the amount and variety of information and its carriers increased so much that mathematical approaches to cope with the information became essential. Mathematicians developed calculation methods for information, applying such mathematical and cybernetics concepts as the *set*, *uncertainty*, and *probability*. The mathematical theory of information was created largely by Ralph Hartly (1928) and Claude Shannon (1948). Shannon expressed the quantity of information using a formula that is now well known. He substituted a binary unit, the bit, for information amount calculations. This introduced the concept and practice of mathematical substitution into the world of information science.

Over the 60-plus years that have passed since Shannon’s discovery, the mathematical theory of information has seen many practical applications in information, as described in Melanie Norton’s digest (2010). All of these applications have circled around such concepts as data, database, data mining, information storage and retrieval, and knowledge management. With all of these concepts, information is rarely regarded as a means of understanding the laws of nature; rather, it is regarded primarily as a commodity, to be used to make new commodities and new goods.

However (and this is very important for the theory of information science), such an interpretation of information, mostly in an everyday sense, obscures its main property as a fundamental feature of matter. We are referring to the physical basis of information, as manifested in natural phenomena. It was in spite of the fact that the cognition of information phenomena from the point of view of physics, philosophy, and physiology had begun much earlier, that information science came into being.

Atoms, molecules, and elementary particles, as well as their corpuscular properties, were known to science before R. Taylor (1966) and H. Borko (1968) introduced the discipline of information science and Karl Steinbuch (1957) proposed the discipline of informatics. For the theory of information science, one very important discovery was that of elementary fermions (quarks and leptons) that form electrons and elementary particles of the atom’s nucleus. The information science community must be thankful to Werner Heisenberg (1952) for the discovery of the fundamental law of microcosm: quantum field theory. This theory led Anton Zeilinger (1999) to take into consideration the features and performance of elementary particles in a quantum field to formulate the principles of information quantization, a theory that describes the connection between elementary particles and information. According to Zeilinger, “*An elementary system represents the truth value of one proposition.* [By “proposition”, he meant a statement, the truth of which can be verified directly by experiment.] ... The truth value of a proposition can be represented by one bit of information with “true” being identified with the bit value “1” and “false” being identified with the bit value 0. Thus, our principle becomes simply: *An elementary system carries 1 bit of information*.” (p. 635) The movement of electrons and protons in accordance with quantum field theory could be proven by experimentation and computation, once they were considered as constituting elementary systems, and the information content of the atoms of any chemical element could be calculated according to atom mass and the wave function of its electrons and protons.

It was a very important discovery for information science and informatics that the information content of any substance composed of atoms could be calculated for all stages of its chemical, thermal, and catalytic processes, as well as for other processes, using the bit values.

The information characteristics of elementary particles are an unquestionable physical argument for the characteristic of information as a fundamental feature of matter, and the bit is the unit that can be applied for the measurement of the information content of complex physical systems that occur in nature.

The preceding statement shows that *uncertainty*, *probability*, and *set*, as concepts of physics, mathematics, cybernetics, and quantum mechanics, serve as the basis for mathematical and physical substitution of information and prove that information exists beyond our conscious awareness. The bit is the universal unit for information measurement.

Our understanding of information as being contained by natural phenomena received much input from the investigations of the relation between information and such concepts as *multiplicity*, *complexity*, and *organization*. Later investigations dealt with features and structures of objects and organisms in biology, geology, geography, topology, and many other disciplines. It was W. Ross Ashby (1956) who showed, for the first time, the inseparability of the concepts of multiplicity and information. He equated the amount of information with the corresponding amount of multiplicity. Multiplicity can be viewed as the number of different elements (as it is with Zeilinger's truth value propositions), and as a base two logarithm of this number of elements (as it is in Shannon's formula for information amount). In this case, this logarithm is the means for the transformation of the multiplicity set of elements into the amount of information of this set, expressed in bits. Ashby introduced the Law of Requisite Variety, which plays a key role in information theory.

In biology, Nicolas Rashevsky (1955) was the first to show that the information content of the molecules of complex biological structures could be represented by the topology of a corresponding set of adjacent values. He illustrated how the structural relationships in a biological molecule could be mapped to a graph, and how the amount of information content in the molecule's structure could then be derived from the graph. Ernesto Trucco (1956) further developed the idea of topological information content, with the theory that the amount of information content represented by a graph can be determined from the vertices and the lines that connect them. According to Trucco, the amount of information in the topology is represented by the relation of the total number of vertices to the number of vertices that are topologically non-equivalent, that is, have a different grade. This relation can be expressed by a logarithm function, for which Trucco used base two. The resulting binary logarithm formula for topological information is practically identical to Shannon's formula for information entropy. Hence, topological information theory provides additional evidence in favor of the claim that information is a fundamental feature of matter, this time on the basis of the structure of complex chemical compounds.

Henry Quastler (1964) investigated the role of information in the organization of biological structures. He based his calculations on the assumption that a cell is a highly organized system consisting of perhaps $2*10^{11}$ atoms containing approximately $5*10^{12}$ bits of information. All healthy cells of an organism are completely functional; if a cell stops being a functional part of the organism, only the information of its atoms remains. Information is contained in the unique coding of the biopolymer molecules of DNA and RNA. In effect, information is communicated in the "memorization", so to speak, of this code and its transmission from generation to generation. Quastler's role in the development of information science also includes the establishment of non-probabilistic methods of uncertainty removal, that is, information creation.

The fact that there exists, in the areas of information science and informatics, the concept that information is a fundamental feature of matter and biological structures, with the bit as the universal unit for information content measurement, has very important implications for the application of those information areas in the investigation of natural systems.

1. The "bit", in its physical sense and as a part of mathematical expressions, is not only a unit for statistical information calculation. The quantity of information in the atom of any chemical element, as well as that of any complex chemical substances composed of an aggregate of atoms of different elements, can also be evaluated.

2. The quantity of information, as expressed in bits, reflects the degree of complexity, organization, and variety of the systems. Regressive decrease of the quantity of information, or its progressive increase, demonstrates the evolution of the system in the form, respectively, of destruction (entropy increase) or creation (entropy decrease).

APPLICATION TO THE INSTANCE OF PETROLEUM AND HYRROCARBONS

These principles of information theory and informatics can be applied to the structure, composition, and development of any complex natural system. Petroleum was chosen as an example because it is composed of more than 500 hydrocarbon compounds that we qualify as hydrocarbon sequences. The calculation of the information content quantity of petroleum and hydrocarbon gases is based on the fact that their molecules consist of atoms, and the information content of any atomic or molecular structure must be calculated in bits. The level of system complexity indicates the quantity of information. The higher the amount of information, the more complex and varied the system is. As noted by Nikolaĭ Amosov, the quantity of information is directly proportionate to the indices of the degree of the matter's development (Amosov, 1964). If one relates this statement to petroleum, the process of its genesis may be considered as a hierarchical structure, with carbon and hydrogen atoms at the base, and with each subsequent level of development characterized by more complex hydrocarbon sequences having more variety and more information content.

The present author has substantiated a method for the information volume calculation of petroleum hydrocarbon sequences, using topological information derived from the conventional molecular structures (R. Seiful-Mulukov 2010, pp. 131-142). Starting with the premise that carbon and hydrogen atoms constitute approximately 99 percent of petroleum, and that those atoms have their normal atomic orbital structure within the upper asthenosphere, we next consider how the process of petroleum generation starts in the Earth's upper mantle. In Table 1, the rows from bottom to top show the successive meshing of hydrocarbon molecular structures, reflecting the formation of the more complex hydrocarbons, with corresponding increases in information volume. The dynamics of information volume changes in the conventional mass of the hydrocarbon molecule -- from the simplest methane to the final full hydrocarbon sequences in petroleum – clearly demonstrate that the formation of the more complex hydrocarbons is in line with a corresponding increase in information volume. The information volume changes provide a basis for us to consider the argument that the petroleum generation process starts in the Earth's upper mantle, with carbon and hydrogen atoms appearing with minimal information volume. Within sedimentary rocks, the simplest hydrocarbons, with an information volume of less than 550 bits, are subjected to catalytic transformation, leading to the formation of more complex hydrocarbons of petroleum, as explained by J. Germain (1969). The main stage of petroleum formation, with its conventional molecule containing no less than 4,860 bits of information, takes place in the Earth's crust, under thermodynamic conditions that cause the catalytic conversion and conservation of hydrocarbon sequences. Petroleum formation ends in reservoirs, under a microstate of minimal entropy that corresponds to a high level of complexity, variety, and volume of information. Such a model of petroleum formation, and of the dynamics of its entropy changes, completely excludes the idea that the chemical composition of petroleum can remain unchanged in a reservoir over hundreds of millions of years.

Table 1. Information content of hydrocarbon sequences in petroleum generation*

Chemical element or hydrocarbon	Formula (gross or empirical ones)	Information volume (bit)
Bitumen	$C_{45}H_{51}O_2SN$	6219
Heavy oil	$C_{37}H_{62}SN$	5228
Light crude	$C_{32}H_{66}OSN$	4862
Butane	C_4H_{10}	547
Butylene	C_4H_8	539
Propane	C_3H_8	424
Propylene	C_3H_6	403
Ethane	C_2H_6	289
Ethylene	C_2H_4	269
Acetylene	C_2H_2	248
Methane	CH_4	154
Carbon	C	109*
Hydrogen	H	10*

*According to the inorganic model of petroleum genesis (After Igor Gurevich, 2009, p. 150.)

CONCULSIONS

We share the following observations:

1.The dominant understanding of information, in the fields of information science and informatics, resulted from its initial, historical sense as *informatio* (data, intelligence, or knowledge); furthering the development of the practical application of information in this sense, including computerized information technologies, remains the main goal of those fields. The view of information as a commodity -- and as goods that enable the production of new commodities, new experience, and new knowledge – is an appropriate and very important perspective in information science and informatics.

2.The structure and thematic scopes of information science and informatics, as well as the names of these disciplines, are historically conditioned. In spite of differences in the name, both areas are dealing with the same phenomenon: information. Any science is developing as long as scientists of different schools of thought are exchanging ideas. In this context, an understanding of the ideas and terminology of the information science community by the informatics community, and vice versa, is very important. This presentation is intended to be one step in this direction.

3. Information should not be considered as merely a commodity or merely goods. Information is a fundamental property of matter. The main unit of measurement for this property is the bit. The application of the information bit for the understanding of the structure and evolution of natural systems gives us new methods and means for scientific investigation that differ significantly from traditional and superficial ones. We believe that such methods of investigation, using an information approach, have a promising future.

REFERENCES

Amosov, N.M. (Nikolaĭ) (1964). “Thinking and Information,” in Problemy myshleniya v sovremennoy nauke (Issues of Thinking in Modern Science) (Moscow: Mysl’), pp. 387-409.

Ashby, W. Ross (1956). An Introduction to Cybernetics. London: Chapman & Hall Ltd., 1956.

ASIS&T Thesaurus of Information Science, Technology, and Librarianship, third edition (2005). (Revised edition of ASIS Thesaurus of Information Science and Librarianship, ed. Jessica Milstead.) Ed. Alice Redmond-Neal and Marjorie M.K. Hlava. Medford, New Jersey: Information Today, Inc.

Borko, Harold (1968). “Information Science: What Is It?” American Documentation, Vol. 19, No. 1, pp. 3-5.

“Informatique” (2011). http://fr.wikipedia.org/wiki/Informatique. Retrieved January 2011.

Germain, J.E. (1969). Catalytic Conversion of Hydrocarbons. London: Academic Press.

Gurevich, Igor (2009). Information Characteristics of Physical Systems. Moscow:11th Format; Sevastopol: Cypress. ISBN 978-5-902030-74-4 (in Russian)

Hartley R.V.L. (Ralph) (1928). “Transmission of Information.” *Bell System Technical Journal*, Vol. 7, No. 3, pp. 535-63.

Heisenberg, Werner (1957). “Quantum Theory of Fields and Elementary Particles.” *Reviews of Modern Physics*, Vol. 29, No. 3, pp 269-78.

Norton, Melanie J. (2010). Introductory Concepts in Information Science. Second edition. ASIS&T Monograph series. Medford, New Jersey: Information Today, Inc.

Otten, Klause and Anthony Debons (1970). “Towards a Metascience of Information: Informatology.” *Journal of the American Society for Information Science*, January-February 1970, pp. 89-94.

Rashevsky, Nicolas. (1955) “Life, Information Theory, and Topology.” The *Bulletin of Mathematical Biophysics,* Vol. 17, No. 33, pp. 229-235.

Quastler, Henry (1964). The Emergence of Biological Organization. New Haven, Connecticut: Yale University Press.

Shannon, C. E. (Claude) (1948). “A Mathematical Theory of Communication.” *Bell System Technical Journal*, Vol. 27, pp 379-423 and 623-656, July and October, 1948.

Seiful-Mulukov, R.B. (Rustem) (2010). Petroleum: Hydrocarbon Sequences – analyses of the models, their origin and evolution. Moscow: 11th Format. ISBN 978-5-902030-80-5 (in Russian)

Sokolov, I.A., ed. (2009). Informatics: State of the Art, Problems and Prospects. Moscow: Institute for Informatics Problems, Russian Academy of Sciences. ISBN 978-5-902030-69-0 (in Russian)

Steinbuch, Karl (1957). Informatik: Automatische Informationsverarbeitung. (Technische Mitteilunger der Standard Elektrik Gruppe. Berlin: SEG-Nachrichten.

Taylor, Robert S. (1966). “Professional Aspects of Information Science and Technology.” In Annual Review of Information Science and Technology, Vol. 1, ed. Carlos A. Cuadra. New York: John Wiley & Sons, pp. 15-40.

Trucco, Ernesto (1956). “A Note on the Information Content of Graphs.” The Bulletin of Mathematical Biophysics, Vol. 18, No. 2, pp. 129-135.

Zeilinger, Anton (1999). “A Foundational Principle for Quantum Mechanics.” *Foundations of Physics*, Vol. 29, No. 4, pp. 631-43.

MODEL TRANSFORMATION CHAINS AS STRATEGY FOR SOFTWARE DEVELOPMENT PROJECTS

Héctor A. FLOREZ FERNANDEZ
Universidad Distrital Francisco Jose de Caldas
Bogotá, Colombia

ABSTRACT

This paper presents a description about the software development based on model transformation chains. This process begins with the creation of a domain metamodel. Later, a model-to-model transformation is applied to this metamodel getting architecture metamodel. The final step in the process consists in the application of a model to text transformation getting the final source code of the application. Also, this paper presents a case study where is demonstrated the complete process about the model transformation chain.

Keywords: Model, metamodel, model transformation, UML, software engineering, model to text transformation, code generation.

1. INTRODUCTION

In different contexts involving the development of software products, it is possible to see that generally, most of the functional requirements posed around the same context, satisfy the same set of features. Typically, software developers make a code re-use generated in previous projects that resolves certain requirements in order to improve development time of a particular product. This code re-use is an immediate solution that does not necessarily solve the problem on a new project due to the specific situations in the new project. These situations in almost all cases have differences with the implementation made in previous projects, yet this kind of solution demands a lot of time and effort by the development team. In this way, is a better alternative build metamodels allowing the abstraction of the general features of a particular context to make transformation processes.

An initial proposal but enough is by creating a context metamodel which includes concepts related to specific domain entities to be resolved, a metamodel architecture which must be platform independent and would abstract the different architectural elements that may have a solution for the domain, and a metamodel related to the platform on which the final implementation of the software product is made. Based on the metamodels and transformations, when a develop for a new project is required, would be enough to make a design for a new metamodel based on the context, which would apply different transformations necessary to provide results in the source code that implements elements, features and general services for that particular project context. In this paper, a description of the processes of models transformation is presented, using EMF (Eclipse Modeling Framework) for modeling, ATL (Atlas Transformation Language) to model transformation and Acceleo for model to text transformation. Finally, the results of a case study in the domain of maze games is presented, which it is possible to verify the final product fully functional, based on a model of initial context.

2. MODELS AND METAMODELS

A model is a set of statements about a system which is under study. These statements are relating to the abstraction of an element that can be considered true or false. In software engineering, a model refers to an artifact developed in a modeling language, which describes a system through different types of diagrams. Stachowiak stated that a model requires three key features: *mapping*, indicating that a model is based on an original concept; *reduction*, indicating that the model reflects only a selection of original and *pragmatic concept*, indicating that the model needs to be used taking the specific purpose of the original concept [1]. On the other hand Bran Selic states that an engineering model must contain the following characteristics: abstraction, indicating that the model is always a reduced representation of a system that it represents; understanding, indicating that the model must maintain a way that directly appeals to the intuition; precision, indicating that the model should provide a true representation of the interesting features of the modeled system; predictive, indicating that the model is to predict the interest but not the obvious properties of the modeled system and low expression, indicating that the model must be economically significant for the construction and analysis. [2]

A copy of the original concept is not a model. A model is a description of something. The software engineering models are created in the analysis phase and are used to describe the problem. Additionally, all models can be transformed and not everything is necessarily convertible to a model [3]. The modeling is a human activity, because every action is preceded by a construction of a model [4].

There are two different kinds of models are:

- Token model. These models capture elements with specific and particular aspects of an item.
- Type model. Capture the universal aspects of the original concept. Models can be considered as a model and is not appropriate to consider them as metamodels.

The relationship between a model and its metamodel can be characterized as an instance; however, can be expressed best as a model conforms to the metamodel. The interpretation of a model can be obtained by mapping the model elements with elements from the original concept, which can determine the level of accuracy in the statements with the elements.

Model Driven Engineering (MDE), not only as a model using simple documentation. MDE uses models as an input and / or formal output computer-based tool implementing specific operations [4]. As a consequence, engineering frameworks have evolved progressively model MDA (Model Driven Architecture) defined by the OMG (Object Management Group).

The main use of a metamodel is to facilitate the separation of concerns of a domain. When several models have been extracted from the same system with different metamodels, these models are related [4].

The following figure shows the classical architecture of the OMG. In the low level, M0 layer is the actual system. The level M1 represents the model. The M1 level model is conformed to the metamodel defined at the M2 and it is conformed to the meta metamodel of the level M3. The meta metamodel is conformed to itself.

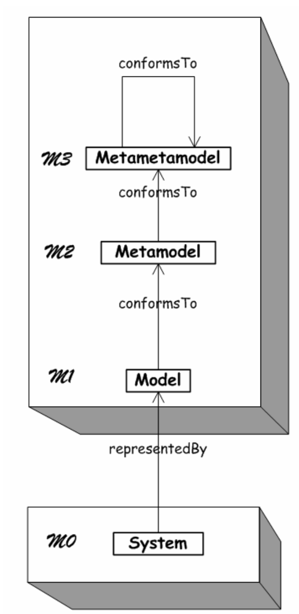


Fig. 1. MDA Levels [4]

3. RESTRICTIONS

At the time of building a metamodel, the model that conforms to this must attend all characteristics defined. However, in many cases, the metamodel cannot specify certain characteristics as constraints obtained by abstraction of original concept that is modeled.

Object Constraint Language (OCL) is a language that allows set restrictions on metamodels. The model that tries to be consistent with the metamodel that contains such restrictions must satisfy these restrictions. Thus, applying to the metamodel a few OCL sentences, it is possible to define a set of additional features to the metamodel.

OCL is a modeling language and not a programming language. Cannot write a program with some logic or flow control in OCL. As modeling language, all implementation issues are out of reach and cannot be expressed in OCL. OCL queries defined, on the one hand, as a set of elements that satisfy a restrictive expression, and secondly, it is possible to apply special operators to interact with the relations [7].

The OCL expressions necessarily must be connected with some element of the metamodel. To do this, there are two keywords that are "context" that indicates the context in which the expression is applied and "self" making reference in an expression context.

An OCL expression, which can be invariant, should be true for all models conform to the metamodel to which is applied the restriction. Likewise, an expression can contain preconditions and postconditions in a condition associated with a method. In this way, it is necessary to satisfy the condition specified in the postcondition if the precondition is met.

OCL contains a large set of functions and operators that allow setting any restrictions on a metamodel that must be satisfied by the model conformed to the metamodel.

4. MODEL TRANSFORMATION

A model transformation, allows converting a model Ma that is conformed to a metamodel MMa to a model Mb that is conformed to a metamodel MMb. The transformation is defined by a transformation model Mt that is conformed to a transformation metamodel MMt. MMa and MMb metamodels must have direct relationships between concepts to allow the solution to the requirements established for a particular project. The three metamodels must be conformed to with a meta metamodel such us MOF (Meta Object Facility).

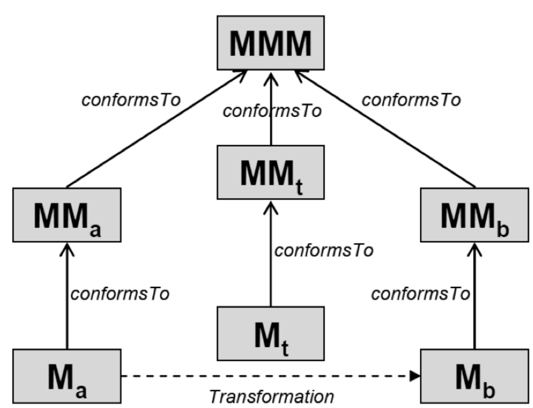


Fig. 2. Model transformation [8]

To perform transformations between models, it is necessary to use a specialized language for this task. Two common languages are Atlas Transformation Language (ATL) and Query View Transformation Operational (QVTO).

ATL allows transformation by defining modules. A module must contain a name, an output metamodel and an input metamodel. Furthermore, a module can define rules. Each rule must contain a name, the class from the input metamodel and one or more classes from the output metamodel. Each rule allows assigning the attribute information of each class of output metamodel.

QVTO allows transformations by creating functions of mapping models. It is necessary to include metamodels through the keyword “modeltype”. The execution of the transformation is initiated through a main function in which it can appeal to the mapping functions required for transformation.

5. MODEL TO TEXT TRANSFORMATION

Model to text transformation allows generating source code in any language. The transformations generate text files with the desired content and desired extension. The model to text transformation always has an input model and an input metamodel. The output is the set of source code files generated in the language selected to deploy the final application.

Acceleo is a project that allows this kind of transformations. With this tool a transformation requires to include the metamodel with which generating the text. Acceleo allows to build a set of "templates" that are the components where is implemented the decisions of the transformation. It also has a statement @main indicating which templates will be run. If a template does not contain the statement @main, it must be called since another template for its execution.

6. MODEL TRANSFORMATION CHAIN

The strategy used for building a model transformation chain is shown with the next figure.

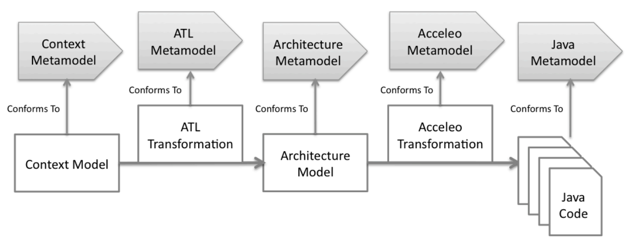


Fig. 3. Model transformation chain strategy

To build the model transformation chain is necessary to create the next artifacts:

- Context metamodel. This metamodel contains all elements related with de specific domain.
- Architecture metamodel. This metamodel contains all elements from the context metamodel adding some elements related with the architecture pattern Model View Controller.
- Context model. This model has to be conformed to the context metamodel. This model models a specific project in the domain defined in the context metamodel
- ATL Transformation. This artifact has the context metamodel as input and generates the architecture model which is conforms to the architecture metamodel. This transformation has the necessary rules to convert the elements from the context model adding the additional characteristics defined in the architecture metamodel.
- Acceleo Transformation. This artifact has the architecture model as input. This artifact has a set of rules that convert the elements in the architecture model to a source code in any language.

With this propose, the context model and the architecture model are platform independent, yet the files generated by the acceleo transformation are platform dependent. As a consequence, it is necessary to create a different acceleo transformation for each platform in the case of generates the project in different programming languages.

7. CASE STUDY

There are some games that share certain characteristics that allow propose a transformation chain for getting a semiautomatic construction of those games. As a case study is taken maze games, which have a single player, passive elements such as block and active elements such as enemies.

Based on the above context metamodel poses a maze game that includes the general common characteristics to this kind of games.

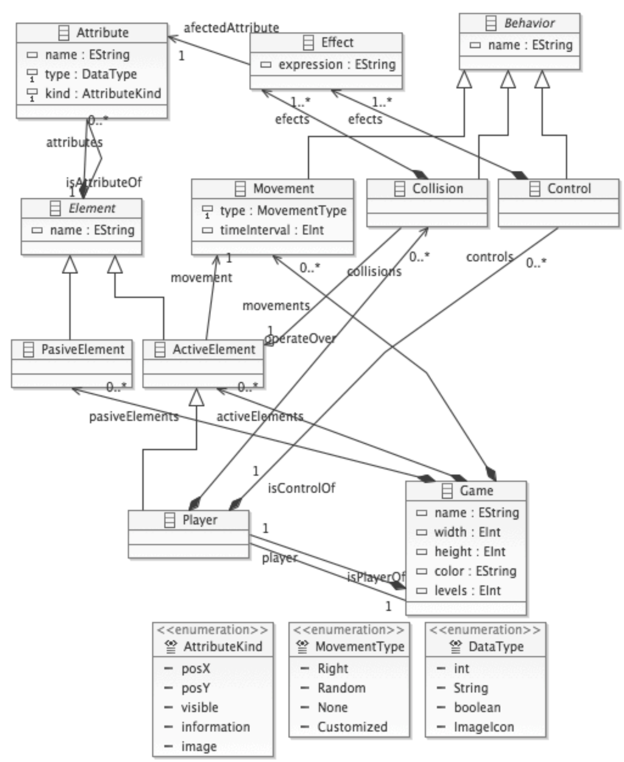


Fig. 4. Context metamodel for maze games

This context metamodel contains the next EClasses:

- Game. It is the root EClass. It has containment relations with the other EClasses.
- Element. It is an abstract EClass. It has a containment relation whit attribute.
- Attribute. It has a data type that could be int, String, boolean or ImageIcon. It has an attribute kind that could

be posX indicating that this attribute is related with the x coordinate, posY indicating that this attribute is related with the y coordinate, visible, information indicating that this attribute has information about the element in the game and image indicating that this attribute has the path of the ImageIcon.

- PasiveElement. It inherits from Element
- ActiveElement. It inherits from Element. It has one relation with the EClass movement.
- Behavior. It is an abstract EClass
- Movement. It inherits from Behavior. It has one type that could be Right which implies that the active element will turn right when it touch a passive element, Random which implies that the active element will turn in some direction when it touch a passive element, None which implies that the active element will do nothing and Customized which implies that the movement can be programmed in the final application.
- Collision. It inherits from Behavior. It makes an effect when the player collides with an active element. It generates an effect in one attribute of the active element collided.
- Control. It inherits from Behavior. It makes an effect when the user presses a button.
- Effect. It changes the value of one attribute of one active element.

Based on the context metamodel and applying the architecture pattern Model View Controller, the next architecture metamodel is generated.

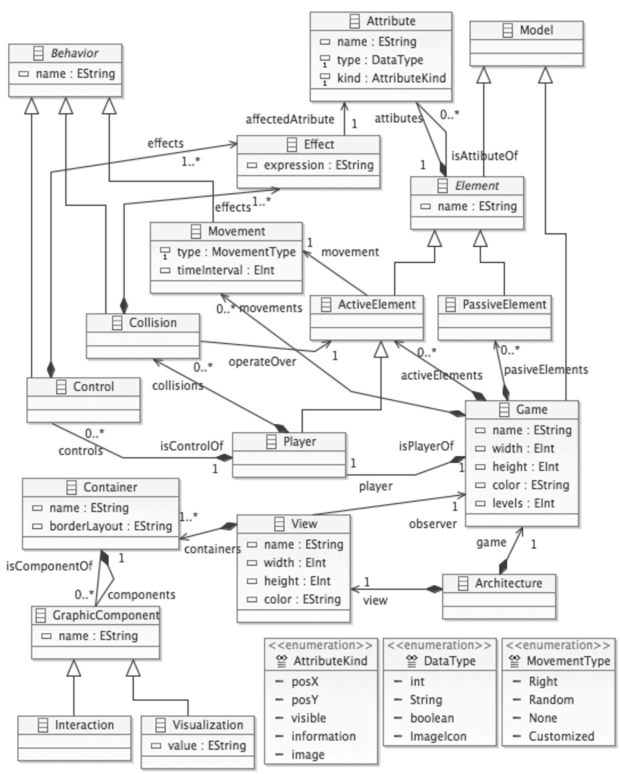


Fig. 5. Architecture metamodel for maze games

This metamodel has the same elements than the context metamodel, yet it has the next additional EClasses:

- Architecture. It is the root EClass. It has containment relations with the EClass Game and View.
- View. It has containment relations with the EClasses related with the view in the pattern Model View Controller.
- Container. It makes reference to an graphic container to deploy the final application
- GraphicComponent. It is an abstract EClass.
- Interaction. Is inherits from GraphicComponent. It makes reference to an interaction element such us a button
- Visualization. It inherits from GraphicComponent. It makes reference to a visualization element such as a label.

Based on the context metamodel, the model of the figure 6 is created. This model represents the game pacman. There are the next elements in this game:

- Blocks. This is a passive element. It represents the set of blocks in the game.
- Pacman. It is the player. It has the next attributes:
  - xPlayer. It is the variable, which contains the x coordinate of the player.
  - yPlayer. It is the variable, which contains the y coordinate of the player.
  - Image. It is the variable, which contains the path of the imageIcon of the player.
  - Points. It is the variable, which contains the number of points. One point is equivalent of on cookie eaten.
  - Lives. It is the variable, which contains the number of opportunities that has the player. When this is 0, then game is over.

  The player also has the next behaviors:
  - Up. It is the control, which the user can move the player up.
  - Down. It is the control, which the user can move the player down.
  - Right. It is the control, which the user can move the player right.
  - Left. It is the control, which the user can move the player left.
  - GhostTouch. It is the collision that occurs when the player touches an enemy. This collision decrements the number of lives.
  - CookieTouch. It is the collision that occurs when the player touches a cookie. This collision increments the number of points.
  - FruitTouch. It is the collision that occurs when the player touches a fruit. This collision increments the number of lives.
- Ghost. This is an active element. It represents the set of ghosts in the game.
- Cookie. This is an active element. It represents the set of cookies in the game.
- Fruit. This is an active element. It represents the set of fruits in the game.
- Right. It is a preconfigured movement that can be applied to an active element. This movement turns right the active element when it collides with a passive element or with a boundary of the game.
- Random. It is a preconfigured movement that can be applied to an active element. This movement turns wherever the active element when it collides with a passive element or with a boundary of the game.
- None. It is a movement that can be applied to an active element. When an active element has this movement, it does not have any variation in the position.

In this case, there is a properties file, which includes the coordinates and the image of each block, ghost, cookie and fruit in the game.

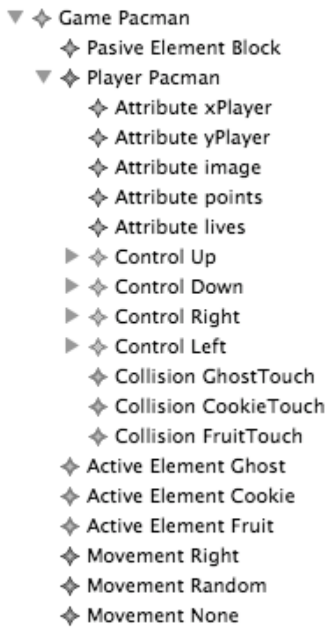


Fig. 6. Context model for pacman game

Applying the model transformation and the model to text transformation, the source code in java is generated. For this case, executing the final application the game is totally ready to play. Also, the game generated has two levels with different elements. The figure 7 shows the level one and the figure 8 shows the level two. The elements of each level depend of a configuration file that describes the coordinates and images.

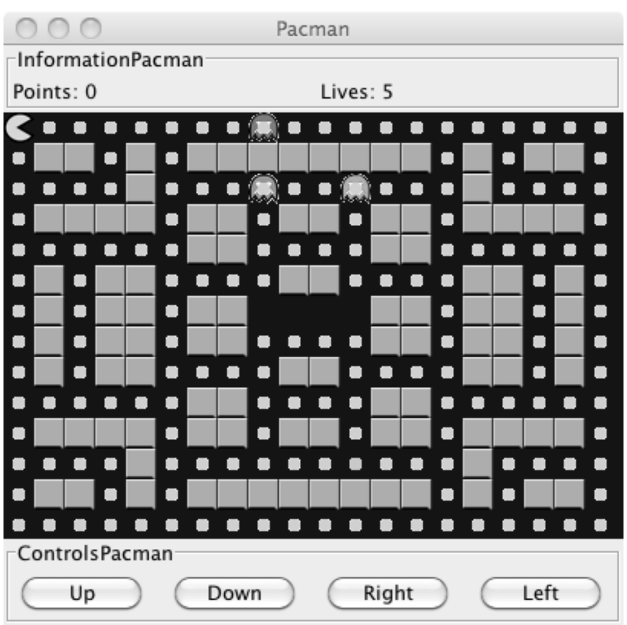


Fig. 7. Final project of the Pacman game, level one

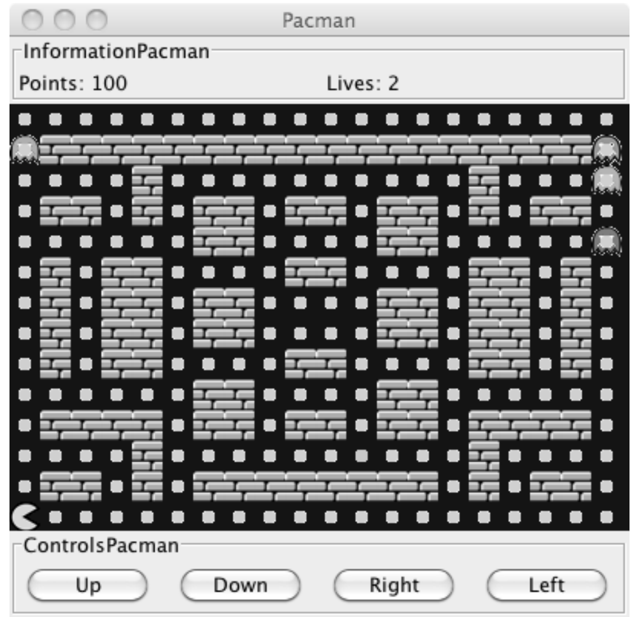


Fig. 8. Final project of the Pacman game, level two

For this project, there is a context model related with the pacman game with 21 elements. After applying the model transformation chain, the result is a source code with 10 java files and 825 code lines in total.

8. CONCLUSIONS

A model transformation chain offers advantages related to the development time of software products that are part of the same context, allowing the development team to concentrate on the analysis and design of the product and the particular characteristics that must be implemented based on the infrastructure generated by the chain.
A model transformation allows adding the project a major set of components that make reference not only context but also refers to the visualization and control.

Model to text transformations allows the generation of plain text, which can be interpreted as source code, providing the opportunity to generate the same final product in different languages without affecting the initial design based on modeling from the context of project.

The application of a transformation chain allows semi-automatically generate a final product of software, because based on the generated elements it is necessary to implement the specific characteristics of each project.

9. REFERENCIAS

[1] Kuhne Thomas. What is a Model?

[2] Muller Pierre-Alain. Modeling Modeling

[3] Ed Seidewitz. What models mean

[4] Bezivin Jean. On the Unification Power of Models

[5] Bezivin Jean. MDA Components: Challenges and opportunities

[6] Steinberg Dave, Budinsky Frank, Paternostro Marcelo, Merks Ed. EMF: Eclipse Modeling Framework

[7] Casas Juan, Arenas Mercedes. OCL (Object Constraint Language)

[8] ATLAS group. ATL: Atlas Transformation Language

[9] Laarman Alfons. Achieving QVTO & ATL Interoperability, An Experience Report on the Realization of a QVTO to ATL Compiler.

Pattern-Based Development of Enterprise Systems: from Conceptual Framework to Series of Implementations

Dr. Sergey V. Zykov, Ph.D.
National Research University Higher School of Economics, Moscow, Russia
szykov@hse.ru

Abstract

Building enterprise software is a dramatic challenge due to data size, complexity and rapid growth of the both in time. The issue becomes even more dramatic when it gets to integrating heterogeneous applications. Therewith, a uniform approach is required, which combines formal models and CASE tools. The methodology is based on extracting common ERP module level patterns and applying them to series of heterogeneous implementations. The approach includes a lifecycle model, which extends conventional spiral model by formal data representation/management models and DSL-based "low-level" CASE tools supporting the formalisms. The methodology has been successfully implemented as a series of portal-based ERP systems in ITERA oil-and-gas corporation, and in a number of trading/banking enterprise applications for other enterprises. Semantic network-based airline dispatch system, and a 6D-model-driven nuclear power plant construction support system are currently in progress.

1. Introduction

The objective of the paper is a systematic outline of the new technology for developing large-scale integrated heterogeneous applications. Currently, the multinational enterprises possess large, geographically distributed infrastructures, aimed at the same business goals. Each of the enterprises has accumulated a tremendous and rapidly increasing data burden, comparable to an avalanche. In certain cases, the data bulk exceeds petabyte size, and it tends to double every five years. Undoubtedly, management of such data is a serious challenge. The problem becomes even more complicated due to heterogeneous nature of the stored data, which varies from well-structured relational databases to non-normalized trees and lists, and to weak-structured multimedia data. The technology presented in the paper is focused at more efficient heterogeneous enterprise and uniform data management procedures. The technology involves a set of novel mathematical models, methods, and the supporting software engineering tools for object-based representation and manipulation of heterogeneous enterprise data.

2. Heterogeneous Enterprise Systems: Concepts, Problems, Models and Tools

Brute force application of the so-called “industrial” enterprise software development methodologies (such as IBM RUP, Microsoft MSF, Oracle CDM etc.) to heterogeneous enterprise data management, without an object-based model-level theoretical basis, results either in unreasonably narrow “mono-vendor” solutions, or in inadequate time-and-cost expenses. On the other hand, the existing generalized approaches to information systems modeling and integration (e.g., category and ontology-based approaches, Cyc and SYNTHESIS projects – [1-4] – do not result in practically applicable (scalable, robust, ergonomic) implementations since they are separated from state-of-the-art industrial technologies (CASE, RAD etc.).

Thus, the suggested technology of integrated development and maintenance of heterogeneous internet-based enterprise software systems has been created. The approach is based on rigorous mathematical models and it is supported by software engineering tools, which provide integration to standard enterprise-scale CASE tools, commonly used with software development methodologies. The approach eliminates data duplication and contradiction within the integrated modules, thus increasing the robustness of the enterprise software systems (ESS). The technology integrates a set of ESS development levels: data models, software applications, “industrial” methodologies, CASE, architecture, and DB management. The novel technology elements are: (i) conceptual framework of ESS development; (ii) a set of object models for ESS data representation and management; (iii) CASE tools for semantic-oriented ESS development (ConceptModeller) and intelligent content management (ICMS); (iv) ESS implementations [5]. For adequate modeling of heterogeneous ESS, a systematic approach has been developed, which includes object models for both data representation and data management [5,6]. The general technological framework of ESS development provides closed-loop, two-way construction with re-engineering.

The general technological framework of ESS development contains stages, which correspond to data representation forms for heterogeneous software system components, communicating in the global environment.The object nature of the "class-object-value" model framework provides compatibility with traditional object-oriented analysis and design approach (OOAD), as well as with other certain promising approaches (such as D.S.Scott's variable domains [7], V.E.Wolfengagen's conceptual method [8]) and helps to extend the mentioned approaches to model the ESS internet-based environments. The following technological transformation sequence is suggested: (i) a finite sequence object (e.g., a lambda calculus term); (ii) a logical predicate (higher order logic is used); (iii) a frame (as a graphical representation); (iv) a XML object (class definition generated by the ConceptModeller engineering tool); (v) an UML diagram (CASE tool data scheme) in the ESS (meta)data warehouse [5].

Therewith, the warehouse content representation is based on semantic network situation model, which provides intuitive transparency for problem domain analysts when they construct the problem domain description. The model can be ergonomically visualized through a frame-based notation. Warehouse content management is modeled as a state-based abstract machine and role assignments, which naturally generalizes the processes of similar engineering tools, such as (portal page template generation, portal page publication cycle, role/access management etc. Therewith, the major content management operations (declaration, evaluation, personalization etc.) are modeled by the abstract machine language. The language has a formal syntax and denotation semantics in terms of variable domains. The transformation sequence of the model is: (i) a term of variable domain algebra (D.S.Scott's computations theory is used)[7]; (ii) a domain-based function (higher order logic is used) [7]; (iii) a frame (a graphical notation); (iv) a XML object (a template for a ICMS portal page); (v) HTML code (ICMS portal page code) of the ESS portal.

The architecture of the integrated heterogeneous enterprise content warehouse provides unification due to generalized object association-based relationships at the data at metadata levels. Uniform heterogeneous ESS content management is based on a uniform portal foundation, which serves a meta-level enhancement over the enterprise data warehouse. Assignments act as code scripts; they change ICMS machine states, and provide dynamical, scenario-driven content management.

The ConceptModeller tool assists in semantically-oriented visualized development of heterogeneous ESS data warehouse scheme [6]. A semantic network-based model is suggested, which works in nearly natural-language terms, intuitively transparent to problem domain analysts. Model visualization is based on frame representation of the ESS data scheme. Deep integration with mathematical models and ESS CASE tools provides a closed-loop, continuous lifecycle with reengineering. The ICMS tool is based on an abstract machine, and it is used for problem-oriented visualized heterogeneous ESS content management and portal publication. ICMS features a flexible scenario-oriented management cycle and role-based mechanisms. ICMS provides a unified portal representation of heterogeneous (meta)data, flexible content processing by various user groups, high security, ergonomics and intuitively transparent complex data object management.

3. Pattern-Based Development for Enterprise Systems

The general ESS development framework [5] potentially allows the following benefits: (i) applying a "spiral-like" lifecycle to the general ESS development framework; (ii) ESS "tuning" by applying a "spiral-like" lifecycle and subsequent verification; (iii) requirement "tracing"; (iv) building a repository of ESS "meta-snapshots", with which the system and/or warehouse could be "reincarnated" to virtually any previous state using component-wise strategy; (v) building a "pattern catalogue" [9] for heterogeneous ESS, based on the integrated repository of various ESS state "meta-snapshots"; (vi) developing a repository of "branches" for "cloning" slight ESS variations for the "basis; (vii) developing a formal language specification (e.g, a DSL technology-based one) [11]; (viii) adjusting" the existing ESS "meta-snapshot" repository components to match requirements; (ix) reuse of the desired components.

The preferable ESS development framework tends to be iterative; in certain cases waterfall is an option.

An essential feature of the general ESS development framework is its two-way organization. The approach provides reverse engineering possibility both for ESS in general, and their components in particular. The practical value of the approach is provided by the verifiability of heterogeneous ESS components at the uniform level of the problem domain model, which is practically independent upon the hardware and software environment of the particular component. Therewith, a major theoretical generalization is a possibility of mathematically rigorous verification of the heterogeneous ESS components by a function-based model [7,10]. The ESS engineering models are oriented at a promising "pure" objects approach, which is a strategy of .NET and Java technologies, where any program entity is an object.

An essential benefit of the approach suggested is a possibility of adaptive, sequential "fine tuning" of ESS heterogeneous component management schemes in order to match the rapidly changing business requirements.

Such benefit is possible due to the reverse engineering feature of the integrated general iterative framework of ESS development. The reverse engineering is possible down to model level, which allows rigorous component-wise ESS verification. Thus, conventional reengineering and verification can be enhanced by flexible correction and "optimization" of the target ESS in strict accordance with the specified business requirements. This is possible due to the suggested model-level generalization of the iterative, evolutionary ESS development framework. Another benefit of the suggested ESS development framework is a possibility of building a "catalogue of templates for heterogeneous ESS", which is based on an integrated metadata warehouse, i.e., a "meta-snapshot" repository. Thus, the software development companies get a solution for storing relatively stable or frequently used configurations of heterogeneous enterprise software systems. The solution potentially allows avoiding the integration problems of "standard" ESS components and/or combinations, which have been obtained previously. The approach allows serious software engineering project savings for clients, provided the ESS developer's "meta-snapshot" repository already stores a similar or an analogous integrated solution to the system required. The above consideration clears the way for "meta-snapshot" repository development, which stores the chronological sequence of ESS solutions as a tree with the "baseline" and slight variations of ESS "branches".

This is similar to version control CASE tools. The approach allows a reasonable selection of most valuable deliverables of the ESS lifecycle phases, and organization of similar solution "cloning". Therewith, the "clones" may be created both for different client enterprises, and for different companies of a single enterprise.

Further discussion could cover the prospective areas of "meta-snapshot" repository development. First of all, to describe the metadata warehouses and the related enterprise-level business requirements it seems reasonable to develop new DSL-type problem-oriented meta-languages. Let us call them the MetaWarehouse Description Language (MWDL) and the Requirement Specification Language (RSL) respectively. Further, the formal models, outlined in the paper and given a more detailed coverage [5], allow interrelating the RSL and MWDL entities. Semantic-oriented search mechanisms assist in revealing ESS "meta-snapshot" repository components, which provide the closest matching to the new requirements. The approach potentially allows terms-and-cost-effective and adequate transforming of the existing ESS components in order to match the new requirements with minimum corrections effort and, consequently, with minimum labor expenses. Therewith, the global perspective it becomes possible to reuse certain ESS components for current or new clients. Selection criteria for such "basic" components may be percentage of reuse, ease of maintenance, client satisfaction, degree of matching business requirements etc.

4. Implementations: Results Obtained and Projects-In-Progress

ITERA Oil-and-Gas Group: a Portal-Based Solution. The suggested methodology has been practically approved by development of Internet and Intranet portals in ITERA International Group of Companies. During the design stage, problem domain model specifications are transformed by the innovative ConceptModeller SDK to UML diagrams, then by Oracle Developer/2000 integrated CASE tool – to ER diagrams and, finally, into target IS and enterprise content warehouse storage schemes.

Using the suggested data model, the architectural and interface solution has been customized for enterprise resource management IS with content personalization for a wide spectrum of user and administrator types.

To provide the required industrial scalability and fault tolerance level, the integrated Oracle design and implementation toolkit has been chosen to support UML and business process reengineering.

A set of models have been constructed including problem domain conceptual model for enterprise content dynamics and statics as well as a model for development tools and computational environment in terms of state-based abstract machines, which provide integrated object-based content management in heterogeneous enterprise portals. For the model set, a generalized development toolkit choice criteria set has been suggested for information system prototyping, design and implementation. A set of SDKs has been implemented including ConceptModeller visual problem oriented CASE-tool and the CMS. According to the approach, a generalized interface solution has been designed for Internet-portal, which is based on content-oriented architecture with explicit division into front-end and back-end sides. Portal design scheme is based on a set of data models integrating object-oriented methods of management of data and metadata (or knowledge).The major implementations of portals in ITERA Group were: CMS for network information resources, official Internet site, and enterprise Intranet portal.

Distributed Trading Company: a Domain-Driven Messaging System. A trading corporation used to commercially operate a proprietary Microsoft .NET-based message delivery system for information exchange between the headquarters and the local shops. The system was client-server based. The client included a local database and a Windows-based messaging service, while the server side consisted of a Web service and central database. The operation/maintenance challenges were: complicated client-side code refactoring; difficult error localization/reduction; inadequate documentation; and decentralized configuration monitoring/management for remote shops. To solve the problems mentioned, an approach based on domain-driven development [11] and Domain Specific Languages (DSL) has been suggested. The approach included problem domain modeling and problem domain DSL development. A XML-based DSL was used, which extended the scope of the ESS programming language. The methodology instance included DSL scope detection, problem domain modeling, DSL notation/constraint development, and DSL testing. The lifecycle model was iterative with reengineering support. The Windows service containing a DSL parser is a part of the application. The DSL parser input is a current message transfer map.

The DSL scope included message transfer rules/parameters, and adding new types of messages. The next methodology stage was building semantic model of the objects handled by DSL. Three types of the objects were revealed: messages, message transfer channels and message transfer templates. DSL describes object metadata, i.e., configurations and manipulation rules. Templates were core elements of the model, and channels were links between template instances. Templates and channels together make message maps. DSL described the maps, i.e. the static part of the model, while messages referred to its system dynamics and store the state. Templates define actions with messages, i.e. transform or route them. Based on DSL class model and implementation, messaging maps were built, which were later used by parser to generate system configuration. At this stage, DSL syntax and semantics were built. Messaging map was built as an XML document, defining system configuration and containing templates for routing, message processing, transfer channels and their relationships. The parser creates channel objects based on DSL channel descriptions. Then it configures the messaging system by creating message processing objects similarly. Finally, it instantiates the I/O channels, and creates relationships between channels and message processor.

Thus, the DSL-based refactoring resulted in an enterprise trade management system with transparent configuration and a standard object-based model. The DSL developed solved the problem of messaging management. Since changes are chiefly localized within the transfer configuration /map, the change management has been dramatically simplified. The DSL-based methodology instantiation assisted in conquering complexity, made the proprietary system an open, scalable, and maintainable solution. The approach can easily be customized to fit a broad class of similar proprietary systems.

Air Transportation Planning System. Air traffic planning system is an area of work-in-progress.

The problem is to develop remote access to the planning data. An operating solution currently exists. However, it is based on an outdated TAXXI-Baikonur technology, which is no longer evolving after early 2000s. The technology involves component-based visualized assembling of the server application. The ready-made VCL library components from Borland had been integrated with proprietary TAXXI components. The client side is TAXXI Communicator, i.e. an XML browser, which is a "thin" client. The TAXXI technology is limited Windows framework, which is the only possible basis for both client and server-side applications. According to the State Program of Planning System Updates, the Main Air Traffic Management Centre is going to create the new remote access solution. The internet-based architecture is to be implemented in Java technology and to operate on the Apache web server platform. The solution is to query Oracle-based data centre, process the query output and retrieve the results of the air traffic planned capacities to an intuitive and user-friendly GUI. The practical application of the solution is the global integrated ESS, which is providing a uniform and equal information access to all of the international air traffic participants. The suggested pattern-based and component-wise approach is going to unify the issues of the architecture-level update and application migration in Russia. The methodology will also simplify the integration challenges of the global air traffic management software solution.

Nuclear Power Plant: Approaching a 6D-Model Based Implementation. Another challenging aspect of the methodology implementation is related to high-level template-based software re-engineering for nuclear power plants (NPP). To provide worldwide competitiveness, it is necessary to meet the following requirements: (i) meeting quality standards throughout the lifecycle; (ii) high security under long-term operation; (iii) term-and-cost reduction for new generation facilities development. The approach combines state-of-the-art production potential, advanced control methods, and software engineering tools. Each stage of the NPP lifecycle is mapped into a set of business processes, where not only people, but also enterprise systems are interacting.

Identifying operation sequences, the systems form business process automation standards. For example, workflow mechanisms can assist in building enterprise standards on electronic documents validation and approval. During a certain NPP lifecycle, the enterprise systems acquire information on it. Finally, each of the enterprise systems reveals certain NPP aspects: design, technology, economics etc. Thus, various objects, the systems together describe NPP as a huge object. Heterogeneous nature of the data objects, and a million number of units make NPP a high complexity information object. A major competitiveness criterion in nuclear power industry is a set of electronic manuals, which helps to assemble, troubleshoot, repair NPP etc. Such manual set provides transparent information models of NPP (units), which allow getting information on the object without directly contacting it. Such a versatile description, combined in a single data model is often referred to as a 6D model, which includes 3D-geometry, time and resources for operating the plant. Since mechanisms for information searching, scaling, filtering and linking, should provide complete and non-contradictory results, the information models should have well-defined semantics. The uniqueness of data entry assumes information model data acquisition by the enterprise systems throughout the lifecycle. While a single information model can be derived out of a single system, the 6D model should combine information models of a number of systems. The methodology for building a 6D model suggests portal-based system integration, which can be based on a “platform” capable of entire lifecycle support. The further information model development assumes monitoring system state changes and their influence to the other parts of the system. This helps to immediately react on critical issues in NPP construction, which can be used for decision making. A major nuclear industry challenge is a typical optimized nuclear reactor. The idea is to select typical invariant units for rapid “template-based” development of a set of slightly varying versions. Applying the methodology to the 6D information model of the nuclear reactor, is promising for pattern-based component-wise NPP series development.

5. Conclusion

Implementation of the suggested approach allowed to developing a unified ESS, which integrates a number of heterogeneous components: state-of-the-art Oracle-based ERP modules for financial planning and management, a legacy HR management system and a weak-structured multimedia archive. The implementation of internet and intranet portals, which manage the heterogeneous ESS warehouse content, provided a number of successful implementations in diversified ITERA International Group of companies (approx. 10,000 employees in over 20 countries). The systematic approach to ESS framework development provides integration with a wide range of state-of-the-art CASE tools and standards of ESS development. Other implementations and work-in-progress areas include: air transportation planning system, messaging system for a trading enterprise, a nuclear power plant and banking solutions. Each of the implementations is a domain-specific one, so the system cloning process is not straightforward, and it requires certain analytical and CASE re-engineering efforts. However, in most cases the approach reveals patterns for building similar implementation in series, which results in substantial term-and-cost reduction of 30% and over. The series can be applied both to subsidiaries and to different enterprises.

References

[1] Guha R., Lenat D.: *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project*. Addison-Wesley, 1990

[2] Lenat D., Reed S. Mapping Ontologies into Cyc. *In: AAAI 2002 Conference Workshop on Ontologies for the Semantic Web*, Edmonton, Canada, 2002

[3] Birnbaum L., Forbus K., et al.: Combining analogy, intelligent information retrieval, and knowledge integration for analysis: A preliminary report. *In: ICIA 2005*, McLean, USA, 2005

[4] Kalinichenko L., Stupnikov S.: Heterogeneous information model unification as a pre-requisite to resource schema mapping. In: *ITAIS 2009*, Springer, 2009, pp. 373-380

[5] Zykov S.: Integrated Methodology for Internet-Based Enterprise Information Systems Development. *In: WEBIST2005*, Miami, FL, USA, May 2005, p.p.168-175

[6] Zykov S.: Pattern Development Technology for Heterogeneous Enterprise Software Systems. *Journal of Communication and Computer*, 7 (4), 2010, pp.56-61

[7] Scott D.: *Lectures on a mathematical theory of computations*. Oxford Computing Laboratory Technical Monograph. PRG-19, 1981, 148 pp.

[8] Wolfengagen V.: Event Driven Objects. *In: CSIT'99*, Moscow, Russia, 1999, p.p.88-96

[9] Fowler M.: *Analysis Patterns: Reusable Object Models*, Addison Wesley, 1997, 223 pp.

[10] Curry H., Feys R.: *Combinatory logic*, Vol.1, North Holland, Amsterdam, 1958

[11] Evans E.: *Domain-Driven Design: Tackling Complexity in the Heart of Software*. Addison Wesley, 2003, 560 pp.

Learning to Program: a game or a boogeyman

Adão C. CAMBRAIA
Instituto Federal Farroupilha – Campus Santo Augusto, cambraia@sa.iffarroupilha.edu.br
Santo Augusto, RS, Brasil

and

Maria Angélica F. OLIVEIRA
Instituto Federal Farroupilha – Campus Santo Augusto, mariaangelicafo@gmail.com
Santo Augusto, RS, Brasil

ABSTRACT

This paper presents the problem of the programming learning and presents the use of Scratch programming language to ease the process of learning to program, linked to the knowledge production of other disciplines, using Piaget as the background for the reflections. The Scratch language was first used in Programming I discipline in the Integrated Computer Science Course at Instituto Federal Farroupilha - Campus Santo Augusto and, as a result of the initial work, a huge involvement of the students is perceived, making learning more attractive and changing that "bogey" look of the program.

KeyWords: Programming Learning, Scratch, Integration, professionalizing education, logical reasoning.

1. INTRODUCTION

The teaching and learning process of programming is usually seen by many students of the computer sciences as the "boogeyman", because it is a subject that requires logical reasoning, critical thinking and concentration, which makes the learning stressful, difficult and frustrating to some students.

At Instituto Federal Farroupilha – Campus Santo Augusto, the Integrated Computer Science Course is that one with technical subjects of computer science and also the common issues. According to the Institutional Development Plain (PDI) of the Institute, the articulation between the knowledge areas of basic and professionalizing education is necessary to develop the integrated curriculum; and these, on the other hand, with the research and extension area. "The integration demands the relation between the general and specific knowledge to be continually built throughout the formation, under the work, science and cultural axis"[9].

The integration of the knowledge areas is not that easy referring to the common issues, for example: Portuguese, Mathematic, and Physics. The perfect integration of them with the computer science area can be presumed. In the technical aspects, the teaching-learning program presents difficulties: a) for the students – to understand the new language, to assimilate problems; to realize the connection between the abstract and the concrete (we usually use Mathematics to introduce to the programming); b) for the teachers: the languages are commonly used in connection with Mathematics, which makes it difficult to relate with the concrete aspects. We then get into a vicious circle: on one hand, the teacher unable to work, even though he/she increases and diversifies the ways to teach; on the other hand, the student unable to learn. These aspects are responsible for the high failure and evasion rates of the Computer Science Course, especially the subsequent[1] courses.

In the hunt after a solution to this complex situation, a research with the second year students of Programming I subject in the Integrated Computer Science Course was made. With this research, we intend to understand and make practical a change to the teaching of programming into something more interesting for the student and, thus, to make the understanding more efficient. This way, the paper proposes the use of Scratch programming language to reach the goal, facilitating the concepts organization and allowing the students to check their results as the programming grows. We then start to discuss some important aspects in our research.

This paper is organized as follows. In the second section we discuss the teaching of computer programming. In the third section the Scratch language is presented. In the fourth one, the proposal of using Scratch in the Integrated Computer Science Course and its first results are presented. Finally, in the fifth section some considerations about this paper and future claims are presented.

2. PROGRAMMING TEACHING

The programming teaching is focused on giving the students the opportunity and the knowledge to develop programs and solutions to the everyday problems. To [8], learning to program is a difficult process to many

[1] Subsequents are technical courses for those who have done or are doing the High School. Their curriculum consists only subjects of the technical area.

students, which is noticed in the subjects that introduce to the programming teaching, such as Algorithms, when we realize that many of them can not understand and learn simple abstract concepts. Facing the problems, they cannot find a solution; but, when the teacher proposes a solution and gives them explanations, they pretend they understand. However, the same difficult appears again in the next exercise. To [2], this inefficiency in the process creates a lack of motivation that can result in the students giving up. However, the Instituto Federal Farroupilha reality creates another situation, because the Computer Science Course is integrated with the Secondary School, and for this reason the students do not evade, but only continue the course because of the secondary school. In this way, we observe that this lack of interest in the programming is also reflected in other technical subjects, damaging the students' performance in the course.

The programming subject, by itself, is considered the "boogeyman" by most students of Computer Science. Authors such as Rocha (1991), [3] emphasize the teaching-learning failure which have been recurrent in many institutions. Facing the truth, a certain number of efforts have been developed in order to reduce this problem, which reinforces the weaknesses mentioned, presenting ideas to revert this scenario [8]; [2]. However, there are only a few positive and effective results in solving these problems.

Even with the difficulty of learning a programming language, Seymour Papert, in his book "Logo: computers and education", had idealized programming for children. The author writes that, when he was a child, he used to play with gears in his grandparents' garage, as he could easily assemble and disassemble differentials. Once at school, Mathematics was easier for him than it was for his mates. Later, when he studied Piaget and how children learn, he could understand why he had more facility at school. He understood the when he received a mathematic task, for example, he associated it with the gears and was then able to comprehend the meaning of the subject, solving it easily. According to the author, children can learn just by the use of this gear system. However, this system had an affective meaning to Papert but could mean nothing to any other teenager.

When he wrote about Logo, a primitive version of Scratch, Papert (1986) said that the motivation is not the creation of a final product, but a process: the mistake found can be used as a challenge to reach the goal. This way, people with different skills and interests can discuss with each other about the mistake. The idea is to find the reason of the mistake, and not put it out and forget it. Depending on how the computer is used, people get happy when they can delete mistakes without leaving the trace. In the case of Scratch, the opposite attitude is suggested. The mistakes are seen as members of the learning process, because they help us to understand what just happened. Teaching the computers to "think", children start a journey to discover how *they* think. Thinking about the way they think turns children into epistemologists, an experience that few adults had.

3. SCRATCH

Scratch is a programming language, created at MIT (Massachusetts Institute of Technology), inspired by LOGO and Squeak (Etoys), which intend to be easier and more intuitive. Because it does not require a previous knowledge in other programming languages, it is ideal for those who are beginners in programming. It uses a graphical interface that allows structuring programs as assembling blocks. Each language block contains a separate command, which can be put together if they fit. And the commands can be changed through the menus. It was developed to help in the computing concepts learning and also in other knowledge areas. With Scratch it is possible to create cartoons, games and other interactive programs.

There is a variety of initiatives in activities that show the use of Scratch programs in different areas. Some example projects can be found at www.scratch.mit.edu or http://kids.sapo.pt/scratch/. In these sites it is possible to provide the programs made by students and find many others already made or under developed. The idea is to reproduce the programs and improve them. The reproduction is stimulated because it improves the understanding about computing and the subject. Therein, the student learns how to run a particular function, for example, and how to make a character move on the screen and identify parts of a cell. The students begin to cooperate to achieve their goals. Even a program that is considered to be ready can be improved by another person. In this context, learning communities can be set, so everybody can share information and, at the same time, challenge themselves to do their best.

In most cases the game is used as an understanding tool to learn the program, which leads to the construction of the computing knowledge. During the time, the students have been using this knowledge to develop projects in different areas. To illustrate this issue [1] reports the Scratch use in a Secondary School, as he describes how the Science teacher developed a work that enhanced the production of knowledge in her area, using Scratch. One of the projects can be found at http://scratch.mit.edu/projects/telle/54060.

In this work, we propose the development of activities in the programming area for the students to know the software. At this point, it is possible to create games, simulations and cartoon, improving the knowledge about computing and, at the same time, the students are proposed to develop concepts in others areas, usually in extra time. The student chooses the area and the concepts he wants to work with and establishes a dialogue with the teacher to clarify eventual doubts.

4. LEARNING THROUGH THE SCRATCH INTERFACE

The concern is now to understand the interface as a learning possibility Mathematics and Physics. [4] says that this interface is "a contact surface, of translation, of articulation between two spaces, two species, two different reality orders: from a code to another one, from analogical to digital, from mechanic to human…

Everything that is translation, transformation, transition, it is from the interface order".

Many Physics concepts involve interface systems; the materials and equipment provide the interfaces with the physical world. This research proposes the use of Scratch as a program to study physics, where programming concepts can also be explored in an integrated teaching process. Though this proposal is complex, it can promote an integrated study between different knowledge areas in the classroom. Facing the necessity of controlling, learning to program and to be good using Scratch, it is interesting that this process gets a meaning in the student's concrete life; for example, by simulating the construction of electric circuit.

The human action and interpretation contribute to new ways to represent the physical world. Concepts and symbols allow us to understand the physic world. So, a concept such as electric current is only a representation, which demands an equipment to enable a quantitative analysis to be understood.

It is necessary to comprehend the development to understand the learning process. To [5], development explains learning, because it is recognized as a process in connection with the totality of knowledge structures. On the other hand, learning is caused by situations and a psychological experimentation. Therefore, it is essential that the developing person observes the *operation* development, because knowing demands the change, the transformation and the comprehension of this transformation process. [5] believes that the knowledge begins when the child assimilates something from the physic or social environment. This assimilated content causes a number of perturbations to the person because it carries something new to the assimilating structure, something unknown up that moment. This rebuilding is arrangement. It is this movement, this action of getting the lost balance back that creates a new person, because new assimilations occur. [5] names as "the main equilibration" the mechanism of evolution or development of the organism. This is the knowledge working up.

To [5] the development is related to the following factors: maturation, experience, physical environment effects (social transmission) and the equilibration. Piaget believes that children are not miniature adults. Children have different physical and cognitive structures that change during their development and pass by various stages. A set of cognitive structures correspond to each one, a schema which consists in a pattern of behavior or thought that emerges from the mixing of simpler and primitive integrated units in a more organized and complex scenario. The experience, although widely used by empiricists, does not have the same meaning in the Piaget work. The experience is associated with the constructed knowledge, not with the object but with the relationship established between the person and the object. The social transmission is made through education, the social and observation of the others. The equilibration is an internal mechanism that permits the balance between adaptive processes of assimilation (that sets the new information in existing cognitive structures and systems) and the arrangement (that modifies or changes the existing structures and systems). This way, the equilibration is understood as the mechanism to keep the dynamic balance of the person, a self-regulation. It is a game of adjustments and compensations to reach people progressively more complex coherent structures.

Piaget considers four periods in the evolutionary process of the human specimen, which is recognized by what the individual can best perform in age group. These are stages of intellectual development for children and teenagers, which are: sensorimotor (up to 2 years old), preoperative (2 to 7 years old), concrete-operative (7 to 11 or 12 years old); formal-operative (12 years old onwards).

The intelligence in [6] is the organism's adaptation ability to new situations, the construction of new structures. It is related to the external world and the biological adaptation. To the neuroscience, the intelligence is the object of researches that investigate how brain can learn and how the memory works in the molecular and cellular level in the cortical areas. In the education process, researches intend to understand factors such as: learning, memory and emotions; neural plasticity, activities that stimulate cerebral cortex areas; the brain answers to primitive heritage, like pictures, images and symbols.

Associated studies with learning and the human brain are highlighted in the Gestalt Psychology Theory. Through behaviorist psychology, the emotions, perceptions, perform important role in the learning process. In this theory the brain is a dynamic system in which an interaction occurs between the elements that participate in the perception and stimulus received by the senses in favor of learning. Skinner tried to incentivate the study of human nature, behaviors and controls, operant conditioning, listing concepts and providing models to teaching/learning. However, it is in Piaget genetic psychology that we find the explanations about the learning genesis.

To Piaget, intelligence is an individual capacity, where the human being is a *homo faber* of himself/herself, who is building his/her intelligence while reality representation is established. For him, intelligence is an adaptation way. While structuring his theory, he organized observations, experimentations and conclusions about the development of human intelligence, putting it as an individual surviving ability, starting in the newborn reflexes. The educator has to establish a dialogue between theoretical and practical knowledge and promote an explanatory speech, of the symbolic expressions and representations involved in the process of human development and learning.

5. SCRATCH PROPOSITION IN THE INTEGRATED COMPUTER SCIENCE COURSE: FIRST RESULTS.

From the studies run about the Scratch language and its positive effects that can be reached with the programming, for example, the frustration which comes from the failure in advancing in program implementation that is replaced by the motivation of building the first program. The results can be seen even though they are very simple. On a picture, the command "effects" is applied, followed by some interact instructions with the mouse; effects are obtained as the user moves the mouse.

This facility to program is also a result of Scratch interface organization that is divided in the following parts: category commands; scripts area, stage; sprite. The following commands can be found in the area: movement, appearance, sound, pen, control, sensors, numbers and variables. Inside, each item has different commands that can be dragged to the script area forming the program. The stage is where the programming results can be seen and the *sprite* is an object found in the middle of the stage. If the goal is to move the sprite on the screen, which category of commands should be used? Just click on the movement category that a series of commands will appear to move the sprite, just choose the one that best fits with the problem; it is a completely intuitive language.

Besides, the language provides a series of resources that allow changes in the stage and sprite. A list of sprite and background options is available to the programmer to add on the stage, it is possible to only change the sprite or to insert new ones. With this device, students can create a diversity of activities: dialogues, stories involving many sprites (characters) and at the same time they change the stage (background). Multimedia animations.

With a little more knowledge it was possible the development of a game proposed by Paddle2See, a highly regarded "scratcher" in Hampden, Maine, USA, the author of this idea; "There is a flock of sheep with fractions written on the back. There are three or four fences and, preferably, a sheepdog. The student must command the dog to push the sheep against the fence, separating them in equivalent groups. If the sheep gets to the fence it does not belong to, the other ones inside the fence run away. In the end, the points got by the student are added". It is a complex game which involves lots of programming concepts that students implement by cooperating with each other.This article proposes the use of Scratch, which enables the learning of a programming language and also allows the interaction between the subjects that make up the curriculum of Integrated Computer Course.

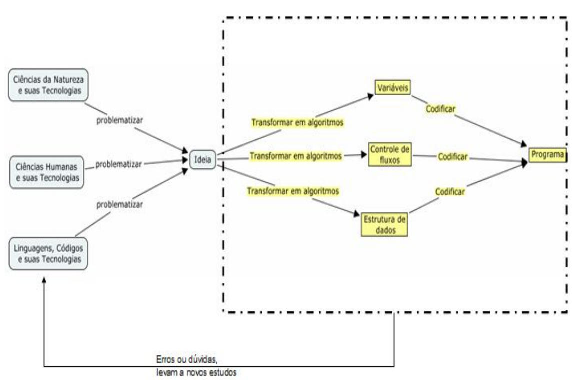


Figura1: Algorithms Knowledge Construction

In figure 01, the relation between the three areas (defined by the National Curriculum Parameters) and programming is showed. On one side the areas can be visualized; Languages, codes and technologies; Human Science and its technologies and Nature Science and its technologies. The students are proposed to develop some concepts of one area; they have to have an idea which they want to implement. With this idea, they start thinking about the instructions, a series of tests area directly run in Scratch – as this is an intuitive language, it helps them to choose an appropriate command for the situation.

As the programming goes on, the mistakes can be found and corrected by returning to the program to review some command that was mistakenly used; it can then be replaced by another one that is more appropriate. This relationship with Scratch eases the identification of mistakes in understanding some concepts, which provides a return to the knowledge areas to enable explanations. With this feedback the understanding in the area gets better and the student learns to program.

In these first Scratch studies, we noticed that while the program is under development, the students have the possibility to see the results, and it enhances the confidence and the sensation of creating a product. It may look simple, but these impressions help the students to like programming better.

6. CONCLUSION

No doubt the programming subject is one of the most terrifying in the computer science course and for this reason it may explain the high failure and evasion rates in this course. This way, the dissatisfaction with this reality has forced the teachers of the Computer Science Course into a research for alternatives to change the scene. Scratch appears in this context.

Scratch is a programming language considered to be more accessible because it does not require previous knowledge of any other programming language, which makes it ideal for teaching and for those who are beginners in programming. Its interface allows building programs using building blocks, which represent a programming command and makes it possible to develop and work the language concept in an interactive and animated way, as if it were a game; this is its main attraction.

Up to the moment, Scratch has been used as an experiment in the Programming I subject, i.e., exclusively for programming teaching without concerns about the other teachers using it in their classes. The first impressions of using Scratch were positive, which has created more interest from the students in the subject. The involvement and the wish of learning were visible, different in comparison with the frustration expressions, before the use of the new language.

To be effective, this research needs to expand its use to the teachers of other areas, which permeates an information work and comprehension of Scratch efficiency also in producing knowledge in other areas. This work was developed in a second year class of the Integrated Computer Science Course, but we intend to use it also in the first year, in algorithms subject and that the teachers of the common issues adopt it as an additional device to the educational computing classes. So, in the second year the regular subjects of the common issues can continue the use of Scratch combined to the Programming I and in the other areas of language programming. It is interesting to emphasize that in the Computer Science Course we plan to develop some extension work with Scratch in elementary schools around the region, with the purpose to introduce the students to the program, so they can join the integrated courses with a brief knowledge of scientific and technological production.

Using Scratch in the dynamics of knowledge construction means enhancing the equilibration process (Piaget, 1976). From the moment that the student chooses the subject to create the algorithm up to encoding the program, he passes through lots of unsettlement process. When he chooses a subject, an assimilation process occurs (he understands the concept, according to his perception) and he starts programming and better understands the concept; then he notices that it is not exactly how he thought before; it takes him back to the books until the settlement process occurs.

Obviously, this is not only about incorporating a technology to the classroom and imagining that the construction knowledge process becomes more efficient. It is entering a new culture in the educational space. A culture in which the teacher is not the only one who knows things and the student puts himself as the subject of building his knowledge. With Scratch the child programs the computer and not the computer programs the child.

7. REFERENCES

[1] José duarte. **O Scratch dá o que falar: dos usos domésticos aos usos na escola**. Revista Educação e Matemática. Janeiro/Fevereiro.1996.

[2] E.L.França;, Z. C.Feliz; M.S. Souza; T.B.Carneiro;, P.R.C. Souza, C.A.P.D. Filho. **Utilização de Sistemas Tutores Inteligentes para o ensino da Programação**. VII Simpósio de Excelência em Gestão e Tecnologia.2010.

[3] A. Gomes and A. Mendes. **Suporte à Aprendizagem da Programação com o Ambiente SICAS**. Actas do V Congresso Ibero-Americano de Informática Educativa, Viña del Mar, Chile. 2000.

[4] P. Levy . **As tecnologias da inteligência: o futuro do pensamento na era da informática**. Rio de Janeiro. RJ. Editora 34.1995.

[5] J. Piaget. **Desenvolvimento e aprendizagem**. In: LAVATELLY, C,S e STENDLER, F. Reading in child behavior and development. New York: Hartcourt Brace Janovich.1972.

[6] J. Piaget.. **O nascimento da inteligência na criança** . Trad. Álvaro Cabral. Quarta edição. Rio de Janeiro: Editora Guanabara S.A.1987.

[7] J. Piaget**. A equilibração das estruturas cognitivas: problema central do desenvolvimento**. Traduzido por: Marion Merlone dos Santos Penna. Rio de Janeiro: Zahar Editora.1976.

[8] E. P. Pimentel; V.F. França; N. A. Omar. **Caminho de um ambiente de avaliação e acompanhamento contínuo da aprendizagem em Programação de Computadores**. Workshop de Educação em Computação e Informática do Estado de Minas Gerais - WEIMIG.2003.

[9] M. Ramos. **Possibilidades e desafios na organização do currículo integrado. In: Ensino Médio integrado: concepções e contradições**.2005.

Enterprise Adoption of Model Based Systems Engineering

Forrest HOWIE
Missile Systems and
Wayne O'BRIEN
Intelligence and Information Systems, Raytheon Company,
22260 Pacific Blvd., Sterling, Virginia 20166, America

ABSTRACT

This paper is based on research to assess the current state of adoption of Model Based Systems Engineering (MBSE) in the global enterprise that employs the authors. The assessment used a detailed questionnaire completed for twenty-three programs across our company. The results provide a close look at where and how MBSE has been applied in our company, as well as insights into how to make it more extensively and effectively used.

Keywords: Model Based Systems Engineering, Model Based Engineering, use case.

1. INTRODUCTION

Model-Based System Engineering (MBSE) is an approach to engineering that uses models as an integral part of the technical baseline to express the requirements, analysis, system design, and verification of a capability, system, and/or product throughout the customer's acquisition lifecycle.[1] Our definition of a model is an organized collection of concepts, obeying certain syntactical and methodological rules, from which documents or diagrams can be expressed.[2] Appendix 5.2 provides further definitions, observations, and descriptions for context.

This paper presents a summary of the assessment of the current state of adoption of MBSE throughout a global enterprise. The assessment was performed in 2012 by a team of engineers including the authors, Robert Healy, and Bobbi Young, based on detailed questionnaires completed for programs in the five business units of the enterprise. The approach used to collect the information is described fully in Appendix 5.1.

The goal of MBSE in our company is to leverage use-case driven, model-based development employing a range of modeling methods and tools to enable systems to be produced faster, cheaper, more predictably, and to a higher quality, as well as be more easily evolved and made interoperable. Use cases represent the ways of using a system to provide observable results of value, so the system's architecture flows from the relationships among them.

2. MOTIVATION AND BACKGROUND

MBSE is in many respects a mature technology that has been applied successfully for years in our company, yet it is not used extensively. When MBSE is applied fully, the resulting model is the design (throughout this paper, when we use *model* in this context, we mean one or more models as discussed in section 3.1.2.3.)

Visual modeling emphasizing diagrams helps understand and manage complex systems, including systems of systems and geographically dispersed systems. Visual models allow the manipulation of semantically rich diagrams for development and maintenance, rather than traditional specifications and code that do not readily reveal the overall understanding of a system. Modeling tools allow the detail to be elided from diagrams as appropriate for the purpose while providing more detailed views once the higher-level views are understood.

When Model Driven Software Development (MDSD) is combined with MBSE for software intensive systems, the resulting model is the application,[3] as well as the design. One way of looking at such visual modeling is as another layer of abstraction added to the increasing number of layers that has occurred in software development since the use of register switches and machine-language programs became obsolete. Developers may view this additional layer as simply increasing complexity and isolating them from more direct involvement with the details provided by traditional specifications and code. Training and experience might overcome this negative view, but at least in some cases, it may depend on the scope of the system – some systems may be too small to warrant extensive use of MBSE-MDSD.

The above considerations provide a context for the remainder of this paper. The information from the questionnaires, summarized in the next section, provides insight into why MBSE adoption is so limited, in spite of its mature state and successes.

3. ASSESSING THE INCORPORATION OF MBSE INTO THE DEVELOPMENT PROCESS

This section considers objective factors that describe the use of MBSE. We developed a common analytical framework based on interviews in order to evaluate these factors in programs that applied MBSE. The interview questions are repeated as

[1] Internal definition based on the Final Report Model Based Engineering (MBE) Subcommittee (National Defense Industry Association, 2011) 7.
[2] Internal definition.

[3] Andrew Lyons, UML for Real-Time Overview (Rational Software Inc., 1998) 7

subsection headings 3.1.1.1 through 3.1.2.4. Each subsection summarizes the responses to the corresponding questions.

The results of the interviews were analyzed to identify patterns in the current practice of MBSE throughout the company. As a point of additional reference an impression of the perception of the value of MBSE to those interviewed was also collected. The following subsections describe our process, discovery, assessments and results. Appendix 5.1.2 describes a supplement to the questionnaires to further asses the breadth and depth of use of MBSE on each program

3.1 Factors Contributing to the Successful Implementation of Model Based Systems Engineering

Success in the implementation of MBSE has both a technical and programmatic aspect. For technical success MBSE must produce the technical baseline in adequate detail to enable the technical teams to build their products directly from the model with traceability and without any interpretations or changes. Program success in MBSE means that it provided the same or better quality in less time and/or at lower than a traditional Systems Engineering (SE) effort with the same scope. To achieve such technical and programmatic success the level of rigor in the execution of MBSE is key.

In practice there are often inconsistencies or compromises in the application of MBSE. The establishment of a model-based system development capability requires a minimum set of interdependent factors to be in place. If one of the factors is compromised or applied inconsistently then the negative impact is likely to be significant. These factors, along with related measures, discovered through our team's activities and prior knowledge, are discussed in this section.

3.1.1 Degree of Cultural Adoption: Cultural adoption includes those factors that are not strictly technical that are essential to widespread application of MBSE. The interview findings related to such adoption are summarized below under the headings of the questions asked.

3.1.1.1 Did you use any models? This question simply confirmed that the program used one or more models in one or more disciplines and phases of the lifecycle. Of the twenty-three programs identified, all twenty-three confirmed the use of at least one model in at least one phase.

3.1.1.2 What is your definition of MBSE? Having a common definition of MBSE is a measure of how widely accepted its use is. Of the twenty-three programs reviewed, nineteen had varied though somewhat similar definitions, three had no definition, and one used our standard definition. As confirmed by answers to other questions, this suggests that MBSE is not well enough established to have a common understanding among engineers.

3.1.1.3 How much training did your team receive prior to using MBSE? The amount of training received prior to using MBSE indicates the level of commitment by the program to its use. Eight had little prior training, seven had moderate prior training, seven had none, and one had substantial prior training.

MBSE today is best done using tools specifically intended for systems modeling, rather than a general purpose tool like Excel, although Excel is often a secondary tool even when tools intended for MBSE are used. Understandably, the skills, experience and training in the use of the MBSE tools directly affect the efficiency and effectiveness of their application, especially given the complexity of some of the tools. On the positive side, the responses to this and the following questions suggest that MBSE can be accomplished through relatively small investments in training prior to beginning the program. On-the-job training, provided by consultants, vendors, or other program staff appears to be a reasonable option.

Organization of the effort, configuration management of the models, the training needed for adding or modifying content versus to reading and using content are all part of the implementation decisions. These factors would significantly affect execution of MBSE and suggest that a degree of standardization of the tools and methodologies used across the organization would be of great benefit.

3.1.1.4 Were any members of the SE team familiar with MBSE prior to the start of the project? Eighteen teams had members with prior MBSE familiarity, five had none. The high number with some familiarity runs counter to the lack of training and a common definition, indicating that MBSE is in the awareness stage of adoption.

The considerations in the previous section apply here, but in addition, the importance of understanding just the essential MBSE concepts and process is at the root of the question. MBSE is a dramatic departure from the traditional SE process and the degree to which the underlying concepts and process are understood and advocated is of major importance to the success of an MBSE effort.

3.1.1.5 Where in the lifecycle are you using MBSE? Thirteen of the programs used MBSE in the early stages, five were mixed in that at least part of more than one phase used it, and five used it for the full lifecycle. Use of MBSE across more of the lifecycle suggests a stronger commitment by the program and more systematic use, improving the results as well as measuring the degree of adoption. The high proportion of limited use across the lifecycle is another indicator MBSE is still in the early stages of adoption, with ad hoc application within a program.

3.1.1.6 Number of Technical Disciplines Actively Using the Model: SE and Software Engineering (SWE), combined with at least one other discipline in using MBSE on ten of the programs (i.e., at least SE, SWE, and one other discipline used MBSE on each of the ten), SE used MBSE combined with SWE (but no other disciplines) on four programs, on two programs only SE used MBSE, two used MBSE for SWE only, two had SWE plus at least one other discipline (not SE), two had SE plus at least one other (not SWE), and on one program, all disciplines used MBSE.

The use across the project disciplines (e.g., systems, mechanical, software, electrical, hardware) indicates the degree of adoption by the program. Also, a substantive indicator of perceived value is the degree to which the disciplines chose to integrate their processes and models with other disciplines. We used the number of models as the indicator for the level of integration; the fewer the models, the greater the integration (discussed in section 3.1.2.2 below).

3.1.1.7 Were the Software Requirements Specification (SRS) replaced or greatly supplanted by the use of MBSE? Fifteen of the programs surveyed did not replace or greatly supplant the SRS, six used models to supplement the SRS, and two replaced the SRS with models.

Acceptance by the customer is a crucial determinant, perhaps decisive in the long run, in the widespread adoption of MBSE. Generally stated any modeling work that did not have a customer who recognized the value of such work was supplemental at best. Our measure for the degree of customer adoption of MBSE was its willingness to replace the traditional SRS. The ability of two of the programs to replace the SRS with models demonstrates the feasibility of eventually transitioning completely to MBSE. The technical side to this question is the degree to which the model is the technical baseline, or the model is the design, which is discussed in section 3.1.2.3 below.

The lack of acceptance by the customer is both an assessment of the current state and potentially an indication of the need to invest in the education of the customer about MBSE. The core of such education would be the purpose, requirements, outcomes and pay-off (requires metrics, see section 3.1.2.4) of the use of MBSE, in business as well as technical terms. This gives a very direct opportunity to set expectations. It is incumbent on the systems engineering function to ensure that the funding and technical execution match the expectations being communicated.

3.1.1.8 Was the use of MBSE successful or not? The person interviewed (a technical lead or equivalent) on fourteen of the programs concluded that MBSE was successful, six indicated mixed results, and three viewed MBSE as a failure, largely for lack of broader and/or deeper use within in the program. This is one of the greatest obstacles to adoption, because degree of use and success are interdependent.

3.1.1.9 Would you use MBSE on another program? Respondents for twenty (the sum of the successful and mixed outcomes in the previous section) of the programs said they would use MBSE on another program, two were unsure, and one did not respond. That is, even those who viewed it as only partially successful would use MBSE again, suggesting the perceived value is higher among practitioners than the adoption of MBSE indicates. This in turn points back to the need to invest in education as described in section 3.1.1.7

3.1.1.10 Why did you use MBSE on this program? Fifteen programs used MBSE because of staff familiarity, five because it was mandated (prime program management), and one because of customer familiarity (two gave no reason). The connection between this finding and the success of the MBSE effort is discussed in the conclusion.

3.1.2 Technical maturity of MBSE: The technical maturity of MBSE also affects its adoption, along with the cultural factors discussed above. These factors are intended to measure the effectiveness of the technology itself, both tools and process, rather than how it is perceived and understood by its practitioners and other stakeholders.

3.1.2.1 What Perspectives of the system model, using the SysML taxonomy, were developed?
MBSE can be considered, using the Systems Modeling Language (SysML) taxonomy, as comprised of four perspectives covering both description and analysis. These perspectives – requirements, behavior, structure and parametrics – are typically required in some variation to develop a system. The numbers of these perspectives modeled suggest the penetration of MBSE into the technical development of the system.

Eight programs used MBSE for structure and behavior only; six for structure, behavior, and requirements; four used it for all four perspectives; three used it for behavior only; and two used it for behavior, requirements, and parametrics. This limited use of the four perspectives reinforces the indications from the cultural questions that MBSE use is largely ad hoc, for whatever reasons.

3.1.2.2 Were all SE artifacts maintained in one model or more than one? Thirteen programs used multiple models, nine programs used one model, and one did not respond to this question.

MBSE tools support a range of modeling notations, sometimes including a specific methodology. When the notations and methodology included do not match those with which the team members, especially across disciplines, are familiar, the team may attempt to impose methods not directly supported by the tool, causing ongoing reconciliation issues as the tools are applied. In particular, different disciplines or groups within a discipline may develop separate models. For a given program, the effort required to maintain the development models and how successful the results are depend on what is expected from the tool, the skill of the SEs, and tool capabilities. Given the nature or the tools and the modeling required, skill in using the tools may require considerable training and experience. The matching of tools to a program and system methodology also requires technical knowledge of both, as well as planning to achieve success. A specific example where experienced use can help is in understanding tool idiosyncrasies for geographically dispersed collaboration (even within the same facility). Though most of the tools on the market have mechanisms for group use, each may require tool specialists and tailoring, both of which need to be accounted for in schedule, cost, and risk management. All of this is compounded by the development and maintenance of multiple models, even if they are in the same tool. Initially, the use of separate models may be expeditious, but eventually they or their results must be integrated, a highly error prone process.

3.1.2.3 Is this project an example of employing a system model as the technical baseline, in lieu of deliverable documents? Sixteen programs said no, seven said yes. There is ongoing discussion about how many models – from one integrated model to multiple models – are needed for an optimal model-based technical baseline. For simplicity, this paper refers to the optimum number as the system model. Multiple perspectives (as discussed above in 3.1.2.1) are likely to be required if the system model is to be the technical baseline (see section 3.1). The broader the implementation of the system model across the entire development lifecycle (see section 3.1.1.5 above) and all technical disciplines and perspectives, the more probable it is that no requirement escapes will occur. Escapes, in this context are those requirements that are either misallocated or omitted and later construed as errors in the engineering. These errors cause rework, the cost of which rises exponentially as the development progresses. There is a tradeoff in labor cost with the reduction in errors, related to how much engineering time is needed to capture the technical baseline in the system model. The consequence of error and/or failure is potentially high as previously discussed, but the likelihood can be difficult to monetize. Context appropriate experience can help set this level and the nature of the elaboration captured in the model.

3.1.2.4 Were any measurements of the gain or loss collected by the project? Eighteen of the programs said no measures were taken, and five used miscellaneous measures. Along with use for at least a major portion of the program (see section 3.1.1.8), this is one of the chief obstacles to widespread adoption of MBSE. Technical and program leadership as well as the customer need to have a means to assess whether there is value added by using MBSE.

3.2 Future Directions

Based on the above analysis in section 3.1 Factors Contributing to the Successful Implementation of Model Based Systems Engineering, the following follow-on activities would be useful:

- Analysis of interviews and the Characterization Matrix for MBSE Use (see Appendix 5.1) to determine when and where MBSE should be used on a program
- Determine how do we make adoption systematic
  - Identification of quantitative measures of the impact of MBSE on cost, schedule, and quality
  - Increasing customer buy-in
  - What impact on training would wider customer acceptance have, e.g., would it replace the need for overview awareness of MBSE for program leadership?
  - What's the balance between internal and external buy-in?

- Determine how to identify a reasonable assortment of tools and methodologies to cover a wide range of engineer knowledge and program variation?
- Determine how to develop a methodologist position that would be a champion, trainer, and mentor for MBSE starting with Early Phase Systems Engineering?
- Determine more specifically what role customer interest played in those programs that were successful?
- Develop general training for the principles and concepts of MBSE versus specific tools and methodology for all engineers, to focus on general methodological and tool considerations
- Determine why parametrics are seldom used
- Determine how to improve the transfer knowledge across programs

4. CONCLUSIONS

MBSE use was clearly *ad hoc*, rather than systematic, e.g., only five programs used MBSE for the full lifecycle, thirteen used it only in the early part of the program, and five were mixed (see section 3.1.1.5 Where in the lifecycle are you using MBSE?. Further, the breadth and depth of use were limited, e.g., the total number of programs using MBSE was small and only four programs used MBSE for all four SysML perspectives, and only six used parametrics (see section 3.1.2.1 What Perspectives of the system model, using the SysML taxonomy, were developed?. Only five programs used MBSE because it was mandated, none explicitly by the customer (the mandate was internal in four cases and from the prime in one case). In two cases it appeared that the customer may have influenced the decision. Of those where MBSE was mandated, one said the results were mixed and one said it was unsuccessful; both were internally mandated, suggesting the top- down approach alone is not sufficient. The two programs where the customer may have influenced the decision to use MBSE were both rated successes, suggesting the importance of even interest on the part of the customer (see next paragraph).
Also, while customer mandate or acceptance was limited, usefulness for internal purposes was still almost universal, as suggested by the nearly 90% of programs where the use of MBSE was rated at least partially successful (see section 3.1.1.8 Was the use of MBSE successful or not?). The implication is that while customer acceptance would make MBSE more effective, we should consider its use for internal purposes as well.

One key finding was the almost total lack of quantitative measures of the effectiveness of MBSE. Eighteen of the programs surveyed did not have any measure and the measures used by the other five were weak. Lack of customer acceptance as indicated by only two programs where the SRS was replaced by MBSE and six where it supplemented the SRS (see section 3.1.1.7 Were the Software Requirements Specification (SRS) replaced or greatly supplanted by the use of MBSE?) also stood out as a major obstacle to adoption of MBSE. Lack of metrics and customer acceptance are related in that it is easier to make the case to the customer when convincing metrics are available.

The major reason for using MBSE was staff familiarity (see section 3.1.1.10 Why did you use MBSE on this program?), indicating that support from program leadership is not strong. As with engineers (next paragraph), this may not point to increased effort to educate leadership, because acceptance by leadership is likely to follow naturally from customer acceptance.

Engineers had a favorable view of MBSE – nearly 90% viewed its use as at least partially successful, as noted above, and staff familiarity was the reason for using MBSE in 65% of the programs. The SE teams on eighteen of the programs had at least some prior familiarity with MBSE. This positive view on the part of engineers occurred even though the team in only one program had substantial training before the program started (see section 3.1.1.3 How much training did your team receive prior to using MBSE?). Engineers were able to find the tools they needed and willing to learn what was necessary after the program started. This suggests that MBSE can be accomplished through relatively small investments in training prior to beginning the program. On-the-job training, provided by consultants, vendors, or other program staff appears to be a reasonable option. Tool availability and engineer training and awareness may not be as big obstacles as generally thought.

While not a primary obstacle to adoption of MBSE (as noted above), training would still help in improving adoption, because staff familiarity was such a big factor in choosing MBSE. Given the willingness of engineers to learn on the job, training may only need to focus on concepts with general methodological and tool considerations, rather than the far more expensive in-depth training on specific tools and methodologies. Further, it may suffice to train a relatively small number of engineers within each discipline, who would then serve as champions and mentors. Existing training courses within the enterprise should be updated to include MBSE.
Engineers seem to find benefit in behavior, structure and requirements modeling, but less for parametrics (only used for six programs).
Models are generally not being used as part of the deliverable or shared with the customer, largely because of customer reluctance to replace the traditional deliverables.

5. APPENDICES

5.1 Collection Method for Interview Results

5.1.1 Interview Objectives: To obtain information for the assessment described in this paper, we selected programs based on the following criteria:

- Executed on contract
- Unclassified
- Large enough to require a team to execute
- Known as a possibility by one of the members of our assessment team

We grouped the selected programs by business unit and assigned a person from our team to contact each program. Team members conducted interviews with individuals for each program who were verified as the most knowledgeable available for that program to answer our questions.

5.1.2 Characterization Matrix for MBSE Use: In addition to the responses to the interview questions summarized in the body of this paper, we completed a matrix (Figure 5.1) that captured the nature of the work performed in terms of a taxonomy using the four perspectives described in section 3.1.2.1. The breadth of the use of MBSE for a given perspective was indicated by placing an X in one or more of the columns adjacent to the rows indicating level of use (Component, Subsystem, System, or Environment) with Environment representing the greatest breadth. The depth of use was indicated by placing a weight (values of 1-5, with 5 being the highest level of use) in one or more of the columns for the four perspectives. The weighting factors were a qualitative representation given by those most knowledgeable of the models' range of purpose and completeness. A follow-on effort would make use of the matrix, in conjunction with the questionnaire results summarized in the body of this paper, to identify criteria for deciding the extent to which MBSE should be applied at various levels, across the lifecycle, for the different perspectives.

<Bus. Unit>
<Program>
<Tech.

Component				
Subsystem				
System				
Environment				
	Structure	Behavior	Require-ments	Parametrics
User Needs	<weight*>			
Mission				
Solution Concept				
Analysis of Alternatives				
Engineering				
Integrate, Verify & Validate				
Manufacture				
Operations				
Category Total				
Overall				

Figure 5.1 Characterization Matrix for MBSE Use

5.2 Modeling Terminology and Overview

- UML is a general purpose modeling language for visualizing, specifying, constructing and documenting the artifacts to describe the behavior and interfaces of software-intensive systems. SysML is a domain specific language for systems engineering originally initiated by the International Council on Systems Engineering (INCOSE) with the further release by the Object Management Group (OMG) of version 1.2 in 2010. The OMG definition of SysML is as follows:

  "SysML is a general-purpose graphical modeling language for specifying, analyzing, designing, and verifying complex systems that may include hardware, software, information, personnel, procedures, and facilities."

- Experience has shown that the SysML can provide a common engineering language for systems, software, and hardware engineering disciplines as SysML is an extension to a subset of Unified Modeling Language (UML) creating a common language across the disciplines.

- The use of a modeling language such as the Unified Modeling Language (UML) or Systems Modeling Language (SysML) and associated diagrams can be used to specify less ambiguously than text alone what a system must achieve

- UML and SysML do not require a particular process. It is up to the engineering team and the customer to use these languages in the most beneficial way

- The architecture can be specified using a traditional functional approach, an Object Oriented (OO), or a hybrid of both approaches. The system architecture is increasingly, but not always, represented using the SysML Block Definition Diagram (BDD) and the Internal Block Diagram (IBD). UML is strictly OO, while SysML represents functionality separately from the BDDs and IBDs, then allocates it to them.

- SysML BDD and IBD diagrams are extensions to the UML. The BDD is the SysML counterpart to the UML Class diagram and describes the system hierarchy and system classifications.

- The BDD and IBD are structure diagrams. System Use Cases may be realized by blocks in the BDD and an instantiation of the blocks and their associated parts can be described with the IBD. The blocks and parts can contain both attributes and behaviors (like Software classes and objects).

- The IBD diagram also serves to show connections between the parts of the system and associated data and/or message flows. A more rigorous representation of the interconnections between the system blocks can be realized using ports and interfaces (with the IBD) and their associated interface contracts.

- Tool challenges (e.g. traceability, interoperability, scalability, and cross- discipline applicability) are common, but engineers demonstrated a willingness to work through them on the job

- Finding a consistent and quantitative way of measuring engineering productivity in an MBSE environment is a current challenge facing the industry

- The appropriate capture and subsequent delivery of the evolving model is substantively different than the culture of deliverables specified in Contract Data Requirements Lists (CDRLs). Interactive use of the model in its environment by the customer as well as the development engineers means a challenge for the program's configuration and data management function. Offsetting such potential challenges would be substantial reductions in CDRL production and presentation costs, if the customer's milestone schedule and lifecycle management in general align with the MBSE approach.

- System design of the requirements resides in the model as visual representations (e.g., Block, Class, Sequence, Activity, State and Performance diagrams that specify how the requirements are achieved.

- MBSE models provide a huge step in improving communication by bringing the engineering disciplines, customer, and other stakeholders to a common understanding of what and how the system needs to perform, thereby reducing downstream defects and integration costs

- Improved communication translates to improved program performance.

- Requirements traceability and configuration management are much improved through the use of coherent systems models. High level system requirements (what the system must do) are maintained in a requirements management tool (e.g., DOORS) and/or the model and traced to the use cases and other model elements that satisfy those requirements

- Once the requirements are understood, high-level UML use cases can be identified. Such use cases are created by identifying and describing the ways the system will be used to produce observable results of value

- Since a system's use cases represent the ways of using it to provide observable results of value, the system's architecture flows from the relationships among them. Usually not all use cases can be examined in detail in the time available so only the architecturally significant ones are considered initially. There are techniques available to help in the selection of the use cases including the generation of a set of abstracted use case surveys. These are designed to cover more functional dimension to less depth in order to point to most valuable area for more detailed elaboration.

- The relationships among use cases are the starting point for the next level of architectural detail, providing system structure, behavior, interfaces, message flows, requirements traceability, and parametrics.

- High-level use cases (Operational Use Cases in DoDAF terms) are decomposed further by identifying system use cases (same sense as DoDAF) and through the use of include or extend use cases. Each use case is a collection of sequences of actions (flows of events) that directly satisfy the customer requirements. Use cases can also be a semantically rich vehicle to get customer buy-in relating to the major ways the system will be used. Their simplicity should not overshadow their value in communicating with the customer.

- Interface Design Documents (IDDs) can be generated through the use of the interface contracts between the system under development and the external actors.

- At a minimum, it is recommended that the following artifacts should be generated to adequately describe the system architecture:

  - Structure View – (Block Definition, and Internal Block Diagrams)

  - Behavior View – (Use Case Diagram, Sequence Diagram, and Activity or State Diagram)

  - Requirements View – (Requirements Diagram, Performance Diagram or Requirements Table for Traceability

A Case Study in USA Rehabilitation Service Delivery Using a Classification Regression Tree Analysis to Reduce Balance Impairments and Falls in the Older population: Impact on Resource Utilization and Clinical Decision-Making.

Lucinda Pfalzer, PT, PhD, FACSM, FAPTA[1], Amy Yorke, PT, MPT, NCS[1], Jennifer Ellis, PT, DPT, MS[2], Laurie Otis, PT, MBA, MHA[2], Bridget Gorski, BS[1]
[1]University of Michigan-Flint, Flint, Michigan;[2] Gentiva® Health Services, Atlanta, Georgia

ABSTRACT
***Background/Purpose**:* Over 1/3 of adults over age 65 experiences at least one fall each year. This pilot case report uses a classification regression tree analysis (CART) to model the outcomes for balance/risk of falls from the Gentiva® Safe Strides® Program (SSP). ***Case Description**:* SSP is a home-based balance/fall prevention program designed to treat root causes of a patient's balance problems. ***Methods/Outcomes**:* Analysis starts with the cohort of patients enrolled in SSP having a Berg Balance Assessment completed (n=165) and sequentially divides it into subgroups, creating a regression tree model. Descriptive statistics were calculated to summarize demographics, self-reported pain measure, foot sensation, and initial exam and discharge scores for tests of impaired balance/risk of falls. ***Results/Discussion**: A*verage (±SD) age for fallers was 82.6 (±8.3) years with 39.8% (n=37) male and 60.2% (n=56) females, and 77.3 (±10.1) years for non-fallers with 43.5% (n=30) males and 56.5% (n=39) females, respectively. 43% of patients demonstrated improved balance on discharge from home health. The CART yielded a tree model after 12 partitions. The best discriminating variable was BBS score of < or ≥ 33 on initial examination. ***Conclusion**:* This pilot case analysis enables Gentiva® and policy makers to improve efficiency and effectiveness of service delivery.

Keywords: balance, falls, CART analysis, clinical decision making, physical therapy

INTRODUCTION

One-third of older adults over the age of 65 will experience at least one fall during the year [1, 2]. Of those older adults that fall, 20-30% sustain a moderate to severe injury making it difficult for independent mobility and increasing the risk of an early death [3]. For older adults age 65 and older, medical costs for falls treated in emergency departments in 2005 totaled $6.3 billion, with $5.8 billion spent for patients that were subsequently hospitalized [4]. Falls are the leading cause of fatal and non-fatal injuries for older adults age 65 and older [5]. Complaints of imbalance [6], risk of falling and sustaining an injury secondary to a falling [7-8], increase with age. Age related changes have been demonstrated in common tests and measures of balance impairments [9], normal movement patterns [10], and self-report of imbalance [11].

Balance is achieved through a coordinated effort of body systems including sensory and motor systems [12-14]. The external environment provides information through a person's visual, somatosensory, and vestibular systems. Sensory information is subsequently interpreted in the central nervous system. In response, an appropriate motor response is developed, and a motor strategy is activated in order for a person to remain upright [12-14]. Preserving an upright position includes maintaining, achieving, or restoring the body center of mass relative to the limits [14] of stability. The functional goals of balance include maintaining a position, successfully transitioning between movements or positions, and reacting to an external perturbation in order to stay upright[14-15].

A fall is commonly defined as unintentionally coming to the ground or other lower surface not caused by a loss of consciousness, sudden paralysis, seizure, or a strong perturbation [16]. Falls are typically caused from a multitude of factors including both intrinsic and extrinsic risk factors that can be often be corrected or managed [17]. Intrinsic risk factors (within the person) such as weakness, visual impairments, gait and balance problems, depression, urinary incontinence, orthostatic hypotension, cognitive function, sensory deficits, and comorbidities increase a person's risk of falling. Extrinsic risk factors (outside the person) consist of environmental factors (e.g. inadequate lighting, loose rugs) and polypharmacy also lead to increasing a person's fall risk. Intervention focused on environmental changes, balance, strength training, functional mobility, decrease in medications with management of visual changes and orthostatic hypotension have been recommended as effective treatments to decrease fall risk [18]. Treatment of balance dysfunction in older adults has been shown to be most effective when the approach is multifaceted and individualized [18].

MATERIALS AND METHODS

Gentiva® Safe Strides®. Home health care services are commonly available to patients that are unable to leave their home to receive outpatient medical services. Most often, home care is provided to Medicare beneficiaries over the age of 65. Gentiva® Safe Strides® (SSP) is a comprehensive fall risk reduction program developed for patients over the age of 65 years who are home bound and have a history of falls, and/or present with ≥ two modifiable fall risk factors. Physical therapists were trained and verified on how to systematically collect the SSP data including neuropathic pain ratings, the Berg Balance Scale (BBS), and Performance Oriented Mobility Assessment (POMA) and the Dynamic Gait Index (DGI). All clinicians recording sample data had successfully completed Gentiva's Outcome and Assessment Information Set (OASIS) C documentation and respective specialty clinical training. Gentiva® collected SSP outcome information on initial evaluation and at discharge. At discharge, paperwork was reviewed for quality assurance and keyed into a companywide secure database per standardized processes. Data retrieval was achieved through database query, extrapolating and matching all completed cases, and de-identified through an honest, independent data-broker.

Balance Assessment. The purpose of utilizing clinical balance assessment tools is to identify the presence of a balance problem and assist in determining the underlying cause of a balance problem [15]. Measurements utilized in the SSP program are as follows:

Tinetti Performance Oriented Mobility Assessment (POMA) was developed as a measure to screen older adults for balance and gait impairments [19]. The POMA consists of 16 items including 9 balance activities and 7 gait items. Lower scores on the Tinetti balance significantly predicted occurrence of falling and ADL decline [20]. Minimal detectable change (MDC) has been determined to be 5 in older adults [21]

Berg Balance Scale (BBS) consists of 14 functional activities including sitting, standing, and postural transitions [22, 23]. Score of < 45 on the BBS have been associated with increased risk of falling [24].

Dynamic Gait Index was developed as a measure to assess and document a patient's ability to respond to changing task demands during walking [25]. It consists of eight items that vary the walking task by changing walking speeds, walking with head turns, turning and stopping, walking over and around obstacles, and ascending and descending steps [25]. Scores < 19 have demonstrated increased risk of falling in older adults [26].

Modified Clinical Test for Sensory Interaction on Balance was developed to systematically test the influence of visual, vestibular, and somatosensory input in standing balance without the use of computerized equipment. Standing balance is assessed under four different somatosensory and visual conditions including firm surface eyes open and closed, and foam surface eyes open and eyes closed [27]. SSP documented results of the mod CTSIB if a patient successfully completed 30 seconds of each of the four trials.

Sensation is evaluated utilizing the Semmes-Weinstein Monofilament (SWMF) testing [28]. Sensation on the plantar surface of the foot was tested in 5 locations bilaterally (10 total). A score < 7/10 served as a trigger for the implementation of monochromatic infrared energy (MIRE) to be used in branches where it was available.

Neuropathic pain was assessed utilizing an 11 point visual numeric rating scale of 0 (no pain) – 10 (intolerable pain). For this analysis, initial and discharge neuropathic pain measurements were utilized.

Subjects. Participants in an IRB approved trial were 455 with 165 patients in the BBS sub group enrolled in the SSP program from January 1, 2009 to December 31, 2009 from 11 Gentiva® branches across the USA.

Statistical Analyses. Analyses were performed with IBM SPSS® version 19 (Chicago, IL) and SAS STAT® system Version 9.1 (SAS Institute Inc., Cary, NC, USA) and JMP® Version 9 (Cary, NC). The tests of significance between the two groups (MIRE and No MIRE) by time (baseline and discharge) was a two-way repeated measures ANOVA of Group x Time with the alpha level set at $p \leq 0.05$. The second analysis was on a subset of the balance impairment data from the BBS with groups that were determined by discharge BBS Scores of less than 45 (group 1- balance impaired) or equal to greater than 45 (group 2 – no balance impairment) and assessment of clinical factors, self-reported neuropathic pain, foot sensation, and other balance performance measures at baseline and upon discharge from the Safe Strides program. A model was developed for these two groups for the CART analysis. Descriptive statistics were used to summarize patient demographics, baseline and discharge assessments of foot sensation, neuropathic pain, sensorimotor test of balance (Modified CTSIB), and other balance impairment measures (POMA, DGI) (Table 1). Means and standard deviations were calculated for the continuous variables; medians and interquartile ranges were calculated for the ordinal variables, and frequencies and percentages were calculated for the nominal variables. Binary logistic regressions were used to compare differences between the two groups BBS impairment groups versus no BBS impairment at discharge from the SSP for each independent variable. Logistic regression analyses were used to calculate 95% confidence intervals and odds ratios for the associations between demographics, neuropathic pain, foot sensation and balance impairment measure (95% CI and OR shown in Table 2). Results were considered significant at $p \leq 0.05$, 2-sided. In subsequent analyses, CART was conducted to develop a model for identifying factors and measures from baseline and discharge assessments that characterized the 2 groups (balance impairment vs. no balance impairment). The CART determines the best combination of demographics, physical performance measures and self-reported symptoms to yield a more effective classification model than logistic regression models.

CART analysis is a nonparametric statistical procedure that will classify subgroups of patients with no balance impairment and those with balance impairment within a cohort/sample, each with its own set of risk factors and cut points. The analysis starts with the entire cohort and sequentially splits the dataset into 2 subgroups that are the most different with respect to the balance impairment, creating a tree model. The CART model relies on statistically optimum recursive splitting of the patients into smaller and smaller subgroups, based on the critical levels of the predictor variables (Figure 1). The best discriminating variable is selected first and provides the first partition. After this, the remaining variables are examined to determine whether they can provide further discrimination, and this process continues until no further significant discrimination (partitioning) is possible. CART analysis splits a continuous variable into two groups (balance impairment versus no balance impairment) based on an exhaustive search aiming to find the split (including nonlinear splits) producing the largest improvement in goodness of fit for the model created. The percentages of patients with no balance impairment vs. balance impairment were calculated for the base branches of the regression tree.

RESULTS

Demographics. SSP cohort (n=455) were males (n= 200, 44.0%), female (n= 255, 56.0 %) with a mean age of 79.7 ± 10.4; patients were included who had both baseline and discharge assessments of balance

Table 1: Demographic and Balance Characteristics of SSP cohort			**P value**
Subjects	MIRE (Group 1) (n=183)	No MIRE (Group 2) (n=272)	
	n (%)	n (%)	
Male	76 (41.5 %)	124 (45.6%)	
Female	107 (58.5%)	148 (54.4%)	
	Mean ±Stand. Deviation	Mean ±Stand. Deviation	
Age (years)*	78.3 ± 10.6	80.6 ± 10.4	0.022
Number of visits*	14.5 ± 5.7	16.2 ± 8.5	0.021
Berg Score Pre – Treatment	27.6 ± 11.2	27.5 ± 10.0	
Berg Score Post-Treatment**	39.1 ± 13.3	39.7 ± 10.3	0.564
mCTSIB Score Pre-Treatment*	0.9 ± 0.9	1.0 ± 1.0	0.538
mCTSIB Score Post-Treatment	2.4 ± 1.2	2.4 ± 1.1	
Pain Score Pre-Treatment (VAS)*	3.4 ± 3.6	1.4 ± 2.7	0.000
Pain Score Post-Treatment (VAS)	1.5 ± 2.4	0.8 ± 1.9	
Foot Sensation Score Pre-Treatment (SWM)*	2.5 ± 2.1	3.2 ± 2.2	0.001
Foot Sensation Score Post-Treatment (SWM)	5.4 ± 3.5	5.2 ±3.2	

* significance set at p ≤ 0.05 from one-way ANOVA (Group)
**significance set at p ≤ 0.05 from two-way repeated measures ANOVA of Group x Time

CART model for Balance impairment versus no balance impairment on discharge. Table 2 shows the two BBS balance impairment groups' means and standard deviations or frequency and percent for each independent variable, with associated 95% confidence intervals and odds ratios. Logistic regression analyses in Table 2 show that the group classified with balance impairment (Group 1) which was older and had more women had significantly greater balance impairment on initial BBS score and mCTSIB score. Only statistically significant (p≤ 0.05) explanatory variables from the binary logistic regressions were entered into the final CART BBS model.

Table 2: Demographics for Subgroup of Subjects Who Were Tested With BBS (n=165)

	All BBS n=	Group 1 (<45 BBS) n=	Group 0 (≥45 BBS) n=69	95% CI (OR)*
Male	67 (40.6%)	37 (39.8)	30 (43.5)	
Female	98 (59.4%)	56 (60.2)	39 (56.5)	0.578-2.022 (1.081)
	Mean ± SD			
Age	80.3 ± 9.4	82.6 (8.3)	77.32 (10.1)	1.028-1.109 (1.068)
Number of treatments	16.4 ± 8.4	16.7 (8.2)	16.4 (8.7)	0.972-1.048 (1.009)
Berg Score Pre - Treatment	27.5 ± 10.4	22.3 (8.5)	33.9 (8.6)	0.810-0.896 (0.852)
Berg Score Post-Treatment	39.4 ± 11.5	31.7 (9.3)	49.3 (2.9)	
mCTSIB Score Pre-Treatment	1.0 ± 1.0	0.7 (0.9)	1.5 (1.0)	0.287-0.590 (0.412)
mCTSIB Score Post-Treatment	2.3 ± 1.1	1.8 (1.0)	3.0 (0.7)	
Pain Score Pre-Treatment (VAS)	2.2 ± 3.2	1.9 (3.0)	2.6 (3.4)	0.835-1.015 (0.921)
Pain Score Post-Treatment (VAS)	0.8 ± 1.7	0.7 (1.7)	0.8 (1.7)	
Sensation Score Pre-Treatment (SWM)	3.0 ± 2.2	2.7 (2.1)	3.2 (2.2)	0.768-1.024 (0.887)
Sensation Score Post-Treatment (SWM)	5.3 ± 3.4	4.8 (3.4)	5.8 (3.1)	

*95% Confidence Interval (Odds Ratio)

Of the 17variables that were entered into the CART analysis for the 2 groups, the JMP program selected 6 for the tree for balance impairment versus no balance impairment. The fit of the CART model for BBS groups during splitting is shown in Figure 2 with an r=0.648. The accuracy of the CART model for BBS groups was expressed in terms of sensitivity (=0.935), specificity (=0.935) (See figure 3), and overall accuracy and misclassification rate (Table 3).

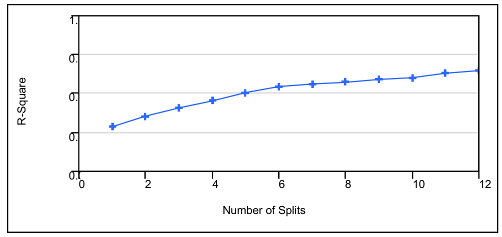
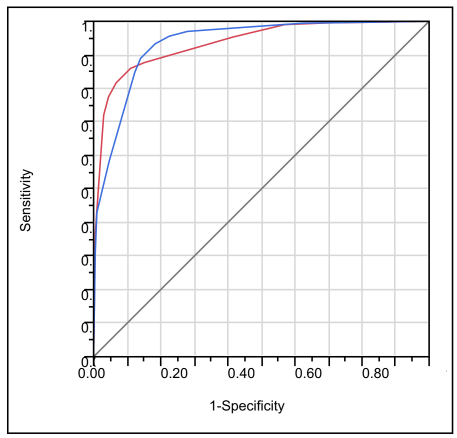


Figure 2. Split History for CART Model for BBS Groups

Table 3 Fit Details of CART Model for BBS Groups

Measure	**Training**	**Definition**
Entropy RSquare	0.6475	1-Loglike(model)/Loglike(0)
Generalized R-Square	0.7880	(1-(L(0)/L(model))^(2/n))/(1-L(0)^(2/n))
Mean -Log p	0.2405	∑ -Log(ρ[j])/n
RMSE	0.2670	√ ∑(y[j]-ρ[j])²/n
Mean Abs Dev	0.1611	∑ |y[j]-ρ[j]|/n
Misclassification Rate	0.0926	∑ (ρ[j]≠ρMax)/n



Legend	BBS Impairment Groups	Area
red line	0	0.9353
blue line	1	0.9353

Figure 3. Receiver Operating Characteristic Curves for Group 0 (Unimpaired balance on BBS) and Group 1 (Impaired balance on BBS)

DISCUSSION

Among older adults over the age of 65, falls are the leading cause of nonfatal injuries, hospital admissions secondary to trauma, and death due to injury [29]. Falls may occur in an older adult secondary to intrinsic and extrinsic risk factors [17]. An individualized multifaceted intervention focused on environmental modifications, balance, strength training, functional mobility, medication management, appropriate vision care, and management of orthostatic hypotension have been recommended as effective treatments to decrease fall risk [18]. SSP is a comprehensive fall risk reduction program delivered by Gentiva® that incorporates individualized care based on the results of a comprehensive assessment.

Currently, there is a lack of evidence that assists physical therapists in determining frequency and duration of treatment [30]. The CART analysis completed on the SSP outcomes provides information that may assist a physical therapist in making decisions regarding determining treatment parameters. The results in table 2 and CART analysis shown in figure 1 demonstrates that the SSP was effective in improving a person's fall risk. The analysis provides the clinician information that may assist in determining intensity of services and identifying individuals at greater risk from baseline clinical and demographic data routinely collected by physical therapists. BBS was able to explain 81% of all the risk that would place a patient into risk or no risk of falling. Patients that score >39 on the BBS intervention may simply require education on decreasing fall risk. Patients that score in the range of < 33 seem to require a higher frequency and intensity of care that the current SSP program is completing. Patients that score between 33-39 on the initial BBS appear to need services that are less intense in frequency and duration as those patients who score <33. The Tinetti POMA as compared to the BBS was less robust at an r^2=0.73. The DGI was slightly less at r^2=0.69. For the mod CTSIB, the classification was dependent on the number of conditions that were scored as impaired.

Misclassification of a patient may have negative effects. If a patient is identified as being high risk but is actually low risk, there is an increased cost to the health care system. A more devastating misclassification occurs when a patient is identified as being a low fall risk when in actuality the patient is at high risk, and sustains a fracture secondary to a fall. In the 9 patients that were misclassified utilizing the Berg, all patients were over the age of 82 and were misclassified as low risk when in actuality they were at high risk of falling. Based on this finding, the BBS may not sensitive enough and further balance testing using the POMA, mod CTSIB, or DGI is recommended in order to determine fall risk.

CONCLUSIONS

This pilot case analysis serves to improve health care service delivery outcomes and resource utilization in a fall risk reduction program. Information gathered from CART enables Gentiva®, physical therapists and policy makers to maximize appropriate referral and utilization of services, and improve efficiency and effectiveness of service delivery. This model addresses a critical health issue for the aging population where this type of analysis is not widely utilized.

REFERENCES

[1] Hausdorff JM, Rios DA, Edelber HK. Gait variability and fall risk in community-living older adults: a 1-year prospective study. **Archives of Physical Medicine and Rehabilitation.** 2001;82(8):1050–6.
[2] Hornbrook MC, Stevens VJ, Wingfield DJ, Hollis JF, Greenlick MR, Ory MG. Preventing falls among community-dwelling older persons: results from a randomized trial. **The Gerontologist.** 1994:34(1):16–23.
[3] Alexander BH, Rivara FP, Wolf ME. The cost and frequency of hospitalization for fall-related injuries in older adults. **American Journal of Public Health.** 1992;82(7):1020–3.
[4] CDC. Cost of fall injuries in older persons in the United States, 2005. Located at: Accessibility verified June 26, 2011. http://www.cdc.gov/homeandrecreationalsafety/Falls/data/cost-estimates-figures1-2.html.
[5] CDC. Fatalities and injuries from falls among older adults-United State, 1993-2003 and 2001-2005. **MMWR** 2006;55:1222-1224. Located at: Accessibility verified June 26, 2011http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5545a1.htm. .
[6] Colledge NR, Cantley P, Peaston I, Brash H, Lewis S, Wilson JA. Aging and balance: the measurement of spontaneous sway by posturography. **Gerontology.** 1994;40:273-278.
[7] Campbell AJ, Spears GH, Borrie MJ. Examination by logistic regression modeling the variables which increase the relative risk of elderly women falling compared to elderly men. **J Clin Epidemiol.** 1990;43:1415-1420.
[8] Rubsenstin LZ, Powers C. The epidemiology of falls and syncope. **Clin Geriatr Med.** 2002;18:141-158.
[9] Steffen TM, Hacker TA, Mollinger L. Age-and gender-related test performance in community-dwelling elderly people: Six-minute walk test, berg balance scale, timed up and go test, and gait speed. **Physical Therapy.** 2002;82(2):128-137.
[10] Lin SI, Liao CF. Age-related changes in the performance of forward reach. **Gait & Posture.** 2011:33:18-22.
[11] Sixt E, Landahl S. Postural disturbances in a 75-year-old population: Prevalence and functional consequences. **Age and Ageing.** 1987;16:393-398.
[12] Horak F. Clinical assessment of balance disorders. **Gait Posture.** 1997;6(1):76-84.
[13] Horak F. Clinical measurement of postural control in adults. **Physical Therapy.** 1987;67:1881-1885.
[14] Pollock AS, Durward BR, Rowe PJ, Paul JP. What is balance? **Clin Rehabil.** 2000;14(4):402-406.
[15] Mancini M, Horak FB. The relevance of clinical balance assessment tools to differentiate balance deficits. **European Journal of Physical and Rehabilitation Medicine.** 2010;46(2):239-281.
[16] Gibson MJ, Andres RO, Isaacs B, Radebaugh T, Worm-Petersen J. The prevention of falls in later life. A report of the Kellogg International Work group on the prevention of falls by the elderly. **Danish Medical Bulletin.** 1987;34(Suppl 4):1-24.
[17] Guideline for the prevention of falls in older persons. American Geriatrics Society, British Geriatrics Society, and American Academy of Orthopedic Surgeons Panel on Falls Prevention. **J Am Geriatr Soc.** 2001;49:664-672.
[18] Summary of the Updated American Geriatrics Society/British Geriatrics Society Clinical Practice Guideline for the prevention of falls in older persons. American Geriatrics Society, British Geriatrics Society Panel on Falls Prevention. **J Am Geriatr Soc.** 2011;59:148-157.
[19]Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. **J Am Geriatr Soc.**1986;34(2):119-126.
[20] Lin MR, Hwang HF, Hu MH, et al. Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. **J Am Geriatr Soc.** 2004;52:1343-1348.
[21] Faber MJ, Bosscher RJ, van Wieringen PC, Faber MJ, Bosscher RJ, van Wieringen PCW. Clinimetric properties of the performance-oriented mobility assessment. **Phys Ther**.2006;86(7):944-954.
[22] Berg K, Wood-Dauphinee S, Williams JI. Measuring balance in the elderly: Preliminary development of an instrument. **Physiother Can**. 1989;41:304.
[23] Berg K, Wood-Dauphinee SL, Williams JI, Maki BE. Measuring balance in the elderly: Validation of an instrument. **Can J of Pub Health.** 1992;83:S7-11.
[24] Bogle Thorbahn LD, Newton RA. Use of the Berg Balance Test to predict falls in elderly persons. **Phys Ther.** 1996;76(6):576-585.
[25] Shumway-Cook A, Woollacott MH. *Motor Control: Theory and Practical Applications*. Baltimore, Md: Lippincott Williams & Wilkins; 1995.
[26] Shumway-Cook A, Baldwin M, Polissar NL, Gruber W. Predicting the probability for falls in community-dwelling older adults. **Phys Ther**. 1997;77:812–819.
[27]Shumway-Cook A, Horak FB. Assessing the influence of sensory interaction of balance. Suggestion from the field. **Physical Therapy.** 1986;66:1548-1550.
[28] Mueller MJ. Identifying patients with diabetes who are at risk for lower extremity complications: use of Semmes-Weinstein monofilaments. **Physical Therapy.** 1996;76:68-71.
[29] Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. Web–based Injury Statistics Query and Reporting System (WISQARS) [online]. Accessed February 20, 2012.
[30] Field-Fote, Edelle C. Does the dose do it? **Journal of Neurologic Physical Therapy.** 2009;33(4):177-178.

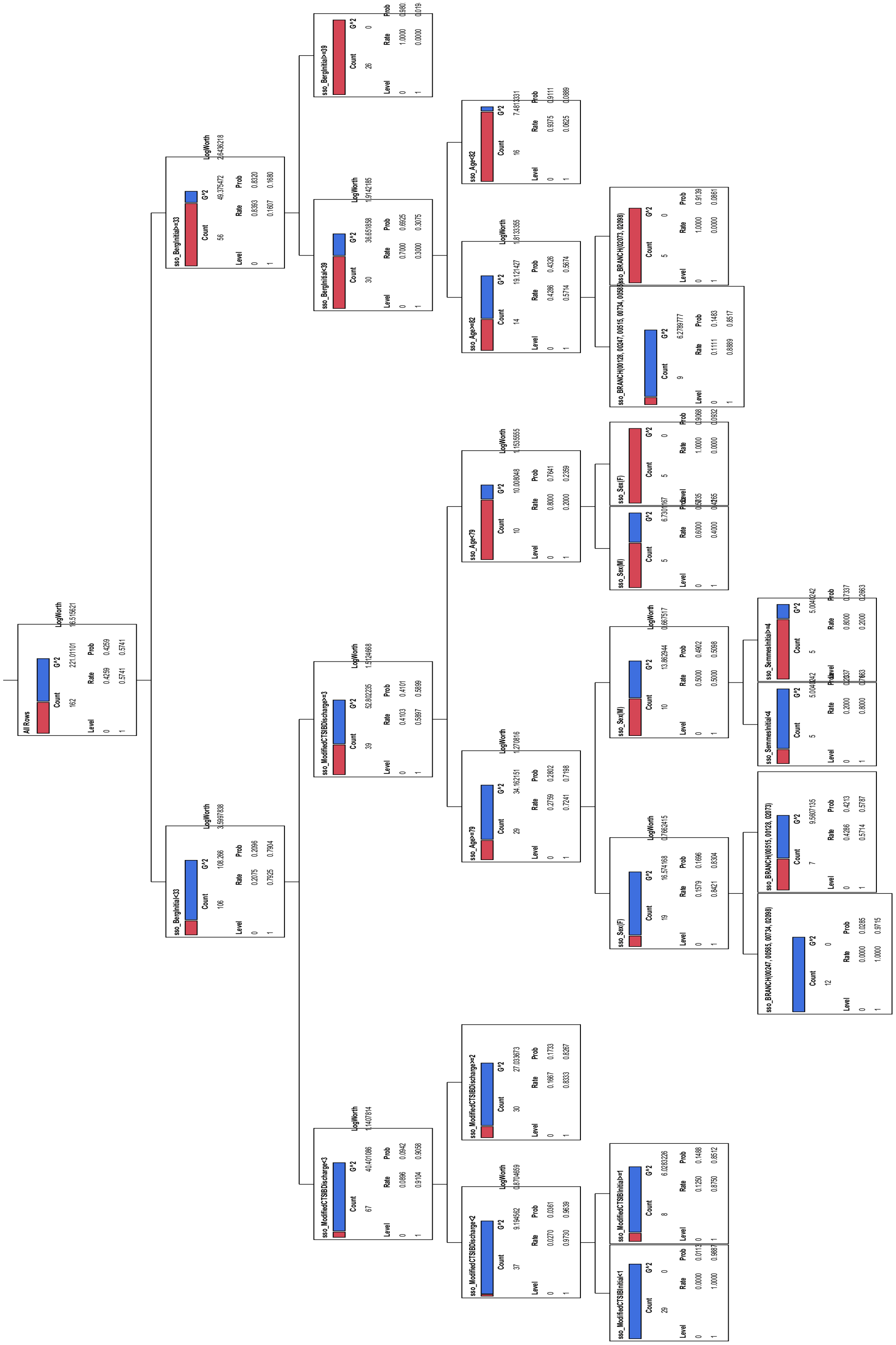


Figure 1: Classification Regression Tree Analysis Model of Falls from Berg Balance Scores.

A Design Pattern Language for Oldschool Action Games

Daniel CERMAK-SASSENRATH
Interdisciplinary Unit, Faculty of Design & Creative Technologies
Auckland University of Technology
Auckland, New Zealand

ABSTRACT

This article discusses the application of an Alexandrian pattern language to the design of interactive systems. It grew out of an University course titled *A Pattern Approach to Action Game Design*, which was offered as an elective in the Creative Technologies program at Auckland University of Technology, NZ, in 2011. We sketch out the idea of design patterns and describe our experiences with the process of using them for designing oldschool action games, that is, finding patterns, making a language, using it for creating several game designs and realizing one of these designs collaboratively. We discuss the concept of the course and present our pattern language and the game we made. While the language is arguably more like a patchy pattern collection, the various game designs quite loose and the realized game unfinished, the process was challenging and intense, and offered students a new perspective on design. In the spirit of design patterns, we only did what the task at hand required, not artificial exercises. We attempted to connect theory and practice in a natural, direct way as we presented, discussed and used everything we did in order to continue our journey. Our course was not aimed at fixed or frozen products, but on a process that is constantly in flux through collaboration by people who interact and share a common pattern language, use, test, revise and refine it while moving on.

Keywords: Pattern design, Methodology, Game Design and Teaching.

INTRODUCTION

Media, artifacts and processes of any complexity are structured by patterns. This observation was the starting point for our course on design patterns. In the course we explored how patterns can be used generatively to inform creative processes for the design of interactive systems: Design patterns "are dynamic. They have force. They are generative." [3]

While Alexander created his pattern language for the domain of architecture ([3], [2], [1]), we aimed at computer action games. The focus was on game play, not on e.g. graphics, physics or coding. We analyzed classic 8-bit and 16-bit computer games, mostly C64 and Amiga games as a large number of these is readily available through emulators (*VICE* for C64 on Mac, *WinUAE* and *WinFellow* for Amiga on PC). The patterns are also identical in old and new action games, it appears, but the patterns are easier to spot in games that are technically limited to the essentials (i.e. interaction), than in the latest quite elaborate and complex games. We limited the scope of our approach to certain types of action games, and we chose to include jump n' run (e.g. *Mario*, *Great Giana Sisters*), shoot 'em up (*Xenon 2*, *R-Type*) and maze games (*Gauntlet*), but excluded sports (*Kick Off*, *Projectile*, *Speedball*), race (*Super Cars*), karate (*IK+*) and sniper games (*Cabal*).

The concept of Alexandrian design patterns has been applied to a range of different domains. They also have already been applied to interaction and game design (e.g. [13], [7], [6], [9], [4], [10], [11], [12]). We could have used (part of) an existing language for computer game design [5] but we preferred to attempt to cover the whole process of finding patterns, formulating a language and using it ourselves. This was done to create a more engaging and challenging situation to learn, to facilitate understanding of the theory, and also to strengthen the feeling of identification with what we do with a sense of discovery. While risking making a pattern collection instead of a complete language and not reaching a very high grade of abstraction or depth, we valued the process more than the product.

THE IDEA OF DESIGN PATTERNS

Following Alexander [3], patterns are seen here as rules of thumb. For instance, when making a barn, build it "in the shape of a rectangle, 30–55 feet wide, 40–250 feet long, the length at least 3x feet, where x is the number of cows the barn has to hold." Include a wide double door for the hay wagon. "Devide the inside of the barn into three parallel aisles: two cow milking aisles down the outer sides, and a central hay-storage aisle [etc.]." While there are myriads of variants and there is constant flux in all systems, they retain a certain kind of invariant "character, a 'thing,' a 'structure,' which remains the same". This invariant structure of entities and relations between them is the area in which design patterns work. They control and (trans-) form it. In effect, what we call a church is a selection of patterns in certain relationships with each other.

A pattern language is an explicit way of notation for design principles. Patterns attempt to express invariant concepts which apply to specific problems in certain situations. They try to "relate [...] context, problem, and solution, in an unchanging way". Each pattern is formulated as a "*three-part rule [...] which establishes a relationship between a context, a system of forces which arises in that context, and a configuration which allows these forces to resolve themselves in that context*".

Patterns are basic, deep, potent, simple, ordinary and easy to understand; they are not mysterious or complicated or to be used only by specialists. Every pattern "is so concrete, so clearly expressed as a rule, and as a thing, that anyone can make one, or conceive one, in the buildings where he lives, or in a building which is going to be created." It takes only little time to "*design a building in this way*. [...] The speed is the essence. It takes time to learn the language. But it takes no more than a few hours or days to design a house." The power to create "lies [...] in the simple mastery of the steps in the process, and in the definition of these steps." But when selections of these simple elements are combined and integrated, they "generat[e] an entirely unpredictable system of new and unforeseen relationships" and complex systems.

We are always dealing with a system of patterns; "patterns are not isolated" but "interdependent, at many levels". Patterns interact with each other, in a system of relationships. "Each [pattern] is incomplete, and needs the context of the others, to make sense." Large patterns give small patterns a place and put them in a certain relation to the whole, and the small patterns realize, facilitate and support the large patterns.

The realisation of patterns depends on the given context. Patterns are relations of relations in many variations; they are not just physical parts, stackable objects or building blocks which are repeated identically. Patterns need to be fitted to specific settings for best results. A system which is alive or anything beautiful cannot be made "merely by combining fixed components" or "*by adding preformed parts*". It can only be generated by a process in which "each part is modified by its position in the whole", and is "different every time [it] occur[s]". While each realisation is uniquely tailored to a specific situation, "[t]he patterns repeat themselves because, under a given set of circumstances, there are always certain fields of relationships which are most nearly well adapted to the forces which exist." The focus on situated activity and the appropriation of specific settings and adaption to certain situations and circumstances, and the unity of action and space appear to point to parallels in phenomenology (e.g. [14]).

The process of adaption to specific local circumstances also favors a wholistic approach of planning and making, giving up the division of mind and matter. It attempts to connect reasoning and acting in a natural way. Planning informs the making and the practice feeds back into the planning. "The person who draws a working drawing cannot draw each window, or each brick, differently, because he has no basis for knowing the subtle differences which will be required. These only become clear when the actual building process is already under way." [3] A design can be created on location, as close as we can get to the actual situation and to the shared, collectively experience.

When creating a design using a pattern language, every single pattern is to be made as intense as possible. "*There is no reason to be timid.*" This process is not about compromise but about creating one strong pattern at a time and relate it to all the other patterns that are already in place in the system, to "go all the way with it". Multiple patterns in one place take not away from each other but complement, enrich and balance each other.

The act of putting a pattern into a system is an act of integration, not addition; it is a fluid process, in which each pattern has the power to "transform [...] the whole design created by the previous patterns". Patterns "are not parts, which can be added – but relationships, which get imposed upon the previous ones, in order to make more detail, more structure, and more substance". This is obvious in game design; a single feature transforms the whole game. The parts the design is composed of "overlap and interlock to such an extent that the oneness of all things becomes more marked". The whole design is transformed with each new pattern which is introduced, and, in turn, each new pattern is also transformed by the patterns and the structure which is already is place. The observation that every act is to be seen in relation to what has already happened connects to Heidegger's notions that acting comes first, and of being thrown into the world, and not being able to step back.

The concept of design patterns sees design as a process in which the whole precedes the parts. The design stays whole during the entire process, while it is being differentiated, or rather, while the patterns in it differentiate themselves. While keeping the system whole in this process, "structure is injected into the whole by operating on the whole and crinkling it, not by adding little parts to one another." The design starts and develops as a single entity. "The form of the whole, and the parts, come into being simultaniously."

The process of using design patterns for building has been discussed for making new buildings so far. But "there is a second, complementary process which produces the same results, but works piecemeal, instead. When a place grows, and things are added to it, gradually, [...] the gaps are filled, the small things that are wrong are gradually corrected, and finally, the whole is so smooth and relaxed, that it will seem as though it had been there forever." This process is the same process at work as before, when making something new, but "stretched out in time".

A pattern language is to be "*morphologically and functionally complete*" for a specific task (e.g. building, blues music [6]). "It is morphologically complete, when the patterns together form a complete structure, filled out in all its details, with no gaps. And it is functionally complete when the system of patterns has that peculiar self-consistency in which the patterns, as a system, generate only those forces which they themselves resolve – so that the system as a whole, can live, without the action of self-destroying inner conflicts." [3] In creating a design, you then only need to follow (invariant) internal requirements and logic, not (changing) outward and external images, trends, styles or pressure. You can let go of your control over the design and "let the pattern[s] do the work." In this process nothing is to be added "except just what the patterns demand". This brings out the "natural, necessary order of a thing". If we already have all the answers before we start our work we cannot listen to what the design asks for. We "must start with nothing in [our] mind[s]", and be "comfortable with the void, [...] confident that the laws of nature, formulated as patterns, [...] will together create all that is required." Such a system which follows internal rules only is free from contradictions that weaken it; it can be pure and strong and true because it is at peace with itself, "in tune with its own inner forces".

A town or a neighborhood is always in flux, constantly changing, not a finished or frozen product. Even more, "[t]here *is* no product [...]: the building and the town, which live, are that incessant flux, which, guided by its language, constantly creates itself." (emph. added) Such systems are alive because they are tested and refined in use.

People can make, adapt and share pattern languages "*for any building task [they] face.*" Anybody "*with a pattern language can design any part of the environment*" and is entitled to do so, as "*it is essential that the people do shape their surroundings for themselves*" because they as users know best: "[W]indows must be shaped by people who are looking out". Large systems as towns are made up of "*millions upon millions of these tiny acts, each one in the hands of the person who knows it best, best able to adapt it to the local circumstances.*" This applies to large systems and small: "Each detail has meaning. Each detail is understood. Each detail is based on some person's experience, and gets shaped right, because it is slowly thought out, and deeply felt."

A pattern language arguably provides a group a people with a means of effectively communicating, "*almost as if they had a single mind*" to collaboratively make a whole, single and integrated structure because, "with a [pattern] language, the assumptions are almost completely explicit from the start". Patterns invite discussion, because they "are not fragile – they are as solid that they can be talked about, expressed quite clearly", challenged and questioned. A pattern like the ENTRANCE TRANSITION [2] "can be shared, precisely because it is open to debate, and tentative." [3] But using patterns is not a mechanical process, guaranteeing anything or a magic bullet for success. "Pattern languages are the source of beauty and of ugliness." The patterns are only as capable as the people who use them.

To test a system or pattern, Alexander argues for querying and trusting people's feelings as humans and users, and not for asking experts' opinions or blindly following fashions. Everybody involved in the process of design "can decide for himself whether [a pattern] is true, and when, and when not, to include it in his world." To judge a pattern he suggests to go to a town, building or place, where the pattern in question is implemented, "and [to] see how [we] feel there", to ask why we like something or not, and to try to identify and isolate the the core of this experience. This will accurately tell us all we need to know about the pattern. This is not asking for our opinions or tastes but purely for feelings. Alexander claims a very high rate of agreement in the cases he did this experiment with the WINDOW PLACE pattern.

Using patterns is not teaching us anything new. But "*they only remind us of what we know already*", in our hearts, "old feelings", what we have forgotten and cannot access. A "*[pattern] language, and the processes which stem from it, merely release the fundamental order which is native to us.*" It helps us "to come more into touch with the simple reality of things, and thereby become egoless and free" and "to be [ourselves]", to "act as nature does". Using patterns is not a goal in itself, and patterns are not cooking recipes, that can or should be followed to the letter; they are concepts, that need to be applied in the spirit in which they were conceived. When we have rediscovered the process which lets us get in touch with our ordinary, deep and "innermost feelings", the use of pattern languages has reached its end.

THE CONCEPT OF THE COURSE

The concept of the course was centered around the idea of combining theory and practice in a natural way. The practical work should be carried by theory, and the discussion of theory should be informed by practical experiences. Exercises should not be artificial or detached from the design process, but everything we do should be presented and discussed, and feed into the next step of the process. All participants work collaboratively on a whole range of tasks differing in scope, difficulty and priority. This provides ways of engaging everybody, giving all participants ample opportunity to identify with the process and to make it their own project. Everybody can discover what he/she can contribute, and try out new things, learn, take risks. The participants were aware that a course like this can quite easily go wrong; it was conceived as an exploration into the unknown, and this definitely added a sense of discovery, surprise and thrill.

While the teaching time was formally devided into lectures, tutorials and lab sessions, in practice, the distinctions were fluent. In the lectures, theory, examples and our experiences were discussed, and students presented the results of the exercises. The theory was mainly focussed on Alexander's *Timeless Way* and *A Pattern Language*, but we also looked at, for instance, Borchers' example of a pattern language for blues music. A number of articles on game design patterns were also referred to.

The collective work in the lab included playing classic games with emulators, finding patterns, creating our own game designs, coding and testing. We started by implementing our own *Pong* and *Tron* versions (Figure 1) to get going. Usually, we would start to work together on an exercise right after the lecture. This should provide an immediate positive hands-on experience. It makes people feel part of the process and helps to reduce both the distances between the topics of the lecture and the own work, and between the participants. It also helps to reduce the amount of time before people actually start to engage. The practical work should trigger the need for theory, create questions, and offer experiences that can be discussed in the lecture. This makes the theory appear less artificial and give the practice background and value, and also provides it with reasons and goals. We favored group work over individual work to not only to make people collaborate with and motivate each other, but also learn from one another and challenge each other's ideas.

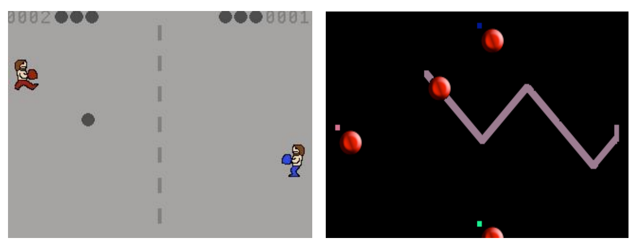


Figure 1: Boxing cavemen *Pong* and diagonal 4-player *Tron*

There were ten theoretical and practical exercises, nine of which had to be done to pass the paper. Some of the exercises were to be done individually, the rest in 3-person teams. Part of most exercises was a presentation in class and a hand-in (i.e. a pdf or source code). Usually, the topics of the lectures trailed the exercises by one week to enable students to first make their own experiences to which they then could relate when discussing theory in the lecture. The paper relied heavily on students' participation, so expectations in this department were high.

DISCUSSION

At the end of the semester, students were invited to give verbal feedback about their experiences and opinions on the use of design patterns. Additionally, views expressed in the final excercise, the reflective statement, are collected here. Students commented on different aspects of pattern design, some specific to our course, some quite general. The discussion of how the concept of the course worked in practice is centered around the idea of creative and generative use of design patterns, not on everyday teaching and learning issues.

Students found the process of playing classic action games, looking for patterns and identifying their essential properties enjoyable and rewarding. It provided them with a sense of discovery and ownership. Many people said it opened their eyes to the concept of patterns. "[...] I thought it was going to be difficult to find the patterns. Instead, I found that once I started looking for the patterns, they were absolutely everywhere." (Reflective Statement) Students also commented that this was the point they began to understand what we were talking about in the lecture. "Much of the class came up with similar patterns so we classified them and worked towards building a final pattern language. This helped us all tune in with each other on what a pattern language actually was and how things should be categorised. It worked well." The lecture clearly benefitted from this experience. We used a free real-time multiuser online text editor [8] for working on the pattern language. This facilitated a feeling of an onging process among the participants, because everybody could always access the latest version of our pattern language, use and change it. "With the online collaboration tool Etherpad we were able to alter and read the document in real time as edits were being made, and see the formation of a document. We were able to influence each other and be influenced as the document took shape." We were constantly on the move on this trip.

Students enjoyed creating their own game designs. At that point, our pattern language was still very loose, and a number of patterns were certainly added to the designs as afterthoughts. Nevertheless, people were aware of patterns, and used them to some degree. And it was fun. "The most interesting part was when we all had to come up a game design idea to present for an action game." Several game designs were created, presented and discussed. All the game designs were for jump n' run games, which was surprising; apparently *Xenon*-style shoot 'em up and *Gauntlet*-style maze games are out of fashion at the moment. There were no really radical designs, but more detail tweaking, copying popular games, some transformation and sadly no multiplayer. Among the game designs were *Dragon Eggs*, a *Mario*-esk "Action Platformer" (Game Design Document) with "Medieval and Fantasy themes"; *Radical Hamster Force*, "Kind of like a mix of Alex the kid, Mario, and Kirby combined but with optional weapons"; *Krystal in the Hood*, a "classic platform game" in which the player has to "move from left to right and from top to bottom" through multiple levels; and the later realized *Super Bush! Chronicles*. We voted for one of these game designs to be realized and integrated features (i.e. patterns) from the other designs. The voting was appreciated, and taken quite seriously.

The single most successful aspect of using patterns was arguably their benefits for group work. "Everyone managed to work together in teams and made the best out of their abilities to achieve the game". The use of patterns helped quite a heterogeneous group of more than ten people to identify and actively engage with a single project, and facilitated collaborative decision-making in our numerous meetings. Contributions could be were very specific and to the point, e.g. in discussing the game designs. "I don't think it would have been a bad or significantly different game if we hadn't taken the process of developing and deciding on patterns prior to writing up a game design document. But it helped the process and I think it allowed for a more concrete development of the project. Using patterns you know what you want, then design around that idea." Students came to see patterns as an interesting design methodology. They described them to be a very useful tool, a "powerful [...] development technique".

However, while patterns seemed to work well for creating game designs, it was different during the implementation phase in which "getting the game working became the focus instead." Students remarked that implementing the game simply had not much to do with the idea of using patterns. "The [pattern] language helped the game designers to not miss important parts of a game out and to think about how parts interacted. From there however, we stopped thinking about patterns." Coding *C++* proved to be hard as most people were inexperienced coders. Getting the basic functionality right was challenging enough, and possibly hindered access to the process on a higher, more abstract and interesting level. "While [*C++* is] fast and powerful there are certain low level elements that one cannot avoid using." Trying to get e.g. the sound working "was a waste of time that I didn"t have to waste." Students suggested using a game engine instead. An advantage was that everybody working on the code knew what, why and when something needed to happen, which was very helpful for collaboration. "Whoa, I thought, that's an interesting approach." People could be quite specific about what they wanted to do, and what they wanted other people to do. "At first I thought that the patterns weren't going to be necessary, but as we progressed I found them to be quite essential in terms of laying out the game – because we knew what we wanted, where to implement it, how it worked and it was all written down and discussed with the group." Despite the difficulties, people commented that making the game was a fun and very intense experience, and that it felt quite magic to see how the patterns came to life.

Some students felt that patterns "restricted [their] creative freedom". As this was our first use of patterns, the process may have been quite mechanical, and not as fluent, spontaneous and radical as it could be. "The world is chaos and unorganized, putting everything into an organized list makes life boring and ends opportunity for innovation and creative thinking." Our language was not very deep and powerful. "[...] I feel my common sense and knowledge of how games work being much more helpful to me than using a set of rules. Creative decisions yielded better results and just experimenting until it feels right". While the question of creativity and patterns was addressed in one of the lectures, it was apparently not discussed clearly enough. Patterns have nothing to do with "what I would call cloning", and a person using a pattern language is certainly not "playing it safe". Patterns are not only for people who "can't or don't need to think creatively". Of course, there is some truth to the matter. Patterns need a closed system to work, e.g. an engineering problem to solve. It would appear absurd to have a pattern language for creating art, for example.

Initially, people were sceptic of the generative force of patterns. "How could that possibly work? Sure, it was fascinating to look at already completed games and see how they could be broken down into patterns, but I didn't completely believe that the reverse could be done – taking patterns and creating a game from them." During the semester, students were questioning the idea of letting go and not trying to control the whole design top-down. They were surprised by the drive and the immediacy with which patterns asked for action. "[...] honestly I didn't feel like it was going to work. I thought that something bad was going to happen that would stop the progress of the project and slow it down for everyone [...] but fortunately I was mistaken."

AN ACTION GAME DESIGN PATTERN LANGUAGE

In our course, we created a pattern language for action games, i.e. jump n' run, shoot 'em up and maze games. Excluded are, although these are arguably also action games, sports, race, karate one-on-one and sniper games. Our patterns roughly follow the format of Alexanderian patterns. Each one has a name, a certain context or situation, a short description, and is connected to larger patterns (above) and smaller patterns (below). In many cases examples of an occurance of a particular pattern in a game are given. All participants worked collaboratively on this language. They wrote, edited, moved, revised and deleted patterns. While a number of patterns was identified and described, the result is more like a collection than a complete language. The overall number of these patterns still needs to be reduced, the hierarchy needs to be revised, and patterns need to be linked and related to each other.

The titles of our patterns are IN-GAME OBJECTIVES, MORALITY, SOMETHING TO DO FOR THE PLAYER, ACTION CONSEQUENCES, REWARD FOR RISK, CHARACTER, SPECIAL ABILITIES, WEAPONS, ARMOR, ENEMIES, CIVILIANS, LIMITED LIVES, HEALTH BAR, MANAGE CHARACTER, POWER-UPS, ITEMS TO COLLECT, SHOP, A SETTING FOR THE GAME, LEVEL THEMES, SECRETS, HIDDEN CHAMBERS, INVISIBLE GOODIES, SHORTCUTS, CHEATS, TRAP, COMPETITION BETWEEN PLAYERS, QUICK MOVEMENT, LINEAR FLOWING GAMEPLAY, EVERCHANGING ENVIRONMENTS, and GAME GETS HARDER. Because the pattern collection is too large to be included here completely, three example patterns of different abstraction are given.

REWARD FOR RISK

One of the most common patterns in action games is REWARD FOR RISK. It differs from most other patterns in that it is an abstract pattern – it describes a style of gameplay rather than an actual object in the game.
Why is REWARD FOR RISK such a useful pattern to implement in action games? Because it creates a psychological hook for the player. The human brain is wired so that if we successfully complete something risky, we get rewarded with a short burst of positive endorphins, along with an immense feeling of relief and satisfaction. Very quickly, the player gets addicted to this short emotional high, and is willing to invest significant time in a game to experience it. While this pattern is characteristic of the gambling genre, it is also an essential pattern for Action games as it keeps the player engaged with the game.
Examples of the REWARD FOR RISK pattern include: Having to risk your life against a difficult boss to beat a level, being able to cross a dangerous lava pit with the potential reward of an extra life and fighting more challenging monsters to get better loot.
The risk is something that has to be balanced carefully if the reward is critical to the main gameplay; if a game is too hard to complete then players will rage quit. If a game is too easy then players will get bored. However, if the reward is something that the player does not necessarily need to finish the game, then the risk can be as high as you want.
Goes with patterns: ENEMIES, TRAP, INVISIBLE GOODIES, SECRETS

SHOP

In a game with many different enemies and/or levels, it might be interesting to offer the player the possibility to decide about what weapons he wants to have. Players can buy and sell weapons and other equipment, and if the prices vary between shops, they can even trade with them. Shops might be localized, e.g. vary in offer and price. Shops can be located anywhere in a level, but most commonly between levels or at the halfway point.
Most shops in games of this type contain a very basic interface. There is usually a basic, easy-to-navigate scrollable item menu – either filling the screen, or over a graphic depicting a shop counter. Sometimes, however, a shop will only appear as an options screen or dialog box after a particular action has been completed, asking you whether or not you want to buy or upgrade something. Upgrades to weapons, armour or vehicles are usually available, and better enhancements cost more. Other items that can often be bought from game shops include ammunition, damage boosters, and health items. Items from the shop are usually paid for with items collected in levels, or using an in-game currency that collectables or score can be exchanged for.
Therefore: Put a SHOP into your game when you want to add an element of strategy to the action game, and give the player control over his abilities/equipment. Vary offer and price between shops to enable trade. Place them at the end of levels or at the half-way point.
Goes with patterns: ITEMS TO COLLECT, POWER-UPS, WEAPONS, ARMOUR, SPECIAL ABILITIES
Examples: *Xenon 2*, *River City Ransom*

TRAP

Is part of the patterns: SECRETs
In addition to enemies, traps can be dangerous to players. They can be easy to see or hidden. There are many different types of traps. Most traps are part of the level and cannot be defeated or destroyed, simply avoided.
Put TRAPs in your game to add a sense of discovery to it. Players will then carefully observe every detail in your level design. Traps should be visible (as in *Rick Dangerous*), and not only be found by trial-and-error (as in *Lost Vikings*). Traps can also be dangerous for enemies, therefore enabling the player to use them for his advantage, adding a twist to the game beyond shooting at everything that moves. There should be a reason for the trap, and a payoff for defeating it, e.g. a bonus.
Goes with patterns: ENEMIES, REWARD FOR RISK
Examples: *Rick Dangerous*, *Lost Vikings*

ACTION GAME: *SUPER BUSH! CHRONICLES*

Super Bush! Chronicles (Figures 2 and 3) is a single-player jump n' run game. It is about a panda bear defending its jungle against fierce goblins who want to build a town at this location to "support their gambling needs" (Design Document). The game includes a number of patterns from our pattern language: SOMETHING TO DO FOR THE PLAYER, ACTION CONSEQUENCES, IN-GAME OBJECTIVES, REWARD FOR RISK (score, goodies), CHARACTER (panda bear), MORALITY (helping a good cause, defending the forest, liberating caged jungle animals), SPECIAL ABILITIES (jumping very high and bamboo stick kendo, eat weapon upgrades (bamboo stick) to boost health), ENEMIES (goblins with axes, chainsaws or guns, boss goblins in construction vehicles, i.e. bulldozers), LIMITED LIVES (three), HEALTH BAR (for player and goblins), MANAGE CHARACTER, ITEMS TO COLLECT (nuts as currency), SHOP (buy armour and weapon upgrades), COMPETITION BETWEEN PLAYERS (through saved hiscores), GAME GETS HARDER (increasing number of enemies and traps), TRAPS, A SETTING FOR THE GAME (conflict over natural resources) and LEVEL THEMES (five levels with slightly different themes).

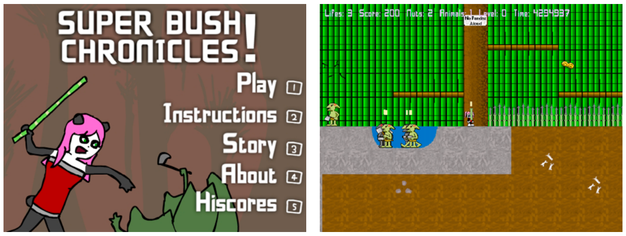


Figure 2: *Super Bush!* title screen and in-game screen shot

We recorded our own sound and drew original graphics. The game is controlled with the keyboard. It was implemented using *C++* and the *SDL* library. A PC download is available at www.dace.de.

Figure 3: Weapon and armour shop, goblin sprite sheet

CONCLUSION

The course on design patterns was quite ambitious and challenging. We did not have much time to discuss theory, play games, find patterns, make a language, use it for designing action games and realize one of the designs. It was an intense trip aiming to combine practice and theory, experience and reflection. The theory was in many cases the subject of the exercises, or informing the practical work to a considerable degree. The practical work relied on an understanding of the theory, in a direct and natural way. We only did what the task at hand required and followed the internal logic of the process. We presented, discussed and used everything we did. Everybody was always aware why he/she was doing something and why this was necessary. The students were engaged and interested in the new perspective the concept of patterns could bring to their practice, and everybody was curious if it would work for us.

Of course, we had some problems that held us back. Most students were not experienced in game (or interaction) design, and many were novice coders. Only few had an overview of classic action games. The students did not have a solid base of design knowledge that they could apply through the new perspective of design patterns. In creating and discussing the game designs, students did not feel the intrinsic necessity to strictly follow-through with design patterns; this points to our language being not complete, as new ideas were constantly suggested at all stages of the design process.

On the other hand, there were solid benefits in using patterns. The collaboration was working well, and discussions were very specific and to the point. Students were keen to participate and many invested a lot of time and energy. Everybody could quite easily make a relevant contribution, and this got people deeply involved and interested. Students had the feeling of genuine discovery and ownership. We did everything by ourselves, and all of us shared the design and the process. The process of finding, describing and using patterns could have gone seriously wrong, and this definitely add some thrill to it. The course was not over-prepared. We faced real questions, issues and problems, and needed to find solutions. To include the possibility of real failure opened up the possibility for real success.

And it was fun. We gained more from the process than we invested, it appears. The process developed a kind of momentum of its own. All participants saw how our various, loose and general ideas for a game were transformed into a coherent design, and then, towards the end of the semester, how the design was turning into a working game, literally in the course of a few days (and nights). This was a quite impressive and also a strange experience.

REFERENCES

[1] C. Alexander, M. Silverstein, S. Angel, S. Ishikawa, D. Abrams, **The Oregon Experiment**, New York: Oxford University Press, 1975.

[2] C. Alexander, S. Ishikawa, M. Silverstein, M. Jacobson, I. Fiksdahl-King, S. Angel, **A Pattern Language**, New York: Oxford University Press, 1977.

[3] C. Alexander, **The Timeless Way of Building**, New York: Oxford University Press, 1979.

[4] S. Björk, S. Lundgren, J. Holopainen, "Game Design Patterns", **Level Up: Digital Games Research Conference**, 2003, pp. 4–6.

[5] S. Björk, J. Holopainen, **Patterns in Game Design**, Hingham: Charles River Media, 2005.

[6] J. Borchers, **A Pattern Approach to Interaction Design**, Chichester: Wiley, 2001.

[7] D. Church, "Formal Abstract Design Tools", **Gamasutra**, July 16, 1999. www.gamasutra.com/view/feature/3357/formal_abstract_design_tools.php (16/6/2011).

[8] **Etherpad**, www.ietherpad.com (Nov 4, 2011).

[9] B. Kreimeier, "The Case For Game Design Patterns", **Gamasutra**. March 13, 2002. www.gamasutra.com/view/feature/4261/the_case_for_game_design_patterns.php (16/6/2011)

[10] B. Kreimeier, "Content Patterns in Game Design", **Game Developers Conference**, 2003.

[11] P. Lemay, "Developing a pattern language for flow experiences in video games", **Situated Play**, Proc. DiGRA 2007 Conf., 2007, pp. 449–455.

[12] R. Nystrom, **Game Programming Patterns** / Getting Started / Introduction. gameprogrammingpatterns.com/introduction.html. Last modified on July 08, 2010 (2/6/2011).

[13] Z.B. Simpson, "Design Pattern for Computer Games", **Computer Game Developer's Conference**, Austin, TX; San Jose, CA. Nov 1998, May 1999. www.mine-control.com/zack/patterns/gamepatterns.html (6/7/2011)

[14] L.A. Suchman, **Plans and Situated Actions: The Problem of Human-Machine Communication**, Cambridge: Cambridge Univ. Press, 1987.

Parametric Analysis of Acoustical Requirements for Lateral Reflections: Melbourne Recital Hall Case Study

Design and Modeling Methods and Methodologies

Erica Claustro and Jacquelyn Haas
Masters of Architecture Candidates
College of Art and Design, New Jersey Institute of Technology
Newark, New Jersey
Email: ejc8@njit.edu and jh233@njit.edu

This paper is an investigation of the Melbourne Recital Centre as a case study to define the parameters necessary for good acoustical quality as it relates to the Binaural Quality Index and determining the intimacy of the hall by its initial time delay gap. The Melbourne Recital Centre, designed by Ashton Raggatt McDougall Architects, is a significant case study, as its design was driven by the acoustic requirements of reflection and diffusion through Odeon Acoustical Software. It achieves the same acoustical quality of older, ornately designed shoebox concert halls, from the perspective of contemporary design and fabrication tools and techniques. The sleek design of the Melbourne Recital Centre successfully reflects sound waves in low, mid, and high frequencies due to corresponding wall panel differentiation in the corresponding scales, as engineered by Arup Acoustics.

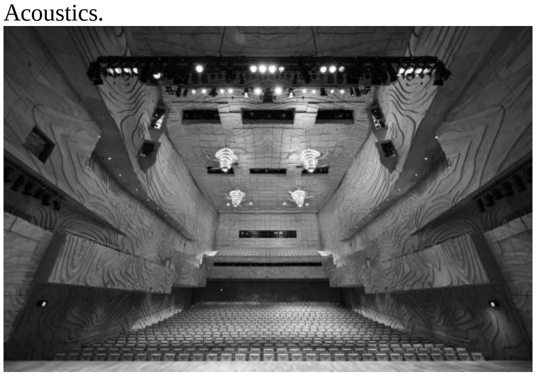


Figure 1: Melbourne Recital Centre and Melbourne Theatre Company buildings by ARM (Photo: John Gollings)

The analysis of Melbourne Recital Centre consisted of visual analysis of photographs and orthographic drawings of the concert hall. After which, the Melbourne Recital Centre could be modeled in Autodesk Inventor with specific dimensions parametrically controlled. Modeling of the Melbourne Recital Centre was for two purposes: to first understand acoustical mechanisms of the hall, and to parametrize certain dimensions and test the effects of changes on certain parameters. In its simplest expression, a parameter could simply be the dimension of a line controlled by the parametric software; it is a distinguishing factor that can be variable either numerically, geometrically, or algorithmically.

Modeling of the Melbourne Recital Hall is achieved through extrusions in both plan and section. In both directions, outlines of the elements were drawn in sketches; in sectional elements, parameters are controlled in the planar directions, while planar elements are controlled sectionally. From the sketches, constraints and driven dimensions must be carefully considered, in order to force the sketches to respond to parameter changes as desired. The sketches were then extruded to differing dimensions, which are not parametric dimensions, although they do differ between
extrusions.

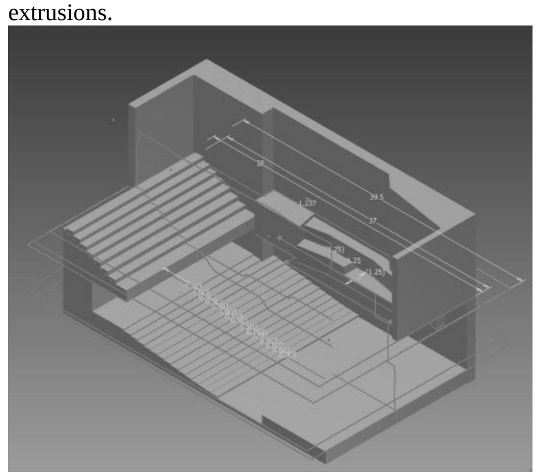


Figure 2: Screenshot of Autodesk Inventor model of Melbourne Recital Centre with visible parameterized dimensions

These parameters are significant, because it is these dimensions that control the amount of protrusion in to the original "shoebox" shape, i.e. balconies, side walls and panels. In particular, these sets of parameters reflect specific ratios between the amount of protrusions between the balconies. Increasing or decreasing these parameters should affect the ratio of Surface Area to Volume of the space, another important acoustic performance parameter. The hypothesis of this paper is to investigate the assumption that a higher Surface Area to Volume ratio would indicate more protrusions into the classic "shoebox" concert hall shape, therefore affecting lateral reflections, ultimately influencing BQI and ITDG. Additionally, these specific ratios of the depth of protrusions could affect lateral reflections, and therefore BQI and ITDG. Testing the lateral reflections would require two additional elements; the accurate measurement of Surface Area, and connection to the stress analysis tool results, explained below. Achieving an accurate measurement of Surface Area is problematic due to the nature of the model construction; because all of the elements of the hall were modeled in "parts", due to the variations in parameters, there is much Surface Area that is overlapping or redundant. An overall measurement of Surface Area would produce a much higher value than occurs in reality due to this phenomenon. Therefore, the main challenge in calculating surface area would be how to only take into account visible surfaces.

The Binaural Quality Index is one of the most accurate predictions of a hall's acoustical quality, especially for shoebox halls. The BQI measures the difference in lateral reflections as it reaches a listener's ears; the more difference between one's ears, the better the perceived acoustical quality. The BQI only takes into account lateral reflections; sound waves from directly ahead of a person, such as those direct sound waves from a stage, do not affect BQI.

The Initial Time Delay Gap is examined when determining the acoustical intimacy of a space. A recital hall is considered to be intimate when one listens to music in a large space but experiences as if in a smaller room. ITDG is the interval between the arrival of direct sound and the first reflection at listener. The smaller the ITDG, the more intimate the space. For intimate recital halls, it is necessary to aim for time delays for first reflections of 20 milliseconds or less.

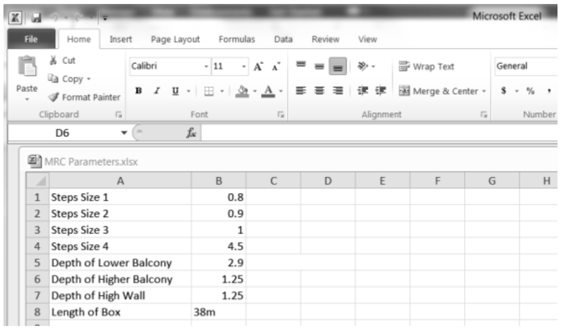
Wall depth, soffit, and balcony dimensions were recorded in an Excel chart, with the numerical values analyzed and transformed into algorithms based on dependent relationships between the architectural elements. The benefit of using a program such as Autodesk Inventor is the ability to use secondary programs such as Excel to analyze and create formulaic relationships between data; this forms the basis of the research needed to generate a design process for alternative forms. By setting one of the balcony depths as a base value, the other values of balcony depths could be re-written as equations in terms of the base balcony depth, thus setting up proportional relationships between the depth values. This base value could then be modified, with the other depth values updating correspondingly. Furthermore, these algorithms could be changed to reflect different proportions. Additionally, the Excel file contains dimensions relating to overall volume calculations. Once a method for measuring Surface Area has been calculated, these values can then be integrated into the Excel file as well.

It is hypothesized by using the force analysis tools in Inventor, the magnitudes of both indexes can be visualized. Part of the research involves generating new surfaces using existing parameters of Melbourne Recital Centre, such as a spherical network that could be used as a concert hall interior surface. By applying these parameters derived from the Melbourne Recital Centre case study, optimized for lateral reflections, BQI, and ITDG, one can analyze the sound wave reflections off these created surfaces and predict hypothetical acoustical quality. In separate part files, three spheres were created, and then given the three parameters corresponding to the lateral wall protrusions. These spheres then take on the proportional ratios of the balconies, as determined by the algorithms in the Excel file.

Ray tracing of lateral reflections was done in Inventor. A sound source point was created at the center of the edge of the stage and linked to particular points on the wall surfaces in a 3D sketch. In this case, the spherical surfaces. A plane normal to the point on the curved surface and a plane through the source point, the point on the surface, and a third point was created in order to produce a work axis through the intersection of those two planes. The angle of reflection and thus the angle of incidence was found by measuring the angle of the line from the source point to the point on the surface to the work axis. This process was repeated for each spherical surface.

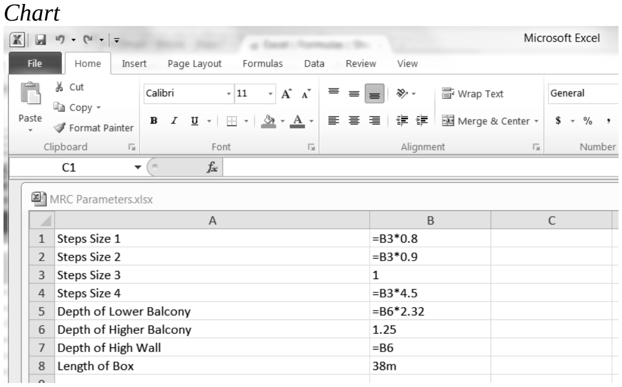
In order to graphically visualize the sound pressure of the ITDG, the analysis tool was tested by using the Remote Force option. The ITDG Limit line was created by creating a line perpendicular to the reflection and extended to intersect with the line of incidence. Any point along the line of incidence before the point where this line intersects with the ITDG limit line is along the ITDG. Coordinate points along the ITDG and approximate magnitude of sound generated by an orchestra (pascals converted into pounds per square foot) were input into the Remote Force Analysis tool. A vector is produced under this tool which represents the magnitude and direction of the force. In this case, the sound pressure is the force being represented. Because the ITDG is dependent of the sound source and listener position, the surface representing the chair was tested. The vector was positioned in the direction of the chair surface and ran an analysis. All three spheres were tested. The visual results produced by the analysis tool showed that the sound pressure was insignificant on that chair surface. Thus, the ITDG was too long. If the ITDG was in the preferable range of 12 to 25 milliseconds, hypothetically, the sound pressure would arrive on the surface and show up visually by the analysis tool. In theory, the analysis tool could produce graphically visual results; however, the scale of the magnitude of force (sound pressure) that was being tested versus the scale of the magnitude of forces Inventor (i.e. tension and compression) was created to measure, the results do not seem visually significant. If there was a way to change the scale, perhaps the visual results would vary.

It was hypothesized that the sphere surfaces, having extreme curvature, thus a wider angle of reflection, would create a longer ITDG, proving that these surfaces would not be preferable for an intimate hall. For an intimate recital hall, ITDG should ideally range from 12 to 25 milliseconds. Although the depth of the spheres share the same parameters of the balconies of the Melbourne Recital Centre, the angle of curvature produces significantly results in terms of striving for intimacy. If the balconies of the MRC were tested in the same way, hypothetically, the results would show the sound pressure on the chair surface because the recital hall is considered to be intimate, as well as the curvature of the balconies and side walls are not as extreme as the produced spheres.



	A	B
1	Steps Size 1	0.8
2	Steps Size 2	0.9
3	Steps Size 3	1
4	Steps Size 4	4.5
5	Depth of Lower Balcony	2.9
6	Depth of Higher Balcony	1.25
7	Depth of High Wall	1.25
8	Length of Box	38m

Figure 3: Screenshot of Excel Parameter Chart



	A	B
1	Steps Size 1	=B3*0.8
2	Steps Size 2	=B3*0.9
3	Steps Size 3	1
4	Steps Size 4	=B3*4.5
5	Depth of Lower Balcony	=B6*2.32
6	Depth of Higher Balcony	1.25
7	Depth of High Wall	=B6
8	Length of Box	38m

Figure 4: Screenshot of Excel Parameters with embedded algorithms

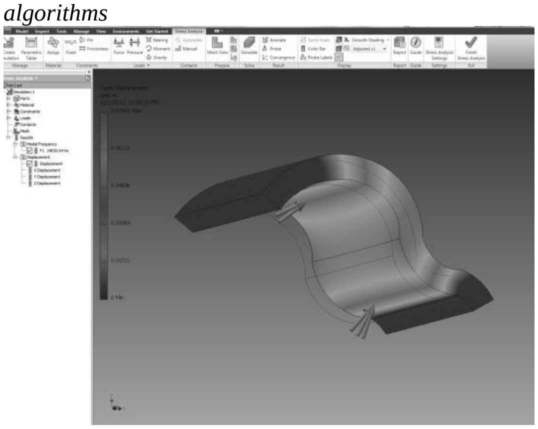


Figure 5: Screenshot of Autodesk Inventor force analysis tool

This evaluation of the Melbourne Recital Centre creates an opportunity to generate rules and algorithms applicable to the design of concert halls that strive for intimacy and optimized acoustical quality. This overall process can be repeated in other categories of design as well, in which certain ratios and relationships are vital to the performance of other architectural systems. A benefit of this process is that it can be done relatively quickly and inexpensively. This is a mechanical and manual way of testing the acoustics on a portion of a surface. The leading industry software, Odeon Acoustical software, is a highly complex software designed to measure many acoustical qualities in many different kinds of environments, from train stations and airport terminals to restaurants, concert halls, and even outdoor spaces. Odeon is capable of importing 3D models, or one can model surfaces in Odeon itself. Additionally, Odeon can incorporate the absorption data of surface materials.

Compared to the Odeon Acoustical Software, the leading software program in the industry to test acoustics, this can be done by a student for free using Autodesk's free three year trial for students, and the Microsoft Office Suite included on most computers. A university license for Odeon Acoustical Software for one seat starts at over 5,000 euros, with basic commerical packages jumping to over 14,000 euros for one seat..

Research was conducted in conjunction with a graduate architectural elective course, Arch 662, entitled, the Design of Parametric Space, taught by Assistant Professor, Rhett Russo, at the New Jersey Institute of Technology, in the Fall of 2011.

References

Beranek, Leo. "Concert Hall Acoustics—2008." *JAES* 56.7/8 (2008): 532-44. *Leo L. Beranek*. Web. <http://www.leoberanek.com/>.

Beranek, Leo L. *Concert Halls and Opera Houses: Music, Acoustics, and Architecture*. New York [etc.: Springer, 2010. Print.

Beranek, Leo Leroy. *Acoustics*. New York, NY: Published by the American Institute of Physics for the Acoustical Society of America, 1986. Print.

Leonid Makrinenko, *Acoustics of Auditoriums in Public Buildings*, Acoustical Society of America, Woodbury, NY, (1994).

"Melbourne Recital Centre." *Catalyst*. ABC. Australia, 02 Dec. 2009. *Abc.net.au*. Web. 22 Sept. 2011. <abc.net.au/catalyst/stories/2488923.htm>.

Orlowski, Raf. "The Acoustic Design of the Elisabeth Murdoch Hall, Melbourne, Australia." Acoustics '08 Paris, Euronoise. Web. 20 Sept. 2011. <intellagence.eu.com/acoustics2008/acoustics2008/cd1/data/articles/001069.pdf>.

"The Building: Melbourne Recital Centre." *Melbourne Recital Centre*. Web. 27 Sept. 2011.

<http://www.melbournerecital.com.au/aboutus/the_building>.

Parkin, P. H. *Acoustics Noise and Buildings.* London: Faber and Faber, 1969. Print.

Room Acoustics | Noise Prediction Software | Sound Absorption | Acoustic Modeling | Loudspeaker Cluster | BRIR. Web. 27 Sept. 2011. <http://www.odeon.dk/>.

Parametric Constraints in Acoustical Design: The Case Study of Teatro Degli Arcimboldi

E. Lucia Huerta
College of Architecture & Design
New Jersey Institute of Technology
Newark, NJ 07102 USA

ABSTRACT

This paper investigates the use of parametric constraints in digital modeling for the purpose of acoustical analysis and design development. Parametric associations consist of an arbitrary constant whose value characterizes a set of physical properties whose values then determine the characteristics or behavior of the constituent parts [3]. The primary objective of this research is to focus on the acoustic panel parameters in the design stall wall of the Teatro degli Arcimboldi. The key design parameters in the wall are: the panel's curvature, its spline position, and degree of tilt. His paper will discuss how these parameters affect the acoustic environment of the hall. Through the use of a 3D parametric model constructed in Autodesk Inventor, the following acoustical conditions will be examined: focalization, absorption, reflection, and diffusion.

Keywords: Parametric, Focalization, Absorption, Reflection, Diffusion, Constraints

1. INTRODUCTION

A working design methodology will be explained that incorporates numerical design data from Microsoft Excel and its visualization within the 3 dimensional environment of Autodesk Inventor. Analysis will then be performed through the alteration of this data set and observed in real time in the 3-dimensional Inventor model. Each unconstrained component in an Inventor assembly has six degrees of freedom. Degrees of freedom are defined as 3 for translation and 3 for rotation in each axis XYZ. This definition applies to mechanical transformation. This research adapts the mechanical paradigm of degrees of freedom to the acoustic performance of the stall wall design.

2. DEGREES OF FREEDOM

First Degree of Freedom: Panel Curvature

Even dispersion of sound eliminates "dead spots" and "live spots" known as focalization. When sound waves from a point source strike a plane wall, they produce reflected spherical wavefronts (Fig 1) and focalization occurs when reflected sound waves converge onto one another resulting in intensified overlapping in some areas and sparse convergence in other areas [2]. Generally, concave shapes are used to focalize sound because all rays from the focus of a parabola to its surface will be directed outward as parallel rays, while convex shapes are considered to be anti-focal because they disperse sound multi-directionally [2].

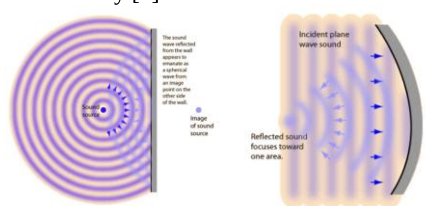


Fig1<http://hyperphysics.phy-astr.gsu.edu/hbase/sound/reflec2.html#c3>

In the case of the concave wall that surrounds the stalls in the Teatro degli Arcimboldi, diffusion of sound waves occurs despite its parabolic shape and thin wood elements, a material usually valued for its reflective qualities. The rigid wood slats are anchored to the concave structure which frames a well dampened space (Fig 2).

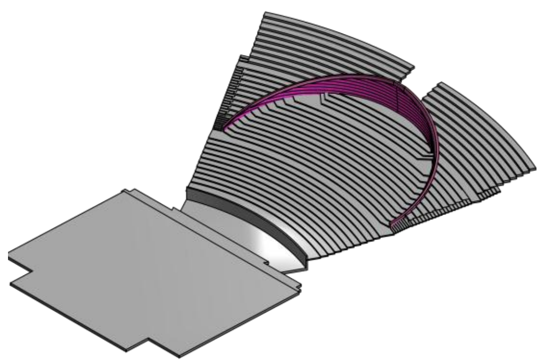


Fig 2 Stall Wall

When the width of a parabola, called the focal chord, is altered, so too, is the angle of tangency to the associated normal planes so that reflected sound rays can be geometrically and numerically informed (Fig 3 & 4).

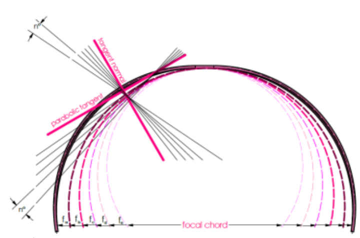


Fig 3 Focal Chord / Parabolic Tangent Relationship

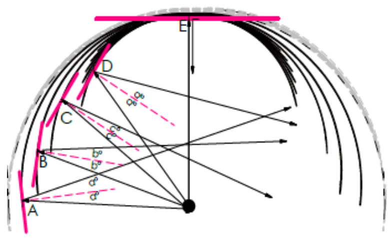


Fig 4 Tangent Normal / Sound Reflection Relationship

This portion of the parametric investigation consists of constraining the diameter of the arched wall in order to quantitatively measure the acoustical responses of reflected sound waves in relationship to varying reflective properties of the parabolic tangent. A point X has been chosen on the stall wall to receive a direct sound ray. The curvature of the wall has been altered by adjusting the focal chord distance, thusly affecting the parabolic tangent and consequently redirecting the trajectory of the reflected sound ray (Fig 5).

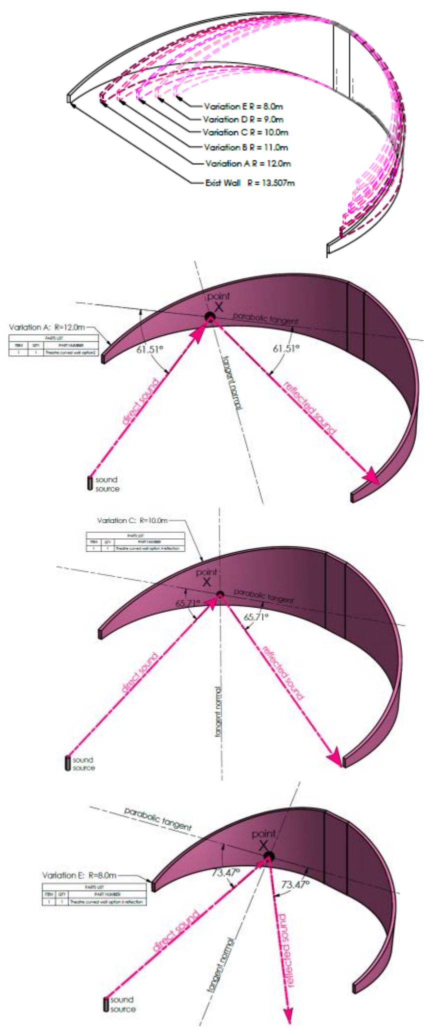


Fig 5 Curviture Variation Influence on Acoustics

Further parameters can be constrained to investigate more derivations of design alternatives. The discrete irregularities within the curve itself can be parametrically controlled so as to amplify or exaggerate distortions along the curved surface. These distortions can then be strung together to create a chain effect parametrically informed by varying degree of focal chord length from wave to wave as well as frequency of occurrence (Fig 6). The result will be the reflection of sound rays concurrent with convexity and alternatively, the focalization of rays correlative with concavity.

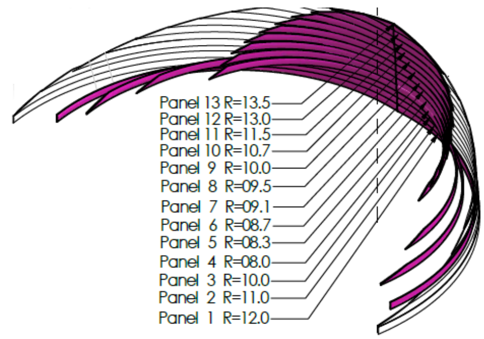


Fig 6 Curviture Variation of Individual Panels

Second Degree of Freedom: Spline Vertex Position

The arched stall wall is a solution to distribute the lateral sound energy towards the stalls in the center of the main hall. The concavity of the wall surface tends to focalize sound waves resulting in unevenly disbursed intensity levels. In addition to disguising the wall cavity, the cylindrical wall conceals within its parabolic lines discrete irregularities which direct sound rays to be either diffused or reflected (Fig 7). The reflection of sound is governed by the *Law of Reflection* which states that the angle of incidence is equal to the angle of reflection. Reflected waves can interfere with incident waves, producing pattern of constructive and destructive interference [2]. See (Fig 8) below.

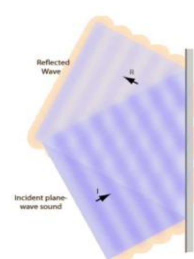


Fig 7 Stall Wall Spline Profile **Fig 8** Constructive/Destructive Waves
(Fig 7 - Photo Courtesy of Proceeding of the Institute of Acoustics)
(Fig 8 - http://hyperphysics.phy-astr.gsu.edu/hbase/sound/reflec2.html#c3)

Further specificity of the aforementioned acoustic principles in relation to the parabolic tangent can be applied to the curvature of the spline of each wood panel. By varying the location of a parabolic vertex along any given point on the spline of the wood element, sound rays from a single source can either reflected or diffused simultaneously and repeatedly within the same panel for the extent of the parabolic path (Fig 9).

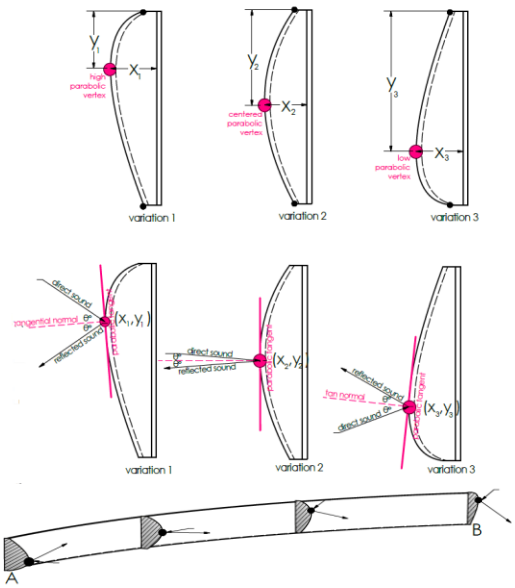


Fig 9 Multi-directional reflections of variant spline cross-sections

Additionally, the amplitude of curvature of the spline determines the thickness of the wood panel and thus incorporates a controllable degree of density. Material density is an additional contributing parameter to the characteristics associated with reflection and diffusion. By thickening or thinning the profile of the wood panels, it then becomes possible to direct sound waves according to their frequencies. High frequency sounds are inclined to travel further and be more reflective than their lower frequency counterparts.

One applicable use of such particularized design control in the case of the Teatro degli Arcimboldi is effectively employed to direct sound rays above the spectator-head- height back toward the stalls to create the perception of sound envelopment while directing lower sound rays into the floor to prevent undesired sound overlaps [1]. See (Fig 10).

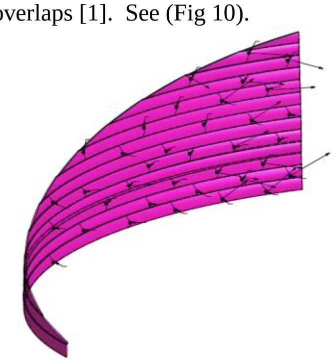


Fig 10 Diffusive Effect of Multiple Varied Spline Conditions

The parametric research for this portion of the study was conducted by examining the acoustical effects of variant parabolic vertices at multiple cross-sections within a single panel. The vertices of the end profile-splines were assigned almost inverted (x,y) coordinates in order to locate each end-spline vertex at opposing high/low points. The intermediate cross-sections identify the transitional vertex points between these profiles. Sound reflection trajectories were calculated using the same aforementioned parabolic tangent technique. The results yielded the ability to reflect and diffuse sound several times within a single panel, despite the concavity of the panel path (Fig 11).

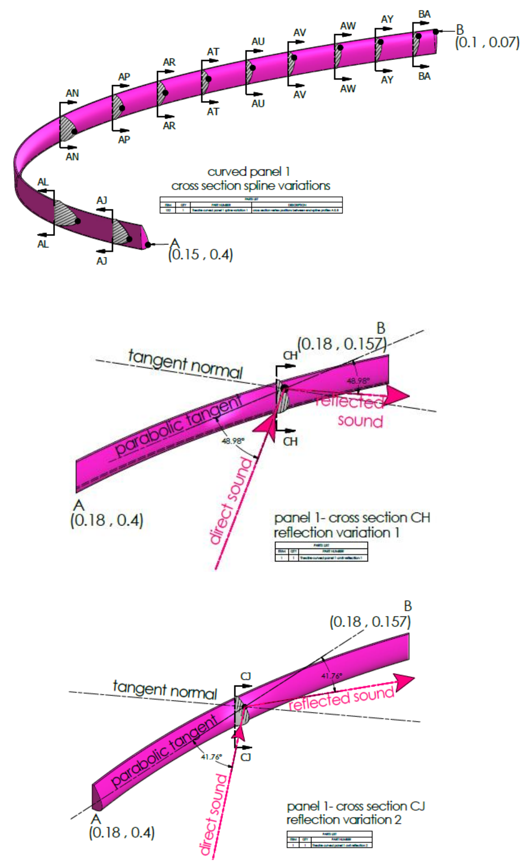


Fig 11 Acoustic Implications of Variant Spline Vertex Position

Third Degree of Freedom: Panel Tilt

A sound ray incident upon a reflective surface will be reflected at an angle equal to the incident angle. Both angles are measured with respect to the normal to the surface[2]. Varying the angle of the reflective surface alters the angle of the normal plane and thusly alters the angle of incidence with respect to the normal which results in a variant output of the reflective angle. Large flat parallel areas reflect sound rays in opposition to the path of direct sound while angled surfaces spread the sound and contribute to even dispersion [2]. A variant angle of tilt allows for the ordinarily focal tendency of a parabolic shape to become diffusive by orienting reflective rays away from direct sound. The avoidance of sound wave collision results in multi-directional dispersed sound which envelops the auditor. The degree of tilt toward the stalls contributes toward the overall diffusive effect of the stall wall. When considered as a single entity, the purpose of the wall is to control the reflection of sound rays in order to prevent unwanted secondary waves from behind the spectators seated in the stalls. The angle at which the entire wall tilts is intended to redirect sound waves down into the ground and away from the audience (Fig 12).

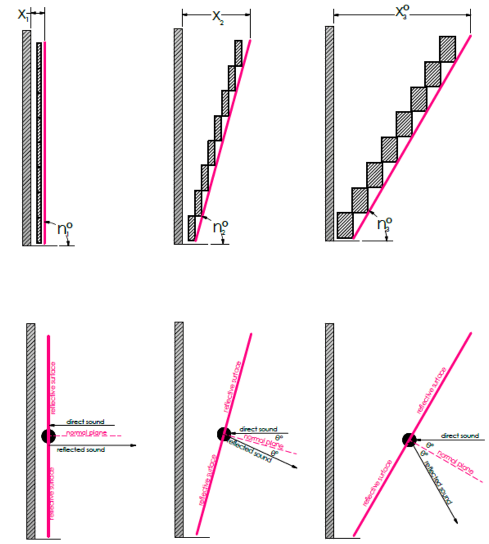


Fig 12 Angle of Tilt/Angle of Incidence Relationship

The parametric investigation consists of quantitatively controlling the angle of tilt by placing a dimensional constraint on the thickness of the first panel and setting the thickness of the remaining panels to follow suit. The wood panels are stacked away from the framing wall so that at the apex of the parabolic stall wall, the distance between the top wood panel and the anchoring wall, is at the greatest distance (Fig 13).

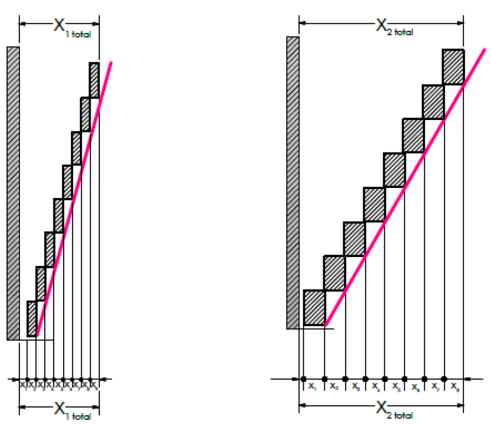


Fig 13 Panel Width/Angle of Tilt Relationship

As the/thickness of the panels increase or decrease, the angle of tilt will vary respectively. The proportional relation between panel thickness and angle of tilt creates a measurable qualitative acoustical response in the variant reflected sound rays correlative with the corresponding angle of incidence in association with the varying normal plane. Constraints were placed on the distance from the wall to be proportional with the relative angle of tilt directed away from the frame.

This parametric association can then be quantitatively controlled to affect the qualitative acoustical response required to accommodate the dual role of the Arcimboldi Theatre. The reflected surface is used to calculate the trajectory of reflective sound rays using the same methodology employed in the previous two techniques.

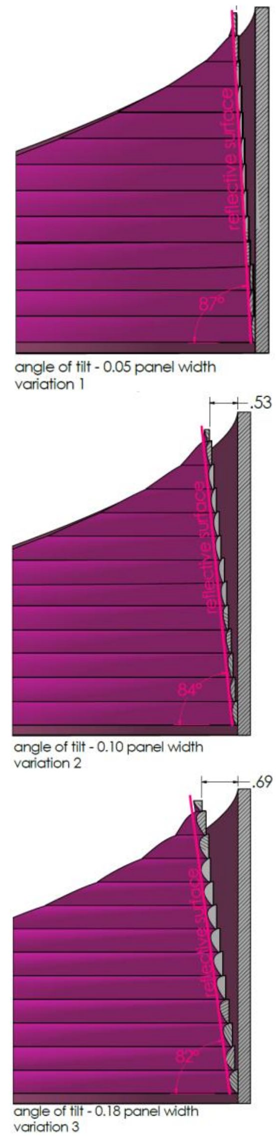


Fig 14 Panel WidthTilt Claculations

3. CONCLUSION

The proper distribution of diffractive elements within the components that comprise the stall wall at the Teatro Degli Arcimboldi can be numerically and geometrically informed for optimal performance. The use of physically based parametric constraints facilitates the ability to fine tune the acoustical response of design elements in architecture. This type of information based modeling in all aspects of design development can contribute to more cohesive and responsive design alternatives.

4. ACKNOWLEDGEMENT

This research was conducted in conjunction with a graduate architectural elective course (Arch 662.011) entitled The Design of Parametric Space, taught by Assistant Professor, Rhett Russo, at the New Jersey Institute of Technology, in the fall of 2011.

5. REFERENCES

1. Commins, D. " Acoustics of Teatro Degli Arcimboldi in Milano Design, Computer and Scale Models, Details Results" *Proceedings of the Insititute of Acoustics* Vol 4 Pt 4 2002.

2. Nave, Carl R., *Hyper Physics Web* Dec. 2011 <http://hyperphysics.phy-astr.gsu.edu/hbase/hframe.html>

3. "Parabola." *Wikipedia, the Free Encyclopedia*. Web. 01 Dec. 2011. http://en.wikipedia.org/wiki/Parabola

Note: Teatro degli Arcimboldi was designed by Vittorio Gregotti working with architects Mario Botta and Eliabetta Fabbri

A Parametric Workflow: from *Grasshopper3D* to *Autodesk Inventor*

Eli MELTZER
School of Architecture, Newark Institute of Technology
Newark, NJ 07102, USA

ABSTRACT

This paper seeks to explore the parametric relationships guiding the design of seating arrangements in concert halls, specifically relating to ergonomics, sight lines, means of egress and placement of balconies, and their relation to the wider considerations of the Time Delay Gap and Reverberation Time. The exploration uses computational parametric tools, primarily the *Grasshopper* plug-in for *Rhino3D*, and *Autodesk Inventor*, to graphically model, display and test these parameters. Interoperability between these two software platforms is, therefore the primary area of study for this paper, as we seek to marry the open-ended design capabilities of *Grasshopper* with the robust production tools of the *Inventor* suite. Research focuses on the sharing of data through the *gHowl* add-on for *Grasshopper*, to export parameter values to *Excel*, which are applied to *iPart* factory tables in *Inventor*, to generate the various seat arrangement iterations. The optimization and specialization of both software platforms is a key goal of this research. Specifically, we are seeking to leverage the iterative and recursive parametric modeling capabilities of *Grasshopper*, to supplement the "piece-by-piece" logic of *Inventor*; while *Inventor* will be used primarily to produce BIM (Building Information Modeling), and presentation quality drawings.

Keywords: Parametric, Grasshopper, Inventor, Rhino, Fabrication, Workflow

1. CONTEXT

The integration of information based design tools is continuing to revolutionize the design disciplines, in particular, the fields of architecture and industrial design. Amongst the tools developed, primarily within the engineering industry, parametric software platforms and BIM are the most widespread and influential. The parametric refers to a design process "based not on fixed metric quantities but on consistent relationships between objects, allowing changes in a single element to propagate corresponding changes throughout the system" (Meredith, 4). BIM, on the other hand, refers to a data model capable of capturing "domain-specific information about entities … that is constructed around building entities and their relationships to one another [of which] geometry is only one of … these building entities" (Cheng, 491).

It is important to understand that traditionally the fields of architecture, engineering and construction each have their own method of organizing information and hence their own tools and softwares. With the introduction of information-based modeling, it is possible for the designer to include any aspect of the design or construction process into one virtual master model, which accounts for every aspect of each individual part of the building. Within the field of parametric tools themselves, there have developed two primary vehicles: algorithmic modelers, and constraint-based modelers. The algorithmic is defined as "a method of generation, producing complex forms and structures based on simple component rules," and often relies on the use of complex computer scripting tools (Meredith, *ibid*). A constraint-based modeler, on the other hand, relies primarily on the application (often visual) of "dimensions and constraining geometries" upon predetermined 2D sketches or 3D models, and these dimensions and constraints are the "parameters, or input points, that you [the designer or engineer] would then change to update or edit the [model]" (Waguespack and Tremblay, 1). To date, primarily constraint-based modelers integrate with BIM, while algorithmic modelers are mostly reserved for form finding, and the formal integration of design constraints. The goal of this paper is to explore new design methodologies and methods that unify these two related, and yet disparate, design tools. The development of an effective and efficient workflow that moves between the algorithmic, the constraint-based, and BIM is the primary purpose of this design research.

2. AREA OF RESEARCH: DESIGN METHODS AND METHODOLOGIES

For the purpose of this paper, we have selected *Grasshopper*, "a graphical algorithm editor tightly integrated with [McNeel] *Rhino*'s 3-D modeling tools", developed by David Rutten at Robert McNeel & Associates (www.grashopper3d.com), as an example of an algorithmic modeler, and *Autodesk Inventor* as an example of a constraint-based parametric and BIM modeler. Communication between these two platforms occurs exclusively at the level of *Microsoft Excel* spreadsheets, through the sharing of data within a common domain. The *gHowl* add-on for *Grasshopper* and the *iPart* and *iFactory* tools within *Inventor* are the specific tools of data input and output for these two programs. The optimization and specialization of both software platforms is a key goal of this research. Specifically, we are seeking to leverage the iterative and recursive parametric modeling capabilities of *Grasshopper*, to supplement the "piece-by-piece" logic of Inventor; while Inventor will be used primarily to produce BIM and presentation quality drawings. The introduction of *Excel* in to the design workflow reduces the focus of design parameters to quantifiable pieces of data, or mathematical relationships, and exposes the designer directly to the underlying data model that drives his or her project.

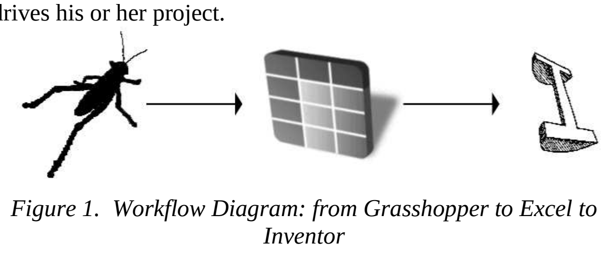


Figure 1. Workflow Diagram: from Grasshopper to Excel to Inventor

In general, the proposed workflow begins by establishing a series of user-defined parameters in Grasshopper either through direct panel input, interactive dynamic sliders, or culled from direct manipulation of referenced Rhino geometry. Through the application of *both* mathematical and geometrical means, the parameters are processed and analyzed, which generates the parametric design, or space, as well as a new series of driven parametric dimensions. These combine with the design parameters to form the data model. A key difference to note exists between a "user-defined parameter" and a "driven parametric dimension". A user-defined parameter constitutes the input and first buildings blocks of the parametric design process, which determine the basic geometric principles involved. A driven parametric dimension, on the other hand, comes about as a result of measurements and analysis performed on the generative design. These driven parameters may drive further generative principles within the design process. Finally, and perhaps most importantly, the designer must deal with the primary "design parameters": those pieces of data by which the designer can *evaluate* the performance and efficacy of his or her design.

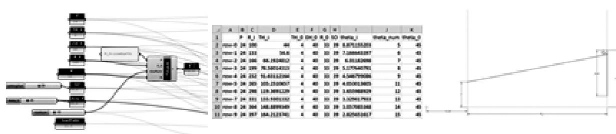


Figure 2. Left: User Defined Parameters in Grasshopper; Middle: Driven Parametric Dimensions in Excel; Right: a base Sketch in Inventor, showing the desired sight line

3. CASE STUDY

For our research, we have chosen to focus on the parametric relationships guiding the design of seating arrangements in concert halls. To expand further on the distinctions made above (Sec. 2, Par. 2), it will be helpful to look at the following case study. In this case, the specific *user-defined parameters* are primarily ergonomic, and many come directly from the building code. For example, in determining the placement of seats to achieve unobstructed sight lines of a particular "viewing point" (a *design parameter*), a typical eye-height from the floor (and from the eyes to the top of the head), and a row-to-row distance can be established by the designer (*user-defined parameters*). These parameters, using simple trigonometry and arithmetic series, can determine the ideal seat locations within a coordinate system (*driven parametric dimension*). Although eye or head heights and row widths are human dependent, and therefore, in a sense, not truly "user" defined, designing further parameters into the process will add layers of designer specificity. In this case, the distance from the first seat to the viewing point changes, and will alter the preferred seat rake (see Fig. 3). Even changes in seat type (with or with-out padding) and manufacturer can alter the outcome.

While designers examine the problem of sight lines primarily in section, they must design the layout of the rows themselves in plan. Seat width is of primary concern, and the most freedom occurs in the choice of a "setting out point", or the center point for the arcs of seats. This can be identical with the viewing point or anywhere in front or behind it. This will affect the depth or shallowness of the rows' arcs. Here, the primary design parameters may be financial, as the owners look to pack as many seats in the hall as possible, while maintaining adequate site-lines. Similar considerations will be necessary for the placement and design of balconies, as well as gangways and paths of egress.

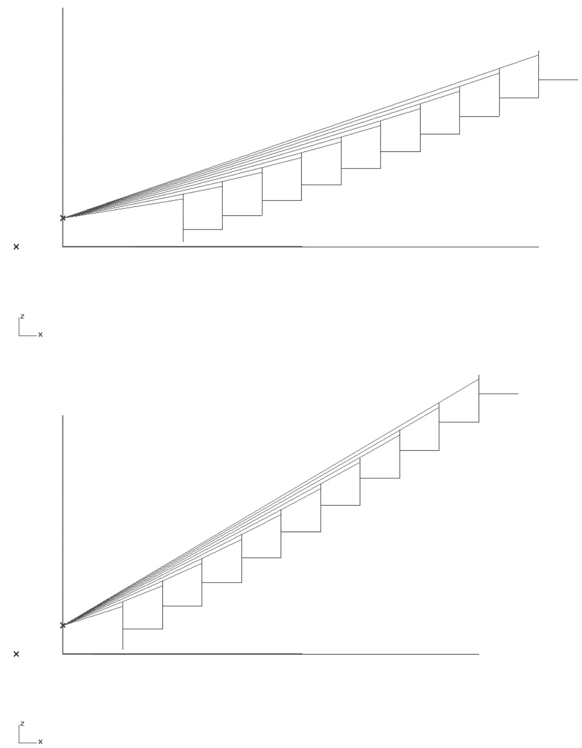


Figure 3. Sightlines in Grasshopper, updating parametrically, as the depth of the first seat row decreases, increasing the rake

Most importantly, in the design of concert halls, all of these design decisions will have an impact on the larger design parameters of Time Delay Gap and Reverberation Time that are essential in determining the quality of these spaces (see Beranek). The interplay between the design parameters of the seating arrangement and of the acoustical performance of the hall quickly becomes a very complicated design problem. For example, in a certain style of seating now popular, known as terrace-style seating, blocks of seating are placed throughout the hall and raised up slightly above the floor plane and the seats in front of them (Fig. 4). This provides a reflective surface on the front (and sides) of these terraces that can be effectively leveraged in the acoustical design of the hall. Of course, changing the height, pitch, or depth of the arc of these terraces will have wide-ranging effects on the seating design parameters. While traditional design methods might make determining these effects a tedious process, with parametric computational tools, they can be analyzed in real time, for both their acoustical and seating arrangement performance.

Figure 4. The Disney Concert Hall in Los Angeles, with its terraced-style seating; designed by Gehry Partners and Nagata Acoustics

Unfortunately, there are no current software platforms capable of effectively analyzing the complex nature of this problem, and interfacing with the design and production software actually capable of managing such a project. An algorithmic solver, such as *Grasshopper*, could analyze the problems, and even provide graphical feedback, but to maintain that adaptability over the course of an entire project would be quite tedious. A BIM program on the other hand, such as *Inventor*, would be flexible enough to adapt to the various parameters, however, to perform the actual analysis within the software would be quite difficult, and likely require some supplemental source coding. This led to the development of our proposed workflow, necessitated by the iterative and recursive nature of the seating problem. Specifically, the placement of any individual seat in the layout requires prior knowledge of the location of all other seats in the plan that may obstruct that seat's view. In the typical "piece-by-piece" Inventor workflow, this would require the tedious placement of each row in order. Even to use parametric *iParts* would require manually adding a new line to the *iFactory* table for each row. In *Grasshopper*, however, we reduced the relationship to a single equation describing the placement of seats within a mathematical series, which drives the generation of the seating geometries. We can update the number of rows, as well as all of the other parameters, through either numerical input, with a dynamic slider, or even by direct manipulation of Rhino geometry, and the geometry updates in real time. The data model driving the updated geometry can be exported to a properly formatted *Excel* spreadsheet and through a simple copy and paste in to the *iPart* table, the entire series of parts (or rows in this case) are generated; they merely need to be placed in an assembly file. At this point, we begin to leverage the great strength of *Inventor*, because the base assembly file can be layered with any number of details, from actual seats, to material finishes, to the form and structure (and even structural framing) of the hall. Furthermore, we can push this data back out to Grasshopper, or any other suitable program for further analysis. This would be particularly useful, for example, in looking at the finishes, and the absorptive qualities of the various materials.

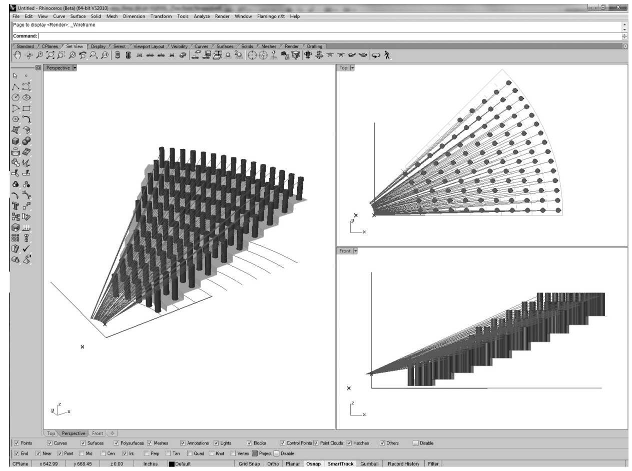


Figure 5. Graphical feedback of the Grasshopper algorithm in the Rhino environment

While similar effects could be achieved by using *Excel* spreadsheets alone (and propagating the mathematical series) to drive *iFactories*, the geometric and graphical feedback provided by *Grasshopper* makes this workflow much more fruitful in the design process. Additionally, this workflow requires no special knowledge of *Excel*, or *VB.NET*, scripting. The use of *Grasshopper* becomes even more essential when the designer must reference existing conditions, such as the floor plan, from a DWG, or other architectural drawing. Only through the *Grasshopper* workflow can the designer translate those conditions into a data model accessible in *Excel*, and other tools down the line.

4. CONCLUSIONS

Research will continue to focus on the exact place of each tool in the workflow, considering the possibility that more than one can likely do each task, and that one or the other may be more properly suited to the task-at-hand. For example, in our case study, to parametrically model the section for a seating arrangement, the software must incorporate some iterative or recursive functions. Such capability typically lies in the realm of an algorithmic modeler. To solve the plan, on the other hand, requires only trigonometric and geometrical analysis, because seats are generally spaced at equal intervals along a fixed path, a simple process for constraint-based modelers

In general, the proposed process fills two large gaps in the *Grasshopper* workflow, especially regards the difficulty in extracting drawings from *Grasshopper* to a level of detail that might be required of shop or working drawings. Further investigations include the ability to translate free-form NURBS geometries modeled in *Rhino* in to the *Inventor* BIM environment, extending its formal capabilities. For example, one could write a *Grasshopper* script that takes a NURBS surface and tessellates it in some controlled manner, then exports the vertex points to *Excel* for import in to *Inventor*. From these works points, the designer could model a series of panels in *Inventor* to populate the surface. Finally, *Inventor* can produce the shop drawings and BOM (Bill of Materials).

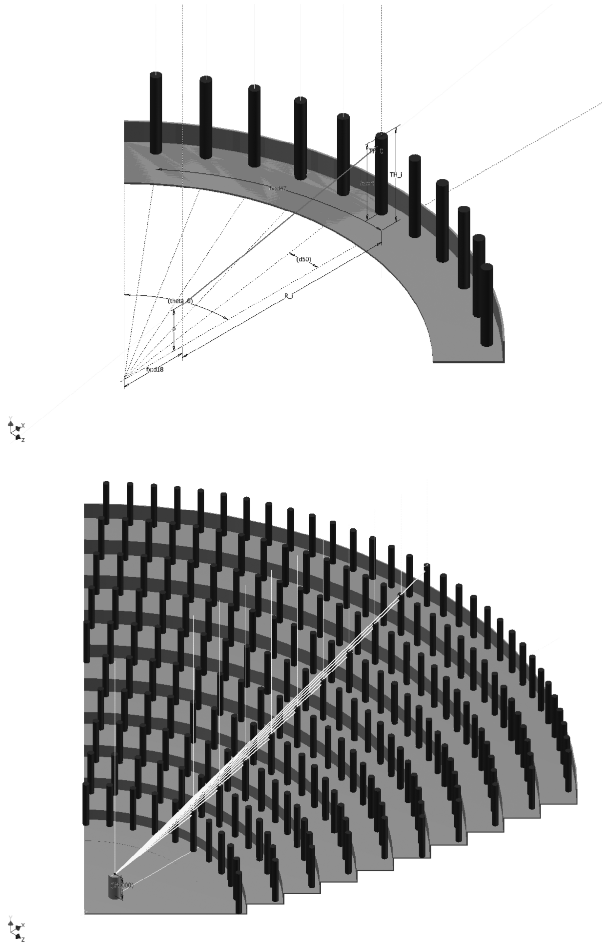
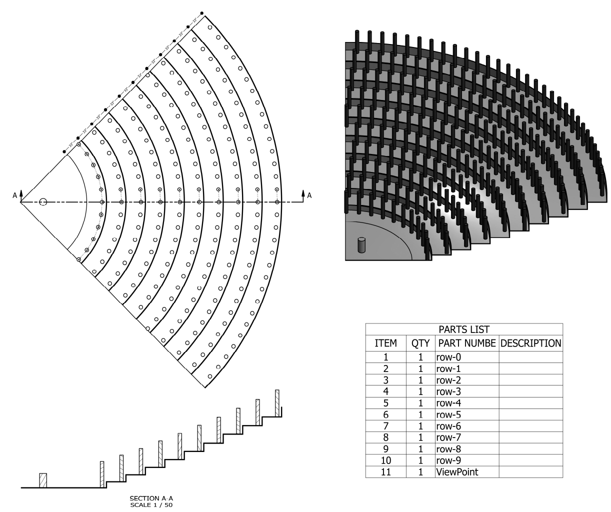


Figure 6. Part and Assembly files in the Inventor parametric environment

While constraint-based modelers, including *Inventor*, typically have difficulty with NURBS geometry, this workflow bypasses this obstacle altogether by using an algorithmic modeler to "translate" the geometry into *Inventor*'s "native" language. By reducing the NURBS geometry into a raw data model, *Inventor* can re-create the geometry based off work points, planes, and axis defined entirely within *Inventor*. Other workflows require that models pass through one or more conversions between file types, which inevitably leads to data loss and corruption, and lack of accuracy in the final product. For example, to get NURBS geometry out of *Rhino* in to *Inventor* would one might save the native *Rhino* .3DM file as an .OBJ or .3DS file, and then import this into *Inventor*. Because *Inventor* has no native support for NURBS these file types will have likely converted the geometry to a polygonal mesh model, which provides, at best, a close approximation of the original design. Even at this point to turn this polygonal mesh into a satisfactorily adaptable parametric model will take a great amount of work, and may require essentially starting from scratch with the imported model as nothing more than a guideline.

With the workflow presented in this paper, however, the designer has complete control over the manner in which the NURBS geometry translates into the *Inventor* design environment. Most importantly, this process adds an entire layer of sophistication into the design process that directly relates to fabrication. For example, the way in which a designer chooses to tessellate a surface (and provide work points for *Inventor*) depends on the number of panel types, the framing system (or the absence thereof), the method of connection to one another (and the frame), and any other number of design parameters. Therefore, by solving an efficient way to transfer a design between the algorithmic and the constraint-based parametric environments, between *Grasshopper* and *Inventor*, the designer must also grapple with design issues related to the actual construction and completion of the project. The design of a data model that allows for efficient parametric modeling and detailing in *Inventor* goes hand in hand with a design logic that allows for efficient and economical fabrication and construction of the project, a boon for designers and those that employ them.



PARTS LIST			
ITEM	QTY	PART NUMBE	DESCRIPTION
1	1	row-0	
2	1	row-1	
3	1	row-2	
4	1	row-3	
5	1	row-4	
6	1	row-5	
7	1	row-6	
8	1	row-7	
9	1	row-8	
10	1	row-9	
11	1	ViewPoint	

Figure 7. A drawing file produced by Inventor, including Bill of Materials (BOM)

ACKNOWLEDGEMENTS

The research in this paper was conducted in conjunction with a graduate architectural elective course, ARCH 662, entitled, *The Design of Parametric Space*, taught by Assistant Professor Rhett Russo, at the New Jersey Institute of Technology, in the Fall of 2011.

REFERENCES

[1] Leo Beranek, **Concert Halls and Opera Houses: Music, Acoustics, and Architecture**, Springer, 2004.

[2] Renee Cheng, "Computing Technologies: Building Information Modeling (BIM)", **Architectural Graphic Standards, Eleventh Edition**, Hoboken, New Jersey: John Wiley & Sons, Inc., 2008.

[3] Michael Meredith, **From Control to Design**, Actar, 2008.

[4] Curtis Waguespack and Thom Tremblay, **Mastering Autodesk Inventor 2011 and Autodesk Inventor LT 2011**, Indianapolis, Indiana: John Wiley & Sons, Inc., 2010.

The Research of Acoustic Design of Concert Halls

Xing Su
College of Architecture and Design
New Jersey Institute of Technology, 07102
xs28@njit.edu

ABSTRACT

The qualitative analysis of sound is a primary concern in the design of concert halls. The acoustic phenomena of a concert hall depends largely on its geometry, based on the case study of Seiji Ozawa Hall (Architects: William Rawn Associates; Acoustician: R.Lwarence Kirkegaard & Associates), the research for this paper will focus on analyzing the geometry of Seiji Ozawa Hall and investigating some basic geometric parameters of concert hall design in regard to its acoustic phenomenon.

Key words: acoustic phenomena, geometry

A. General Methodology

The software used during the research is Rhino Grasshopper and Autodesk Ecotect Analysis. Rhino Grasshopper is a plug in for Rhinoceros 3D, which is a strong tool for parametric design and analysis. Autodesk Ecotect Analysis is an environmental analysis tool that allows designers to simulate building performance, and the result of the analysis can be shown as diagrams in a direct way. In this research, a digital model of the concert hall will be created in Rhinoceros 3D using grasshopper; with geometrical constrains controlling the relationships between different parts of the hall (stage, ceiling, wall, seating area, and balcony). Then the acoustic analysis can be done in Autodesk Ecotect. Within the established constrain, the spatial relationship between different parts of the concert hall will remain the same. And By changing parameters and analyze its corresponding results generated by Autodesk Ecotect Analysis, the conclusion can be draw both qualitatively and quantitatively.

B. Geometry Analysis

The general geometry of Seiji Ozawa Hall is shoebox, There are simple variations on a shoebox: hall width, hall length, ceiling height, balcony height, source point position. They are all building in a parametric control with constrains in Rhino with Grasshopper.

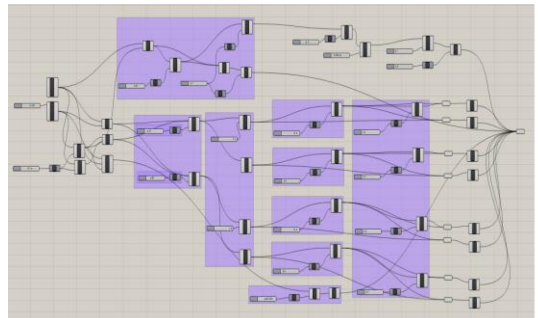


Figure 1. The definition of grasshopper

Figure1 shows the general logic of modeling the concert hall with grasshopper, first, the footprint of the hall is defined as a rectangle, and the balcony and stage are constrained with the volume. In this case, the changeable parameters are set as the following: hall length, hall width, ceiling height, lower balcony height, upper balcony height, stage length, and stage height. All of these parameter can generated as excel form by grasshopper (see appendix). In doing so, by simply changing one parameter, other parameters will change accordingly, that the nature of parametric design.

1) The plan geometry analysis:

The plan analysis is based on the principle that the plan geometry is the only variable parameter, the height of the hall and the height of the balcony will stay the same, this method of analysis will apply to the following analysis as well.

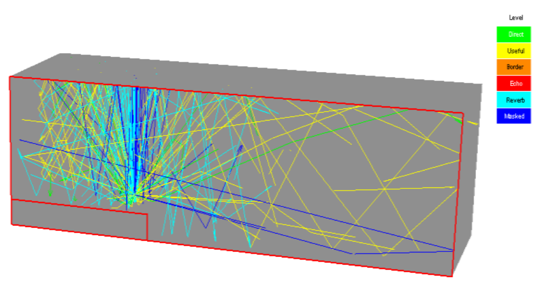


Figure 2. The sound reflection diagram generated by Autodesk Ecotect

Figure2 shows the sound reflection diagram within the concert hall, the model is the model of Seiji Ozawa Hall, the colored lines represent the sound path, different color indicates different acoustic quality, as it is shown in the diagram, green line is direct sound, yellow line is useful sound, orange line is border sound, red line is echo, light blue line is reverb line, and blue line is masked sound, by analyzing at the diagram, we did not find the red line, which means there is no echo. And the useful sound is covered the whole volume.

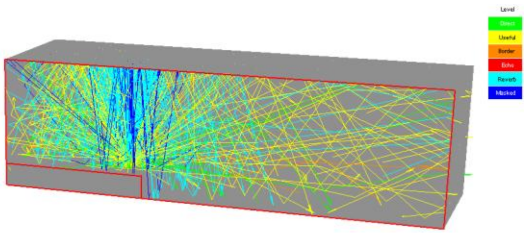


Figure 3. The sound reflection diagram generated by Autodesk Ecotect

In the case of Figure 3, the source has been strengthened by three times, the acoustical quality can be visualized easily, by analyzing Figure 2 and Figure 3, we can get the conclusion that the closer we are from sound source, the stronger sound we get. At the same time, the possibility of receiving noise becomes higher.

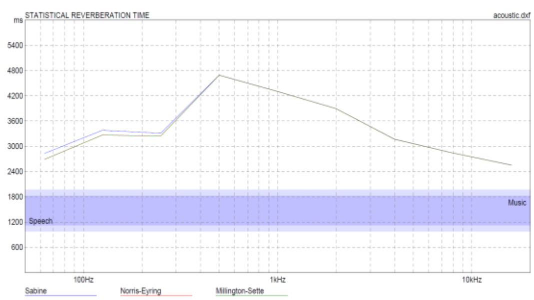


Figure 4. statistical reverberation time analysis

Figure 4 shows the statistical reverberation time of the concert hall, the purple area indicates the acoustical quality of this volume is suitable for speech and music.

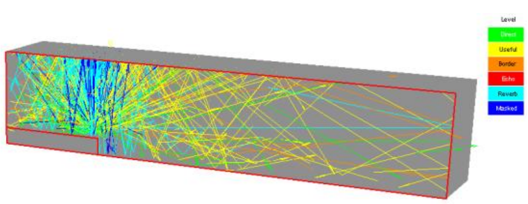


Figure 5. The sound reflection diagram generated by Autodesk Ecotect

In the case of Figure 5, the volume of the concert hall has been elongated to a very long and narrow space, we can visualize that the back of the hall receives little sound, which mean bad acoustical quality, to conclude, over elongated shoebox should be avoid during the concert hall design.

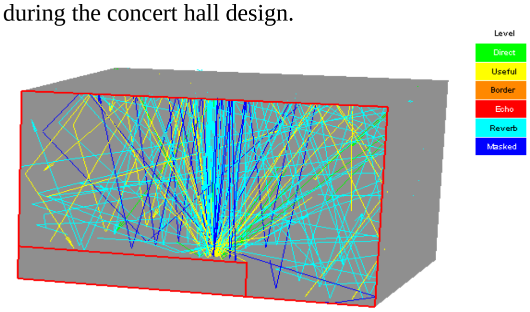


Figure 6. The sound reflection diagram generated by Autodesk Ecotect

In the case of Figure 6, the volume of the concert hall has been shorten to a very long and narrow space, we can visualize that the whole volume receives too much of different type of sound, which mean bad acoustical also, which means over short-fat shoebox should be avoid during the concert hall design.

In general, we should consider an appropriate proportion for the shoebox geometry of concert hall.

2) The ceiling height and stage height analysis:

In the study of ceiling height and stage height, the same methodology was conducted as the plan geometry study.

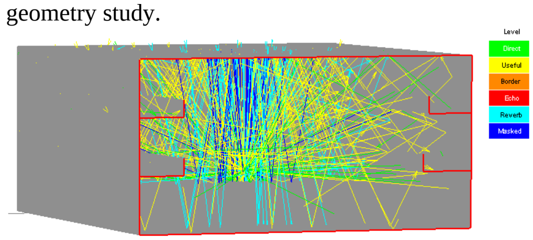


Figure 7. The sound reflection diagram generated by Autodesk Ecotect

Figure 7 shows the reflection situation in a very low ceiling volume, we can visualize that the whole volume receives much more direct and useful sound than the other case; typically, the upper level balcony receives too much strong sound, which indicates a bad acoustical quality.

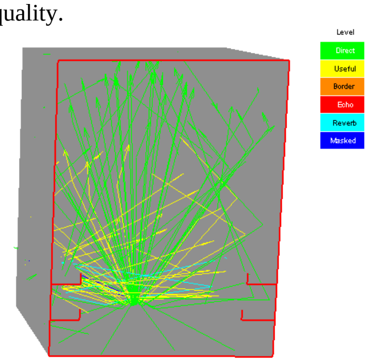


Figure 8. The sound reflection diagram generated by Autodesk Ecotect

Compare to Figure 7, Figure 8 shows the reflection situation in a very high ceiling case, we can also visualize that the high ceiling volume absorb a lot of direct sound, leading to the result that the upper level balcony receives little reflected sound, which indicates a bad acoustical quality also.

In conclusion, similar to the plan geometry analysis, an appropriate ceiling height also has a huge effect on the acoustical quality of the concert hall.

3) The balcony height analysis:

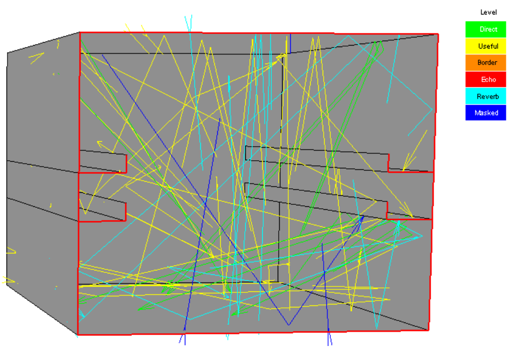


Figure 9. The sound reflection diagram generated by Autodesk Ecotect

Figure 9 shows the situation which the two levels of balcony are very close to each other, we can visualize that the lower level of balcony receives some of the reflected sound from the bottom of the upper level balcony, but receives little reflected sound from the side wall, it is because the distance between the two level of balcony are close, the upper level balcony blocks a huge amount of sound. In this case, the acoustical quality is low.

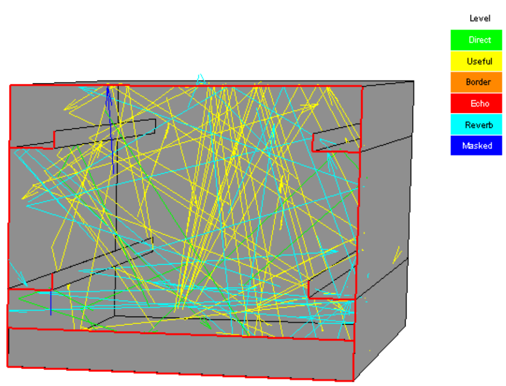


Figure 10. The sound reflection diagram generated by Autodesk Ecotect

Figure 10 shows the situation which the two levels of balcony are very far from each other, we can visualize that the lower level of balcony receives some of the reflected sound from the side wall, but receives little reflected sound from the upper level balcony, as the situation of Figure 9, it is also an undesirable acoustical approach.

C. Conclusion

By analyzing the geometrical design of concert hall, we find out that the volume and the different parts of the concert hall are designed in a logical way which gives the audience the best quality of sound. In a word, besides the outside factor (site, cost, and so on), the geometry design of the concert hall has a tight relationship with the acoustical concern.

D. Acknowledgement

This research is conducted in conjunction with a graduate architectural elective course (Arch 662 011) entitled “the design of parametric space”. This course was taught by Assistant Professor, Rhett Russo at New Jersey Institute of Technology, 2011 fall.

Appendix

Hall length	Hall width	Lower balcony height	Upper balcony height	Ceiling height	Stage height	Figure number used
136	61	10	20	40	7	2
136	61	10	20	40	7	3
136	61	10	20	40	7	4
193	45	10	20	40	7	5
80	45	10	20	40	7	6
193	45	10	20	30	7	7
193	45	10	20	80	7	8
136	61	16	27	50	7	9
136	61	16	40	50	7	10

Note: this excel sheet is generated by rhino grasshopper, indicating the method of parametric design

References

[1] Beranek. Leo. Concert Halls and Opera Houses: Music, Acoustics, and Architecture. Springer, 2004

[2] Leonid Makrinenko, Acoustics of Auditoriums in Public Buildings, Acoustical Society of America, Woodbury, NY, (1994)

Analysis of Computing Curriculum Standard J07 using ICT Common Body of Knowledge

Tetsuro Kakeshita and Mika Ohtsuki
Department of Information Science, Saga University, Japan
{kake,mika}@is.saga-u.ac.jp

Abstract

Information Processing Society of Japan (IPSJ) has announced the Computing Curriculum Standard J07 which is compatible with the Computing Curricula 2005 (CC2005) Series proposed in the United States. Both J07 and CC2005 are composed of five major domains, CS, CE, SE, IS and IT, each of which is developed by different community so that relationship among these domains is not concrete. In this paper, we analyze each body of knowledge (BOK) of the domains and map them into the ICT common body of knowledge (ICTBOK). We estimate the degree of importance and the requirement level in terms of the 155 ICTBOK areas for each domain by utilizing the mapping. As a result, the relationship among the domains is clarified. We also compare the analysis result with the requirement data collected from 58 business persons.

1. Introduction

Quality assurance of higher education is a global trend. To this end, Japanese universities are building systematic education system. Currently many universities and departments have their own diploma policy to demonstrate the ability of their graduates.

The Japanese ministry of education has built a public quality assurance framework. The framework is composed of the standards for establishing universities, the establishment approval system, and the periodical accreditation system.

Industry and the Japanese ministry of industry both realize that ICT (Information and Communication Technology) is necessary for global competition. The Japanese government and the local governments want to increase business efficiency by developing and operating digital government. Large number of high level ICT professionals is necessary to achieve these goals.

The Information Processing Society of Japan (IPSJ) has supported these activities from many perspectives. Among such activities, IPSJ announced the Computing Curriculum Standard J07 [4] as a guideline for college level ICT education. J07 is essentially compatible with the Computing Curricula 2005 (CC2005) [3] developed by ACM, IEEE Computer Society and AIS. Considering the diversity of the computing disciplines, the following five major domains are defined both in J07 and CC2005.

- CS: Computer Science
- CE: Computer Engineering
- SE: Software Engineering
- IS: Information Systems
- IT: Information Technology

Each domain of J07 and CC2005 is developed by different community for historical reasons so that the body of knowledge (BOK) of each domain is described using different terminology. As a result, relationship among these domains is not concrete and clear not only for the society including industry and government but also for faculty members of ICT departments. This situation is not desirable for many stakeholders of the computing curriculum.

In this paper, we analyze the BOKs of the five domains of J07 to map them into the ICT common body of knowledge (ICTBOK). ICTBOK is composed of 6 categories, 23 fields and 155 areas; and covers a wide range of knowledge and skill required for ICT professionals. Thus clear and concrete analysis becomes possible in order to clarify the relationship among the five domains.

Detail of ICTBOK is described in Section 2. Structure of J07 is summarized in Section 3. We shall overview the analysis plan in Section 4. In Section 5, we map BOK of each domain into ICTBOK. Then we estimate the degree of importance of the 23 fields of ICTBOK for each domain. The requirement level of each ICTBOK area is also estimated for each domain.

We are currently executing the J07 follow up survey project which is composed of achievement and requirement surveys. Through the achievement survey, we collect achievement levels of the 155 areas of ICTBOK from department and graduates of ICT colleges. The requirement survey is carried out for business persons of ICT industry of various positions [5]. Through the requirement survey, the requirement levels and the degrees of importance of the 155 areas are collected in order to analyze the detailed requirement. The analysis of the J07 domains is a part of this survey project to clarify the relationship among various activities of education and professional development in ICT.

The analysis result of the J07 follow up survey will be distributed to the survey contributors in order to promote mutual understanding among university, industry and government. It will also be utilized to further improve the J07 curriculum standard.

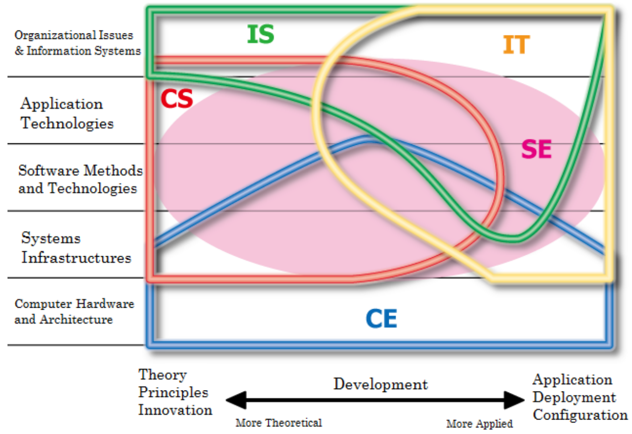
2. ICT Common Body of Knowledge

ICT common body of knowledge (ICTBOK) is developed to uniformly represent knowledge and skill of ICT professionals belonging to various job categories and levels [1]. The ICTBOK can also be used to represent outcomes and requirement of ICT education program. ICTBOK is composed of 7 categories, 23 fields and 155 areas. Analysis of J07 is carried out using the common form representing ICTBOK so that comparison among the analysis results becomes possible.

IPA (Information Technology Promotion Agency) of the Japanese government announces three types of skill standards for ICT professionals [2]. They are (1) ITSS (skill standards for IT professionals) for people working

Table 1 Categories and Fields of ICTBOK

Category	Field
Foundations of Computer Science	Fundamental Theory (7) Mathematics, Applied Mathematics (5) Computer Architecture (9) Hardware (7) Operating System (5)
Media, Human Computer Interaction	Multimedia Data Processing (8) Human Interface (3) Usability (2) Intelligent System (6)
Network and Security	Telecommunication System (9) Computer Network (6) Web Technology (8) Security (6)
Software Development	Database (10) Algorithm and Data Structure (5) Computer Programming (5) Software Engineering (11)
Information System	Project Management (6) System Operation and Evaluation (8)
Business	Business and Administration (8) Technical Communication (5) Society and Ethics (4)
Competences	Competences (12)
Others	*Additional fields or areas can be defined by the users, if necessary.*



(Initially published in CC2005 and updated in J07)

Figure 1 Conceptual Relationship of J07 Domains

for IT services industry, (2) ETSS (embedded technology skill standards) for embedded software development engineers, and (3) UISS (user's information system skill standards) for information system users. We have analyzed the three skill standards and have defined ICTBOK. ICTBOK also covers the teaching domain of college level ICT education by analyzing and integrating J07 and CC2005.

Table 1 represents categories and fields of ICTBOK. The number associated to each field is the number of areas corresponding to the field. In addition to the BOK explained above, ICTBOK also contains the category "competences" expected as a society member. The areas of the competences are developed by the ministry of economy, trade and industry of the Japanese government. The "others" category is defined so that users can add areas during the achievement or requirement survey, if there is a missing area in ICTBOK.

3. Computing Curriculum Standard J07

J07 is composed of six domains, CS, CE, SE, IS, IT and GE [4]. Here GE is a curriculum standard for general ICT education for non-ICT students. Other five domains correspond to CC2005 domains with some modification to adopt situation in Japan. Conceptual relationship of the domains is illustrated in Figure 1.

Curriculum of Japanese university is typically composed of three parts: one-year general education, two-year technical education for specific major domain, and one-year graduation research project. J07-GE is designed for non-ICT students as a part of their general education. Other J07 domains are designed for ICT students as a part of their technical education. Since ICT discipline is rapidly growing and J07 is proposed as a guideline for ICT curriculum development, IPSJ does not want to strictly "define" ICT curriculum. Instead, J07 defines core units for one-year technical education in ICT. Although ICT departments are expected to choose one of the five domains, they can freely design the remaining three-year part of their curriculum.

Among the domains of J07, CS, CE and IT have the same structure. For example, the BOK of the CS domain (J07-CSBOK) is composed of 15 knowledge areas and 138 units. Minimum core coverage time, topics and learning objectives are defined for each unit. If the minimum core coverage time of a unit is more than 0, the unit is called a core unit. The expected outcomes of the student are defined as learning objectives. The topics describe the teaching topics of the unit.

The structures of SE and IS are different from the above.

J07-SEBOK is composed of 17 knowledge areas and 297 units. J07-SE is based on SE2004 [3] except that core time is not defined for each unit. Instead, each unit is assigned to some course whose syllabus contains the weekly plan. The weekly plan also contains the learning objectives for the corresponding unit.

Main component of J07-IS is the learning units (LU). J07-IS is based on IS2002 [3] developed by AIS. 209 LUs are defined in J07-IS. Among them 108 are the core LUs. LU#, level, teaching and learning objectives are defined for each LU. The level of a LU is between 0 and 5; 0 stands for “does not know”, 1 for recognize, 2 for explain, 3 for utilize, 4 for apply, 5 for proficient.

4. Analysis Plan of J07

ICTBOK has three-level hierarchical structure composed of category, field and area. We collect the achievement and requirement levels of each area through the J07 follow up survey. The achievement and requirement levels are defined as shown in Table 2. Basically, the levels are defined according to the Bloom’s taxonomy. Level 2 is the level to explain the knowledge area or to execute the knowledge with detailed instruction. Level 3 of the skill is the level to execute the area with simplified instruction. Typically extensive training at a laboratory is required to achieve level 3 while learning or exercise at a 15-week course is required to achieve level 2. Education to earn a master degree is usually expected to achieve level 4.

We also collect the weight of each area representing the degree of importance of the area through the requirement survey.

The analysis of J07 is carried out in order to estimate the requirement level and the weight of each ICTBOK area for each domain of J07.

Among the domains of J07, CS, CE and IT have the same structure. Thus we shall explain the common

Table 2 Achievement and Requirement Levels

Level	Knowledge	Skill
0	Do not know	Cannot execute
1	Understand or know	
2	Explain	Execute with detailed instruction
3	Join discussion using the knowledge	Execute with simple instruction
4	Utilize the knowledge for problem solving	Proficient in the skill

analysis plan for these three domains. The remaining two domains, SE and IS, will be discussed in Section 5.

Assuming the above BOK structure explained in Section 3, we first make correspondence between the core units and the areas of ICTBOK. If a unit corresponds to a single area, the core coverage time of the unit is assigned to the area. In case that multiple areas correspond to a unit, the core time is proportionally divided among the areas considering the learning objectives and topics of the unit.

When we have built the correspondence, the weight of each area is computed by the sum of the core times of the area. The weight of each category and field can be computed similarly.

The requirement level of each area is basically determined from the learning objectives of the corresponding core units. The learning objectives are described using various verbs summarized in Table 3. Knowledge and skill requirements can be evaluated by these verbs. However, the requirement level is set to zero, if the assigned core time is zero since the teaching of the area is not guaranteed. The requirement level is set to one if the assigned core time is an hour or less, since it is expected to be difficult to teach an area within an hour so that students can explain the area.

Table 3 Examples of Verbs to Describe Learning Objectives

Knowledge	Skill
Explain, describe, discuss, illustrate by example, state, compare, understand, differentiate	Compute, perform, use, solve, develop, design, implement, test, debug, create, apply, choose, analyze

5. Analysis Result of J07 Domains

Figure 2 illustrates the distributions of the importance levels (weights) of the fields of the J07 domains. The weights of the fields are normalized so that the sum of the weights is equal to 100% for each domain.

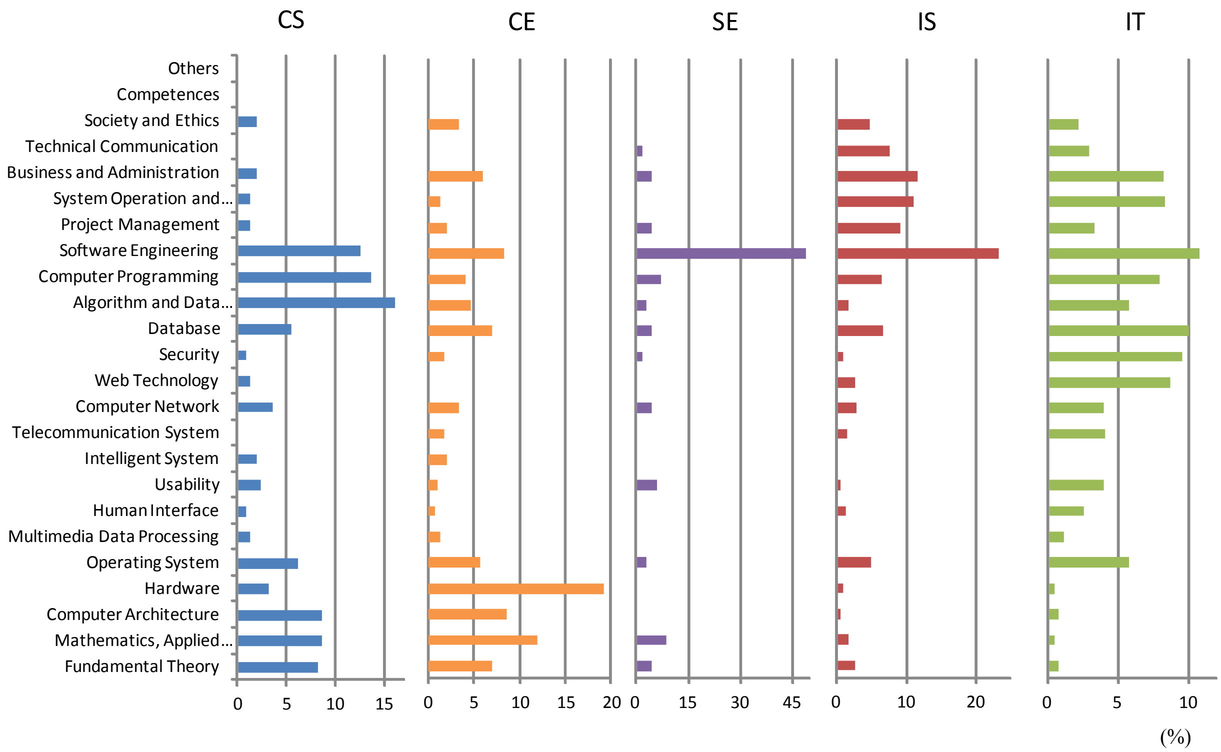


Figure 2 Importance Level Distribution of J07 Domains

Requirement Level	CS		CE		SE		IS		IT	
	Knowledge	Skill	Knowledge	Skill	Knowledge	Skill	Knowledge	Skill	Knowledge	Skill
0	105	128	94	133	115	135	82	133	86	86
1	7	3	11	6	18	11	50	0	11	41
2	43	24	50	16	22	9	23	20	58	28
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0

Table 4 Distribution of # of Areas for each Requirement Level

Table 4 represents the numbers of ICTBOK areas for each J07 domain categorized by knowledge or skill requirement levels. No area has a requirement level 3 or more because level 3 is typically achieved by extensive training through graduation research project which is outside of the scope of J07. Thus the main part of the requirements is represented by the number of level 2 areas.

5.1. Computer Science (CS)

Analysis of J07-CS is carried out as described in Section 4. In Japan, majority of ICT departments teach computer science, which focus on the theory and modeling issues. However, CS is not only a theoretical domain. CS also focuses on computer programming and software engineering as can be seen from Figure 2. The number of level 2 areas is on the average among the five domains.

5.2. Computer Engineering (CE)

Analysis of J07-CE is carried out as described in Section 4. CE is the second largest community of ICT departments in Japan. J07-CE focuses on hardware, architecture and embedded software development. Compared with CE2004 [3], software issue is more emphasized in J07-CE. This can be observed in Figure 2 by the weights assigned to software engineering, computer programming, and database. The number of level 2 areas is the second largest among the five domains.

5.3. Software Engineering (SE)

Taking the structure of J07-SE, explained in Section 3, into consideration, we assign each week of the J07-SE courses to ICTBOK areas. The weight of the area is defined by the number of courses. For example, weight of an area becomes 2/13 if two weeks of a 13-week course correspond to the area. The knowledge and skill requirement levels are estimated according to the following criteria.

- Knowledge and skill requirement levels are zero if the area has no weight.
- Knowledge requirement level is one if the weight of the area is less than or equal to 2/13; two if the weight is more than 2/13.
- Skill requirement level is computed in a similar way as the case of knowledge requirement level. The only difference is that the weight is defined by the number of exercises instead of lectures.

J07-SE strongly focuses on the software engineering field as illustrated in Figure 2. This is because J07-SE intends to develop highly skilled software developer. However the number of level 2 areas is the smallest among the five domains.

5.4. Information Systems (IS)

Analysis of J07-IS is carried out using the core LUs. We first assign the every LU to the area. If there are more than one areas corresponding to an LU, the number is proportionally assigned to each area. The weight of each area is computed by the number of core LUs corresponding to the area.

Requirement levels are determined using the following rules.

- Knowledge and skill requirement levels are zero if there is no core LU corresponding to the area.
- Knowledge requirement level of an area is one if the levels of corresponding core LUs are either 0 or 1, two if the level of a corresponding core LUs is 2 or above.
- Skill requirement level of an area is one if the levels of corresponding core LUs are 3, two if the level of a corresponding core LU is either 4 or 5.

Information systems focus on business administration, system operation and project management as well as software engineering. This is because that J07-IS ultimately aims at developing chief information officer (CIO). The number of level 2 areas is on the average among the five domains. However the number of areas having knowledge level 1 is the largest. This shows that J07-IS requires graduates to have a wide range of knowledge as a CIO candidate.

5.5. Information Technology (IT)

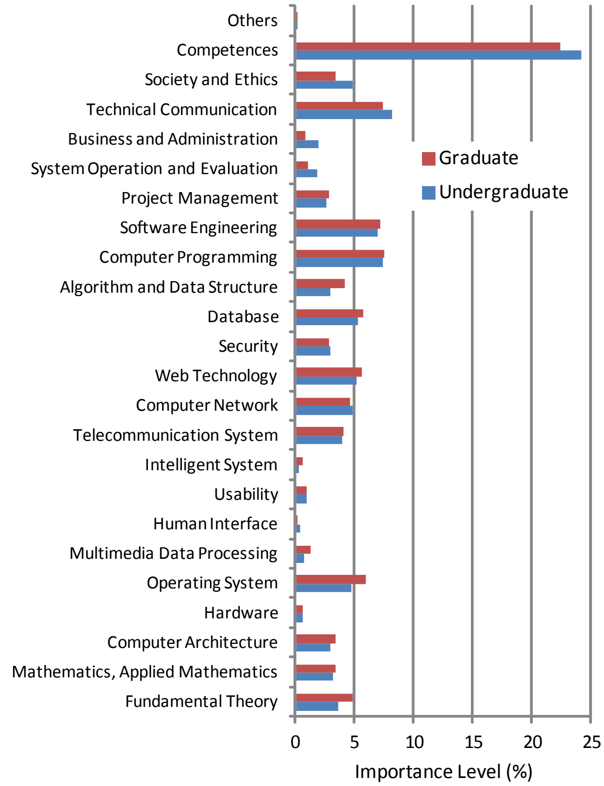


Figure 3 Importance Level Distribution of Industry's Requirement

Analysis of J07-IT is carried out as described in Section 4. J07-IT is partly similar to J07-IS, but J07-IT focuses on education of computer administrator. This can be observed from Figure 2 by the weights assigned to the fields such as database, security and web technology. Compared with other J07 domains, the weight of each field is more widely distributed. The number of level 2 areas is the largest among the five domains. This implies that J07-IT requires practical knowledge and skill as computer administrators.

6. Comparison with Industry's Requirement

6.1. Comparison of Importance Level Distribution

As explained in Section 1, we have conducted J07 follow up survey in order to collect and to analyze industry's requirement in a quantitative manner [5]. Currently 58 business persons with different background joined the requirement survey. Figure 3 illustrates the importance level distribution of 23 fields of ICTBOK. Each importance level is computed by the requirement data collected by the requirement survey. Although the opinions of the business persons are widely distributed, little difference can be found for the importance level

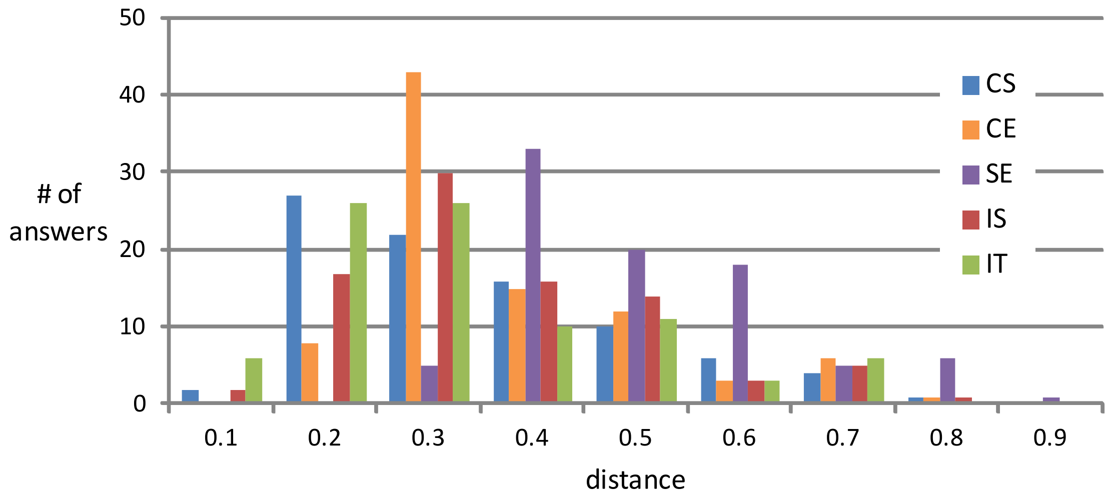


Figure 4 Distribution of Distance between Respondents' Answers and Requirements of J07 Domains

distribution for graduate and undergraduate students. Among the 23 fields, the competences field collects the highest importance. On the other hand, none of the J07 domains mention the requirement for the competences field as can be seen from Figure 2. This is because the competences are out of scope of current J07. However, the Japanese government and JABEE (Japanese accreditation board for engineering education) strongly require competences for Japanese universities.

The competences field is composed of the following 12 areas.

- Ability to act positively
- Ability to set goal and to act
- Ability to listen carefully
- Ability to express one's opinion
- Ability to analyze current status
- Flexibility
- Ability to keep rule or premise
- Ability to manage stress
- Planning ability
- Ability to influence people
- Creativity
- Ability to understand circumstances

We think that requirement for these areas can be achieved simultaneously while learning other skill or knowledge. Such knowledge or skill may not belong to ICT. However the remaining weight (more than 75%) is assigned to ICT related areas so that it is a reasonable solution to teach competences during ICT education.

The readers can observe that industry wants generic skill as well as ICT-specific knowledge and skill. Such tendency happens because the number of ICT engineers is much more than the number of graduates majored in ICT. For example, the number of ICT engineers is about 1 million in Japan; while the number of ICT graduates is about 10,000 per year which is only 1 % of the number of ICT engineers. Thus industry has to hire a large number of non-ICT graduates as ICT engineers.

6.2. Distribution of Distance

We have collected 88 answers from the 58 respondents of the requirement survey. Many of the respondents answered the requirements for both of undergraduate and graduate students so that the number of answers exceeds the number of respondents.

Profiles of the respondents are as follows. 34 respondents are corporate executives or managers, 19 are IT engineers, and 5 are employment section staffs. 38 respondents mainly engage in projects smaller than 1 Million US$ (relatively small size ICT project), the remaining 20 engage in project of more than 1 Million US$. 49 respondents are between the age of 40 and 59, while 8 are younger than 40.

We compare each answer with the requirement of each J07 domain by computing the distance between two requirements. In order to compute the distance, we first compute the importance level distribution of 23 fields of ICTBOK for each requirement data. Next the importance values are normalized so that the sum of the values is equal to 1. The distance of two requirement data is defined by the Euclidean distance of the two vector data.

Figure 4 illustrates the distribution of the distances between the answers and the requirements of J07 domains. The distance category 0.1 represents the interval between 0.1 and 0.2. The remaining categories are defined similarly. Although there are few answers whose distance is less than 0.2 for each J07 domain, more than 73% of the answers have distance less than 0.5 for CS, CE, IS and IT domains. The average distance for SE domain is larger compared with other domains because SE strongly focuses on the software engineering field. Although CS is often argued to have a large mismatch from the industry's requirement, such tendency cannot be observed from the distance distribution. The authors think that this is because that CS places weights on computer programming and software engineering.

7. Conclusion

It is often said that there is a mismatch between industry's requirement and university's achievement in ICT education. Our research focuses on this point and aims at clarifying the mismatch. In this paper, we analyze the J07 curriculum standard to compare requirements of J07 domains with industry's requirement. Relationship among J07 domains is clarified using importance weights and requirement levels. It is observed that the J07 domains should also focus on competences so that the graduates of the domain can achieve competency requirements.

We are currently executing the outcomes survey of ICT departments and graduates. 24 departments and 88 graduates have joined the survey. We are going to analyze (1) mismatch between industry's requirement and achievement of ICT department, and (2) comparison of intension of ICT departments and perception of actual graduates. The analysis result of the achievement survey will be reported in the succeeding papers.

Acknowledgement

This research is supported by the Japanese Grants-in-Aid for Scientific Research. The J07 follow up survey is a joint work with IPSJ education committee.

References

[1] Tetsuro Kakeshita and Masashi Yamamoto, "A unified BOK for IT professionals based on various IT skill standards and an analysis of educational programs using the BOK", *IPSJ Journal*, Vol. 49, No. 10, pp. 3377-3387, 2008 (in Japanese).

[2] IPA, Common Career / Skill Framework, http://www.ipa.go.jp/english/humandev/forth.html

[3] ACM, AIS and IEEE-CS, Curricula Recommendations, http://www.acm.org/education/education/curricula-recommendations

[4] Information Procession Society of Japan, Computing Curriculum Standard J07, 2008 (in Japanese), http://www.ipsj.or.jp/12kyoiku/J07/J0720090407.html,

[5] Tetsuro Kakeshita and Mika Ohtsuki, "Follow up survey of computing curriculum standard J07: requirement level analysis of industry", Proc. of IASTED Int. Conf. on Technology for Education and Learning (TEL 2011), pp. 138-145, Oct. 2011.

Are Chinese Students Satisfied With Higher Education

Su ZHANG
School of Economics, Central University of Finance and Economics, China
and
JingFeng DING
School of Finance, Central University of Finance and Economics, China
and
ShaoHong LU
School of Finance, Central University of Finance and Economics, China

ABSTRACT

The aim of this study is to find factors significantly influencing the satisfaction degree of Chinese college students to higher education, particularly in mainland China. 313 students' responses are measured through a questionnaire on a 5-point likert scale. The study evaluates higher education satisfaction in five dimensions: teaching, campus atmosphere, facilities, management and research. By probit models, the results reveal that some attributes such as teachers, courses, facilities etc. are significantly influential factors.

Keywords:Probit Model,Student Satisfaction,Higher Education

1. INTRODUCTION

Identify factors affecting student satisfaction of higher education is of great importance because higher education institutions are increasingly recognizing that higher education is a service organization, and should place greater emphasis on meeting the expectations and needs of their participating ones, that is, the students. Factors significant influential in other countries might not be effective in Chinese universities for the differences existing among countries with different economy and cultural background. Additionally, new factors will come out with time going by.

This paper divides the satisfaction of universities education into five dimensions: the satisfaction of teaching, atmosphere, facilities, management and research. Our research goes through two parts. Firstly, we study the separate effect of different parts of higher education. Secondly, we examine the overall satisfaction of higher education.

In particular, this paper tries to find the variables that have a significant impact on Chinese student satisfaction of higher education. Besides, we aim to provide suggestions on how to improve student satisfaction.

2. BACKGROUND

Pervin and Rubin (1967)[1] stated that student satisfaction was highly related to student retention. Korman (1970)[2] further pointed out that performance precedes satisfaction, which was strengthened by Siegel and Bowen (1971)[3] who tentatively affirmed that self-evaluation was an important moderator in the causal relationship between performance and satisfaction. Additionally, Sturtz (1971)[4] found that age was an important factor in affecting student satisfaction. Concerning with student demographic information, students of differential backgrounds, beliefs, attitudes, and learning styles may show different levels of satisfaction with a course (cited in Celia C. Lo, 2010)[5].

Price et al.(2003)[6] studied the impact of facilities (e.g. availability of computers, quality of library facilities) on undergraduate student, and found they have something to do with students' evaluation of their university. Mai (2005)[7] suggested IT facilities have significant and positive impact on student satisfaction.

Lots of studies found that parents' socioeconomic status and behavior had significant bearings on their children's education outcomes (Astone and McLanahan, 1991; Weinberg, 2001)[8]. Steinberg et al. (1992)[9] found that parental encouragement had a mildly negative effect on adolescents' school performance. While Kamhon Kana and Wei-Der Tsa (2004)[10] found that the effect of parenting through encouragement or award on education outcomes was positive. Besides, mothers play a more important role than fathers in shaping the education outcomes of children, which was reaffirmed by Lillard and Willis (1994)[11] and Tansel (2002)[12].

Cuthbert (l996a)[13] stated that staff and their relationship with students would override other dimensions such as computer facilities, library, and sports facilities.

However, little has been done to find the factors significantly influencing student satisfaction with university in China.

3. DATA

The data comes from our survey of students around the nation by a group in Central University of Finance and Economics.

The demographic information of students investigated shows that they come from different backgrounds and have different features. Their majors cover economics, law, psychology, art, philosophy, biology and so on. Furthermore, the data are distributed almost uniformly for each feature, such as there are 51.88% female and 48.12% male questioned.

Our questionnaire consists of five parts, namely teaching, campus atmosphere, facilities, management and research. Each parts consists a couple of questions which show the sub-dimensions that may have impact on student satisfaction.

To make the results closer to the real level of satisfaction we divide every question concerning with satisfaction into 5 levels regarding above 3 as satisfied. Compared to the method only having two levels, this technique are equal to having expanded subtle differences between the state of being satisfied and dissatisfied.

4. METHODOLOGY

We employ the probit model to measure satisfaction. The utility of satisfaction for individual i is U_i^1 and the utility of dissatisfaction for individual i is U_i^0, X_i includes a vector of demographic variables which denote the various factors that affect the satisfaction with individual i,

$$U_i^1 = \beta_0^1 + X_i\beta^1 + \varepsilon_i^1 \quad (1)$$

$$U_i^0 = \beta_0^0 + X_i\beta^0 + \varepsilon_i^0 \quad (2)$$

Where $\varepsilon_i^1, \varepsilon_i^0 \sim N(0,\sigma^2)$. But U_i^1 and U_i^0 are latent variables, which are invisible. Thus, we make following definition.

$$y_i = \begin{cases} 1, U_i^1 > U_i^0 \Leftrightarrow U_i^1 - U_i^0 = y_i^* > 0 \\ 0, U_i^1 \le U_i^0 \Leftrightarrow U_i^1 - U_i^0 = y_i^* \le 0 \end{cases} \quad (3)$$

Where $y_i = 1$ denotes the utility of satisfaction is larger than that of dissatisfaction, namely satisfied. Then $y_i = 0$ denotes dissatisfied. We can get the equation of y_i^* from our definition.

$$y_i^* = \beta_0 + X_i\beta + \varepsilon_i \quad (4)$$

where $\varepsilon_i \sim N(0,\sigma^2)$. Then

$$P(y = 1|X) = \Phi(\beta_0 + \beta X) \quad (5)$$

where Φ is the standard normal cumulative density function.

Although the question is traditionally a simple yes or no question, we assess the degree of overall satisfaction (e.g. from strongly dissatisfied to strongly satisfied). Besides, we measure the satisfaction using a multi-attribute rating scale $(e.g. Y: y^{(1)}, y^{(2)}, y^{(3)}, \cdots y^{(m)})$ instead of a single-item rating system. They reported in their essay that each student's overall satisfaction is determined by a weighted average of the gap between a student's satisfaction for each educational attribute. Each student can then be classified as 'strongly satisfied' to 'strongly dissatisfied' on the basis of a computed satisfaction score rather than a self-reported score, namely the 5-point traditional Likert scale, which will enable us to estimate student satisfaction more accurately. Thus y_i is thus calculated

$$y_i = \begin{cases} 1, & \sum_{j=1}^{m} y_i^{(j)} \Big/ m \ge 3 \\ 0, & \sum_{j=1}^{m} y_i^{(j)} \Big/ m < 3 \end{cases} \quad (6)$$

The model above will explain the variables significantly influencing the probability of student satisfaction perfectly. Besides, we have to introduce marginal effect to explain the change of probability of satisfaction causing by the change of variable's level. For example, the family income increases a certain level, how much will the probability of satisfaction change. To solve such questions and to make this research more valuable, this article calculates the marginal effect of the factors selected by probit model. As the variables are discrete, the marginal effect can be thus defined

$$\begin{aligned} ME_i &= \Phi(\beta_0 + \beta_1 x_1 + \cdots + \beta_j(c_j + 1) + \cdots + \beta_k x_k) \\ &- \Phi(\beta_0 + \beta_1 x_1 + \cdots + \beta_j(c_j) + \cdots\cdots + \beta_k x_k) \end{aligned} \quad (7)$$

when x_i changes from c_j to $c_j + 1$, ME_i is the marginal effect for individual i.

Additionally, while not the focus of this paper, we would like to examine how the predicted probabilities vary with the independent variable of rural or urban. To calculate the effect of probability we use the estimated coefficients from probit model. We expect to plot the fitted probabilities of satisfaction improvement as a function of some variable for the two values of rural or urban, fixing the values of other variables at their sample means. Presented below is the equation of the proposed approach.

$$\begin{aligned} &P_{j=m}(y = 1|\bar{x}_1 \ldots, x_i' \ldots, x_j{=}m, \ldots, \bar{x}_k) \\ &= \Phi(\beta_0 + \beta_1\bar{x}_1 + \ldots + \beta_i x_i' + \beta_j x_j(=m) + \ldots + \beta_k\bar{x}_k) \end{aligned} \quad (8)$$

$$\begin{aligned} &P_{j=n}(y = 1|\bar{x}, \ldots, x_i' \ldots, x_j{=}n, \ldots, \bar{x}_k) \\ &= \Phi(\beta_0 + \beta_1\bar{x}_1 + \ldots + \beta_i x_i' + \beta_j x_j(=n) \ldots + \beta_k\bar{x}_k) \end{aligned} \quad (9)$$

where when x_j differ from m to n, let the certain variable x_i changes ascendingly, keeping other variables at their sample mean(e.g. $\bar{x}_k$).

5. RESULTS

5.1 Descriptive data

Table 1 Descriptive Data 1

		Mean	Std. Dev.	Max	Min
TEACHING	TD	0.4252	0.4952	1	0
	TM	0.7010	0.4586	1	0
	TE	2.0753	1.1729	4	1
	ABC	2.1327	0.9773	4	1
	AM	2.0986	1.0387	4	1
	DM	2.0000	0.7572	3	1
	TL	3.5136	0.7284	5	1
	TO	2.3596	1.2699	4	1
	TEA	3.1945	0.7979	5	1
	EDL	2.7808	0.7227	5	1
	RTT	1.7551	0.7068	3	1
	CA	1.3027	0.4602	2	1
	TEE	2.8123	0.8965	5	1
ATMOSPHERR	TA	1.3814	0.4866	2	1
	SA	1.1399	0.3475	2	1
	SLA	1.2687	0.4440	2	1
	CAE	3.2517	1.0794	5	1
	SLI	3.1769	1.0721	5	1
	BSP	3.7279	0.8349	5	1
	FSP	3.1224	0.9155	5	1

	MSP	3.1911	1.0457	5	1
	FMA	3.5445	0.9817	5	1
	HPE	3.9048	1.0308	5	1
	RFD	3.0340	0.9593	5	1
	FEDI	3.0990	0.8878	5	1
FACILITY	CL	3.1190	0.9363	5	1
	CE	3.4403	0.9182	5	1
	CCL	1.4708	0.5000	2	1
	SS	2.6541	1.2764	5	1
	SC	2.7755	1.0139	5	1
	CSU	1.5017	0.5009	2	1
	LS	2.7612	1.2618	5	1
	LBU	1.6735	0.4697	2	1
	LE	1.8310	0.3754	2	1
	DS	2.7279	1.0808	5	1
	SF	2.5188	1.0713	5	1
	SER	2.0580	1.1226	5	1
	LCH	1.0952	0.2940	2	1
	LC	3.2253	1.2291	5	1
	SLE	3.0853	1.0962	5	1
	SME	2.8116	1.0198	5	1
	SSE	3.2329	0.9812	5	1
	SIS	3.3707	1.0229	5	1
	LE	0.8151	0.3889	1	0
MANAGEMENT	JDA	0.5205	0.5004	1	0
	LQN	0.3163	0.4658	1	0
	PJI	1.8982	1.3454	5	1
RESEARCH	SAFN	0.6062	0.4894	1	0
	TAAF	0.6920	0.4625	1	0
	ECE	2.8435	0.9859	5	1
	ERM	3.1599	0.9078	5	1
	ALE	3.1361	0.7674	5	1

Notes: ABC, Arrangement of Basic courses; PJI, Part-time Job Income; ALE, Academic Lecture Effect; RFD, Restriction to Free Discussion; AM, Arrangement of Majors; RTT, Relationship between Teaching schedule and Teaching material; BSP, Benefits of Social Practice; CA, Class-hour arrangement; SA, Schoolmate Attitude; CAE, Critical Attitude Effect; SAFN Students learn to Analyze First-hand data or Not; CCL, Campus Cultural Landscape; CE, Campus Environment; SC, Sufficiency of Computers; CL, Campus Layout; SER, Sufficiency of Entertainment Room; CSU, Computer and Software Updating; DM, Division of Majors; SF, Sports Facilities; DS, Dining room Scale; SIS, Sufficiency of Indicating Signs; ECE, Experimental Courses Effect; SLA, School Leader Attitude; EDL, Exam Difficulty Level; SLE, Sufficiency Logistics Equipment; ERM, Effect of Relationship with Mentor; SLI, Support Level of Innovation; FEDI, Freedom to Express Different Ideas; SME, Sufficiency of Medical Equipment; FMA, Frequency of Moral Activity; SS, Sufficiency of Self-study room; FSP Frequency of Social Practice; SSE, Sufficiency of Safety Equipment; HPE, Healthy Personality Effect; TA, Teacher Attitude; JDA, Job Description of Administrators; TAAF, Teacher Ability to Analyze First-hand data; LBU, Library Books Updating; LC, Living Conditions on campus; TD, Teacher Degree; LCH, Living on Campus or at Home; TE, Teacher Expertise; LE, Landscape Facility; TEA Teaching Effects on abilities; LE, Library Environment; TEE, Teaching Effects on employment; LQN, Librarians Quiet or Not; TL, Teacher Level; LS, Library Scale; TM, Teaching Method; MSP, Meaning of Social Practice; TO, Teaching orientation.

Table 1 provides descriptive data on various independent variables in five aspects as we have mentioned, reflecting some status quos of Chinese university students. Data reveals that most variables are not at high level which means there is possibility that educational organizations could carry out adjustments.

We have expected some results such as variables TE, 2.08, which indicates students on the whole are slightly rather than completely satisfied with the quality of teachers, BSP, 3.73, showing that students attach great importance to social practice, and so on. However, some results truly fresh our thinking such as EDL, 2.78, showing exams a little easy for students, and LQN, 0.32, lower than we thought, showing that librarians could not always fulfill their duties.

Other data shows that students lack spirit of innovation most at the lowest rate of 9.66%. The teachers are mostly graduated from 985 or world-class universities at a rate of 40.91%, indicating they are highly qualified in their study areas, however, score for teaching is less than 3,which means teachers' potential may be restricted. Additionally, when teaching students to do research, they emphasis more on data processing. Pitifully, universities focus more on teaching students knowledge from textbooks than cultivating students' creative spirit.

The satisfaction with education falls into five part: the satisfaction with teaching, atmosphere, facilities, management and research. We use 5-point Likert scale to measure the level of satisfaction, as shown in **Figure 1**.

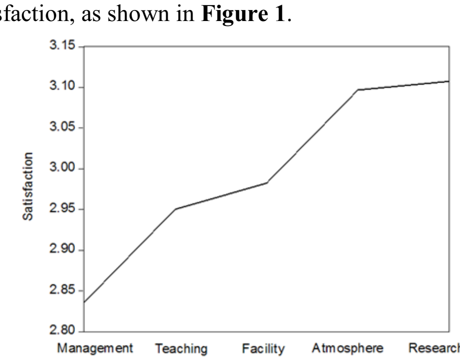


Figure 1 Students' satisfaction of five aspects

Figure 1 shows that mean satisfaction with management, teaching, facility, atmosphere and research is 2.84,2.96,2.98.3.10,3.11, from lowest to highest, similar to 3, wich is the average level in Likert 5. Therefore, the satisfaction with all parts is quiet low and it is necessary for us to carry out the study of finding out the factors that significantly affect the satisfaction with higher education.

5.2 Satisfaction of teaching

Table 2 Satisfaction of teaching

Variable	Coef.	Std. Error	z-Stat.	Prob.	Marginal effect
TF	0.6708	0.1793	3.7405	0.0002	0.2518
RU	0.7110	0.2191	3.2444	0.0012	0.2668
FI	-0.1704	0.0822	-2.0720	0.0383	-0.0640
TE	-0.3260	0.0876	-3.7223	0.0002	-0.1223
ABC	0.3268	0.1248	2.6183	0.0088	0.1227
AM	-0.2419	0.1071	-2.2597	0.0238	-0.0908
TO	0.4071	0.1407	2.8941	0.0038	0.1528
EDL	0.2798	0.1363	2.0529	0.0401	0.1050
RTT	0.3573	0.1368	2.6116	0.0090	0.1341
CA	-0.7551	0.2176	-3.4707	0.0005	-0.2834
C	-2.0828	0.8244	-2.5265	0.0115	
McFaddenR-squared	0.3788	Prob(LR statistic)	0		
LR statistic	144.3486				

Notes: *TF tuition fee; RU rural or urban ; FI family income*
All coefficients are significant at 0.1, similarly hereinafter

Table 2 shows that the marginal effect of teachers' expertise brings 12.23% difference to the probability of satisfaction. The arrangement of class hour is negatively correlated with the satisfaction with teaching and the correlation is quite obvious. The reasonable and importance-oriented arrangement of class hour gains 28.34% higher probability of satisfaction. The highest marginal effect shows that improving the quality of arrangement is a really feasible way to improve student satisfaction. The fewer and easier basic courses results to higher probability of satisfaction. While the relation between the setting of majors and the probability of satisfaction is just the opposite. Tuition fee, help to the development of students' ability, the relationship between teaching schedule, teaching material and the difficulty level of exam are all positively correlated with satisfaction with teaching. The variable changes from one level to another causes the probability of satisfaction to increase 25.18%, 15.28%, 13.41%, 10.50% respectively.

However, factors such as parents' educational background, the teaching method, teaching purpose, quality of textbooks don't have significant influence on student satisfaction.

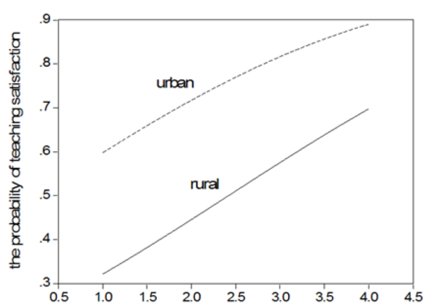


Figure 2 The structure of basis courses

Figure 2 above reveals some results we will briefly discuss, though not the main problems we study. It shows that the improvement of the breadth and variety of the basic courses will promote the probability of teaching satisfaction. Students from urban areas have higher level of satisfaction than those from rural areas and with the improvement of the breadth of a variety of the basic courses, the gap between them shows a decreasing tendency, indicating students coming from cities may adapt better to the university teaching.

This result may be largely due to Chinese practical conditions, that is the disparity between cities and rural areas. As we know all the universities in China are located in urban areas, it may be a challenge for students from rural areas to adapt to a new lifestyle and a new concept towards university education in cities different from their hometown to great degrees, which may reduce the probability of their satisfaction with university teaching.

5.3 Satisfaction of atmosphere

Table 3 Satisfaction of atmosphere

Variable	Coef.	Std. Error	z-Stat.	Prob.	Marginal effect
TF	0.3956	0.1669	2.3706	0.0178	0.1493
TA	-0.3738	0.1949	-1.9177	0.0551	-0.1411
SA	-0.8882	0.2416	-3.6764	0.0002	-0.3353
SLA	-0.6242	0.1977	-3.1575	0.0016	-0.2356
RFD	-0.2088	0.0992	-2.1040	0.0354	-0.0788
C	3.2819	0.5948	5.5181	0.0000	
McFadden R-squared	0.2009	Prob(LR statistic)	0		
LR statistic	55.5941				

Table 3 reveals that the tuition fee has a positive correlation with the probability of satisfaction, while others are all negatively influential. Students who regard the tuition fee cheap have 14.93% higher probability of satisfaction. As is known in China that tuition fee in key university is much cheaper than it in ordinary universities and training schools. Moreover, compared to other universities, key universities have better environment and atmosphere. So it's not surprising for us to obtain such results. Among the attitude of schoolmates, teachers and school leaders, the attitude of schoolmates has the largest absolute coefficient of -0.89, as well as the largest marginal effect of -0.34, which impresses us that the importance to improve students' overall quality overrides that of faculty's. Lastly, we can find that students whose discussion is less restricted have 7.88% higher probability of satisfaction, which matches our view that discussion with full freedom with classmates will contribute to the formation of positive and free atmosphere. So it's advisable for policymakers to put emphasis on developing free discussion between students.

5.4 Satisfaction of facilities

Table 4 Satisfaction of facilities

Variable	Coef.	Std. Error	z-Stat.	Prob.	Marginal effect
M	0.0498	0.0250	1.9915	0.0464	0.0186
TF	0.3042	0.1795	1.6952	0.0900	0.1135
RU	-0.3919	0.2169	-1.8065	0.0708	-0.1463
CL	0.2505	0.1338	1.8728	0.0611	0.0935

CE	0.2989	0.1437	2.0794	0.0376	0.1116
CCL	0.5164	0.2035	2.5370	0.0112	0.1927
LE	-0.5008	0.2700	-1.8549	0.0636	-0.1869
DS	-0.2071	0.1167	-1.7749	0.0759	-0.0773
SF	0.4475	0.1119	3.9993	0.0001	0.1670
LCH	0.6735	0.3328	2.0238	0.0430	0.2514
SLE	0.2282	0.1064	2.1458	0.0319	0.0852
SSE	0.2108	0.1172	1.7987	0.0721	0.0787
C	-4.5287	0.9472	-4.7813	0.0000	
McFaddenR-squared	0.3151	Prob(LR statistic)	0		
LR statistic	102.5088				

Notes: M major

Table 4 shows that campus layout, campus environment, campus cultural landscape, sports facilities, living conditions, safety equipment are all positively related to student satisfaction with facilities. Particularly, the marginal effect of campus cultural landscape is the highest except living in school or at home, which shows the importance of campus cultural landscape, indicating that paying attention to the erection of cultural landscape is fruitful to improve student satisfaction. That distinct majors lead to different levels of probability of satisfaction with facilities. Besides, students feel the tuition fee are too expensive have lower satisfaction. We also find rural students own a higher probability of satisfaction than urban students. That may because students from rural areas have lower demands for facilities. Besides, living in school or at home is the most important factor, as living at home results in 25.14% improvement in probability of satisfaction. In addition, better the dining room conditions are, easier to get books from libraries, lower the possibility of student satisfaction is.

5.5 Satisfaction of management

Table 5 Satisfaction of management

Variable	Coef.	Std. Error	z-Stat.	Prob.	Marginal effect
MEB	0.2299	0.0835	2.7531	0.0059	0.0847
FA	0.2220	0.0991	2.2413	0.0250	0.0818
JDA	0.5585	0.2061	2.7091	0.0067	0.2059
DSS	0.1326	0.0669	1.9839	0.0473	0.0489
AMDI	0.2414	0.1121	2.1536	0.0313	0.0890
LQN	-0.6008	0.2235	-2.6883	0.0072	-0.2215
PPPS	0.1648	0.0859	1.9183	0.0551	0.0607
PJI	0.2132	0.0728	2.9284	0.0034	0.0786
C	-3.0153	0.5460	-5.5229	0.0000	
McFadden R-squared	0.2424	Prob(LR statistic)	0		
LR statistic	81.8815				

Notes: *MEB mother education background FA family atmosphere*

Table 5 reveals that mothers' educational background and family atmosphere is positively correlated with the probability of satisfaction of management, which indicates that students with better growing-up surroundings are more likely to get satisfied with management. The aspects to be improved in scholarship rating system and dormitory service are both positively correlated with the probability of student satisfaction. Whereas both their marginal effect is below 0.1, suggesting that the different aspects don't make evident difference in the probability of satisfaction. That whether librarians keep quiet or not is the most significant determinant correlated with the probability of student satisfaction of management. Librarians keeping quiet contribute 22.15% higher probability to satisfaction, which reflects a fact that librarians are not so quiet in Chinese universities and it is the most important part we should emphasize on in order to improve management service. Additionally, part-time job service and income are positively correlated with the probability of student satisfaction.

5.6 Satisfaction of the cultivation of Scientific Literacy and Research Methods

Table 6 Satisfaction of the cultivation of Scientific Literacy and Research Methods

Variable	Coef.	Std. Error	z-Stat.	Prob.	Marginal effect
TF	0.3513	0.1863	1.8858	0.0593	0.1324
SAFN	-0.7933	0.2401	-3.3040	0.0010	-0.2990
TAAN	0.7667	0.2368	3.2383	0.0012	0.2890
ECE	0.3143	0.1315	2.3896	0.0169	0.1185
ERM	0.5136	0.1481	3.4680	0.0005	0.1936
ALE	0.3271	0.1720	1.9022	0.0571	0.1233
WADFR	-0.3226	0.1432	-2.2537	0.0242	-0.1216
C	-2.6057	0.6586	-3.9565	0.0001	
McFadden R-squared	0.3420	Prob(LR statistic)	0		
LR statistic	109.7568				

Table 6 indicates that tuition fee is positively correlated with the probability of satisfaction while the way to get data is negatively correlated. That is, professionals getting data directly owns 12.16% higher probability of satisfaction. This result is similar to that teachers' being able to use first-hand data to analyze contributes to a 28.90% higher probability of satisfaction. Whereas students learn how to analyze first-hand data are less satisfied. The results suggest that we should encourage teachers to perform first-hand data analysis, while don't let teachers teach students how to use first-hand data to analyze. Maybe learning how to use first-hand data to analyze is a little hard for students and it will put much more pressure on them, hence reduces the probability of satisfaction of research. Above all, it's worthwhile to study further. The influence of experimental courses, the relationship with professionals and academic lecture on students' development of scientific literacy and research methods are positively correlated with the probability of satisfaction. The results suggest that policymakers should enhance the development of these aspects.

5.7 Satisfaction with university education

Table 7 Satisfaction with university education

Variable	Coef.	Std. Error	z-Stat.	Prob.	Marginal effect
M	0.072	0.0317	2.2724	0.0231	0.0274
FA	0.2176	0.1186	1.8352	0.0665	0.0827
ABC	0.289	0.1347	2.1457	0.0319	0.1098

TEA	0.4309	0.1755	2.4554	0.0141	0.1637
CA	-0.9678	0.2988	-3.2388	0.0012	-0.3677
TA	-0.5128	0.2673	-1.9182	0.0551	-0.1948
SLA	-0.8832	0.3333	-2.6498	0.0081	-0.3356
CCL	0.8309	0.2620	3.1708	0.0015	0.3157
DS	0.3194	0.1278	2.4985	0.0125	0.1214
DSS	0.2487	0.0870	2.8576	0.0043	0.0945
PPPS	-0.2494	0.1196	-2.0841	0.0372	-0.0947
TAAF	0.6435	0.2671	2.4093	0.016	0.2445
ECE	0.4513	0.1452	3.1073	0.0019	0.1715
C	-4.2228	1.2853	-3.2854	0.001	
McFadden R-squared	0.5179	Prob(LR statistic)	0		
LR statistic	155.6908				

Table 7 contains factors significantly influencing student satisfaction with the whole university education. Analyzing those results, we could attain following findings.

Factors have significantly effects mentioned above don't necessarily affect satisfaction as a whole.

Effects of some factors have been magnified and others minified based on marginal effect. For example, the marginal effect of basic courses structure decreases from 0.12 to 0.11. While the absolute value of the marginal effect of arrangement of class hour increases from 0.28 to 0.37 as the highest, indicating rationalizing the arrangement of class hour an effective way to improve overall satisfaction.

Some factors which positively correlate with one aspect of satisfaction become negatively related to overall satisfaction.

Many distinctive factors related to China have been found. For example, the campus cultural landscape is a very important factor impacting the probability of satisfaction at a coefficient of 0.83 and a marginal effect of 0.32, showing that schools with creative and distinctive cultural landscape will enjoy much higher probability of satisfaction.

6. SUMMARY

To sum up, the satisfaction with teaching is most sensitive to the arrangement of class hour. Similarly, having schoolmates optimistic and initiative is most likely to contribute to higher probability of satisfaction with atmosphere. With regard to facilities, our analysis shows that school's cultural landscape should be creative. In terms of management, whether librarians keeping quiet or not do make a considerable difference in student satisfaction. The results herein about research show that school should teach students more about first-hand data analysis to increase the probability of student satisfaction. All these obtained results have shed light on the school leaders or the policymakers that what kind of measures they should take mostly to improve student satisfaction.

Furthermore, an important point to highlight is that "environmental" variables, including family, individual characteristics and basic information, are very important in differing student satisfaction. Background or characteristics clearly make a difference in satisfaction with education. Therefore, it's essential for parents and social institutions to improve cultivation system.

Besides, there is much more to study about the detail difference between rural and urban students.

7. ACKNOWLEDGEMENT

We acknowledge the support "Eleventh Five-Year Plan" of the Beijing Educational Science:ADA09039.

8. REFERENCES

[1] Pervin,L. A., & Rubin, P. B. (1967). Student dissatisfaction with college and the college dropout: a transactional approach. *Journal of Social Psychology*, 72, 285-295.

[2] Korman,A.K.(1970).Toward a Hypothesis of Work Behavior. Journal of Applied Psychology, 54, 31-41.

[3] Siegel,J.P.&Bowen,D.(1971).Satisfaction and performance: causal relationships and moderation effects. *Journal of Vocational Behavior*, 1971, 1, 263-269.

[4] Sturtz, S. A. (12,1971). An analysis of age differences in college student satisfaction. *Journal of College Student Personnel*,1971,12, 220-222.

[5] Celia C.Lo1(January 2010).How student satisfaction factors affect perceived learning.Journal of the Scholarship of Teaching and Learning,Vol. 10,No.1,47 – 54.

[6] Price, I., Matzdorf, F., Smith, L. and Agahi, H. (2003). The impact of facilities on student choiceof university. *Facilities*, Vol. 21 No. 10, 212-22.

[7] Mai,L.(2005),"A Comparative Study between UK and US: The Student Satisfaction in Higher Education and its Influential Factors, *Journal of Marketing Management*, 21, pp.859-878.

[8] Astone, N. M., & McLanahan, S. S. (1991). Family structure, parental practices and high school completion. *American Sociological Review*, 56, 309–320.

[9] Steinberg, L., Lamborn, S. D., Dornbusch, S. M., & Darling,N. (1992). Impact of parenting practices on adolescent achievement: Authoritative parenting, school involvement, and encouragement to succeed. *Child Development*, 63(5), 1266–1281.

[10] Kamhon Kana, Wei-Der Tsa (2004). Parenting practices and children's education outcomes. Economics of Education Review, 24 (2005), 29–43.

[11] Lillard, L. A. & Willis, R. J. (1994). Intergenerational educational mobility: Effects of family and state in Malaysia. *Journal of Human Resources*,29(4),1126–1166.

[12] Tans el, A. (2002). Determinants of school attainment of boys and girls in Turkey: Individual, household and community factors.*Economics of Education Review,* 21(5), 455–457

[13] Cuthbert, P. F. (1996a). Managing service quality in HE: is SERVQUAL the answer? Part 1. *Managing SenJice Quality*, 6(2), 11.

Numerical Hybrid Simulation Modeling Verification for a Curved 3-Pier Bridge (Investigation of Combined Actions on Reinforced Concrete Bridge Piers (CABER) Project)

Adel ABDELNABY
Civil and Environmental Engineering, University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

,

Thomas FRANKIE
Civil and Environmental Engineering, University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

,

Prof. Billie SPENCER
Civil and Environmental Engineering, University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

and

Prof. Amr S. Elnashai
Civil and Environmental Engineering, University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

ABSTRACT

Reinforced concrete bridge piers are subjected to complex loading conditions under earthquake ground motions. Bridge geometric irregularities and asymmetries result in combined actions imposed on the piers as a combination of displacements and rotations in all six degrees of freedom at the pier-deck juncture. Existing analytical tools have proven their inadequacy in representing the actual behavior of piers under these combined actions, particularly in their inelastic range. The objective of this investigation is to develop a fundamental understanding of the effects of these combined actions on the performance of RC piers and the resulting system response.

This paper describes a part of the CABER project that verifies the numerical hybrid simulation of the curved bridge. In this part two models were introduced, a whole model and a sub-structured hybrid model. The whole model was established using the Zeus-NL analysis platform, which is capable of performing inelastic nonlinear response history analysis of the whole curved bridge. The hybrid model was divided into three modules which comprised the deck, left and right piers, and the middle pier of the bridge. The three modules were modeled by Zeus-NL as a static analysis module interface. The simulation coordinator (SimCor) software was utilized to communicate between these modules using a Pseudo-Dynamic time integration scheme. Results obtained from both models were compared and conclusions were drawn.

Keywords: Numerical/Experimental Hybrid Simulation, RC Bridges, Combined Actions and Inelastic Dynamic Analysis.

1. INTRODUCTION

The numerical hybrid simulation analysis discussed in the paper was performed as a partial requirement for the completion of the numerical-experimental hybrid analysis of the CABER project. The CABER project involves modeling a curved bridge with three piers of varying heights in unique soil conditions and with unsymmetrical spans along the curve of the bridge deck (Figure 1). The three piers will be tested experimentally at the Multi-Axial Full-Scale Sub-Structures Testing and Simulation (MUST-SIM) facility at UIUC. Two piers comprising one module will be tested on the large strong wall at a scale of 1:3, while the third pier will be tested simultaneously at the small scale facility located at the same site. The deck, abutments, and soil will all be modeled computationally. However, in this paper, the three piers are modeled numerically along with the deck, abutments and soil. This is done in anticipation of encountering various phenomena to explore, as well as challenges in controlling the experimental specimens to overcome prior to conducting the complete hybrid experiment described above.

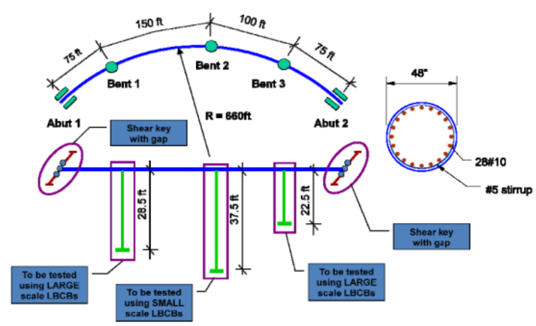


Figure 1: Conceptual drawing of complex curved bridge for CABER project hybrid analysis.

New and advanced considerations that must be addressed in this specific project involve:

- Functional constraints due to the skewed bridge deck and the torsional moments that will develop in the structure as a result
- Geometric constraints due to uneven spans and heights of the three piers
- Structural constraints due to the modeling of the joints and soil-foundation interaction
- Multi-directional motion of the applied earthquake record

Addressing these issues requires significant advances in the control software used to implement the prescribed displacements provided by the simulation coordinator software, SimCor. These actions are applied through the use of the Load and Boundary Condition Boxes (LBCBs) in all six degrees of freedom. In addition to considerations regarding controls, the UI-SimCor platform will require updating for use of the multidirectional event and the functional constraints introduced by the increasing complexity of the structure. The following section provides a brief overview of the concept of hybrid simulation, in order to give the reader insight into the facilities and software used.

2. OVERVIEW OF HYBRID SIMULATION

Hybrid Simulation is a platform for the analysis of structures using any combination of computational and experimental testing methods. Traditional experimental and computational approaches can be performed simultaneously through sub-structuring of the given system into separate modules of interest. Through the application of a simulation coordinator platform, communication is enabled between the experimental or computational modules of the structure being analyzed. This allows for increased flexibility in testing programs along with an improvement of the accuracy and efficiency previously available through traditional standalone experimental or computational testing programs (Watanabe et al. 1999, NSF 2000, Tsai et al. 2003, Kwon et al. 2005, Pan et al. 2005, and Takahashi et al. 2006).

The most distinctive feature of the MUST-SIM facility is the L-shaped post-tensioned concrete strong wall (Figure 2) and its three modular 6-DOF Load and Boundary Condition Boxes, LBCBs, shown in Figure 3. These boxes allow for precise application of complex load and boundary conditions.

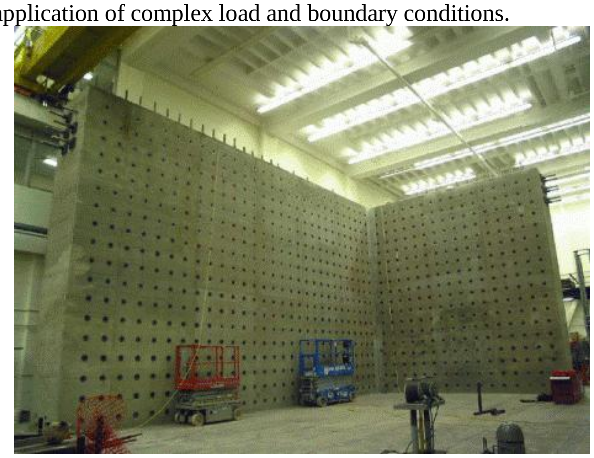


Figure 2: Strong wall at UIUC MUST-SIM facility.

The LBCBs can impose motions on structural specimens that are determined from the results of concurrently running numerical models of the test specimen and the surrounding structure-foundation-soil system through employing hybrid simulation.

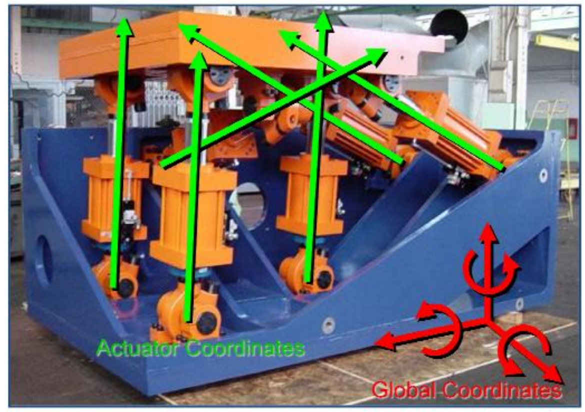


Figure 3: 6-DOF Load and Boundary Condition Box.

The computational tools of most interest in this investigation are the FEA programs compatible with SimCor, the simulation coordinator used at the MUST-SIM facility. The programs currently supported by the SimCor platform are Zeus-NL, OpenSees, FEDEASlab, ABAQUS, and Vector2.

The simulation coordinator (SimCor) provides the communication between modules necessary to perform hybrid analysis. This framework provided allows for the utilization of any combination of analytical platforms and experimental facilities to be integrated and simulate a larger, more complex system. SimCor utilizes pseudo-dynamic (PSD) simulation for distributed analysis and experimentation. This concept involves the substructuring of the complex/whole system into smaller modules that can be solved with individual computers connected through a variety of available communication protocols. The time integration scheme is performed in SimCor and modules perform static analysis based on the information received from SimCor. A more detailed explanation of this process is illustrated in the figure below.

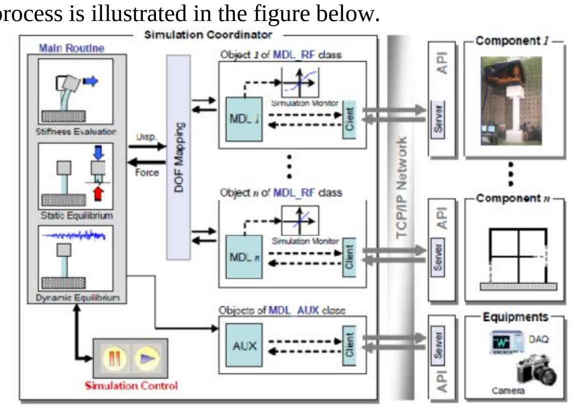


Figure 4: Hybrid simulation framework in SimCor.

3. NUMERICAL HYBRID MODEL

In this section, a whole Zeus-NL model of the curved bridge is introduced (Figure 5). The curved bridge system is also split up into three Zeus-NL modules as shown in Figure 6. The first module consists of the superstructure, and the second module is the inner pier representative of the small-scale experimental structure for the final experiment. Finally, the third module includes both of the outer piers, which represent the portions of the structure that will eventually be tested on the large strong wall.

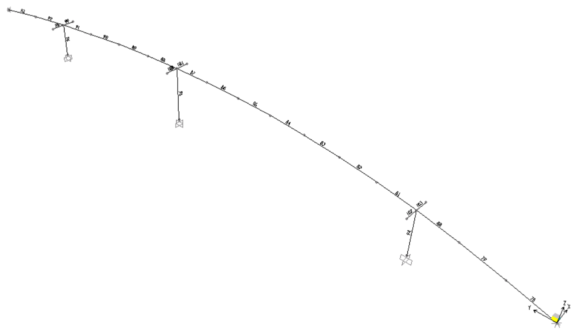


Figure 5: Whole Zeus-NL model.

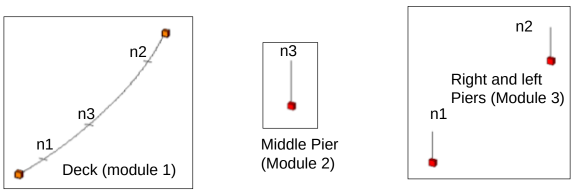


Figure 6: Three modules of the numerical hybrid model.

The bridge deck was divided into sixteen segments; each segment is 25 ft. long. The curvature of the bridge was taken as 1/ (650 ft.). The deck, piers, and transverse beams cross sectional dimensions are shown in Figure 7. In this analysis, the deck and transverse beams were assumed to remain in the elastic range during the analysis, which is usually the case for these types of bridges. Moreover, the strains at critical locations along the deck and transverse beams were monitored throughout the analysis and it was revealed that the strains have not exceeded the cracking strains of concrete. Hence, inelasticity and failure were only assumed at the piers which will be eventually modeled experimentally.

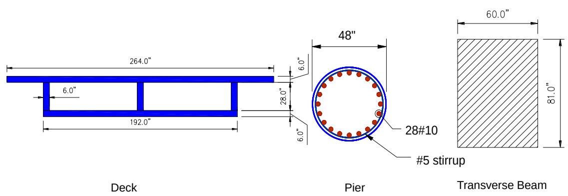


Figure 7: Deck, piers, and transverse beams cross sectional dimensions.

The loading of the bridge included bridge self-weight as well as the weight of pavement, utilities and finishing. Traffic load was also included as the assigned structure live loads. Seismic loads were applied as earthquake ground accelerations at the pier bases. The acceleration records were applied without scaling on the transverse direction of the bridge while at the transverse direction the same record accelerations were scaled by 0.25.

Four scalings of the same synthetic record were applied in series in this response history analysis. The scalings represent earthquake motions compatible with response spectra of MCE-Cracking, -Yielding, -Design Level, and -Failure subsequently as shown in Figure 8.

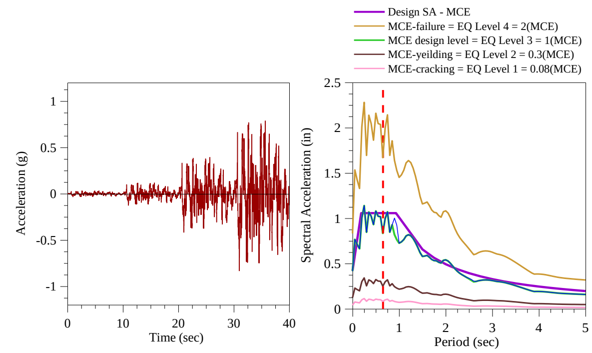


Figure 8: Earthquake record.

The boundary conditions of the bridge are assumed as follows:

1- The abutment-deck interface is simulated by two non-linear spring models as shown in Figure 9. The first spring model simulates the gap between the bridge deck and the abutment and can therefore account for pounding effects during earthquake shaking. The second spring model depicts the hysteretic response of the shear key element of the abutment as shown in the same figure. The two spring models are connected in series to yield an overall response shown in Figure 9.

2- Fixation is assumed at the base of each pier.

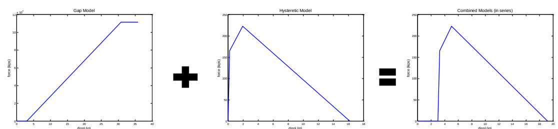


Figure 9: Two spring models in series and the overall spring model behavior.

4. RESULTS AND OBSERVATIONS

The results obtained from the whole model as well as the results of the hybrid analysis are provided in this section to demonstrate the success in running the hybrid simulation using SimCor to communicate between three modules that are representative of an ultimately more complex combined experimental/analytical hybrid test. The results are very satisfactory in the sense that the displacements and rotations obtained from the whole model matched very well with the hybrid analysis results. The deformations at the top of the inner pier obtained from the whole and hybrid models are plotted on top of each other and provided in Figure 10.

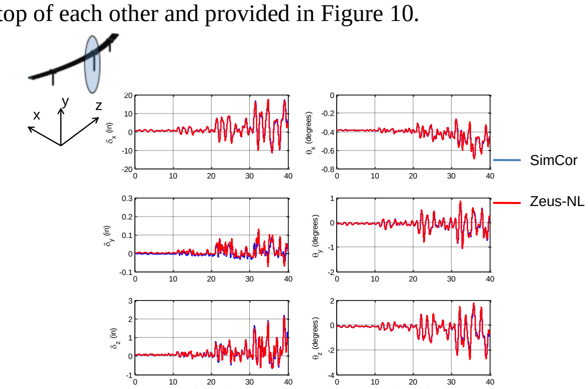


Figure 10: Deformations at the inner pier.

The stroke and force limits of the LBCBs are also checked in this study using the results obtained from the numerical models. The actuators of the large scale LBCBs to be used for the outer piers have stroke capacities of +/- 10 inches in the X-direction and +/- 5 inches in the Y- and Z-directions. The force capacity is -200/+300 for all actuators of the large scale LBCBs. The actuator forces and strokes were calculated using Matlab code developed for the MUST-SIM facility by SunJig Kim. This code is capable of calculating the actuator forces and strokes given the displacements and rotations imposed on the boxes in the 6-DOFs. Figure 11 and Figure 12 show the measured strokes and forces respectively for one of the outer piers that will be tested using the large scale LBCB. The anticipated displacements and forces do not exceed the capacities of the loading boxes. Similar checks were performed for both the other outer pier as well as the inner pier that will be tested using the small scale box.

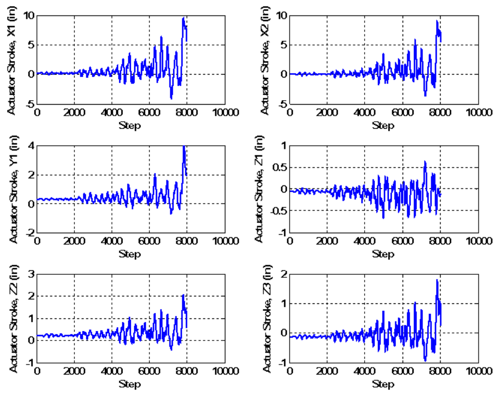


Figure 11: Strokes of LBCB actuators at one of the outer piers.

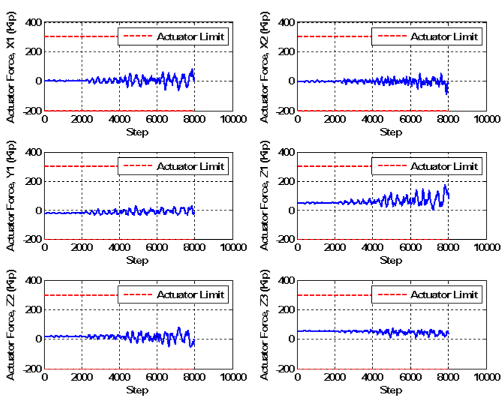


Figure 12: Forces imposed by the LBCB actuator on one of the outer piers.

5. SUMMARY AND CONCLUSIONS

In summary, hybrid simulation has been shown to be capable of providing an accurate and efficient analysis of complex structures not previously attainable through traditional methods. An overview of the methodology of hybrid simulation and the underlying motivation was provided. The experimental, analytical, and simulation coordinator components of the platform specific to the NEES MUST-SIM facility were provided, in addition to the variety of hybrid methods available through the use of these components. These distinct forms of hybrid analysis include mixed experimental and analytical tests, multi-resolution computational analysis, and multi-site geographically distributed experimental testing programs. An example from the literature was provided to study the aspects of each of these methods, and current shortcomings were addressed within the context of future work for a more complex hybrid analysis which the author will be involved in as a part of the ongoing CABER project. Finally, a hybrid analysis of a simple four-span bridge was performed in a test case designed to provide early exposure and orientation to the hybrid analysis capabilities of SimCor.

6. ACKNOWLEDGEMENTS:

Special thanks to Gregory Pluta and Weslee Walton, NEES site coordinators; Thomas Frankie for his infinite knowledge and support; Prof. Oh-Sung Kwon, Assistant Prof. at University of Toronto, for his help in numerical model development and knowledge in using Sim-Cor; and Post Doc. Researcher Chai-Ming Chang for his support.

This research is funded by the NEES at UIUC and NEES Inc. NEES at UIUC is supported by the George E. Brown, Jr. Network for Earthquake Engineering Simulation (NEES) Program of the National Science Foundation (NSF) under Award Number CMMI-0927178.

7. REFERENCES

[1] Elnashai, A. S., Papanikolaou, V. K. and Lee, D. H. “Zeus-NL User Manual”, Version 1.8.7.

[2] Elnashai, A. S., Spencer, B., Kuchma, D., Ghaboussi, J, Hashash, Y. and Gan, Q. (2004) “Multi-Axial Full-Sclae Sub-Structured Testing and Simulation (MUST-SIM) Facility at the University of Illinois at Urbana-Champaign”, 13th World Conference on Earthquake Engineering, Vancouver, Canada, Paper No. 1756.

[3] Ji, J., Kwon, O. S., Elnshai, A. S. and Kuchma, D. A. “Multi-resolution Distributed FEA Simulation of Reinforced Concrete High-Rise Buildings.”

[4] Kwon, O. S., Nakata, N., Elnashai, A. S. and Spencer, B. (2005) “A Framework for Multi-Site Distributed Simulation and Application to Complex Structural Systems”, Journal of Earthquake Engineering, Vol. 9, No. 5, pp. 741-753.

[5] Kwon, O.S., Nakata, N., Park, K.S., Elnashai, A. and Spencer, B. (2007) “UI-SimCor User Manual and Examples”, UI-SimCor v2.6 and NEES-SAM v2.0, University of Illinois at Urbana-Champaign.

Feedback System For Electric Trimmer

M.Sriram[1], Swagat Sarma[2] , K.Selvajyothi[3]

1,2- B.Tech Students, Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram
3- Assistant Professor, Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram

Abstract— **Automation and enhancement of the quality of trimming by an innovative way of constructing a trimmer has been devised in this paper. The improvements proposed here intend to increase the functionality of the trimmer by adding an extra perspective to the trimming process, at the same time optimizing the space constraints of adding new or additional components to do so; this has been done by adding a closed loop feedback system for additional control and precision. This closed loop feedback system requires the fabrication of an innovative electromechanical system which has been proposed in this paper.**

Keywords- electric trimmer; closed loop feedback; microcontroller, gear mechanism, electromagnetic clutch, bearings.

I. INTRODUCTION

Trimming is a day to day phenomenon, whether it is in front of the bathroom mirror, or in a hospital bed before an operation, or while shearing off the wool of a sheep. However, a problem is always faced by the person while negotiating uneven surfaces. It results in the person bending the trimmer in various ways to effectively trim that part. It results in imperfect trimming and can be quite painful if it leads to cuts on the face or other parts.

Secondly, for people with one hand, or a disability in one hand, or a fractured arm in a sling, trimming becomes an arduous everyday task which they are not able to undertake alone, or if they are able to, they are unable to achieve the desired quality of shave. So, this product also has social benefits too and its development is not purely driven by market motives.

Blaustein et al [1] patented a wet razor and dry trimmer combination so that trimmers can be used while submerged in water. That was one of the first patents approved for trimmers after the first patent for the electric trimmer was granted to Poel [2]. However, there has been no study on the usage of a closed loop feedback control system on an electric trimmer for better control and precision.

Eventually, electric razors got a bit more sophisticated. The injector razor made wet trimming comparitively safer. Then, in the 1970s, wet trimming advanced again with the creation of the twin-blade cartridge. The idea had been around for a while, but the revolution that occurred in the 70s happened from a marketing standpoint. Development on the triple-blade razor began in the 1970s, but apparently there were problems with three blades. They caused irritation. In 1998 the solution to this problem appeared. The Mach 3 trimming system from Gillette uses three blades.

In this paper, a closed loop feedback system has been incorporated with a change in the trimmer mechanical design to allow increased functionality during the process. An electromechanical system in the design allows two degrees of freedom of the blades in contact, while minimizing the addition of new components while doing so, thereby keeping the design of the trimmer feasible in size and ergonomic in nature.

While the design of this trimmer does not innovate in the field of blade technology, or 'wetness' factors associated with shaving, as they have already been researched into extensively, it does propose an alternative method to improve functionality of contemporary trimmers while keeping within similar size and cost constraints. The novelty of this design lies in the ability to provide multiple degrees of freedom to the blades of the trimmer, using limited mechanical components.

II. PROPOSED TRIMMER

A. Concept

To improve the entire process of trimming holistically, the approach of maintaining a user defined pressure over the blades in contact of the trimmer has been chosen. This will lead in partial automation of the entire process of trimming, especially over undulating surfaces.

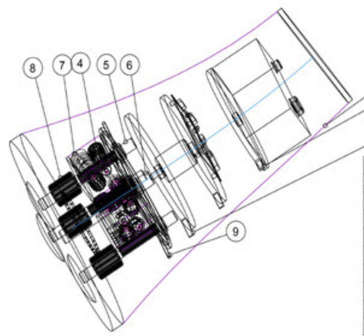
This has been done by using three blade spindles instead of one, on each of which a pressure sensor has to be mounted. The combination of three pressure sensors will sense the pressure being applied to the spindles and the average will be calculated. If any spindle's pressure sensor detects a difference from the average pressure exerted by all the spindles, it will move in a direction so as to offset the pressure difference. So if one of the blades move over the edge, there will be a drop in the pressure on the spindle compared to the other two, therefore the mechanism will force the spindle to move down to compensate for the change in pressure and the blades will stay in contact with the surface to be trimmed.

In case of sudden pressure increase on one of the spindles, it can retract with the same reverse mechanism to reduce and maintain the applied pressure on the spindles.

B. Development

The model hopes to affect an improved solution for effective trimming. A sleek structure has been adopted both for aesthetic look and good handling, hence ease of usage. However, the housing is limited by the inner mechanism and the types of components planned to be used. The three spindles with inner blades are shown in Figure. 1.

The housing also has the function of holding the fixed components like the power supply, gearbox, motor, the printed circuit board etc. in place.



SL NO	PART NAME	MATERIAL
1	HOUSING OUTER	N-POLYPROPYLENE
2	BATTERY SUPPORT	N-POLYPROPYLENE
3	MOTOR SUPPORT	N-POLYPROPYLENE
4	SPUR GEAR	PLASTIC
5	BLADESHAFT	HSS
6	MOTORSHAFT	HSS
7	WORM GEAR	PLASTIC
8	SPUR GEAR 2	HSS
9	SHAFT SUPPORT	N-POLYPROPYLENE

Figure 1. ProEngineer (ProE) model of the Proposed Trimmer.

C. Components of the Trimmer

1) Electronic components

Piezo-resistive transducer is used to provide an accurate, high-level analog output signal that is proportional to the applied pressure. The small form factor and high reliability of on-chip integration make the sensor a logical and economical choice for the desired product. The location of the pressure sensor is between the blades and the housing. The output from each sensor is given to signal conditioning board and controller. Voltage proportional to the normal trimming conditions is given as the threshold limit for the activation of the system. Otherwise the system works normally. When the threshold is exceeded, the corresponding electromagnetic clutch is activated by the control system so that upward or downward motion can be transmitted.

2) Gear Mechanism

The ideal motor to be used is a 24V DC motor with speed rating of 3000 rpm. The main constraint for designing the gear mechanism is the size and ease of holding. Keeping this in mind a single motor drive which provides two degrees of freedom to the blade spindle is provided in the trimmer. As the motor is going to be attached to the single motor spindle, it is needed to transmit motion from the motor spindle to the three blade spindles. The gear mechanism has been designed in Pro/E as shown in Figure 1 with the constrained size and contour of the trimmer.

Gearbox consists of the following parts as shown below in Figure 2 which is discussed in detail in the following subsections. This gearbox provides only the translational motion to the blade spindles and does not include the rotary motion transmitted to the blade spindles from the motor spindle by spur gears connected to both.

a) Bevel Gear System

A bevel gear system (‘A’ shown in Figure. 2) is used to translate the motion to the blade spindle from the wheel of the worm-wheel arrangement.

b) Worm and wheel arrangement

A worm is to be put on the motor spindle which will transmit motion to the wheel (‘B’ in Figure.2). There is supposed to be massive speed reduction in this. A helical cut has been put on the motor spindle in the assembly drawing shown in Figure. 1 to indicate the worm. The speed reduction is done in order to reduce vibration during the linear movement of the blades. The speed reduction in worm wheel is high in order to avoid rapid movement of the blades while trimming.

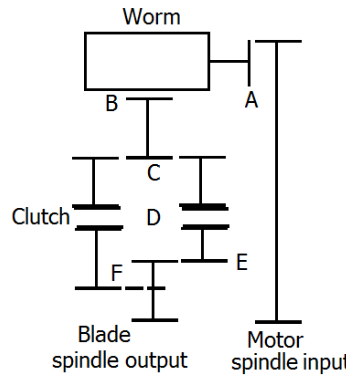


Figure 2. Schematic representation of Gear Mechanism used in the trimmer.

A-Bevel Gear (1:1); B-Worm and worm wheel (1:18); C-Spur gear (1:1:1); D Clutch; E-Spur gear (1:1); F-Spur gear (1:1:1)

c) Spur Gear

Spur gear is being used to transmit the rotary motion from the motor spindle to the blade spindle. The gear is mounted on the spindle using an interference fit so that there is absolutely no slip in the motion. The spur gear is not intended for speed reduction but only speed transmission.

Another spur gear arrangement ‘C’ (as shown in Figure.2) is used to transmit motion to the two electromechanical clutch systems. The spur gear arrangement ‘E’ (as shown in the Figure.2) transmits motion to the blade spindle in the same direction as ‘C’, whereas the spur gear arrangement ‘F’ (as shown in the Figure.2) transmits motion to the blade spindle in the opposite direction as ‘C’. These opposite directions of gear motion make it possible to provide two overall opposing directions (upward and downward) of spindle movement to the blade spindle.

d) Electromechanical Clutch System

The spur gears transmit motion to two spindles connected to an electromechanical clutch (‘D’ as shown in the Figure.2) each. This electromechanical clutch gets activated when an appropriate voltage signal is provided to it by the control system, resulting in the clutch getting engaged and transmitting motion to the rest of the gear system connected to it. It is to be noted that the gear configuration used is such that one of the clutches is connected to a system which provides an upward motion to the blade spindle and the other clutch provides a downward motion to the blade spindle.

e) Rack and Pinion System

A rack and pinion system is used to provide a translational movement to the blade spindle. There is one rack and pinion associated to each of the two electromechanical systems. Each pinion is always engaged to its respective rack. The control system activates and deactivates the electromagnetic clutch depending on the pressure on the blade, is responsible for choosing which rack and pinion system will be used to transmit linear motion in alternate directions (upward or downward).

Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram, Chennai Tamil Nadu, India.

The rack teeth are cut on the deep grove ball bearing which is holding the blade spindle.

3) Ball Bearings

Deep grove ball bearings are used on the blade spindle. The housing of the ball bearing is given a rack cut so that it can act like the rack in the rack and pinion arrangement. This ball bearing performs the function of allowing and transmitting two degrees of freedom for the blade spindle unlike standard ball bearings which provide a singular rotational degree of freedom to its respective spindles.

III. WORKING OF THE TRIMMER

The working of the proposed trimmer is shown in the overall schematic in Figure 3. When the trimmer is switched on, the main motor starts rotating. Rotational motion is transmitted from the motor spindle to the blade spindles using a spur gear at the same speed of rotation as the motor spindle. Speed can be modified for greater or lesser torque.

The pressure sensors detect the pressure applied by the user. At every point of time the average pressure across the three spindles is calculated by the microcontroller and the pressure across each spindle is tested against the average for some set limit, i.e. the threshold value. In case of no variation beyond the threshold, there will be no change in the gearbox configuration and thereby on the position of blade spindle.

In case of a pressure beyond the threshold, the control circuit will send a signal to the power circuit which will activate one of the electromechanical clutch systems attached in the gearbox and start the translation motion for the particular blade spindle either in an upward direction or a downward direction.

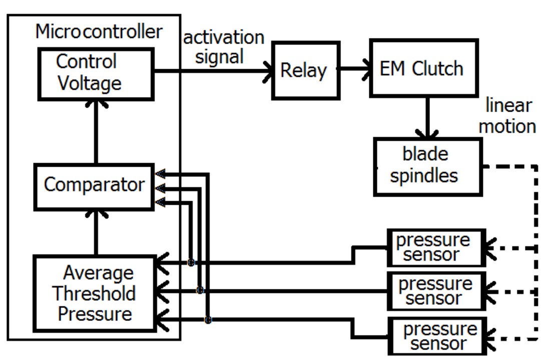


Figure 3. Overall schematic of the proposed Trimmer.

IV. DESIGN INNOVATION

The trimmer has certain innovative features which do not exist in existing state of art electric trimmers. This trimmer does not consider the vibratory motion used by existing trimmers to create a shave as demonstrated in [3]. Neither does it concentrate on the torque transmission or the ability to perform wet trimming as well as dry trimming which have already been done in [4]. The blade to be used by the user of this trimmer can be chosen according to the user's preference itself. However, this trimmer uses a closed loop feedback system to provide an automated stabilization of pressure applied across all blades. This feedback makes it more user-friendly and can be used across different configuration of other types of trimmers as well.

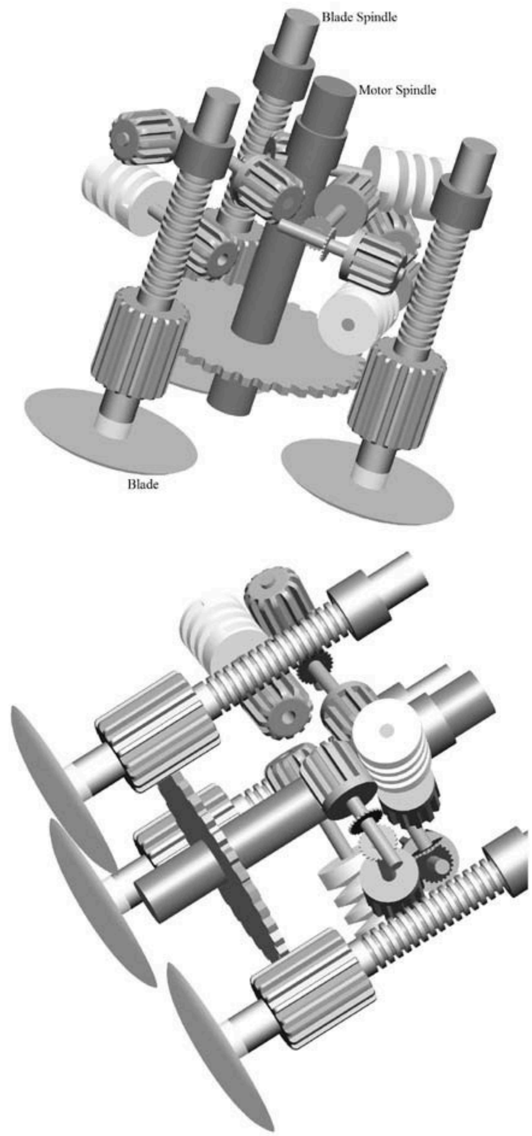


Figure 4. Two different views of Gear mechanism of the trimmer.

Another salient feature of this design is the ability to control and transmit two degrees of freedom using a singular rotary motor. It has been done by designing a gearbox system using electromechanical clutches to control the linear motion of the blade spindles electronically. The extent of the translation of the blade directly affects the pressure applied on the surface on which the trimmer is applied. This can further be modified by using some damping to actuate the motion of the blade; however, it can also be ignored as the motion to be executed is very small in nature.

Thirdly, the threshold pressure which will be selected by the control system to control the blade spindles is modified according to the user on the basis of the current pressure applied. The average of the pressure on all the blade spindles is taken, and the individual pressure on each blade is adjusted on a step basis by extending or retracting the blade spindle so that all the blades move to attain the same pressure. A step basis is chosen so that the increment or decrements in the extension of the blade spindle does not change in spurts, or rather changes eventually. Moreover, after each adjustment to the blade spindle position, the pressure input is reevaluated for continual position change in the blade spindle position, for a very user-friendly experience.

However, the ergonomics of the housing of the trimmer is not discussed here. This design involves an innovation particularly to the functionality of the trimmer.

V. CONCLUSION

An effective trimmer, which can give better trimming than contemporary models is presented in this paper. The pressure sensors are incorporated to enhance the quality of trimming. This automated feedback mechanism will avoid manual feedbacks resulting in cuts and imperfect trimming, thereby resulting in a better trimming experience.

However, this trimmer requires the usage of lots of miniature mechanical components, which might make the trimmer more complicated to manufacture. But this also means that it can be as light as, if not lighter, than contemporary trimmers.

Furthermore this trimmer can be modified for a better ergonomically and aesthetically pleasing housing design. The materials used for the housing can be injection moulded synthetic materials.

ACKNOWLEDGMENT

We would like to thank our director Prof R.Gnanamoorthy and Dr M.Sreekumar for their valuable direction and facilitation for research work. We also would like to thank our undergraduate classmates S.Gayatri, Wanrisa Bok Kharkongor and Nitin.M.Rajeev for their valuable help in the initial stages of the design.

REFERENCES

[1] Lawrence A. Blaustein, Douglas A. Gall, John Willian Nottingham, John Usher, on ‘Wet razor and electric trimmer assembly’, 2010, US Patent No: 7,761,998 B2.

[2] Hendrik Poel, ‘Electric Trimmer’, 1976, US Patent No: 3962785.

[3] Shiro Mizobata, Yoshiro Kawano, Kenzo Itoh, Takeshi Usui, ‘Electric Razor’, 1974, US Patent No: 3842498.

[4] Willem Anema, U.S. Philips Corporation, ‘Dry Shaver’ 1990, US Patent No: D306359.

[5] Hendrik Poel, Johannes A.M. Van Hout, ‘Electric Trimming Appartus, US Patent No: 4930217.

Development of thermal simulation model of supercapacitor module targeting optimal requirement on computational time

Michal FRIVALDSKY*, Peter DRGONA*, Pavol SPANIK*,

and
Norbert GLAPA**
* Department of mechatronics and electronics, University of Zilina
Zilina, 01026, Slovakia

** Panasonic Electronic Devices Europe GmbH,
Zeppelinstrasse19, Lüeneburg, 21337, Germany

ABSTRACT

This paper deals with development procedure of EDLC module's (multi-cell battery system) thermal simulation model and consequential simplification targeting as lowest simulation time as possible, together with achievement of high model's accuracy. On the other side, one of most important requirement is possibility of model's utilization almost on standard personal computer (nor multi-core computation machine). First the problematic according development of supercapacitor thermal simulation model will be described. After it the process of geometrical model development of proposed EDLC module is shown. Consequently the description of optimization steps of simulation model's geometry (simplifications) targeting shortening of simulation time will be described. Finally simulation experiments which are confirming proposed design procedure have been provided.

Keywords: COMSOL, thermal simulation, supercapacitor, computation time.

1.INTRODUCTION

Supercapacitor is an electrochemical capacitor with relatively high energy density. A typical sized electrolytic capacitor displays capacitance in the range of tens of milifarads. The same size EDLC might reach several farads, an improvement of two orders of magnitude. Focusing on future trends in industrial applications this solution presents very perspective alternative for energy storage, whereby its big advantage is that this energy is available to be utilized in a very short time. Due to this ability supercapacitor spreads its utilization into various areas, including consumer electronics as well as industrial. Just second application area (industrial) presents the most suitable possibility of successful use. Considering about mainstream on which each big companies is being focused, we can talk about automotive applications, where high voltage battery module is needed. Just for this case, the supercapacitor module presents most flexible and perspective solution. It can be utilized as main energy storage tank (main energy unit of electric car), or as auxiliary energy storage tank for transient needs (starting of car, supplying of auxiliary electronic systems, etc.).

2.THEORY AND MATHEMATICAL BACKGROUND OF ONE CELL OF EDLC

In principle there exist many types of mathematical models of supercapacitor, which can be then suited for the purposes of thermal simulations. Why is thermal simulation model of supercapacitor so important?

As was already mentioned, it is very perspective solution from the energy storage point of view. But due to fact that each solution has some disadvantages, the negative of its structure lies in low operating voltage and high sensitivity against allowable temperature rise. Due to second mentioned disadvantage, the development of physical sample for specific application could be time consuming. Avoiding this situation can be done through proper design procedure of thermal simulation model, which will be capable to exactly interpret physical behavior (from thermal point of view) of considered supercapacitor. Initial evaluations of thermal simulation model's accuracy can be done through comparisons with experimental measurements. After it, for research of EDLC behavior in different conditions of use, it is sufficient to utilize the thermal simulation model with acceptable complexity.

Due to statements mentioned above it is clear to say that complexity of thermal simulation model is directly related to the mathematical background. Our target was development of model which will be very accurate (novel approach of mathematic modeling was necessary), but also very simple from the computation time point of view. This problematic was well described in our previous works [1] [2] [3].

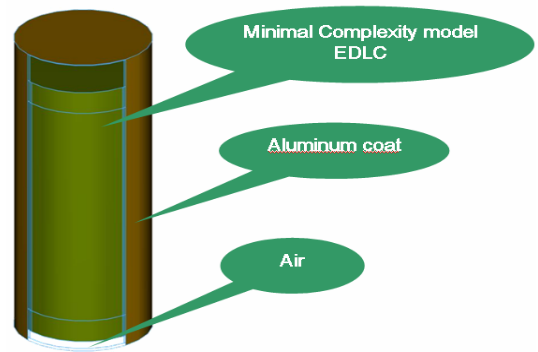


Fig.1. Improved complexity model of 1xEDLC

Also later in [1] three alternatives of EDLC model complexity has been shown. Based on criterions like computational time, relative error, and complexity of model the improved complexity model (fig.1) was selected as best solution for future purposes which is modeling of supercapacitor modules (multi-cell battery system).

3. DEVELOPMENT OF 8xEDLC MODULE´S GEOMETRY

The proposed physical module for practical applications is shown on fig.2. This drawing was made wit the use of

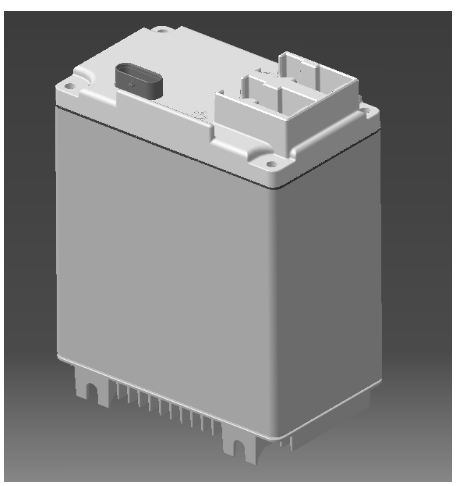


Fig.2. CATIA drawing of physical module

CATIA V5R16 software suited for design and innovations of various devices. Main parts of 8xEDLC module are:

- bottom aluminum cooler
- case of 8xEDLC module (made of PVC or Aluminum)
- top cover (made of PVC)
- gaskets (located inside between capacitors and cooler)
- 8xEDLC module (located inside of case)

Such geometry with all sub-parts is acting as very complex system for thermal simulations based on CFD (Computational Fluid Dynamic) software. Complexity in this case means large geometry, and consequently large computation net with lots of degrees of freedom, what results in our case into unacceptable computation time. Due to fact that for our development purposes we select COMSOL Multiphysics 3.5a software, what is based on CFD, we had to think about optimization task for its possible utilization in selected software. That means, we started to optimize its geometry in order to meet satisfactory results according computation/simulation time.

During development of COMSOL´s geometrical model we were solving several problems. At first we were deciding whether import CATIA model into COMSOL environment or develop a new one with the use of COMSOL drawing tools. The second approach is basically too much time consuming, and second disadvantage is low flexibility of this tool. On the other side, import of created geometry from CAD drawing programs is much comfortable, but also some problems can arise. Generally they can be caused due to fact that during import command a small faces and surfaces of investigated geometry can be damaged, what in final could result in disability of the generation of the computational mesh. The other problem is related with interstitial areas, which exists on the contacts between several components. Also problems could occur when large number of small faces and fillets exists.

Finally, after various attempts and processes, we decided to import CAD drawing and consequently to optimize it with COMSOL drawing and geometry optimization tools (defeaturing tool). In such way, we were able to remove and delete non-important small faces and fillets, thus simplify generation of the computation mesh. The final geometrical model, which is suitable for simulation purposes is shown on fig.3 and consists of these components:

- bottom aluminum cooler (identical to physical sample with aluminum material properties)
- case of 8xEDLC module (identical to physical sample with PVC or aluminum material properties)
- top cover (modified, modification was done due to fact that top cover from physical sample won't have impact on temperature distribution in the module, we have replace it by simple box geometry)
- gaskets (identical to physical sample, with silicon material properties, one is placed between aluminum cooler and case and second between case and top cover)
- 8xEDLC module (located inside of case, the improved complexity model from [1] is utilized)

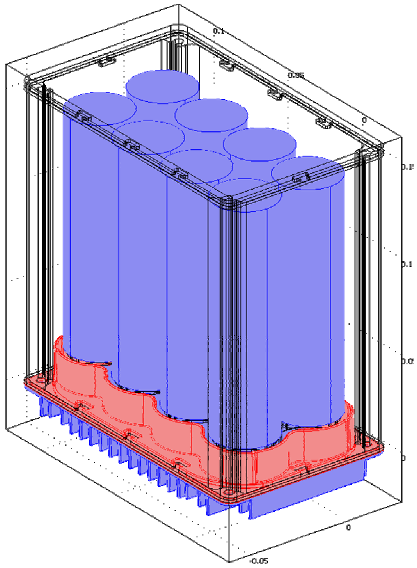


Fig.3. Geometry of thermal simulation model of proposed 8xEDLC module

With this designed geometrical model we processed to design of thermal simulation model, with the use of already mentioned COMSOL v3.5a software.

3. DEFINITION OF PROPERTIES FOR THERMAL SIMULATIONS

Before the process of thermal simulation experimenting can be started, it is necessary to define, what are the targets and aims of such experimenting. As was previously mentioned, the practical use of module is variable, and our proposed composition should be suited as auxiliary energy storage component in automotive application. When to talk about operating conditions, one part (the aluminum cooler) of module should operate at different surrounding then the other one (see fig. 4)

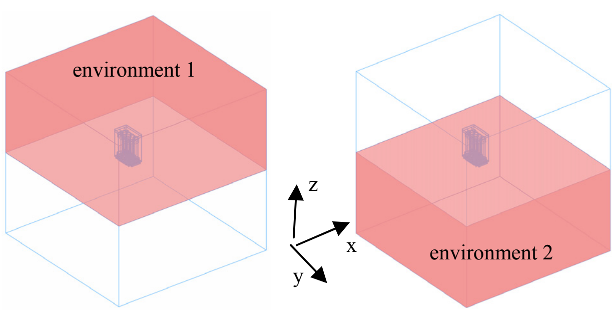


Fig.4. Environments with different simulation properties

Environment 1 acts as internal area, what can be represented as space, in which motor engine is located. That means this surroundings will have temperatures close to operating conditions of gas engine (around 90 ºC). Environment 2 is acting as space, where forced air - flow occurs, so space outside the car engine. Temperatures conditions in this can be related to weather conditions (from -20 ºC up to 40 ºC)

Boundary conditions

COMSOL Heat module supports two boundary conditions:

- Dirichlet condition (constant temperature)

$$T = T_0 \text{ on } \partial\Omega \qquad (1)$$

- Neumann condition heat flux

$$-nq = q_0 + q_r + q_s + h(T_{\text{inf}} - T) \text{ on } \partial\Omega \qquad (2)$$

During experimenting we had used Neumann condition of heat flux.

Heat transfer coefficients

These well-known coefficients have to be utilized during definition of heat processes in COMSOL software. These are:

The Nusselt number: Nu*L*(Re, Pr, Ra) = *h*.*L* / *k*
The Reynolds number: Re*L* = ρ.*U*.*L* / η
The Prandtl number: Pr = η.*Cp* / *k*
The Rayleigh number: Ra = Δ*T* *L*3 ρ.2.*g* β.*Cp* / (η.*k*)

, where

h is the heat transfer coefficient (W/(m2·K)).
L is the characteristic length (m).
Δ*T* is the temperature difference between surface and cooling fluid bulk (K).
g is the gravitational constant (m/s2).
k is the thermal conductivity of the fluid (W/(m·K)).
ρ is the fluid density (kg/m3).
U is the bulk velocity (m/s).
η is the viscosity (Pa·s).
Cp equals the heat capacity of the fluid (J/(kg·K)).
β is the thermal expansivity (1/K)

Nature of the Flow - the Grashof Number

Gr characterizes the flow. It describes the ratio of the internal driving force (buoyancy force) to a viscous force acting on the fluid. Similarly to the Reynolds number it requires the definition of a length scale, the fluid's physical properties, and the temperature scale (temperature difference). The Grashof number is defined as:

$$Gr_L = \frac{g\beta(T_s - T_0)}{\left(\frac{n}{\rho}\right)^2} L^3 \qquad (3)$$

, where:

g is the constant of gravitational acceleration,
β is the fluid's expansion coefficient,
Ts denotes the temperature of the hot surface,
T0 equals the temperature of the surrounding air,
L is the length scale,
η represents the fluid's dynamic viscosity,
ρ is the density.

For the laminar flow with natural external influence, next parameter have to be correctly defined:

$$h_{ave} = 0{,}25 F_{lam} \left(\frac{\Delta T}{L} \right)^{0,25} \tag{4}$$

, what is valid for horizontal location of surface and for bottom side of object.

$$h_{ave} = 0{,}54 F_{lam} \left(\frac{\Delta T}{L} \right)^{0,25} \tag{5}$$

, what is valid for vertical location of surface and for bottom side of object.

For the laminar flow with forced external influence, the previous parameter can be computed using next equation:

$$h_{ave} = \left(\frac{k}{L} \right) \frac{0{,}928 \Pr^{0,33} \mathrm{Re}^{0,5}}{\left(1 + \left(\frac{0{,}0207}{\Pr} \right)^{0,67} \right)^{0,25}} \tag{6}$$

4 SIMULATION EXPERIMENTS

In this section we describe procedure of simulation experiments and consequent evaluation of results. As was previously mentioned, the model contains several surroundings, whereby they can be divided into three sections:

1) surrounding inside of module's box (this means air inside of module, where EDLC cells are located
2) surrounding around module's box (for this subdomain we define just temperature of air)
3) surrounding around aluminum cooler (for this subdomain we define temperature of air and speed of air-flow)

Aim of this experimenting was investigation of operation conditions during which module can still be reliably working. Most sensitive parameter according stable operation of supercapacitor is its temperature of core (electrolyte). Proposed solution of supercapacitor is designed to withstand temperatures of max. 80 ºC. The most important settings are written down in next table:

Table. 1 Setting of variables for different environments

	Ta [ºC]	v [m/s]
environment 1	variable 25 ºC, 45 ºC, 80 ºC	Constant 0 m/s
environment 2	variable 25 ºC, 45 ºC	variable in x-axis 0, 0.1, 1, 5

, where *Ta* is ambient temperature of given environment and *v* is speed of air-flow in give environment and given direction.

During initial simulation experiments we found out that designed computation net (mesh) of thermal simulation model consists of too many mesh elements, what was mainly caused due to high number of small faces and fillets. This problem was solved during development of geometrical model (see chap.3). After optimization we get setting which seems to be optimal and are listed in next table.

Table. 2 Mesh statistics

	1219553
Number of mesh points	153381
Number of elements	911721
Tetrahedral	911721
Prism	0
Hexahedral	0
Number of boundary elements	166992
Triangular	166992
Quadrilateral	0
Number of edge elements	18160
Number of vertex elements	2008
Minimum element quality	0
Element volume ratio	0

With such improvement of number of mesh element (the initial version was consisting o 2135258 elements), what is almost 50% reduction we have decreased simulation time of about 25%. But still computation time was too long for fast purpose verifications. Therefore the second step of optimization process was necessary, i.e. selection of most proper solver, and its settings.

COMSOL Multiphysics includes a number of different solvers for PDE (partially differential equation) based problems. After various attempts we decided to use Direct PARDISO solver which is characterized as highly efficient direct solver for symmetric and nonsymmetrical systems, whereby it utilizes much less memory than other types [7].

Thus almost with all these improvements we targeted requested computation time which was in the range of 1054 seconds up to 3150 seconds (improvement of about amazing 46 times less, initial computation time = 146000 seconds). Even such a reduction of computation time was achieved the differences between initial results compared to optimized were unnoticeable. Next table is showing simulation results for different simulation settings.

Table. 3 Simulation results for various parameter settings

box type	Ta_1 [˚C]	Ta_2 [˚C]	0 m/s	0.1 m/s	1 m/s	5 m/s
Al	25	25	**54,297**	**40,922**	**31,127**	**29,139**
PVC	25	25	**60,257**	**41,754**	**31,086**	**29,127**
Al	25	45	**56,975**	**49,796**	**42,889**	**41,256**
PVC	25	45	**63,719**	**51,555**	**42,937**	**41,134**
Al	45	25	**70,638**	**51,842**	**39,324**	**37,106**
PVC	45	25	**75,003**	**51,929**	**39,333**	**37,258**
Al	45	45	**73,498**	**60,984**	**51,217**	**49,166**
PVC	45	45	**78,499**	**61,735**	**51,168**	**49,149**
Al	80	45	101,92	80,033	**65,871**	**63,428**
PVC	80	45	104,63	80,091	**65,974**	**63,812**

Evaluation of results

The simulation results are listed in table 3, where you can find settings of most important parameters. These were, material properties of EDLC box (aluminum, or PVC), temperature of environment1, temperature of environment2, and speed of airflow in environment 2 in the x-axis.

As was mentioned the critical operation state of module is when temperature of core of supercapacitor cell rises above 80 ºC. From table 3 can be seen, that this situation occurs only in four cases of various settings combination. It can be also said that aluminum type of EDLC box has influence on the temperature distribution in the core of EDLC cell, when no airflow is applied in the bottom side of module (the part of aluminum cooler). Temperature is lower for aluminum type against PVC type of box, but when airflow speed exceeds value of 0.1 m/s the temperature distribution is almost the same, thus temperature of EDLC core is dependent just on the value of air flow.

These simulation results can not be compared with real experimental measurements, because such measurement is very hard to be provided. But according to investigation, which was provided in [1] it can be said that thermal simulation model of EDLC module has to show very high credibility.

5. CONCLUSION

In this paper the formal procedure of thermal simulation model of EDLC module was described. Aim of modeling was achievement of acceptable computation time, together with achievement of high accuracy of interpreted results. Main purpose for such modeling is investigation of allowable operation states of proposed module, due to fact that experimental investigation is often too much time consuming, and also operating condition with various setting are hard to be provided, whereby mostly it is unreal. The critical condition of supercapacitor cell is 80 degrees, whereby through the use of proposed modeling methodology, we were able to design thermal simulation model, by which it was possible to investigate critical and allowable operation conditions. Generally it can be said that proposed procedure with final thermal simulation model represent innovative form for designers of perspective energy storage solutions.

ACKNOWLEDGMENTS

The authors wish to thank to Slovak grant agency APVV for project no. APVV-0138-10 - Research and development of the small power drives with two-phase motors, and for project no. LPP-0366-09 - High frequency power electronic converter. Special thanks is also given to Slovak grant agency VEGA for project no. 1/0943/11.

REFERENCES

[1] Frivaldsky, M., Spanik, P., Čuntala, J., Glapa, N.: **Development and comparisons of the thermal simulation models with various complexity of the electrolytic capacitor for functional verification purposes of super capacitor in different conditions of use**, In: The 2nd International multiconference on complexity , informatics and cybernetics, 27-30 march, 2011, Orlando, FL-USA, pp-283-288, ISBN-13:978-1-936338-20-7

[2] Čuntala, J., Frívaldský M.: 3D **Simulation of thermal field in the core of supercapacitor**, 2009 Technical computing , Prague 2009

[3] Špánik, P., Čuntala, J., Frivaldský, M., Glapa, N., Madleňák,D.: **Thermal simulation of electrochemical double layer capacitor**, ELEKTRO 2010, Proceedings of the 8th international conference, Žilina, Slovak Republic, May 24-25, 2010, ISBN 978-80-554-0196-6. - S. TA2_ 62-66

[4] Sam G. Parler, Jr., Laird L. Macomber: **Predicting Operating Temperature and Expected Lifetime of Aluminum-Electrolytic Bus Capacitors with Thermal Modeling**, Powersystems World, November 1999

[5] Hargas, L., Koniar, D., Stofan, S. **Sophisticated biomedical tissue measurement using image analysis and virtual instrumentation**, LabVIEW - practical applications and solutions, INTECH, ISBN 978-953-307-650-8, 2011 pp.155 - 181

[6] Tiefenbacher, M., Jakubek, S., Kozek, M.: **Modeling and identification of a chain pendulum**, In: The 2nd International multiconference on complexity , informatics and cybernetics, 27-30 march, 2011, Orlando, FL-USA, pp-306-312, ISBN-13:978-1-936338-20-7

[7] COMSOL: **Multiphysics user guide**

[8] Sitar, J., Racek, V., Havrila, R., Gerek, I.: **Input Active Filter for the DC/DC Converter Supplied from the Traction Troley Net**, 11th European Conference on Power Electronics and Aplycations, EPE 2005, 11-14.9.2005, Dresden, ISBN: 90-75815-08-5, IEEE Catalog Number: 05EX1132C

[9] Havrila, R.; Lokseninec, I.: **Soft-Switching High Voltage Converter with Wide Input Voltage Range, Power Factor Correction and Galvanic Isolation**, EPE-PEMC 2004, Riga, Latvia, September 2-4, 2004

A Multi-purpose Design Method for Flexible Mixed-product Lines

Aritoshi Kimura
Department of Information and Systems Engineering
Faculty of Engineering, Kinki University
Higashihiroshima, Hiroshima, Japan

ABSTRACT

If requirements of the production change, the product line should be adjusted to the suitable one. There are many types of optimization method, such as addition or reduction of stations in the production line in succession to the structure of the existing line. In this paper, we propose a multi-purpose design method for flexible mixed-product lines. We have proposed the Fissiparous Algorithm and already shown the effectiveness by designing flexible mixed-product lines. We developed the Fissiparous Algorithm by new operations and used it to design flexible mixed-product lines with various product conditions. The validity of this design method using the Fissiparous Algorithm is shown through solving several design problems of flexible mixed-product lines.

Keywords: Fissiparous Algorithm, Design method, Flexible Manufacturing System, Flexible Mixed-product Line.

1. INTORODUCTION

Companies are required to make their products in the high quality, lower cost and shorter lead time because of the recent severe economic competition. Some flexible product systems are proposed to adjust to such economics. Parallel work stations are added to adapt to fluctuating demand in flexible assembly systems [1]. Product lines which can complete various production plans without extra parallel work stations have product-mix flexibility [2]. We proposed design methods for these production lines [3]-[5]. Using these methods, we can solve larger scale problems. But, if search trees become widely expanded beyond the boundary, it takes too much time to solve and losses their usefulness.

We presented a method for designing which involves algorithms by using new approaches. This algorithm belongs to evolutionary algorithms and imitates fission rules. We named it the Fissiparous Algorithm. Chromosomes in conventional evolutionary algorithms are changed by probabilistic rules. But in our proposed algorithm, unlike these conventional algorithms, evolutions of chromosomes are happen by artificial rules. Our proposed Fissiparous Algorithm is able to apply to various kinds of problems, such as facility layout problems and seat reservation problems and so on, according to change these rules. The validity of the design method using our Fissiparous Algorithm was shown through solving large scale design problems [6], [7].

This paper presents a method for designing flexible mixed-product lines adjusting to various product conditions. We improve the Fissiparous Algorithm by introducing new operation methods of this algorithm. We apply the Fissiparous Algorithm to design flexible mixed-product lines. To examine the performance of our proposed method, we solve some hypothetical problems.

2. THE MODEL OF FLEXIBLE MIXED-PRODUCT LINES

The flexible mixed-product line that we consider here has the following features.

1. Different product models are assembled on the same line.
2. Different production plans are executed on the same line.
3. Quantities of each product assembled in every production plan within a time period are decided.
4. The precedence diagram and combined precedence diagram of each type of product are defined.

The assumptions and design problem of this model are as follows.

2.1 Assumptions

The assumptions of the model are as follows.

1. Workers can complete several kinds of tasks.

2. Workers can use several types of tools.
3. Combinations of a worker and tools which can complete each task are defined.
4. The number of workers per type is defined.
5. The number of tools per type is defined.
6. The number of tools which each worker can use is defined.
7. Time required to complete each task differs by the combination of workers and tools.
8. Types of product are defined.
9. Workers change their tools, if it is necessary. The exchange time is included to complete each task.
10. The upper limit of time to complete all tasks in a station (ULT) is defined. The ULT assigned to stations are same.

2.2 The Design Problem of Flexible Mixed-product Line

The design problem is defined as follows.

Object: Design a line which has a minimum number of stations.

Decision: Decide combinations of tasks and workers at each station.

Constraints:

1. Tasks completed at each station are decided to satisfy precedence relations.
2. One worker is assigned to each station.
3. The sum of workers assigned to each station is less than the number defined in assumption 4.
4. The sum of tools assigned to each station is less than the number defined in assumption 5.
5. The number of tools which are used at each station is less than the number defined in assumption 6.
6. Time to complete all tasks assigned to the *s*th station in the *j*th production plan (ST_{sj}) satisfies the following expression.

$$ST_{sj} = \sum_{i=1}^{I} \sum_{k \in W_s} t_{ks} \delta_{ik} N_{ij} \le ULT$$

$$, \quad s = 1, 2, \cdots, S \quad , \quad j = 1, 2, \cdots, J \qquad \text{Eq. (1)}$$

I: The number of production models.

S: The number of stations.

J: The number of production plans.

W_s : The identification number of tasks completed at the *s*th station.

t_{ks} : The time to complete the *k*th task at the *s*th station.

δ_{ik} : Binary integer, defined as 1 if the *k*th task is needed to make the *i*th model, otherwise 0.

N_{ij} : The quantity of the *i*th model in the *j*th production plan.

3. THE FISSIPAROUS ALGORITHM

The design problem of flexible mixed-product line has widely expanded search trees. Some algorithms are known to solve these large scale problems, but they usually take a long time to output optimal or suboptimal solutions and do not have practical usability. To solve these problems, we propose a new algorithm, the Fissiparous Algorithm (FA). The FA is one of the solution methods based on the idea of genetic engineering and also one of the Evolutionary Algorithm. The FA has the following features:

1. Recombination is not done randomly but artificially.
2. Recombinant operation is done by method imitated asexual reproduction.
3. Crossover is not used.
4. Mutation is not done.
5. The selection is done based on the fitness.

3.1 Encoding of Chromosome

Chromosomes are composed of segments consisting of several loci. Functions or features that chromosomes have are represented by these segments. The number of loci in one segment is not always the same. All of the loci in one segment are not always occupied by genes. The consequences of genes in each segment decide how functions or features work. For example, if a segment does not have any gene in its loci, the function or the feature that is represented by this segment does not work in this chromosome. We basically express genes by numbers and encode chromosomes by a sequence of numbers. Identification numbers of loci are counted from the left locus to the right one. A sample is shown in figure 1.

Chromosomes:

$$\mathrm{CHRO} = \{\mathrm{SEG}_1, \mathrm{SEG}_2, \cdots, \mathrm{SEG}_i, \cdots\}$$

$$, \quad i = 1, 2, 3, \cdots \qquad \text{Eq. (2)}$$

Segments:

$$\mathrm{SEG}_i = \{\mathrm{LOC}_j, \mathrm{LOC}_{j+1}, \cdots\}$$

$$, \text{if } i = 1 \text{ then } j = 1 \text{ else } j = \sum_{l=1}^{i-1} n(\mathrm{SEG}_l) + 1 \quad \text{Eq. (3)}$$

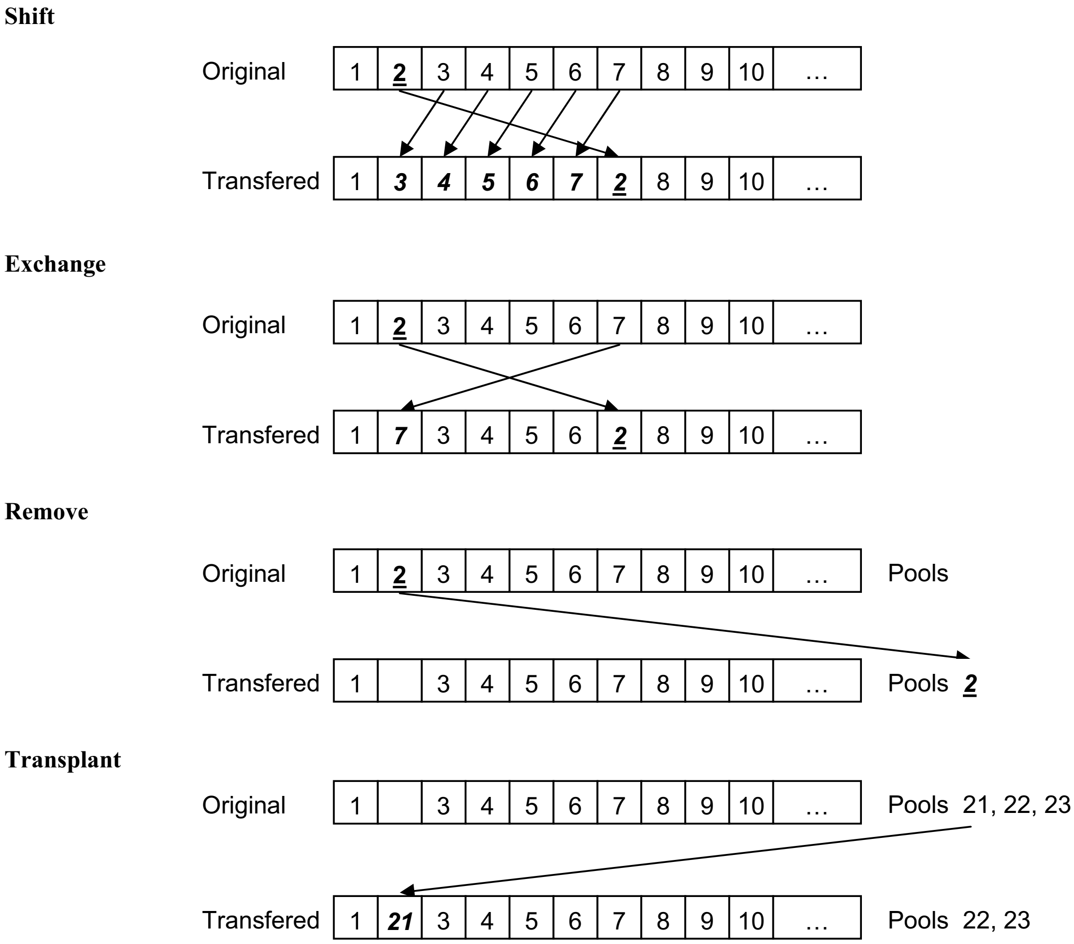
Locus:

$$\mathrm{LOC}_k = \{\mathrm{gene}, \phi\}, \quad k = 1, 2, 3, \cdots \qquad \text{Eq. (4)}$$

Figure1. A Sample Encoding of Chromosome in the FA

Exist a gene(G) or not(ϕ)	G	G	G	G	G	ϕ	G	G	ϕ	ϕ	G	G	…
The number of Locus	1	2	3	4	5	6	7	8	9	10	11	12	…
The number of segment	1			2			3				4		…

Figure2. Samples of Basic Genetic Manipulations in the FA



3.2 Genetic Manipulation

We do not use crossover and mutation operators used in the usual genetic algorithm because the FA is one of the solution methods based on the idea of genetic engineering. Recombination processes are operated by the following steps.

Step1: Nominate loci as Targets

We select some loci as Targets. Genes on each Target are moved to another locus, if possible. Both the number of Targets (U) and the method of selection are decided by a heuristic method in advance. These are able to differ in each problem. We consider that the number of selected loci needs to be large enough so to not be stuck in suboptimal solutions. The notations are as follows.

U : The number of Targets

T_u : The uth Target $u = 1, 2, \cdots, U$

Figure3. A Sample Encoding of Chromosome in This Method

The number of Task	1	6	3	2	4	8	7	5	11	9	10	15	…
The number of Locus	1	2	3	4	5	6	7	8	9	10	11	12	…
The number of station	1			2			3				4		…

Step2: Transfer genes

We change the order of genes by the evolutionary algorithm imitated fissiparous operations.

Step2.1: Copy chromosomes

We copy chromosomes and operate steps 2.2 and 2.3 on only these copies. The notations are as follows.

C(u) : The population of chromosomes operated in step2.

Step2.2: Find loci that genes on Targets can transfer

We find loci, where genes on Targets can move, by some rules that were decided in advance. If several loci are found, we select one locus among them by some rules that were decided in advance. If no locus is found, we skip step 2.3 and accept this copy that has the same order of the original chromosome into C(u). We should make these rules adjusted to each problem in advance.

Step2.3: Transfer genes on Targets

In the FA, we have four types of basic genetic manipulations shown in figure 2. Chromosomes transferred genes are accepted into C(u).

Step3: Selection

We select some chromosomes based on the fitness of each problem. We have to decide the measure of fitness and the number of chromosomes which selected in step 3 in advance.

3.3 Encoding of Chromosome for the Design of Flexible Mixed-product Lines

We represented a gene by a task and encode chromosomes by the number of task. We assign each station to each segment. Stations are arranged from the first station to the last station in a chromosome. Tasks are arranged to all loci in each station in which they are completed. A sample is shown in figure 3.

4. THE REDESIGN PROBLEM OF FLEXIBLE MIXED-PRODUCT LINES

This paper considers the redesign problem as the design of flexible mixed-product lines adjusted to the change of the production ratio after that lines have worked. We made some changes in our current design method by FA. The assumptions and constrains are as follows.

4.1 Assumptions

In the redesign procedure, the assumptions are added to assumptions defined in 2.1.

1. Some workers might be fixed.
2. Some tools might be fixed.
3. Some tasks might be fixed.
4. Some stations might be prohibited to use.

4.2 Constraints

In the redesign procedure, constrains are added to constrains defined in 2.2.

1. The number of stations, including stations prohibited to use, in the existing line is fixed.
2. The arrangement of stations does not change.
3. Do not add new workers and tools.
4. Do not move fixed workers, tools and tasks.
5. Do not assign tasks to stations prohibited to use.

5. NUMERICAL EXAMPLE

We solved several types of hypothetical problems to confirm the effectiveness of the proposed multi-purpose design method. We show numerical examples. Parameters of hypothetical problems are defined as follows.

The number of tasks : 520

The number of production models : 5

The number of production plans : 10

The number of workers : 10 (1 per type)

The number of tool types : 20

The number of stations in the existing production line : 10

[Problem 1 : Redesign the flexible mixed-product line to reduce stations]

The number of stations to reduce : 1

Result (the number of stations in the new production line) : 9

[Problem 2 : Redesign the flexible mixed-product line to extend stations in succession to a part of the structure of the existing line.]

The number of stations to extend to the existing production line : 2

Places to be able to install extending stations (P*):

P11 S1 S2 P12 P13 S3 S4 P14 S5 S6 P15 S7 S8 P16 S9 P17 S10 P18

(S* means existing stations)

Result (the number of stations in the new production line) : 12

6. CONCLUSIONS

In this paper, we present a multi-purpose design method for flexible mixed-product lines. The effectiveness of our design method which uses the Fissiparous Algorithm was confirmed by solving hypothetical problems.

7. REFERENCES

[1] Narumol U. and Carlos F. D., "Impact of Parallel Processing on Job Sequences in Flexible Assembly Systems", **International Journal of Production Research**, Vol.27, 1989, pp.73-89.

[2] Raouf A. and Anjum M. F., **Manufacturing Systems: Flexibility Assessment, Flexible Manufacturing Systems: Recent Development**, Amsterdam: Elsevier, 1995, pp.69-84.

[3] Tenda S., Umezawa T., Okazaki K., "A Design Method for Robotic Mixed Product Lines", **Journal of Japan Industrial Management Association**, Vol.44, 1993, pp.217-228.

[4] Tenda S., "A Method for Designing Mixed Product Lines Constructed of Turret Robots", **Journal of Japan Industrial Management Association**, Vol.47, 1996, pp.91-99.

[5] Tenda S., "A Method for Designing Robotic Mixed Product Lines with the Rules for Changing the Position of Works", **Journal of Japan Industrial Management Association**, Vol.48, 1997, pp.25-32.

[6] Tenda S., Kimura A. and Tsuda E., " A method for designing flexible mixed-product lines using a fissiparous algorithm", **Journal of Japan Industrial Management Association**, Vol.53, 2003, pp.456-465.

[7] Tenda S., Kimura A. and Tsuda E., " Design method-II for flexible mixed-product lines using a fissiparous algorithm", **Journal of Japan Industrial Management Association**, Vol.56, 2005, pp.182-190.

8. ACKNOWLEDGMENTS

I had received advices from Mr. Saburo Tenda, a former Professor at Kinki University in Japan. I gratefully appreciate his help for improving this paper.

Vertical Axis Micro Wind Turbine Design for Low Tip Speed Ratios

Ryan McGowan, Rafael Lozano, and Vrishank Raghav
The Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology
Atlanta, GA 30332, United States

Dr. Narayanan Komerath
The Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology
Atlanta, GA 30332, United States

ABSTRACT

A small-scale vertical axis wind turbine (VAWT) is developed for use in areas lacking adequate energy infrastructure; the materials and methods of construction are selected to minimize cost as much as possible. In order to overcome the self-starting issues associated with VAWTs and low tip speed ratios, solutions such as high solidity, guide vanes, and drag tubes are considered. A computer simulation is written to predict the performance of turbines incorporating these features, and the trends are compared to those from the literature. The mechanics and feasibility of drive systems using motor-generators and bicycles are also examined.

Keywords: Design, Model, Multiple Streamtube Theory, Simulation, VAWT, and Vertical Axis Wind Turbine.

NOMENCLATURE

A_t	swept area
c	chord
c_d	section drag coefficient
c_l	section lift coefficient
C_P	power coefficient
C_t	tangential coefficient
h	height
N	number of teeth
N_b	number of blades
P	power
r	radius
Re	Reynolds number
T	torque
U_t	tangential velocity
U_∞ or V	freestream velocity
α_b	blade angle of attack
β_b	blade inclination from horizontal
θ	angular position
θ_b	blade azimuth angle
λ	tip speed ratio
ρ	density
σ	solidity
ω	angular velocity

1 INTRODUCTION

In recent years there has been a large push toward clean sources of energy; this in conjunction with the advent of the "smart grid" has created an incentive to develop renewable energy technologies scaled for single households. However, these concepts have also found favor in developing nations and areas otherwise detached from an electric grid due to their small scale and autonomy. In particular, vertical axis wind turbines present one promising option and have a number of advantages over the traditional horizontal axis wind turbine. For example, they operate the same regardless of wind direction, and they can be located closer to the ground for easier maintenance. Furthermore, the blades do not require any twist and thus are simpler to manufacture. On the other hand, VAWTs tend to have "dead zones" where the orientation of the blades causes cancellation of forces in the tangential direction, meaning there is no net torque. These are especially an issue when a turbine operates at low tip speed ratios, since it may not be spinning fast enough to escape them. The tip speed ratio λ is the blade tangential velocity divided by the freestream velocity, as shown in Eq. (1):

$$\lambda = \frac{U_t}{U_\infty} = \frac{r\omega}{U_\infty} \quad (1)$$

To prevent the turbine from becoming stuck in these dead zones, it is advantageous to increase the overall torque. Then, if the turbine has sufficient momentum, it will be able to spin through the short bands where the net torque is zero [4]. Guide vanes can be used to change the angle of attack seen by a blade at strategic points along its revolution. By modifying the angle of attack, the tangential force coefficient and consequently net torque can again be increased. In addition, drag tubes may be helpful in starting the turbine when wind speeds are low: Since they are capable of producing a high starting torque, they can accelerate the turbine until the blades take over in providing most of the torque.

Kirke [4] notes that "[s]elf-starting problems can be overcome electrically if the load is a synchronous alternator or a DC generator which can function as a motor and can be used to drive the turbine up to operating speed," although this is more common in grid-connected wind turbines than micro VAWTs. Nonetheless, motor-generators remain a viable option for tackling the self-starting issue and can be replaced with other forms of energy addition, such as a pedal-powered system.

2 CONSTRUCTION

The Georgia Tech VAWT is a scaled-up version of models built by the Experimental Aerodynamics Group and has a diameter of 6 ft. Like previous iterations, it has three arms, each with a pair of straight blades. Those parts which rotate are constructed from aluminum in order to reduce loading on the arms, but the static frame has been built out of steel so as to keep costs low.

The design and construction of the blades are also integral to building a vertical axis wind turbine. The blades should be light and easy to make yet strong enough to withstand high winds. In order to meet these criteria, the blades of the Georgia Tech VAWT have a fiberglass skin and spars as well as foam ribs due to these materials' strength and low weight. Epoxy resin was used to bond the skin, spars, and ribs to each other as is common in the manufacture of composite blades. In choosing an airfoil, a NACA 0015 was decided upon due to its symmetry and widespread use.

3 SIMULATION

To assist in the initial sizing of a vertical axis wind turbine, a simulation based on Strickland's multiple streamtube theory [6] was written. In multiple streamtube theory, the turbine is split into several streamtubes along the horizontal and vertical axes. The streamwise forces on the blade elements are equated to the changes in fluid momentum, and an induction factor is found for each streamtube which describes how the blades slow the air moving through the turbine. A local power coefficient is then computed based on Eq. (2) for each streamtube; these are averaged to determine the overall power coefficient of the turbine. The numerator represents the power produced by the turbine, whereas the denominator represents the total wind power available to the turbine.

$$C_P = \frac{P}{\frac{1}{2}\rho A_t U_\infty^3} \qquad (2)$$

In addition to simply providing the user with a power curve for a given set of dimensions, the simulation gives insight into how changing these parameters can affect turbine efficiency. Though much of the code is adapted from Strickland's DARrieus Turbine (DART) model [6], it is also capable of incorporating the effects of straight blades, variations in Reynolds number, and drag tubes.

Figure 1 shows that for a turbine with curved NACA 0012 blades and a solidity of 0.27, the simulation is in good agreement with Sandia's performance predictions [6]. Similarly, Figure 2 overlays the Georgia Tech performance predictions onto Sandia's performance predictions for a turbine with a solidity of 0.18. The simulation uses the same airfoil data as Sandia's code (that from Jacobs and Sherman [3]); nonetheless, it slightly overpredicts the power coefficient at low tip speed ratios and underpredicts it at higher tip speed ratios. When Reynolds number effects are taken into account, the power coefficient drops as expected: The actual Reynolds number values are for the most part lower than that assumed for the blue curve, so aerodynamic stall would occur earlier [6]. However, the simulation also tends to underpredict the power coefficient when incorporating Reynolds number effects.

The gaps in the green curve result from a lack of available airfoil data at Reynolds numbers below 10,000. These lift and drag coefficients were obtained from Sheldahl and Klimas [5] and at present appear to be the most comprehensive compilation of airfoil data ranging over different Reynolds numbers. However, as not all of this data was obtained experimentally—coefficients at Reynolds numbers below 360,000 were

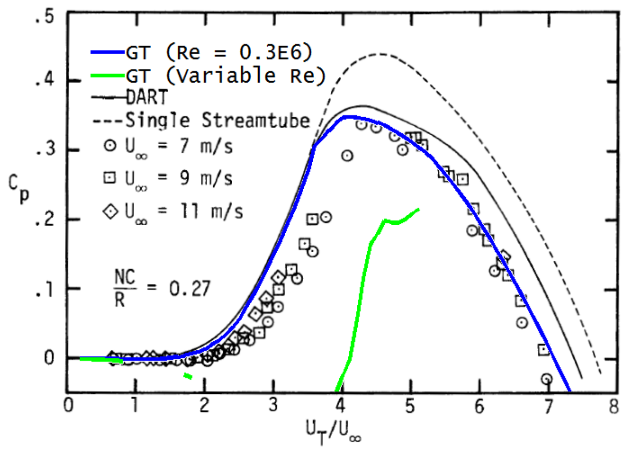


Figure 1: Comparison of Georgia Tech and Sandia performance predictions at a solidity of 0.27.

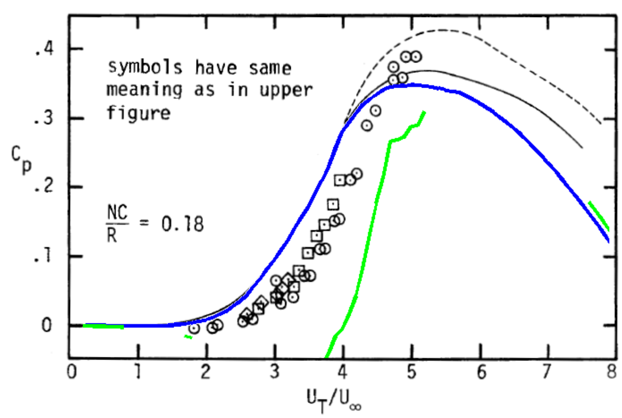


Figure 2: Comparison of Georgia Tech and Sandia performance predictions at a solidity of 0.18.

extrapolated by Sheldahl and Klimas using their own computer program—there are several errors which call into question the validity of the data. In particular, some sets of data for lower Reynolds numbers give a negative lift coefficient at the dip after stall. In any case, Kirke [4] states the "discrepancies between [Jacob's and Sheldahl's] data for the same section under supposedly similar conditions show that it is dangerous to look for trends by comparing the performance of different aerofoil sections at low Re using data from different sources."

Solidity Effects

One extremely important parameter for self-starting is the solidity σ, a measure of the volume being swept by the rotor. This value is calculated as shown in Eq. (3):

$$\sigma = \frac{N_b c}{r} \tag{3}$$

By adjusting the number of blades, chord length, and radius in the graphical user interface, one can see that increasing the solidity pushes the maximum power coefficient to lower tip speed ratios (Figure 3). This is desirable in the case of VAWTs, which often operate at angular velocities lower than those of horizontal axis wind turbines. On the other hand, the plot suggests that increasing solidity too much not only causes the maximum power coefficient to drop but also causes the power curve to have a sharper peak, meaning the turbine efficiency is more sensitive to changes in wind speed. Nonetheless, Templin [7] states, "Although lower values of solidity may widen the useful operating range of the turbine, in terms of the velocity ratio $R\omega/V$, they also reduce maximum available power. Blade centrifugal stresses, will also tend to increase at low solidity, because of the high rotational speeds."

Figure 3 applies to straight-bladed VAWTs using NACA 0015 airfoils, similar to the one built by Georgia Tech. Strickland's DART code has been modified to allow for straight blades by changing the variable β_b in Eq. (4) to a constant $\frac{\pi}{2}$ rad. Here β_b is the angle of a blade segment with respect to the horizontal, h is the height of the blade segment, r is the local radius, and θ_b is the azimuth angle.

$$\beta_b = \tan^{-1} \frac{\Delta h}{r \Delta\theta_b \sin\theta_b} \tag{4}$$

Comparing Figures 1 and 2 to Figure 3, it is interesting to note that the power curves for the straight-bladed VAWT are generally sharper than those for a turbine with curved blades.

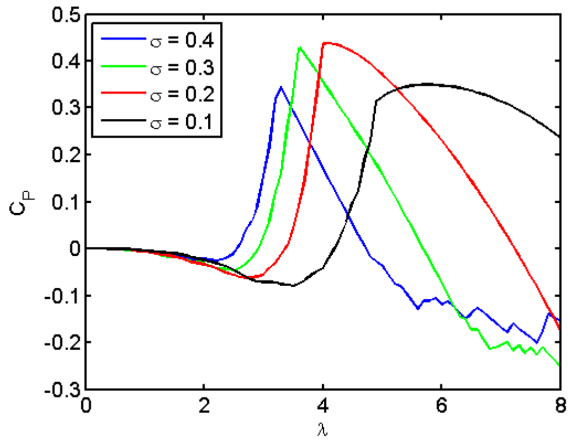


Figure 3: Power coefficient vs. tip speed ratio for various solidities.

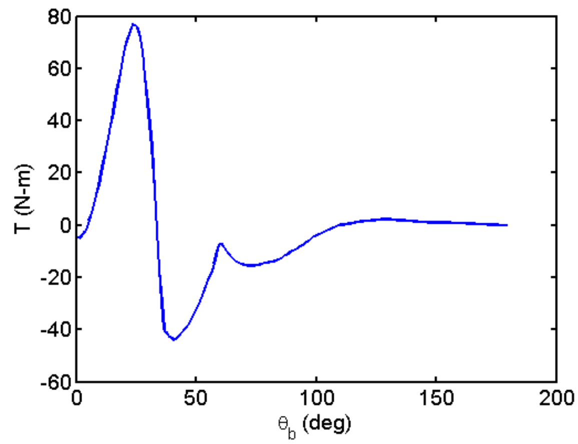


Figure 4: Torque vs. blade azimuth angle.

Guide Vanes

As was mentioned previously, guide vanes can also be used to address the problem of self-starting. Since the torque on a single blade is actually negative for a significant portion of each revolution (Figure 4), it would be beneficial to create positive torque through the intelligent placement of guide vanes. Indeed, not only does a positive torque mean the turbine will accelerate faster in the desired direction, but according to Islam, Ting, and Fartaj [2], if the net torque is also made to be positive for all tip speed ratios up to the operating speed, then this should be sufficient for self-starting.

The tangential coefficient is calculated in Eq. (5). In Figure 5, this value is plotted for a NACA 0015 airfoil at various angles of attack; here it can be seen that the tangential coefficient is maximized when the angle of

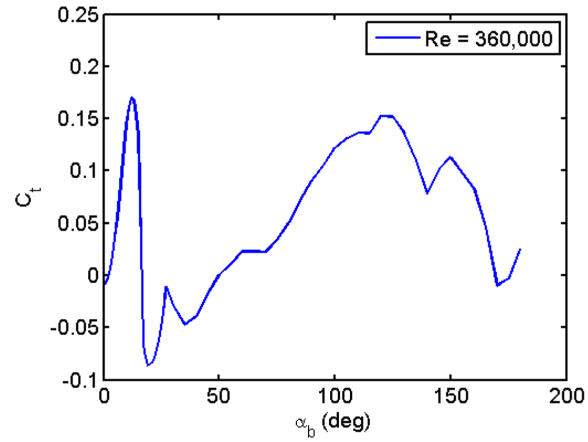


Figure 5: Tangential coefficient vs. blade angle of attack for a NACA 0015 airfoil.

attack is approximately 12°. Theoretically, the guide vanes can also be used to give the airfoils an angle of attack between approximately 100° and 160°—here the tangential coefficient is somewhat lower, but it stays relatively high for a wider range of values. However, the blades typically do not approach this range by themselves.

$$C_t = c_l \sin \alpha_b - c_d \cos \alpha_b \qquad (5)$$

Guide vanes will need to be placed to the sides of the turbine and at azimuth angles where their necessary angle of attack is minimized in order to reduce drag losses. Looking at Figure 4, the greatest benefit will be realized for guide vanes at azimuth angles between 33° and 180°, since torque is either negative or very small in this range.

Drag Tubes

Drag tubes (which operate on the same principle as Savonius rotors) are another potential method of improving VAWTs' self-starting capability and are most effective at low tip speed ratios. Since Savonius rotors are always self-starting if there are at least two scoops at different angles, it is tempting to think of drag tubes as the perfect solution for the self-starting problem. However, according to Kirke [4], a drag tube rotor is only adequate for this job if:

- Its radius is "compatible with that of the Darrieus, so that it not only starts the rotor but continues to provide torque until the Darrieus can take over, and does not then produce negative torque which would reduce the overall efficiency of the turbine in its optimum operating range."

- It is "big enough to produce 'enough' starting torque."

It is apparent that in order to determine whether an auxiliary Savonius rotor is appropriate for a particular VAWT, the effects of the drag tubes must be examined in detail. The simulation assumes that the drag tubes are relatively long with a semicircular cross section, although other shapes can be used.

Whereas the lift and drag coefficients of different airfoils are generally well-documented, it is difficult to find this data tabulated against angle of attack for other geometries. It is known from Hoerner [1] that the drag coefficient of the semicircular tube is 2.30 when it is facing the wind and 1.20 when it is turned away, but this same information is not readily available for other angles of attack. Since it is mainly the drag of the tubes that produces torque, lift effects were neglected. (In any case, there was no lift data to be found, presumably because it is rare to generate lift using shapes other than airfoils.)

Curve fitting was utilized to generate drag data for one full rotation. The curve was assumed to be a sine wave of the form in Eq. (6) since the drag coefficient is minimized when α_b is 0° and maximized when α_b is 180°; A and D are constants.

$$c_d = A \sin \alpha_b + D \tag{6}$$

4 MOTOR-GENERATOR DRIVE

If the turbine is being used for electricity generation, a motor-generator may be attached for the purpose of spinning it up to operating speed. A motor-generator with high current, high voltage, and low "cut in" rpm is desirable since the turbine should produce a useful amount of power despite operating at a relatively low rpm. The battery used should be of the deep cycle variety, deep cycle batteries being better suited than car batteries for repetitive charging and discharging. Since intense fluctuations in voltage due to variations in wind speed can damage the battery, a charge controller is also necessary to regulate the voltage used to charge the battery. On the other hand, the current should bypass the charge controller when the battery is driving the turbine, as the battery should output a constant voltage.

Since it is assumed the VAWT is too large and heavy to start using wind alone, it can be started by initiating a charge or "kick" from a PIC for several seconds

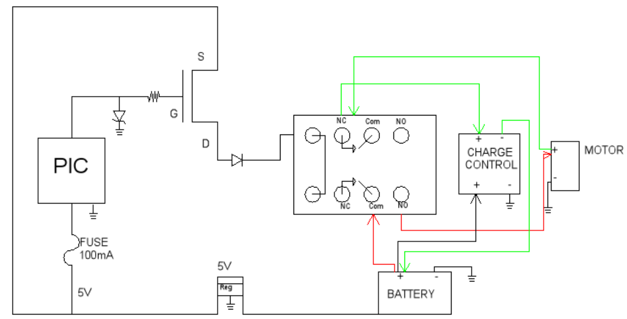


Figure 6: Circuit diagram.

which will run through a MOSFET and into an electrical relay. There, the coil in the relay will guide the charge through the "not closed" and "common" terminals to the battery and motor. This will charge the battery and supply the motor with initial energy to power the turbine. Then, the charge will go all the way around through the battery, through a regulator, and around the PIC back through a diode before reaching the relay once again. Here, the relay will read it in the "not open" and "common" terminal which will then charge the battery and disengage from the motor so that the wind turbine will now be powering the battery. It is important to note that the relay and the motor have to match in terms of maximum voltage and current, and the charge controller has to be hooked up properly. Figure 6 shows this arrangement more clearly.

5 BICYCLE DRIVE

Since one of the main goals of Georgia Tech's turbine is to be used in developing nations, a bicycle drive could provide a cheaper, human-powered solution to self-starting if the mechanism is kept as simple as possible. In conjunction with the simulation and other research on human factors, the goal is to determine how a human can, without excessive exertion, drive the turbine up to speed by pedaling on a bike held in place. Bicycle parts provide the advantage of being relatively cheap and easy to obtain, but challenges to be overcome include transferring power between two different axes (that of the bicycle's wheels and that of the VAWT), optimum gear ratios, whether or not gear shifting will be necessary (and if so, how to accomplish this), and how to disengage the bike from the turbine once it has spun up to speed.

Using only the gears already present on the bicycle frame, it is proven that the operator should be able to switch gears as needed to pedal the turbine up to the required tip speed ratio. From the gear relations given in Eq. (7) and Eq. (8), one can find the rpm that can be achieved at the rear tire where the assembly will be attached. Here θ is the angular deflection of the gear, r is the radius, N is the number of teeth, and T is the torque.

$$\frac{\theta_2}{\theta_1} = \frac{r_1}{r_2} = \frac{N_1}{N_2} \tag{7}$$

$$\frac{T_2}{T_1} = \frac{\theta_1}{\theta_2} = \frac{N_2}{N_1} \tag{8}$$

Assuming a gear ratio $\frac{N_{rear}}{N_{front}}$ of $\frac{14}{52}$ and an average cycling cadence of 60 rpm, the maximum achievable rpm at the rear will be 222.86 rpm. On the other end of the spectrum, a gear ratio of $\frac{28}{39}$ results in a rear angular velocity of 83.57 rpm.

6 CONCLUSION

It has been demonstrated that computer modeling can be a powerful tool for VAWT design, as it is cheaper than building multiple iterations of a turbine and makes it easier to visualize trends. The current state of simulation also shows the need for reliable airfoil data at low Reynolds numbers. A vertical axis wind turbine which is to operate for low tip speed ratios should have a relatively high solidity, but it should not be so high as to compromise its efficiency. It may also benefit from drag tubes and guide vanes, which can serve to increase the net torque or the torque of a single blade when it is in a dead band.

Finally, electrical or mechanical power provides an easy method of ensuring the turbine can start (though at the cost of autonomy). Either the circuit must sense when the turbine is not spinning at its operating speed, or someone must watch the turbine to see if it needs to be pedalled up to speed. Even so, both drive systems are capable of adding the necessary energy for a small-scale VAWT.

ACKNOWLEDGMENTS

This work was supported under the NASA EXTROVERT cross-disciplinary innovation resource development project. Mr. Tony Springer is the technical monitor.

REFERENCES

[1] Sighard F. Hoerner. *Fluid-Dynamic Drag*. Hoerner Fluid Dynamics, Bakersfield, 1965.

[2] Mazharul Islam, David S.-K. Ting, and Amir Fartaj. Desirable airfoil features for smaller-capacity straight-bladed vawt. *Wind Engineering*, 31(3):165–196, 2007.

[3] Eastman N. Jacobs and Albert Sherman. Airfoil section characteristics as affected by variations of the reynolds number. Technical Report 586, National Advisory Committee for Aeronautics, 1937.

[4] Brian Kinloch Kirke. *Evaluation of Self-Starting Vertical Axis Wind Turbines for Stand-Alone Applications*. PhD thesis, Griffith University, 1998.

[5] Robert E. Sheldahl and Paul C. Klimas. Aerodynamic characteristics of seven symmetrical airfoil sections through 180-degree angle of attack for use in aerodynamic analysis of vertical axis wind turbines. Technical Report SAND80-2114, Sandia National Laboratories, 1981.

[6] James H. Strickland. The darrieus turbine: A performance prediction model using multiple streamtubes. Technical Report SAND75-0431, Sandia National Laboratories, 1975.

[7] R. J. Templin. Aerodynamic performance theory for the nrc vertical-axis wind turbine. Technical Report LTR-LA-160, National Aeronautical Establishment, 1974.

Digital Prototyping - RoboCAP

Renata Maria Abrantes Baracho
School of Information Science
Federal University of Minas Gerais
Belo Horizonte, Brazil
renatabaracho@ufmg.br

Marcelo Franco Porto, Gustavo Arruda, Gefferson Silva
School of Engineering
Federal University of Minas Gerais
Belo Horizonte, Brazil
marceloporto@ufmg.br

Abstract **–**The research project "Digital Prototyping – RoboCAP" involves studying, modeling and preparing proposals for the robot development process in a mechatronic project, aiming at project development using two different types of computer software, in order to choose the software most appropriate for the development and simulation of mechatronic projects. The research project considers concepts and techniques of tridimensional simulation with digital prototyping to evaluate and optimize project levels before beginning real robot construction and assembly. It outlines a tridimensional simulation of the RoboCAP team's robot, created by the Robotics Team of the Alto Paraopeba Campus, in the Federal University of São João del-Rei, located in the city of Ouro Branco, Minas Gerais, Brazil. The robotics team is divided into groups with different development phases of the RoboCAP assembly - a robot built and designed for robot competitions. During research, simulations, studies and calculations were performed using the tridimensional model with the objective of applying engineering and design project concepts to the tridimensional model. In making the RoboCAP prototype, a thorough analysis was carried out to compare the two types of specialist software - SolidWorks and Autodesk Inventor – as well as the use of AutoCAD. The analysis method is based on the comparison of temporal metrics and commands, thereby obtaining a relation of the system response to the user in the modeling, assembly and simulation of finite elements. Both types of software are similar when starting to create the piece, whereby they establish the Basic profile and offer specific functions for tridimensional modeling and simulation. By means of prototyping, it was possible to detect an error in the project design that corresponded to the weapons system, which showed a deformity in the screw coupling. After the anomaly was detected, the weapons shaft and toothed wheel parts were modified and adjusted to an equivalent diameter – concomitantly, the screw was adjusted to a different specification. The types of software analyzed have advanced tridimensional modeling resources for the development of robotics projects, with graphic computing technology linked to the database. The tridimensional simulation of RoboCAP with graphic computing resources allows for the use of complex models and calculations to simulate the design and monitoring of the robot, providing for the reduction of errors and an optimized management of processes that involve construction. The benefits of digital prototyping with the tridimensional simulation of the engineering project design are the economy of finances and time. The considerations on the optimization of robot construction have long term benefits that include technological enhancement and economy of resources.[1].

[1] This work is partially supported by Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Governo do Estado de Minas Gerais, Brazil, Rua Raul Pompéia, nº101 - São Pedro, Belo Horizonte, MG, 30.330-080, Brazil.

Computer graphics, digital prototyping, 3D simulation, three-dimensional modeling, engineering design, robot.

I. Introduction

The research project "Digital Prototyping - RoboCAP" involves studying and development model and tridimensional simulation of the robot conceptualized by RoboCAP team´s (Team of the Robotics Mechatronics Course Campus Alto Paraopeba, in the Federal University of Säo Joäo del Rei - UFSJ).

The RoboCAP team´s, a group which studies and researches robotics, was established in 2009 and consists of teachers and students of UFSJ, Mechatronics Engineering. This group was formed to develop projects and participate in robotics competitions and plans to participate internatoinally in "Robot Wars", organized by ROBOCORE - and at a national level Robot Games – the National Meeting of Control Engineering and Automation - ENECA 2012. Currently the focus is to combat the competition, where the team participated in the category with robots that may weigh up to 13 Kg. In competitions of "Robot War", two opposing radio-controlled robots compete in a specially designed arena for a 3 minute round, or until one of two participants destroys or immobilizes the other by means of mechanical force. The robots are subjected to extreme conditions to promote emerging technologies capable of operating under various circumstances. The arena supports large impacts, ensuring the safety of all spectators.

The evolution of computers and processing software focused on simulation and prototyping in virtual media become increasingly available and is frequently used. Among them, two engineering platforms, SolidWorks and AutoDesk Inventor are used nationally and internationally. The project aimed to evaluate the development of the project using the two different computer graphics software, and to develop simulation and tests in the virtual RoboCAP product before it is produced. The goal is to measure in detail the differences in inherent usability, quality, efficiency and applicability of the two software types in contrast to one another. Both software are derived from earlier models of Computer Aided Design - CAD, and both were used in parallel to develop the robot prototype with team RoboCAP.

Development of the project included modeling techniques, simulation of mechanical and electrical systems, and implementation and the optimization of control algorithms. The research project considered concepts and techniques of three-dimensional simulation with digital prototyping to evaluate and optimize various levels of the project before starting the execution of construction and assembly of the robot. The three-dimensional simulation of engineering using computer graphics allows the use of complex models and calculations to simulate the design and monitoring of robot construction providing for fewer errors and optimized management processes that involve the construction, including a class of analysis tools, integrated design, and associated systems based on solid modelers.

The robot is composed of several elements including the wheels, motors, bearings, and the structure of the robot itself. This structure, basically composed of two steel plates to reduce the number of fasteners and welding points, which are large voltage concentrators. One of the plates, which forms the base of the robot, has a thickness of 5mm, and is where all the other preceding elements of the robot are attached.

The following goals have been achieved: perform the simulation of the robot as it would enter competitions, suggesting changes to improve the competitive robot, analyze two software experts in the field of Robotics-Mechatronics Engineering; Autodesk Inventor and Solidworks; record a comparison of time and requirements of each software model as it applies to the modeling and prototyping, use finite element analysis to simulate forces and tensions in order to make the robot more competitive for the robot war, develop a practical application of topics in the course of Mechatronic Engineering School of Engineering with a UFSJ with a interdisciplinary project involving teachers and students.

The research questions were presented, "Investigating between different digital prototyping software which are the most relevant characteristics of each?" and "How the prototype can be used to speed up implementation and to avoid design errors of the robot in order to win the competition?".

Using the simulation it was possible to detect errors in the design of RoboCAP, then rebuild and test new options, resulting in a robot able to participate in competitions.

The benefits of Digital Prototyping with three-dimensional simulation of engineering design involve the economization of finance and time. The considers optimizing building robots, which bring long term benefits that include technological improvement and resource savings.

Currently, it is necessary to develop technologies and products that are increasingly better and more advanced and complex. The digital prototyping stages of construction provides the product construction, directly influencing the choice of production process and logistics industry, and reduce design errors, often unanticipated by the designers.

II. LITERATURE REVIEW

With the need to increase efficiency and profitability of production, increases the use of software for Computer Aided Design - CAD integrated with the database by expanding the concept of digital prototyping.

CAD software emerged for the development of two-dimensional designs - 2D, moving to the possibility of creating three-dimensional models - 3D. Then came more advanced platforms, simpler interfaces, more usable, capable of generating real images, structural analysis and dynamic analysis, enabling the testing of the behavior of the prototype in real conditions.

The software available enables the construction of two-dimensional models or three-dimensional models using the following: real imaging features, finite element analysis, motion simulation with database integration. These combined features make it possible to perform tests and trials on the behavior of the prototype in real conditions.

Autodesk Inventor software provides a set of development tools for mechanical design. During construction of the project is possible to integrate the property of pieces, with restrictions on insert data from a single digital design, [1].

SolidWorks is a design tool that uses parametric solid modeling, based on the characteristics and properties of each element and action, and these properties can be changed at any time of the modeling process, making adjustments to the project. SolidWorks is a CAD software, developed by SolidWorks Corporation, acquired in 1997 by Dassault Systemes SA, [2].

The concept of digital prototyping represents the ability to virtually explore a product before manufacturaing becomes reality. It allows you to create, simulate, validate and manage projects from initial concept to final product, [3] [4].

Digital Prototyping involves the techniques of prototyping in virtual environment which allows changes and optimizes the cost of production. According to H. B. Marri the planning process depends on the processes and decide which machines should be used from beginning to end manufacturing of the product, [5].

There is a consensus of the importance of planning before production. The objective of planning is to determine the dimensions, capturing the designer's intent and to simulate the prototype of the easy modification of the model [6].

The Digital Prototyping helps companies design better products, reduce development costs and get to market faster, [1].

Currently, robotics competitions have reached a nearly professional level with national competitions are organized by RoboCore, and international competitions organized by RoboGames, which occur annually [7]. The robotics competitions held by RoboCore are divided into four categories that explore and enhance the knowledge in programming, artificial intelligence, electronic, mechanical, electrical resistance of materials, among others.

The categories are: Follow Line: autonomous robots must follow a predetermined path, a line drawn on the ground, in the shortest possible time. This objective can be accomplished in different ways, allowing the participation of beginners and experts; Hockey Robot: teams of three radio-controlled robots fight for a match in which the goal is to score as many pointsas possible within a time limit, achieved through number of goals scored with a hockey puck into a net by means of robotic mechanic force; Sumo: Based on ancient Japanese wrestling, robots have to identify it’s opponent and push it out of the arena. This robot may be radio-controlled, autonomous, or even assembled from Lego ® kits dedicated to learning robotics; Combat: the category that attracts more public interest in general, strickly uses radio-controlled robots which face off in a specially designed arena until one of the participants is destroyed or immobilized. The projects developed are subjected to extreme conditions, so that emerging technologies capable of operating under various circumstances can be developed. Battles are conducted in rounds of three minutes each. The arena is built to withstand severe impacts, ensuring the safety of spectators.

Finite Element Method - FEM is to divide the structure to be analyzed in different parts or elements. Each element has the form of a specific structure which is connected to adjacent elements, called nodal points or nodes, which are generated equations for each element. The behavior of the structure as a whole is calculated from the sum of the behavior of the various elements existing, thus obtaining a complete analysis of structure. The FEM is a technique for solving partial differential equations that arise from these nodes, [8].

III. METHODOLOGY

The methodology consists of three-dimensional modeling, simulation and analysis of RoboCAP developed simultaneously in two specialist software AutoDesk Inventor and SolidWorks, for engineering applications, [9] [10] [11].

The first step is the modeling of each part of the robot separately, measuring and calculating with each of the different development software. After, there is a table presenting the comparison of modeling in the software.

The combat robot designed by the team RoboCap is composed of several parts: Pulleys: Pulleys are a part of so-called lifting machinery and transportation. The electric motors usually have a fixed frequency of rotation, however, often the machines are driven by them and need to develop different frequencies of rotation. Because of this, the robots use coupling pulleys, and driven conductive, gears,or bearings. These machine elements have wide responsibility in supporting the robotic equipment. Although the constructive aspect of bearings and springs are very different and have different applications as well, no use of certain equipment could be compromised; Box Reduction: Reducing the speed machines are used to achieve large reductions in transmission, steel structures, plates make up the steel structure or outer shell of the robot, also acting as an armor, so are the parts of the robot more susceptible to impacts. Basically composed of four parts, the plates back, bottom, front and side that supports the locomotion systems and the weapon.

The second step of the modeling process was comparing the use of two different software, Autodesk Inventor and Solidworks, for the modeling process of the RoboCAP robot. Temporal metrics were used to count the number of clicks for each command executed by the program Refog Keylogger. The program was used to count and calculate the total time counter via keystroke. This measure defines the time of use of controls that reflect the performance of use.

The step comprises assembling the robot by combining pieces modeled. During the development of this step it became necessary to amend the initial project because an overlapping system of weaponry was detected that was decreasing the efficiency of the robot, which would directly impact the performance of the robot during the competition. Changes were made and the design by RoboCAP was modified.

The fourth step concerns the analysis of the dynamic forces that act on the system. With the insertion forces on each part of the system generates a graph showing the behavior of the forces. A simulation with video and animation was produced. The simulations showed that the cover design does not support a force with an intensity higher than 4000N, causing plastic deformation and shearing the structure of the front screw.

The fifth stage was the finite element analysis with application software to test the design of the robot. After defining the material, simulated attack situations were made.

IV. DEVELOPMENT

The structure is composed of two steel sheets, Sheet 1 - contoured so as to form the upper case - that it is 2mm thick in reduce welds and fasteners; Sheet 2 - constitutes the base of the robot, where they are connected to all other elements of the robot - 5 mm thick.

Front Sheet - Using SolidWorks software, the design of front plate contains the following steps and the result is shown in Figure 1: created a new project or piece; Sketch 2D command of the profile of sheet with dimensions 220x335 (mm) as the Sketch of the hole where it will be located the weapon of robot and flanges to be bent; Base Flange Sheet command - thickness of 5 mm; Edge Flang command is passed to the flange sheet to be bent with pre-defined geometry, the bending angle of 35 ° and confirms the creation of the fold.

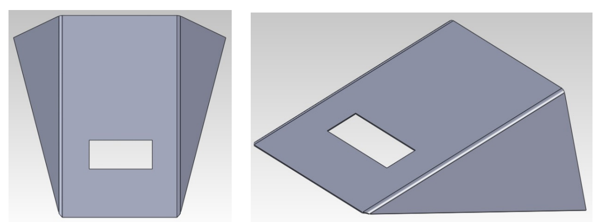


Figure 1 - Front Sheet (SolidWorks)
Source: Developed by author

In Autodesk Inventor, the design of the front sheet contains the following steps and the result is shown in Figure 2: First open a new file SheetMetal (mm). ipt; Sketch 2D command rectangular profile with dimensions 220x335 (mm) and the diameter of the weapon system; Face command for the creation of the sheet (thickness, shape of the bending, angle is defined in advance); Flange command sets the angle of 35 ° and size 187.297mm; Chanfre command to the folds.

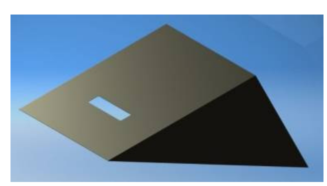


Figura 2 - Front Sheet (Autodesk Inventor)
Source: Developed by author

Wheels - The wheels used in the project are rubber, except for the part where the shaft is fitted (fitting type key), made of plastic. The wheel is divided into external structure of rubber and internal structure of plastic (Figure 3).

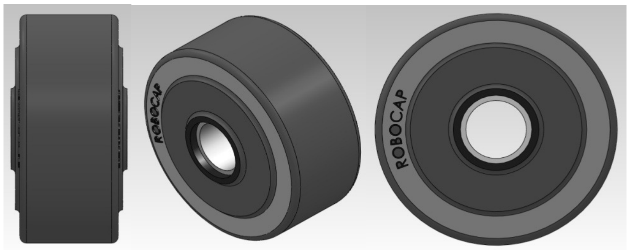


Figura 3 - Wheels (SolidWorks)
Source: Developed by author

Bearing - The bearings also are made of steel (Figure 4). The “Extrude” command sets the width of the bearing and confirms the command for the creation of the piece in 3D; It created an outline of the hole and used the Extrude Cut command, which "subtracts" the created part of the bearing, thus creating the hole.

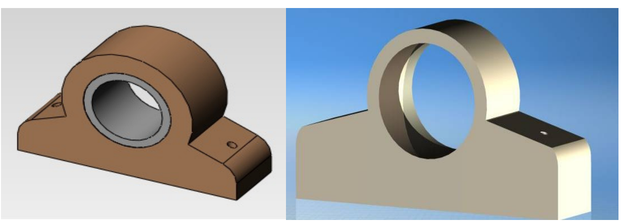


Figure 4 - Bearing (SolidWorks)
Source: Developed by author

Identation Discs - Weapon - The disks are made of steel of high durability to withstand the impact generated by the reaction of attacking an opponent robot. First, a new project or piece was created; Then the plan of the piece was chosen; Sketch 2D command - profile piece, and the central hole through which the shaft and the teeth; Followed by Extrude command - which defines the thickness of the disc and confirms the command to create the piece (Figure 5).

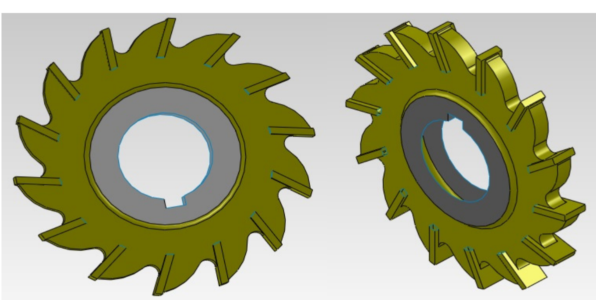


Figura 5 - Identations Discs - Weapon (SolidWorks)
Source: Developed by author

Front Bearing Locomotion - Bearing was used to compose the three supports of the robot and which enable locomotion. The bearing is made of steel (Figure 6). First, a new project or piece was created; then Sketch 2D command - profile piece, two concentric semi-circles of radio 25 mm, 30 mm, semi-circle of 50 mm; finally, Revolve command to create the piece in 3D.

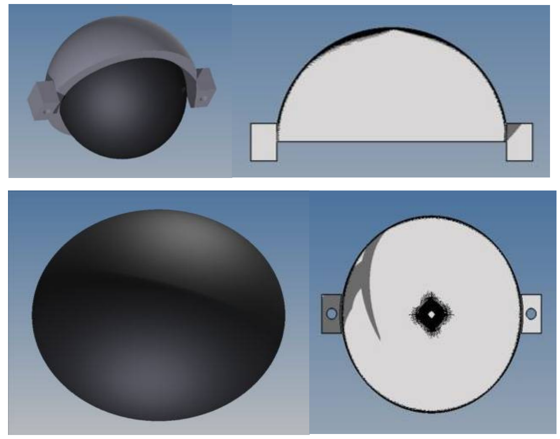


Figure 6 - Front Bearing Locomotion (SolidWorks)
Source: Developed by author

Clamp - SolidWorks software, the design of the clamp contains the following steps and the result is shown in Figure 7: First, a new project or piece was created; then Sketch 2D command - profile piece; followed by Extrude command for creating in 3D; and finally, created an outline of the hole and used the Extrude Cut command, thus creating the hole.

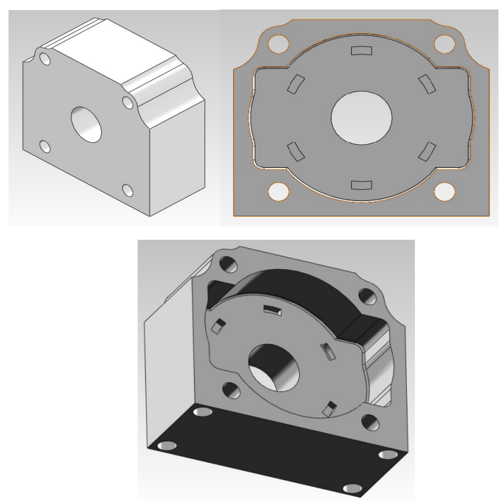


Figura 7 - Clamp (SolidWorks)
Source: Developed by author

Gearbox - SolidWorks software, the design of the gearbox contains the following steps and the result is shown in Figure 8: First a new project or piece was created; then Sketch 2D command - profile piece; followed by Extrude command for creating in 3D; and finally, Extrude Cut command to create the details and contours of the piece.

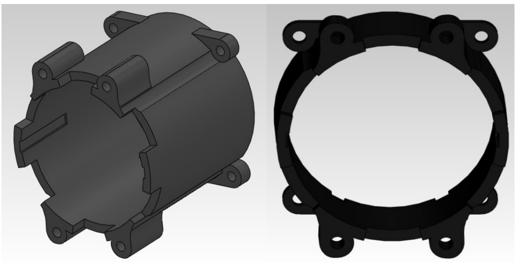


Figura 8 - Gearbox (SolidWorks)
Source: Developed by author

Structure - Single Block - After modeling the structure on several sheets, the simulation was performed using a single sheet, the folds being applied only (Figure 9). The steps for construction of the piece is described below.

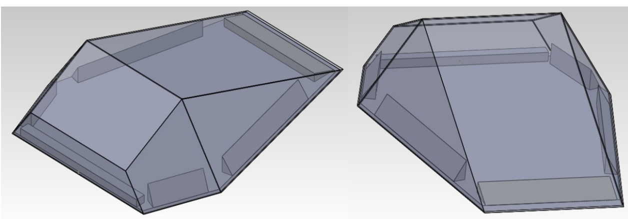


Figure 9 - Structure - Single Block (SolidWorks)
Source: Developed by author

After the first model, the single block design was modified after the development of resistance testing and finite elements.

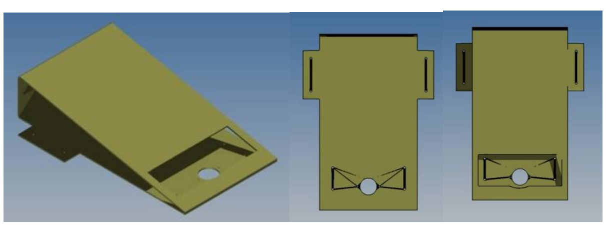


Figure 10 - Carcass (Autodesk Inventor)
Source: Developed by author

Assembly - SolidWorks has its own kind of assembly files - *.asm. It offers a range of options for the complete assembly, as there are several ways of positioning of parts and correlation between the parties. When creating a new assembly is necessary to assign the parts related to assembly. Then, the pieces must be pre-positioned so as to facilitate coupling between them, either by a flat, straight section, among others. Thus, it is possible to establish the correlation between geometric elements such as elements that are parallel, concentric, orthogonal coincident, among others (Figure 11).

Autodesk Inventor has a propriatary format for the assembly environment *.iam . Using the *.iam file, the interface of the software provides tools for assembling the parts in *.ipt. The modeling options are restricted, but the options for positioning parts are complete. When inserting parts in the assembly environment, the browser toolbar adds their physical characteristics, as well as plans for the construction of each piece. Even after assembling, if a single parts is changed, it is updated.

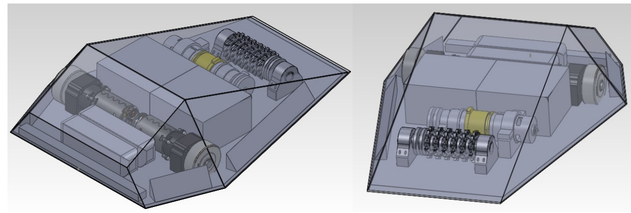


Figure 11 - Carcass (Autodesk Inventor)
Source: Developed by author

V. RESULTS

During the assembly process of the RoboCAP, a a video with the dynamic simulation of the assembly was developed. The sequence of images taken from the video are shown in Figure 12.

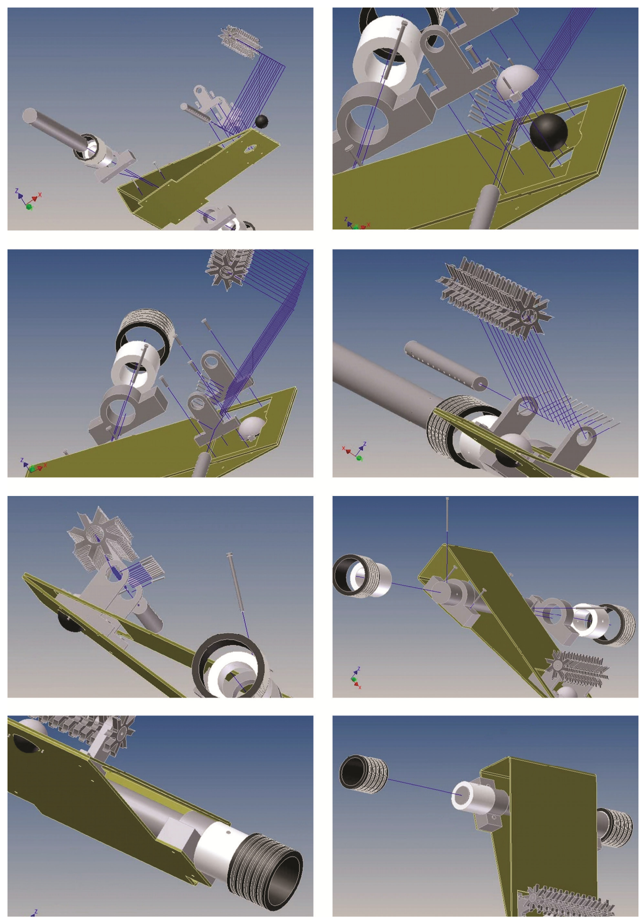


Figure 12 - Sequence of images taken from the video - Part 1
Source: Developed by author

The prototyping process using AutoDesk Inventor and SolidWorks are similar. Both softwares has similarity to start creating the piece, which first establishes the Basic Profile (Sketch). On this platform the size of the figure are established only in 2D format, create the outline of the project without the help of measures, since they are inserted only at the end of the process. The modeling begins after the construction of the Sketch 2D. The 3D modeling tools allow the construction of the piece in 3D and it is available from the menu bar. Sometimes, a user needs to create another profile 2D model on a 3D model in order to achieve the goal of the modeling. The process of alternating modeling between modules of Sketch 2D and Modeling 3D happens in accordance with requests made by the user, from the tool bar in both programs.

A comparison was made of the modeling performed by SolidWorks and Autodesk Inventor to make a thorough comparison of metric advantages and disadvantages inherent in the clash between these two softwares. This research used a careful comparison measures between the use of the two software to perform the same task, such as to "draw" one of the wheels of the robot. Table 1 shows the metric temporal comparison of execution time of each piece in each software. The first column shows the components of RoboCAP, the second column shows the parts, the third column contains the used software (SolidWorks - SW and Autodesk Inventor - INVENTOR), the fourth column shows the time taken to complete the task in each software [12], [13].

Table 1 - Comparison of run time on each software

COMPONENT	PART	SOFTWARE	TIME
		SW	00:03:53
	Structure 1	Inventor	00:16:53
		SW	00:03:25
	Structure 2	Inventor	00:09:15
		SW	00:02:19
	Structure 3	Inventor	00:10:38
		SW	00:03:28
	Base	Inventor	00:04:41
		SW	00:00:56
STRUCTURE	Locking bar	Inventor	00:00:34
		SW	00:04:38
	Bushing	Inventor	00:04:32
		SW	00:00:28
	Axis	Inventor	00:00:24
		SW	Biblioteca
	Disk	Inventor	Biblioteca
		SW	00:00:45
Weapon System	Collar	Inventor	00:00:51

VI. Conclusion

The research objectives were achieved with the development of modeling, prototyping and simulation RoboCAP. It was possible to test, improve the design and performance of the robot in addition to finalizing the design for the production team. During assembly of the RoboCap there was an error in the project because the weapons system had a deformity in the docking of the screw. The error was detected during the stage of assembly. Parts of the axis of weapons and gear were modified to have an equivalent diameter.

Both programs have specific functions for creating sheet metal. The SolidWorks provided easy usability and is very versatile as to their commands and tools. Autodesk Inventor has a privileged view to access the project and their respective bars and tools. The comparison between the two software presented the specifics of each software and the similarity of the modeling process. This fact confirms that the concept of three-dimensional modeling which can be implemented in different software companies leads to the same result.

Digital prototyping made possible a way to simulate real combat situations, find design flaws, make changes and corrections in the design, thereby reducing costs and production errors. The errors found were solved before the production stage and showed direct evidence of increased speed and competitiveness of the robot.

The implementation of this research, involving theoretical and practical application, was a motivator for the team to attempt to rank high in robotic competitions. The development of a mobile robotic system of this type involves technological process, practical and theoretical, which develop a culture of innovation in the team involved in the project.

References

[1] AUTODESK. Autodesk Inc. *Software CAD 3D para projeto mecânico.* http://www.autodesk.com.br/adsk/servlet/pc/index?id=14568941&siteID=100

[2] SOLIDWORKS, SKA. *Validação de Projetos.* http://www.ska.com.br/produtos/recursos_aplicacao/1/solidworks_cad_3d/6

[3] BARBOSA, R. T. ; ALENCAR, Francisco de . *Design & Prototipagem: Conhecimento e uso da prototipagem digital no design brasileiro.* In: 9º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design., 2010, São Paulo/SP. SãoPaulo/SP., 2010. http://blogs.anhembi.br/congressodesign/anais/artigos/70427.pdf

[4] WU G., GUO X. *Digital Prototyping And Simulation of DG Boiler Feedwater Pump*. School of Mechanical Engineering. Anyang Institute of Technology. Anyang ,China, 2011.

[5] MARRI, H. B., GUNASEKARAN, A., GRIEVE, R. J. Computer-aided process planning: A state of art . *The International Journal of Advanced Manufacturing Technology*, 14, 1988.

[6] AULT, H. K.; D.T GIOLAS, D. T. An Investigation of Solid Modeling Practices in Industry. *Engineering Design Graphics Journal*, Vol 69, No 1, 2005. http://www.edgj.org/index.php/EDGJ/article/viewFile/78/75

[7] ROBOCORE. www.robocore.net. Acesso em: 15/07/2011.

[8] ALVES FILHO, A. *Elementos Finitos. A base da tecnologia CA*E. São Paulo: Érica, 2000.

[9] SANTOS, E.; ZUFFO M.; NETTO, M.; LOPES, R. *Computação Gráfica: Estado da Arte. 15.* Simposio Nacional de Geometria Descritiva e Desenho Tecnico, IV International Conference on Graphics Engineering for Arts and Design, São Paulo, Brasil, Nov. 2001. http://toledo.pcc.usp.br/pdf/1341.pdf

[10] PATNAIK, K,; SONI, M. *3D Solid Modeling Of Screw And Scroll Compressor Including Animation*. 2007**.** http://ethesis.nitrkl.ac.in/117/1/10303069.pdf

[11] DUNZHU X., SHENG X., SHOURONG W. *A Digital Prototype Miniature Silicon Microgyroscope.* IEEE International Conference on Nano/Micro Engineered and MolecularSystems(NEMS), 5th Southeast University, Nanjing, China.2010.

[12] PREECE, Jennifer; Rogers, Y.; Sharp, H. *Design de interação: além da interação homem-computador*. São Paulo: Bookman, 2005.

[13] NIELSEN, Jakob. *Usabilty Engineering*, Morgan Kaufmann, Inc. San Francisco, 1993. http://www.labiutil.inf.ufsc.br/hiperdocu-mento/unidade1.html

Facilities Data and Data Models

Deborah L. MacPherson
Cannon Design, Specifications and Research
National Information Exchange Model (NIEM) Business Architecture Committee (NBAC)
National Building Information Modeling Standard (NBIMS-US) Technical Subcommittee
OmniClass Development Committee
dmacpherson@cannondesign.com

ABSTRACT

Building Information Modeling (BIM) needs to be placed in a larger architectural context with data modeling, geography, web infrastructures, public safety, and real estate. A common reference system is needed to ensure fundamental characteristics of building interiors are understood in affiliated software and services. Only a few data need hooks, for example footprints, entrances, floors, circulation, occupancy, spaces and specified contents. Implementation of successful BIM web services will be a symphony of the following: a Data Model (isolate certain International Foundation Classes); a Dictionary (APIs with the buildingSMART Data Dictionary); classification and taxonomies (OmniClass); and Process Descriptions (Information Delivery Manuals and Model View Definitions). A kit of parts, including metadata organization and rules for what is discoverable in registries and repositories, needs to conform to Simple Object Access Protocol (SOAP), Representational State Transfer (REST), Web Services Description Language (WSDL), and new Internet technologies as they are invented. The whole system could benefit from a collection of Tiny Ontologies All Stitched Together (TOAST) branching out from a common core. The US National Information Exchange Model (NIEM) is an ideal home for elements related to justice, emergency management, health and family services - many people could benefit from this work.

Keywords
BIM, GIS, NIEM, OmniClass, Data Structures

INTRODUCTION

As timely and practical as Building Information Modeling (BIM) is, the greatest use may be via web services. For example, to achieve President Obama's 2011 goal for "...a firefighter to download the design of a burning building onto a handheld device" several applications, networks, policies, and Internet layers need to work together **[Figure 1]**. While this might be possible in a closed environment today, for any firefighter to access up-to-date information, BIMs need to be placed in a larger architectural context with data modeling, geography, web infrastructures, public safety, and real estate.

Current BIMs are not scalable to thousands of simultaneous users focused on specific buildings during an emergency, or to seamlessly navigate indoor and outdoor locations. A common reference system is needed to ensure fundamental characteristics of building interiors are understood in affiliated software and services. Only a few pieces of data need hooks, for example footprints, entrances, floors, circulation, occupancy, spaces and specified contents.

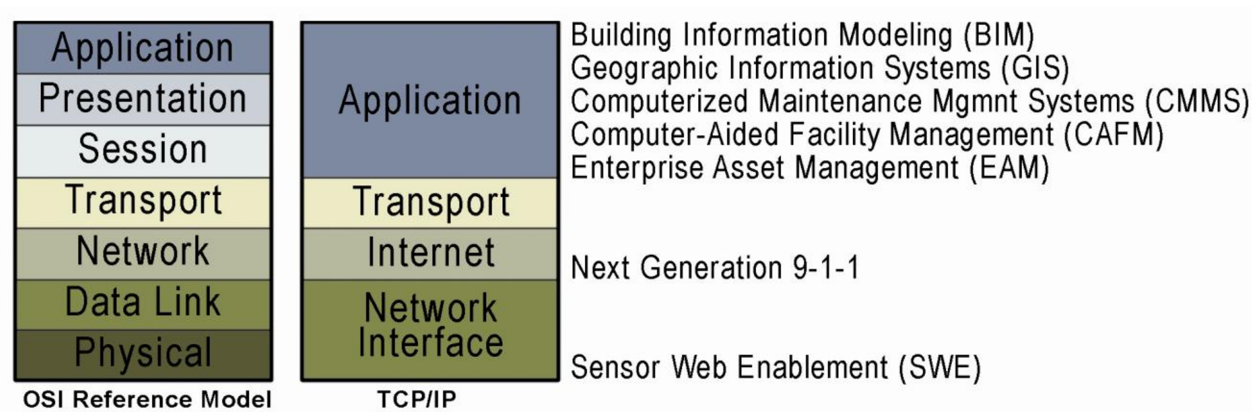


Figure 1- Software, Networks, and Internet Architecture, by Deborah MacPherson

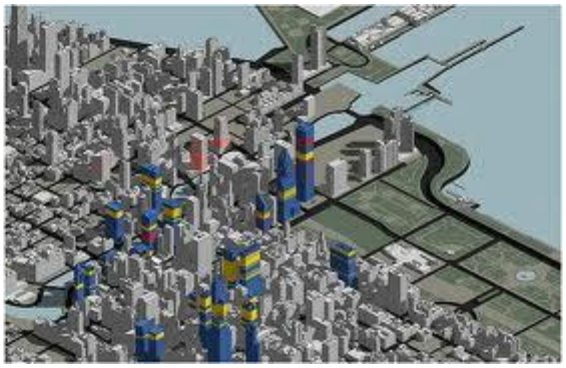


Figure 2 - CityGML, by LiDAR News

OPEN ARCHITECTURE

Standards Development Organizations (SDOs) and vendors should have a summit to understand what everyone else is doing relative to themselves. Significant results could be achieved quickly. Everyone is in a better position to do this now, most Memorandums of Understanding (MOUs) are already in place. Shared concepts, software interfaces, service interfaces, and data structures can be extended in more detail by each stakeholder to understand interrelationships of shared building information through their own data, systems, and viewpoints.

Figure 3 - Two Different Buildings Sharing the Same Layout Colors, by Cannon Design

Different parts of an open architecture for BIM and related technologies will originate, and need to be maintained, by diverse professional organizations. Senders and receivers need to agree on exactly which information to exchange and what it should look like. The whole system needs to be dynamic and cyclical, updating with continuous improvements and strict version control. Defining where each part belongs in the whole world's digital infrastructure will ensure each part is worth doing, will be sustainable, published openly, and owned by people that care about what the information means. The results will not be duplication - but harmonization and increased customer satisfaction.

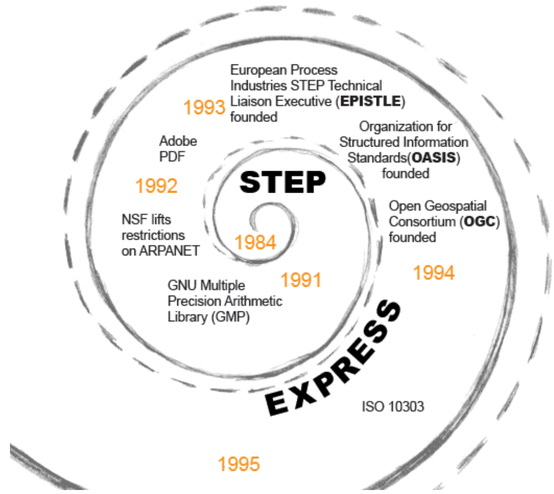


Figure 4 - A Brief History of Facilities Data and Web Standards, by Deborah MacPherson

WEB SERVICES

Web Services are automated machine-to-machine exchanges and data integration processes using open standards over the Internet to simplify complex data into messages.

Successful web services for BIM requires the following: a Data Model (isolate certain International Foundation Classes); a Dictionary (APIs with the International Framework for Dictionaries); classification and taxonomies (OmniClass); and Process Descriptions (Information Delivery Manuals and Model View Definitions). A kit of parts, including metadata organization and rules for what is discoverable in registries and repositories, needs to conform to Simple Object Access Protocol (SOAP), Representational State Transfer (REST), Web Services Description Language (WSDL) and new Internet technologies as they are invented.

Leveraging existing work by others will require specified sets of building data to stay bound together in messages. Message types will define the order data are presented, beginning with the most important sets first, then compiling more detail depending on the contracted web service.

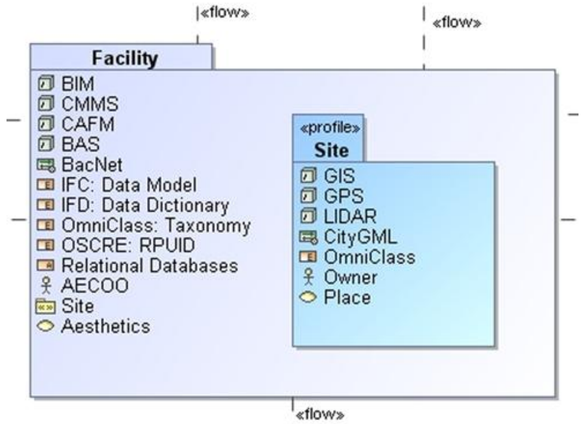


Figure 5 - Facility and Site Packages from NBIMSv2 Proposed Reference Architecture, by Deborah MacPherson and Louis Hecht

ONTOLOGIES

Ontologies are often used for search, decision support, and software design. Matthew West defines ontologies simply as "The things there are and the rules that govern them." BIM ontologies need to be web-based to support workflow functions, provide overviews, relationship structures, necessary attributes, reminders about information to include, and consistency across integration models used with master and reference data. Ronald Reck suggests a collection of Tiny Ontologies All Stitched Together (TOAST) branching out from a common core. The US National Information Exchange Model (NIEM), a data model for data exchanges using a common core, is an ideal home for facility data elements related to emergencies, justice, health, and family services - many people could benefit from a focused collaborative effort.

Large complex models like NIEM and the Open Geospatial Consortium (OGC)'s CityGML are made for different reasons. Ontologies drawing from them can be leaner and fit for purpose. Maintaining linkages between and among data assets, keeping specific data types bound together, and resolving multiple records describing the same assets needs ontologies to tell machines how to interpret the information, what to do with it, and what level of detail to store in each environment without restating the same instructions for each web service or affecting other data models internal structures.

LINKED DATA

Changing a published building layout or occupancy should be able to tap back into open floor plans, linked databases, and public records. Exporting heavy files like typical BIMs are today sends more information than needed and too many copies are generated. Using a Service Oriented Architecture (SOA) approach instead, data remains with one application for reference or modification in other applications by maintaining links and mutually supportive artifacts. Shared elements and a messaging structure created for BIM going outwards could also benefit BIM going deeper into the design process at Architecture/Engineering firms. Some typical software such as BIM and specification writing programs can maintain rolling associations, but they are only two of many **[Figure 2]**. If web services and ontologies could be used to establish semantic connections rather than just software connections, archival flows could extend to knowledge management systems so a critical mass can be reached within firms and across the industry. The number of known answers to repeated questions

Figure 6 - Real Property Unique ID (RPUID), by the Open Standards for Real Estate Consortium (OSCRE)

and combinatorial complexity of geometric objects can be reduced using thematic and logical restrictions. When an OmniClass type usually has a limited range of typically associated components, it will get easier to link back and forth between applications and resources. Repeated instances of OmniClass types, and formulaic buildings like dormitories, will be especially adaptable.

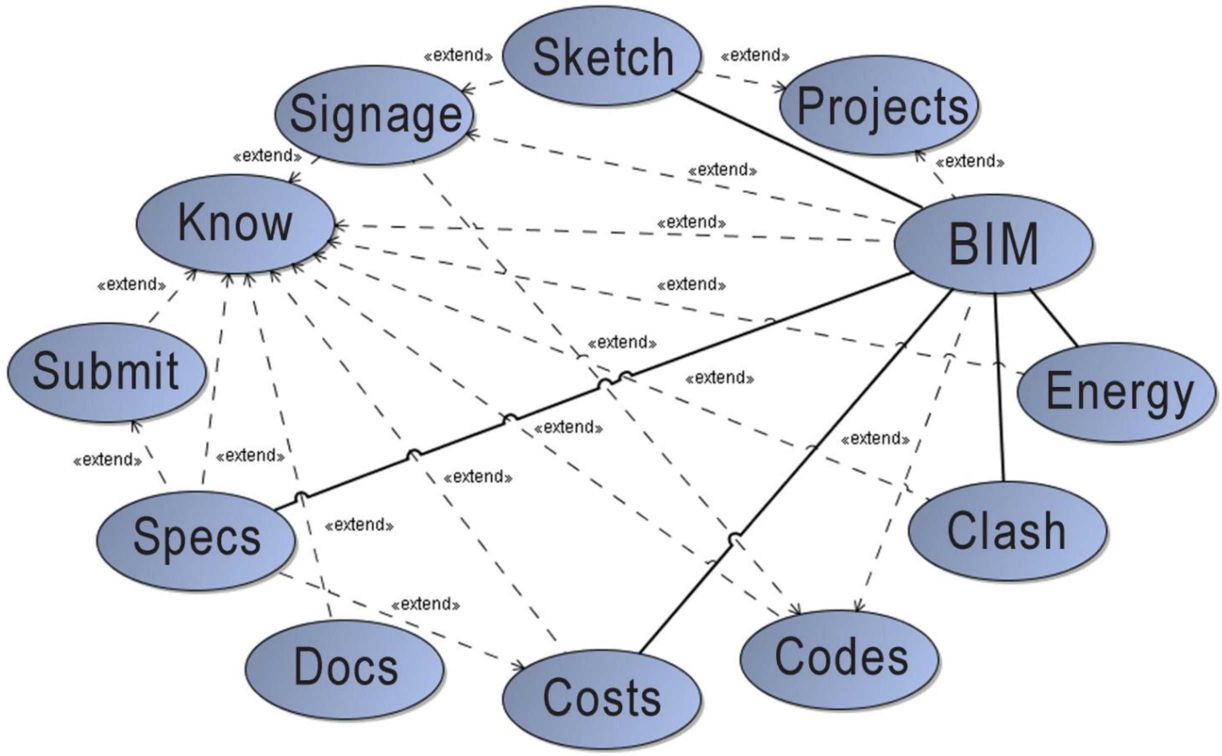


Figure 7 - Typical A/E Software, D. MacPherson

BENEFITS

Live data could be automatically captured in compliance with policy constraints to flow through to first responders, maintenance contractors, real estate agents, energy auditors, financial agents, and other authorized users without duplication and potential human error. Real-time information compared to historical trends will let machines keep track of minutiae and fastest routes to improve scheduling, perform analysis, and assess scenarios.

Several initiatives need to converge so the results will begin showing up in software. Defining a core set of elements with relationships between them and the outside world will let users, or machines on their behalf, periodically submit questions such as "have the regulations for chlorine storage changed in Arlington County?" and get an answer from the local jurisdiction or other authoritative data source.

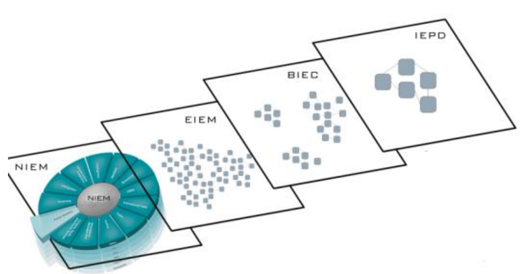


Figure 8 - NIEM Groupings, D. MacPherson

Each user and their machine would only receive the information they need, or are allowed to access, rather than an unwieldy batch of data to extract an answer. The extraction is performed via web services. Invasion of privacy and terrorist planning can be prevented by limiting distribution to public information or authorized web services with elemental traceability. Some web services and ontologies will be free for the general public, some will cost a lot of money. The first and most collaborative versions should be for public safety in public buildings such as schools, hospitals, courthouses, stadiums, airports, hotels, and shopping malls. These are also buildings where anyone may want to look up their current location on a handheld device so consumer demand and Internet service providers can help speed development.

CONCLUSION

Web services and ontologies are progressive; there will never be a final version, finite set of questions and answers, or set of dots to connect. A shared reference architecture needs to celebrate complexity in the world, not dull it down, or the results will not be realistic and effective.

Acknowledgements: Louis Hecht, Ron Reck, Matthew West, Joseph Talamantez, and Christopher Shoup.

Fostering Student Enrollment in Basic Sciences: the Case of Southern Tuscany

Vera MONTALBANO
Department of Physics, University of Siena
Siena, Italy

ABSTRACT

In recent decades it has been detected in Italy a decrease in enrollment in basic sciences, i.e. Mathematics, Physics and Chemistry. The increase in specific orientation in Chemistry, Physics, Mathematics and Materials Science, is strategically crucial to achieve the goal of maintaining and increasing the number of motivated and capable students who enroll in these and other courses scientific degree.

In 2005 the government launched the Scientific Degree Project in which many Italian Universities were involved in implementing activities focused to enhance the interest of high school students towards sciences and scientific degrees. With the purpose of increasing scientific vocations, workshops were organized in high schools and teachers were involved in planning and implementation of laboratories, conferences for scientific outreach, thematic exhibitions, guided tours of research laboratories, summer's schools for students and courses for teachers were realized for developing a cultural enhancement in teaching basic sciences. Initially, the project had been funded for four years, then was refinanced in 2010 and became the National Plan for Science Degree, with the aim of stabilizing the most significant activities and promoting them throughout the country. Particularly significant is the case of activities organized by the Department of Physics of the University of Siena for students and teachers in Southern Tuscany. The methods used in cultural enhancement of teachers and activities designed to support schools with limited laboratory facilities, together with stimulating activities for the more motivated and talented students are allowed to take root for some good practices in physics teaching and orientation to scientific degrees. The project was born in Siena only for Physics in 2005, adding the project for Chemistry in 2007 and in 2009 also the Department of Mathematical Sciences launches its project. Beyond describing the main activities for orientation to Physics, activities done in partnership with chemists, biologists and geologists are reported, as well as an activity in which the Departments of Mathematical Sciences and Physics are both involved in looking for introducing new interdisciplinary methodologies to increase students' understanding in high school of some selected topics in which both Mathematics and Physics give a contribution in the construction of important and mutually reinforcing basic concepts of the two disciplines.

Keywords: Higher education, Science enrollment trends, Student motivation, Orienting, Continuous learning, Physics laboratory, Teaching methods and strategies, Professional development.

1. INTRODUCTION

The enrollment in Physics, Mathematics and Chemistry has declined dramatically almost everywhere in the world [1-5]. The interest of students in learning these disciplines decreases and consequently the results achieved in schools and universities are disappointing. On the other side, the goal of maintaining and increasing the number of motivated and talented students who enroll in courses scientific degrees is crucial in order to achieve a knowledge-based economy more competitive and dynamic capable of sustainable economic growth with more and better jobs and greater social cohesion.

Italy is active since 2005 with a large and detailed plan to remedy this problem funded by the Ministry of Education and Scientific Research [6]. The main purpose is to provide opportunity for students in the last years of high school to learn about issues, problems and processes characteristic of scientific knowledge, also in relation to the areas of labor and the professions, in order to identify specific interests and requirements, making informed and suitable choices in connection to their personal project. Another key objective is to improve the disciplinary and interdisciplinary knowledge of teachers and their ability to interest and motivate students in learning science, and to support them in pre-university orientation process.

In paragraph 2, the enrollment trend in Italy is presented. Actions for inverting this trend are described with a special care to strategy and adopted methodologies. In the next paragraph, the case of southern Tuscany is discussed in detail. Many designed and realized activities are reported. In particular, a summer's school for students is described. Its purpose is to orient students towards physics and to improve teacher practice.

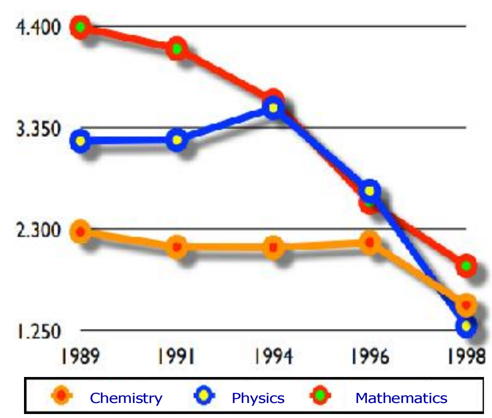


Figure 1: Enrollment in Italian universities before actions were performed for inverting the negative trend [7]

Finally, for teacher professional development, an activity for teachers of physics and mathematics on modeling is presented. A brief account of the effect of these actions on the enrollment trend concludes this article.

2. ACTIONS FOR SCIENCE DEGREE IN ITALY

In recent decades it has been detected almost everywhere in the world a consistent decrease in enrollment in basic sciences, i.e.

Mathematics, Physics and Chemistry. The main consequence is a decreasing of graduates in science disciplines.

The situation in Italy was dramatic: Figure 1 shows that enrollments in basic sciences are more than halved in few years. Since there is a clear association between economic performance and the numbers of engineers and scientists produced by a society, the Ministry of Education and Scientific Research promoted a wide project in order to reverse this trend.

Starting at the end of 2005, the project was named Scientific Degree Project (Progetto nazionale per le Lauree Scientifiche, i.e. PLS) and was financed for four years. During this period and at the end of the project, a large monitoring of all activities was realized in order to identify what actions were more effective and incisive. In 2009 it was launched the National Plan for Science Degree (the same acronym remains: PLS) where some of the most effective methodological aspects were emphasized in new guidelines.

	Scientific Degree Project 2006 - 2009	**National Plan for Science Degree 2010 - 2012**
funds from Ministry	8,5 M€	4,8 M€
funds from Universities	2,7 M€	0,5 M€
Universities	33	42

Table 1: Financial support in Italy from 2006 to 2012 and universities involved in contrasting decreasing of scientific vocations.

The source of main financial support for both actions is showed in Table 1, together with the participating universities.

Scientific Degree Project 2006 - 2009

The project stems from a collaboration of the Ministry of Education and Scientific Research, the National Conference of Deans of Science and Technology and Confindustria, the main organisation representing Italian manufacturing and services companies, and was designed in 2004 with the initial motivation to increase the number of students on degree courses in Chemistry, Physics, Mathematics and Science of materials.

The project focused on three main objectives [6]:

- improving knowledge and awareness of science degree in secondary school, offering students in the last three years of school to participate in stimulating and engaging laboratory activities curricular and extra curricular;
- starting a process of professional development of science teachers in service in the Secondary School from joint work between the School and University for the design, implementation, documentation and evaluation of the laboratories mentioned above;
- promote alignment and optimization of training from the University and the University School for the working world, strengthening and stimulating activities and training workshops at universities, research institutions, public and private companies engaged in Research and Development.

The action for student integrated witn training of teachers was made through more than 100 sub-projects under the responsibility of local referents, located in 33 universities, became 38 in the last period, spread all over the country and organized into four areas of national projects [7], as shown in Figure 2.

Figure 2: Universities in the national project. For each university the activated areas of the project are shown: materials science, physics, mathematics and chemistry (clockwise)

The main novelty compared to previous actions consists in using the same organizational model coordinated at national level, introducing a new model of collaboration between universities and secondary schools and the involvement of new actors in the activity of connecting university – school.

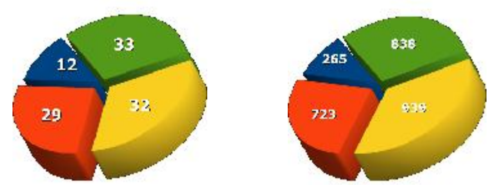


Figure 3: Universities and schools involved in disciplinary projects, starting from smaller, science of materials, chemistry, physics and mathematics (clockwise)

Educational institutions involved in disciplinary projects [7]. are shown at the end of fist biennial period in Figure 3.

National Plan for Science Degree 2010 - 2012

The plan maintains the same purposes of increasing the enrollment in science degrees to which is added the necessity to revise the content and methods of teaching and learning of science in all grades of school, taking into account the new national guidelines for first and second cycle contained in the recent Italian reform of the educational system.

Strategy and methodologies: In order to achieve the above purposes, the Plan pursuits fundamental ideas that have proved effective in trials 2005-2009 [6]:

- orientation does not conceive how a teaching path given to student, but as an action that the student is doing, from meaningful activities that allow to compare problems, issues and ideas of science;
- designing the training of teachers in service by involving teachers in solving concrete problems, developing design and

implementation of educational activities and through comparison with peers and experts;
- pursuing and achieving at the same time the student orientation and training of teachers through the planning and joint implementation by school teachers and university laboratories for students, thus developing also relations between the school system and the University;

Furthermore, a new idea adds: connecting consciously the activities of the Plan with the innovation of curricula and teaching methods adopted in schools, and other contents and methods of teacher training (initial and in-service), for the first and second cycle.

In other words, the main road consists in considering laboratory as a method, not as a place, students must become the main character of learning and joint planning by teachers and university is a mandatory step.

More attention to laboratory is required and different type of laboratory PLS can be proposed:
- laboratories which approach the discipline and develop vocations,
- self-assessment laboratories for improving the standard required by graduate courses,
- deepening laboratory for motivated and talented students.

3. ACTIONS IN SOUTHERN TUSCANY

During the last decade, the Department of Physics is active in orienting students towards Physics and in teachers training in Physics and Mathematics. Our activities are directed to students and teachers living in the provinces of Arezzo, Grosseto and Siena, showed in Figure 4.

Figure 4: the area in Tuscany where Department of Physics performs its activities of orienting and teacher professional training and development.

Early activities
Since 2000, the Department of Physics is engaged in many activities aimed at students and teachers of high schools, such as initial teacher training in the Advanced School for Teaching in Secondary School of Tuscany and various activities for orienting students towards Physics. Furthermore, an advanced courses for teachers of Physics and Mathematics was held. In this course, teachers were engaged in designing learning paths with a particular attention to relationship between physics and mathematics. Physicists (E. Mariotti, V. Millucci, V. Montalbano, S. Veronesi) and expert teachers (R. Benedetti, A. Porri, E. Papi) worked together at that time in order to improve teaching strategy and methodologies in these disciplines..

Scientific Degree Project
The project started in Siena only for Physics in late 2005 with 11 school involved, while the project for Chemistry began in 2007. Physics project was named *Trying and trying again* for remembering to everyone that main activities were focused on laboratories directly performed by students or teachers in the case of professional development. Two main actions, professional development for teachers and orientation to physics for students, were implemented and in the following few examples of both are given.

Advanced course for professional development: We continued a previous experience by designing an annual course for teachers titled *Learning Paths in Physics and Mathematics: Models, Experimental checks, Statistics.* The course was active in Siena from 2005/2006 to 2007/2008 and was attended by thirty teachers.

Updating course for professional development: In the last year of the project, we preferred to realize an updating course for teachers in Arezzo at the local high school titled *The nature of light: from classical physics to quantum physics.* In this course, 17 teachers were enrolled and more than an half of scheduled time was spent in physics laboratory.

Orienting laboratory for classrooms: Didactic laboratories of the department became available for entire classes to perform some experiments in mechanics, optics and electromagnetism, after a joint programming activity with teachers, to support schools where there are no laboratories or where physical experiences can be performed by teachers for demonstrations but not by students for lack of space and resources. The activity took place at the department for schools in the Siena sorrounding, while for the others, activities at the school were performed by using the equipment borrowed from the department.

Laboratory of excellence: The laboratory is designed for developing and strengthening the interest and curiosity in the physics through paths in physics lab (construction and/or characterization of an experimental apparatus), not suitable for teaching in the classroom. Teachers identified among their students a group of suitable work and proposed or chose between our proposed project, that includes realisation and, if possible, the design of an experiment in elementary physics. Small groups of students especially motivated worked in an experimental project coordinated by the teacher out of usual time for school. The activity took place over a period of several months and can also be concluded with the presentation of results in the classroom, or on the day of orientation of the department of physics.

Pigelleto's summer school of Physics: The department took the opportunity of using the accommodation in the natural reserve of Pigelleto in Piancastagnaio on Mount Amiata, thanks to financial support from the Provincial Administration of Siena and the contribution of the Fondazione Monte dei Paschi, to host a select group of students in third and fourth classes. The short training period in the beginning of September included intensive work with the purpose of orientation to physics. Laboratories, problem solving, expert seminars with the aim above all of intriguing, ask questions, be fascinated by the discovery of nature and its laws.

Under the starry sky: This laboratory is addressed to students and teachers of secondary schools of the first and second cycle. The primary goal is to allow children to have direct knowledge of the scientific method and its fundamental aspects, through experiences in practical astronomy: observation, collection of data, its processing, the formulation of hypothesis on mathematical models, the prediction of observable effects and their verification. All that is achieved by activities performed by night at the astronomical observatory of the department.

Physics Olympiads: A preparatory course for preparing interested students to physics olympics, scholastic competition, was held in two schools involved in the project. The department hosted the provincial selection of physics olimpiads.

Conferences and exhibitions: Conferences on physics and surroundings realized in local schools to arouse curiosity and attract attention to the physics and PLS. They were often arranged so that many classes can be present. Exibitions on physics topics were organized in the area and young physicists guided students in descovering the activities.

Orientation day in the Department of Physics: Teachers were invited to bring their class or part of them at the openday for orientation. A full day in the department, starting with some talk on research and the profession of physicist as academic or freelance scientist in industry. In the afternoon, students were guided in a visit in didactic and research laboratories.

National Plan for Science Degree

In the passage from project to plan many activities have been maintained, particularly those which were already organized as PLS laboratories. The request of the Ministry was that, in the first two years of the plan, every local headquarters would organize at least a laboratory of this kind.

The performed PLS laboratories grouped by type are shown in Table 2. More information can be found, such as where they were held and the origins of the students if different from the place of execution, responsible and main contributors in their design and implementation. It is important to underline that there was a complete sharing of responsibilities and working with teachers. The schools inserted in the plan during these two years have been 18.

PLS laboratories can be for entire class or for small groups of students, curricular or extracurricular, realized in school timetable or not. Every year, several hundreds of students and many tenths of teachers are involved in one or more of such laboratories.

The most successful laboratory is the Pigelleto's summer school of Physics, described in details in the next sub-subheadings. In 2011, the Ministry asked to every Regional Scholastic Office to select the best practice in their region for each disciplinary area of the plan. In Tuscany, the best practice for Physics is Pigelleto's summer school.

PLS LABORATORIES	Responsible	Faculty members and expert teachers involved in designing and implementation
Trying and trying again: laboratories in classroom	curricular	
Measures of natural radioactivity	Quattrini*	Montalbano
A learning path on heat in vocational schools	Montalbano	De Nicola* Di Renzone* Frati*
Physics in winery	Benedetti*	Mariotti Montalbano
Laboratories for approaching the discipline and developing vocations	extracurricular	
Under the starry sky Siena, Grosseto, Arezzo	Millucci	Marchini
Toys and physics Arezzo	Porri*	Mariotti
Energy from wind Siena	Montalbano	Valentini* Di Renzone*
Deepening laboratory for motivated and talented students	extracurricular	
Detection and measurement of light curves of astronomical images Siena (from Grosseto)	Marchini	Porri* Millucci
Waves and energy Siena (from Grosseto)	Montalbano	Di Renzone* Frati*
A learning path on spectroscopy Siena (from Grosseto)	Mariotti	Di Renzone*
Sound and surrounding Siena (from Grosseto)	Montalbano	Di Renzone* Frati*
Pigelleto's Summer School of Physics selected like first best practice 2011 for PLS laboratories in Tuscany for physics area (from all provinces)	Organizing commetee	extracurricular
	Benedetti* Mariotti Montalbano Porri*	Gargani* Quattrini*
		5-10 teachers active in lab

Table 2: PLS laboratories implemented in National Plan for Physics in southern Tuscany are presented grouped for type, where stars indicate teachers.

Some interdisciplinary activities have been realized and a brief description for each one is given in table 3. The most significant interdisciplinary experience, in my opinion, has been the laboratory on modeling, whose legacy has been collected from a up-dating course for teachers in which the issues were focused on learning paths in which the teaching of mathematics and physics are strongly correlated and adapted to the demands of the recent reform of Italian high school.

Interdisciplinary Activities	Responsible
Orienting laboratory of Science Faculty: education and research	A. Donati PLS Chemistry
Two-days full immersion stage for orienting students in science degrees (collaboration with Chemistry and Mathematics Plans)	
Laboratory of Modeling	E. Mariotti PLS Physics
Workshop for teachers of physics and mathematics on modeling, clarifying topics common to mathematics and physics (collaboration with Mathematics Plans)	
Master in Educational Innovation in Physics and Orienting, University of Udine	M. Michelini (national) PLS Physics Udine V. Montalbano (local) PLS Physics Siena
Inter-university master for teachers in which 19 Universities collaborate for giving courses in laboratory, often focused on a laboratory performed in the National Plan for Science Degree	
Teaching Mathematics and Physics in reformed secondary school	M. A. Mariotti PLS Mathematics
Updating Course for Teachers at School on adapting mathematics and physics teaching in interdisciplinary designed learning paths for reformed vocational and high schools (collaboration with Mathematics Plans)	

Table 3: Interdisciplinary Activities implemented in National Plan for Physics in southern Tuscany.

The Pigelleto's Summer School of Physics: The summer school has reached the sixth edition and appears to be the activity more interesting and enjoyable for students both for form of organization that for contents.

Figure 5: Students perform an experiment on energy transformation at Pigelleto's summer school.

Forty students from high school are selected to attend at full immersion summer school of physics in the Pigelleto Natural Reserve, on the south east side of Mount Amiata in the province of Siena. The school begins usually in early of September and last for four days. The 2011 edition was titled *Thousand and one energy: from sun to Fukushima,* some previous editions were *Light, color, sky: how and why we see the world* ((2006), *Store, convert, save, transfer, measure energy, and more… (*2007), *The achievements of modern physics* (2009), *Exploring the physics of materials* (2010). Topics are chosen so that students are involved in activities rarely pursued in high school, relationship with society are outlined and discussed. The students are selected by their teachers in the network of schools involved in the National Plan for Science Degree [8].

Figure 6: In the evening, the stars look down…and student are watching them.

In the morning lessons are proposed in which the necessary background for the following activities in laboratory is given. In the afternoon small groups of students from different schools and classes are engaged in laboratories where are forced to take an active role. All groups are supported by one or two teachers that are available to discuss any idea.. Usually we propose different laboratories for each group and the task of preparing a brief presentation for other students is given in order to share with them what they have learned. After dinner, an evening of astronomical observation of the sky is usually expected. If it is cloudy, a problem solving evening is proposed.

Figure 7: A group of students present an experience on mechanical and optical properties of growing spheres

In designing summer school activities, we pay attention to several aspects that can render more effective this action both for teachers that for students. Let me give some example :

- the main topic is related to all activities and must be not trivial;
- when it is possible, laboratories are made with poor materials or educational devices provided by some schools in such way that teachers can duplicate easily the lab in their school;

- almost all laboratories lead to at least one measure and its error valuation;
- methodologies are discussed and selected with the teachers involved in Physics Plan;
- in order to have the best collaboration, students' groups are inhomogeneous and formed by following the teachers suggestions;
- usually in laboratory an expert and a young teacher are involved in order to improve teacher practise.

The summer school is a full interdisciplinary action because several different skills of students are engaged, such as physics and mathematical ones, social behavior for collaborate actively and efficiently with other students, communicative ones for a good transmission of information between groups and in presentation for sharing knowledge. Furthermore, very often the main topic is interdisciplinary from the very beginning, such as materials science, or becomes so in discussing issues relevant to society, for example in the case of energy sources or nuclear energy.

Modeling Since in 2009 also the Department of Mathematical Sciences of University of Siena launches its plan, a workshop for teachers of physics and mathematics on modeling was performed in collaboration with Physics Plan.

This activity is inter-disciplinary for constructions and has continued in an updating course for teachers in which selected topics, named in the same way in both disciplines, are discussed in order to design interdisciplinary learning paths. The purpose is to clarify these topics by using specific tools from physics and mathematics and to outline the similarities and the differences in both contexts.

We believe that this activity can be useful for students, which can acquire a more profound insight on some fundamental concepts, and for teacher professional development.

4. CONCLUSIONS

The actions realized in these years in order to maintaining and increasing the number of motivated and talented students who enroll in courses scientific degrees in Italy are beginning to give some result, as shown in figure 8.

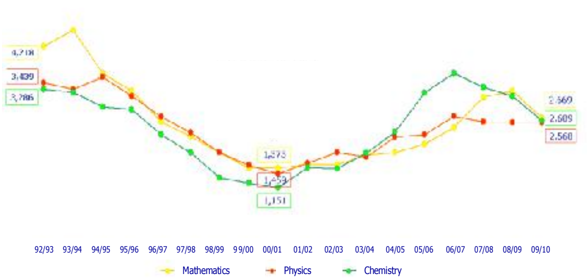


Figure 8: Enrollment in Italian universities in basic sciences[7]

Furthermore, many teachers are interested in activities in which new interdisciplinary methodologies are developed for increasing students' understanding in high school of selected topics in which both Mathematics and Physics give a contribution in the construction of important and mutually reinforcing basic concepts of the two disciplines.

A goal achieved by the National Plan for Science Degree is that a network of schools and teachers has established and it is permanently active in improving the teaching of sciences in high school. A further result is that many teachers, even if forced by the recent reform of the secondary school, are beginning to collaborate actively in mathematics and physics education research.

5. ACKNOLEDGEMENTS

This work is based on activities and experiences which were performed within the National Plan for Science Degree supported by Italian Ministry of Education, University and Research. The author would like to thank the director of department of physics, Angelo Scribano, for the support.

6. REFERENCES

[1] M. Bucchi, F. Neresini, "The scientific vocation crisis and its explanation ", **Observa – Science in Society**, 2004, pp. 1-41, http://chemistry.pixel-online.org/database_review.php?id_doc=70 , accessed 2012 March.

[2] R. Czujko, "Enrollments and Faculty in Physics", pp. 1-11 http://www.aip.org/statistics/trends/reports/june9talk.pdf 2002, accessed 2012 March.

[3] P. J. Mulvey, S. Nicholson, "Physics Enrollments: Results from the 2008 Survey of Enrollments and Degrees", **AIP Focus On**, 2011, pp. 1-10, http://eric.ed.gov/PDFS/ED517370.pdf, accessed 2012 March.

[4] "The Message of the 2004 S&E Indicators: An Emerging and Critical Problem of the Science and Engineering Labor Force ", **National Science Board**, 2007, pp. 1, http://www.nsf.gov/statistics/nsb0407/, accessed 2012 March.

[5] B. Convert, "Europe and the Crisis in Scientific Vocations", **European Journal of Education**, vol. 40, No. 4, 2005, pp. 361-366.

[6] PLS Website, http://www.progettolaureescientifiche.eu/, accessed 2012 March.

[7] N. Vittorio, "Il Progetto Lauree Scientifiche 2005 – 2009. Sintesi delle attività e dei risultati a livello nazionale ", Workshop *Vetrina PLS: da Progetto a Piano*, Siena 4 Nov 2010, http://www.unisi.it/fisica/laureescient/vetrina/20101104_vittorio_pls.pdf, accessed 2012 March.

[8] R. Benedetti, E. Mariotti, V. Montalbano, A. Porri, "Active and cooperative learning paths in the Pigelleto's Summer School of Physics", Proceedings 12th Symposium Frontiers of Fundamental Physics, Udine 21 - 23 November 2011, arXiv:1201.1333v1 [physics.ed-ph], 2012.

Architecture and Service Development for Service Oriented Architectures

Wayne O'Brien
Intelligence and Information Systems,
Raytheon Company, 22260 Pacific Blvd.,
Sterling, Virginia 20166, America

ABSTRACT

Service Oriented Architecture systems, despite the name, often do not include an intentional, overarching architecture. This paper defines and describes a systematic, Model Based Engineering (MBE) approach for creating an optimal, mission-focused overarching architecture for systems or systems of systems with a Service Oriented Architecture (SOA) and developing services to realize them.

Keywords: Service Oriented Architecture, Model Based Engineering..

1. INTRODUCTION

A systematic MBE – Model Based Systems Engineering (MBSE) through Model Driven Software Development (MDSD) – approach for creating a mission-focused SOA and developing services to realize an optimal SOA at the system or system-of-systems (system is used throughout the rest of this paper to include system-of-systems) level would result in architecture-driven services rather than a services-driven architecture. This paper focuses on software-intensive systems where the primary disciplines are systems and software engineering, but the approach is extensible to systems involving other disciplines. Deriving the architecture directly from customer needs would ensure that the services optimally supported the mission. As discussed in Approach, a number of proven enablers are available, but need to be combined in a way that realizes and amplifies their potential.

2. APPROACH

SOA is a de facto implementation for net-centricity and the cloud for the Department of Defense (DoD) and federal civilian agencies in America. SOA is often viewed as the architecture, not a way (style) to implement the architecture. This view may lead to a services-driven architecture rather than architecture-driven services. The architecture is a key reference in determining the capability, granularity, and location of services, yet services are often developed without first having an overarching architecture for such reference. Ironically, there is often no "A" in "SOA."

This paper considers a spectrum of SOA approaches, with the market-based concept of SOA at one end and the internal enterprise SOA at the other. The market-based concept of SOA depends on services that comply with open standards and the assumption that consumers will use a discovery service or service broker to find existing services that meet their needs, as specified in Service Level Agreements. The consumer's specific needs are assumed to be unknown to the service provider (or developer) and the provider unknown to the consumer. The internal enterprise SOA consists of internally developed services or services obtained externally and tested for compliance with internal enterprise standards.

The nature of individual services often means they are coded in an *ad hoc* way. Services are frequently written in isolation spatially and temporally, especially in the market-based approach, offering little or no opportunity to apply conventions for consistency across services solely through collaboration. Even for internal enterprise SOA, completely new development of all services is unusual. Partial or complete legacy applications continue to operate as part of the system along with services (resulting in a SOA-based rather than a pure SOA system), some of which may be built with refactored legacy code.

Services may be developed by individuals or small firms that do not use MDSD tools or small groups within a large firm that hand code the services informally because development of an individual service does not warrant the firm's complex development processes. In large firms, this reflects the lack of an internal enterprise environment with an overarching architecture within which the services could be developed as a large, formally managed project. Large firms taking this informal, *ad hoc* approach are either using the market-based concept formally by requiring compliance with open standards or informally by not requiring such compliance. The result is code within the same system that reflects little standardization or systematic organization and architecture.

An overarching architecture serves as a reference for developing services in terms of their capability, granularity, and location. Such an architecture provides the unifying guidance to systematically determine which capability is needed where and how much capability to include in each service. Such an architecture also serves as a means of capturing, communicating and transforming customer needs [see section ??], whether in an internal enterprise environment or an external market-based environment. Customer needs will drive, through use cases (use-case driven), the development of the architecture, which through Model Driven Software Development (MDSD), can be transformed into a running system. The visual, animated models produced through MDSD serve as a means of communicating with the customer and other stakeholders through the entire lifecycle of the system.

In the market-based SOA environment contemplated by the Defense Information Systems Agency (DISA) of the DoD, the reference architecture is provided in terms of Grid Enterprise Services (GES) and open standards. Developers must comply with both the GES and open standards. Such compliance allows services to be developed in a consistent manner without knowing who all the consumers of a service are or how they may use the service, which may be as an element orchestrated with other services in an application.

In an internal enterprise environment, the enterprise can use the overarching architecture for its SOA applications to cover details of how services are developed beyond open standards, e.g., speed and reliability. An inherently market-based environment such as that created by DISA can also use the internal enterprise approach for mission critical services. As mentioned above, enterprises that want to take advantage of the services available in the market-based environment as well as their primary internal SOA environment can do so by ensuring that their internal architecture requires services to comply with open standards. The internal architecture would be one more standard with which services had to comply. The enterprise's governance program (e.g. The Open Group "SOA Governance Framework") would ensure the compliance of both internally developed and externally acquired services.

The Model Driven Architecture (MDA) specifications from the Object Management Group (OMG) provides guidance on how to develop a model driven architecture and implement it with services through the SoaML Profile, another OMG specification. The approach proposed in this paper builds on the MDA-SoaML guidance by supplementing it with the following:

- A Service Identification and Selection Process (SISP)
- Use Case Realization and Development process
- Identification of Analysis Classes by Affinitizing Actions from use case sequences of actions
- SOA Foundation Reference Architecture
- Model driven code generation for the services

3. A SERVICE IDENTIFICATION AND SELECTION PROCESS (SISP)

Determining useful services for a SOA is a complex tax. Services must be identified, evaluated, categorized (e.g., processing, data management, security), and sized (granularity) in order to successfully implement a SOA.

- Identify candidate services
  - Determine what services are needed to meet top-down requirements, e.g., from business domain analysis, new systems, or pursuits
  - Determine what services are needed from a bottom-up view to support what is already being done by legacy applications and current services
  - Perform middle-out assessment to determine whether services are complete on the one hand (not enough services) or unnecessary on the other (too many services) in terms of whether they support organizational and mission goals – current, anticipated, or contingent). This is a validation of the top-down and bottom-up analysis.
- Evaluate services
  - Business Characteristics
    - Must have business and/or operational benefit
      - Current
      - Transitional
      - Future
    - Must have widespread usage (marketability)
      - Service Reusability (reusable across business organizations and enterprises)
      - Consumer indifferent because eventual users may be unknown at time of deploying service
      - Ownership indifferent because consumers will not have the level of control that comes with ownership
- Categorize services to aid in prioritizing them for selection.
- Determine the appropriate granularity[1] for each service, considering such things as the number of steps of a business process or the number of function points to include in a single service.
- Select services based on the priority determined from a combination of the above considerations. Inter-category priorities need to be resolved based on considerations of cost, schedule, and interdependencies. The granularity may need to be revised to lessen the interdependencies.

4. USE CASE REALIZATION AND DEVELOPMENT PROCESS

This paper views a use case as the completed use case specification represented by any of the standard use case templates described in the literature and used with MDSD processes such as the Rational Unified Process. The specification should be completed through the following steps:

- Identify use cases from the CONOPS or other high level description of the mission and input from subject matter experts (SMEs)
- Complete the specifications in the format of the use case template based on the available sources and continued input from SMEs
- Begin realizing the use cases by relating the use cases to each other (a high-level architecture) in use case diagrams within an MDSD tool
- Continue realizing use cases by transforming the textual flows of events in the use case specifications into activity diagrams (a standard diagram in the Unified Modeling Language [UML], another OMG specification)
- Continue the use case realization process by deriving any other UML or SoaML diagrams needed to fully describe the SOA system, including steps in the next section,

[1] Marks, Eric A. and Michael Bell. Service Oriented Architecture. Hoboken: John Wiley & Sons, 2006.

Identification of Analysis Classes by Affinitizing Actions from Use Case Sequences of Actions

5. IDENTIFICATION OF ANALYSIS CLASSES BY AFFINITIZING ACTIONS FROM USE CASE SEQUENCES OF ACTIONS

The activity diagrams created during the Use Case Realization and Development Process form the basis of this next set of steps.

- Identify actions in the activity diagrams that have some degree of affinity, such as Post an Alert and Distribute Alerts
- Affinitive actions by creating Analysis Classes (high-level classes for decomposition into design-level classes) with responsibilities (to be decomposed into design-level operations) corresponding to actions with an affinity for each other
- Use the object (data) flows between actions to develop collaboration and class diagrams (UML diagrams) for the high-level data model (Data Information Viewpoint 1 or 2 in Version 2.0 of the DoD Architecture Framework)
- Complete the development of design classes and data classes for implementation using an MDSD tool environment

6. SOA FOUNDATION REFERENCE ARCHITECTURE

The SOA Foundation Reference Architecture (SOAFRA) is based on the architecture of the DISA Enterprise Services. The SOAFRA provides the infrastructure architectural components for the overarching architecture of a SOA system.

7. MODEL DRIVEN CODE GENERATION FOR THE SERVICES

When the overarching architecture is developed in an MDSD tool environment as described above, the tools will also generate the code for the services, ensuring that the code complies with the architecture. During the lifecycle of the system, services can be modified or added at the model level rather than the software code level, maintaining the integrity of the architecture over time and greatly reducing development and maintenance costs. The *ad hoc* nature of services, even when developed separately, will be avoided, both in terms of how they interact with other services (the architecture) and how the software code for the services is developed.

8. CONCLUSIONS

The Model Based Engineering process described in this paper for developing SOA systems provides the following benefits:

- An overarching, mission-focused, use-case driven architecture that derives from the problem space (CONOPS and customer needs)
- Analysis classes derived directly from the use cases that drove the architecture
- Services based on responsibilities for the analysis classes
  - Responsibilities correspond to actions affinitized into analysis classes
  - Assures that services reflect the overarching architecture

The resulting SOA system would consist of services that are tailored for the underlying architecture, which in turn captures the CONOPS and customer needs through use cases, rather than the bottom-up architecture that emerges from connecting otherwise unrelated services. The architecture will guide the capability, granularity and distribution of services, geographically, among servers, or in the cloud, implementing a system or system-of-systems with an optimal, mission-focused SOA.

Since the services themselves would be captured in a Platform Independent Model (PIM, part of the OMG MDA specification), they could readily evolve and be targeted for new platforms and technology, while not breaking the architecture.

Debt Environment in Computer Science Education

Tomas Cerny
Dept. Computer Science and Engineering, Czech Technical University
Charles Square 13, Prague, 12135, Czech Rep.

and

Bozena Mannova
Dept. Computer Science and Engineering, Czech Technical University
Charles Square 13, Prague, 12135, Czech Rep.

ABSTRACT

Courses in computer science are often designed with a consideration of interesting topics and materials as well as with a significant amount of individual work. In order to encourage studies on complex exercises and projects, the aspects of competition and collaboration can be applied. Students are often motivated at the beginning of the course, but unfortunately their interest decreases as the semester continues. In this paper we suggest a debt environment approach that keeps students in a continuous pace to meet deadlines, which effectively distributes motivation throughout the course. A study implementing the approach with other methodological aspects is provided in a bachelor computer science course. Students work in teams and also as individuals and compete for their final grade. Applied methodological aspects are discussed regards to students' course evaluation. Evaluation of students' feedback shows positive effect of the approach regards students' motivation. As indicated in the feedback, our approach is significantly more motivating factor for students than competition. The approach can be applied as a complement to competitive and collaborative techniques to improve students' motivation.

Keywords: Competitive learning, Collaborative learning, Pace environment, Computer Science Education, Debt environment.

1. INTRODUCTION

Computer science experienced fast evolution in the last two decades. Many of it's disciplines are relatively young and with a close relation to mathematics. The difference is that the knowledge, which students gained 5 years ago, may not seem sufficient enough nowadays. Education in computer science should prepare experts in the field. They should learn to compete with each other in order to maintain and evolve their knowledge. The industry seeks for experts who, in addition to good technical knowledge, have the capability to build a team and compete with others. In addition, they need to manage pace environment, meet deadlines, handle stress, and use their time effectively.

This paper aims to improve education for computer scientists and prepare them in order to be become experts in their field. Debt environment approach is presented as a complement to competition and collaboration. Students experience the need to meet deadlines and effectively use their time to avoid short-time debt, which may result in lower final grade or failure from the class. A case study employing such environment is provided in this paper. Evaluation of the study is concluded from students' feedback. It is evident that students feel positive impact of the environment on their motivation.

This paper is organized as follows. Section Related Work describes related research, such as competitive and collaborative learning, programming Olympiads and research in the area. Proposed environment is described in section Debt environment. Case Study section employs the debt environment in our course. The Study Evaluation discusses the students' evaluation of the course. The last section concludes the paper.

2. RELATED WORK

Several researchers address improvements to education process. They focus mostly on students' interaction, competition, collaboration and teamwork, students' motivation and technological support. Their results and achievements are significant, but with the exception that the adoption of their results differs among various regions. For example, it is common for North American education system to have a strong competition among the students [7]. On the other hand in Central Europe the education focuses primarily on knowledge (content-driven) and on students as individuals [3].

The way students interact with others and with the materials can be seen as communication patterns [7]. Teacher-student interaction is commonly used in courses but student-student interaction is often ignored. In fact, there are 3 patterns of student-student interaction: competition, individual work and cooperation. Competition [4] is a social process that takes place when rewards are given to people on the basis of their performances when compared to the performances of others doing the same task or participating in the same event. Competition interaction pattern can be seen as the key factor among student-student interaction [7], although under certain circumstances collaboration might be seen more productive. In fact, competition is often applied together with collaboration where student teams compete with each other [2]. The benefit that competition gives is seen in increased motivation in order to do better than others, self-confidence, larger workloads for students, or simply a good preparation for future employment. This technique increases motivation for some, but not necessarily for others [13]. Motivation often relates to winners rather than losers [12]. That could on the other hand affect the results among the peers.

Collaboration [4] is a social process through which performance is evaluated and rewarded in terms of the collective achievements of a group of people working together to reach a particular goal. Davis [5] states in his research that students who collaborate in a group seem to be more satisfied with their course as they feel more involved in their educational process. He suggests also that team collaboration should assign a role to each member. Cooperation with a large group may result in an inadequate separation of the assignments or imbalance in the workload among different members [13]. It is important that team members cooperate well, because it impacts the whole group. In order to indentify problems in cooperation, dysfunctional index (DFI) method can be used. This index is defined as the mean test score of the team divided by the standard deviation of team members' test scores [6]. A team with a small DFI is likely to be dysfunctional. This allows the teacher to identify such a team and locate the issue. Another way to avoid dysfunction [13] suggests that the teacher assign groups based on students' availability to meet outside the class schedule. On the other hand Oakley et al. [11] see self-selection of members more effective as it does not cause interpersonal conflicts. Different approach to mitigate dysfunction of the cooperation is to build small teams. A small team has positive impact on individual accountability [7]. To encourage cooperation the teacher should reward students for the participation in a team [10]. Recognition for both to the individual member and to the whole team should be evident [2]. Team cooperation in a competitive business-like environment prepares students for professional employment [2]. "The team cooperation encourage each other to do the assigned work, and learn to work together regardless of ethnic backgrounds or whether they are male or female, bright or struggling, disabled or not" [7]. Collaborative learning suits well for problem solving and solving techniques [9], which for example is computer science.

Both competition and collaboration techniques were successfully applied in the senior secondary school in the field of mathematics that was situated in Nigeria [8]. The results present the strategy of cooperative learning as more effective than competitive strategy. Furthermore, male students performed significantly better than female students in both learning strategies. Both techniques were also applied in bachelor class and the results were compared with the standard education [3]. From the research [3] students saw both competition and cooperation beneficial through motivation, concentration or increased attendance in voluntary lectures. Application of the techniques also supported active participation of students in the education process.

The motivation for students is brought by competition as well as by collaboration. As next, additional factors to increase motivation should be applied. Students should always be rewarded for their good work. Furthermore, they should feel that they contribute to the education process, so that they can share their knowledge with the entire class, influence the topics, discussion or peer review others. In addition, motivation grows in pace environments.

Another area besides educational institutions emerged impacting students' skills and capabilities. This area focuses on competitions among schools. In these competitions the students compete with each other or within groups and in the end, the best results are rewarded. In the area of computer science, multiple countries organize Programming Olympiads [1], [15], [16]. Among them, the most known is the ACM International Collegiate Programming Contest (ICPC) [14]. In these competitions, students build teams and compete with other teams when solving small algorithmic problems. The team that solves the most problems within the shortest possible time is the winner. These competitions are very popular among the students. For example in ICPC, tens of thousands of students from around 2000 universities compete each year. In fact, these contests employ multiple techniques such as competition, collaboration, pace environment or problem solving. Student attendance in such contests shows that the techniques applied are successful and should be employed in educational system due to the fact that the students often invest their personal time to train and prepare for the competition to do better than others. In addition, students learn to solve programming problems on their own as part of a game, although their primary focus is not the education, but preparation for the competition.

3. DEBT ENVIRONMENT

In the previous section we introduced competition and collaboration techniques. These techniques work greatly for programming competitions and in education bring forward multiple benefits. There are additional factors that may influence the success of these techniques such as well-selected materials, topics, examples and assignments. Another influence is the size of the class. It is easy to motivate a small class, but hard to motivate a large class. In addition, we can often experience that students' motivation decreases over time, mostly for large classes [13]. For example, students are greatly motivated the first month, but then they fall into a stereotype. Students who win competitions often keep up with motivation compared to the ones who are not doing well. As a result a certain subset of students loose their motivation and focus. Later in the course these students fail to deliver homework assignments and slow down the tempo of the rest. At the end of the course some of the students try to bring delayed work.

In order to address the above issues we present a debt environment that distributes motivation throughout the semester, motivates students to submit their assignments on time and introduces fair reduction of the grade for students who do not deliver their work on time. This environment complements competitive and collaborative approaches and is suits well for large classes.

The debt environment puts students or student groups into a role of profitable organization. This kind of an organization starts with no resources. Organization can gain capital by delivering assignments on time, passing the test, quizzes or submitting a project. The whole amount of the capital will be saved in a bank that holds the monopole and charges a large amount of capital for its services each month. The goal of every organization is to stay in positive balance and gain as much capital as necessary to survive the bank monthly charges. If an organization fails to pay charges, a loan is provided. A dept is placed on the organization, so that the loan doubles the forthcoming month. If the monthly loan is not covered the organization bankrupts.

Table 1. Imaginary dollars reward per assignment

Task ID	Task	iDollars per task	iDollars total	Applies to
a	Presentation of a SW design pattern	10	10	Team
b	Article on the SW design pattern	10	10	Team
c	Architectural programs	4x3	12	Team
d	Architectural program documentation	5	5	Team
e	A project applying SW patterns	10	10	Team
f	SW Challenges (first solved wins)	7x3	21	Team
g	Research paper discussion	5x1	5	Individual
h	Tests/Exams	2x20	40	Individual

The environment is complementing collaboration, as the organization can be a group of students. It also complements competition as the organization with the highest balance may gain the best grade. Each student (or a group) must properly manage his/her time and plan to deliver enough work in order to avoid debt. This also gives a student the benefit to make up poorly graded test or to plan his/her workload throughout the semester.

4. CASE STUDY

In order to evaluate our approach, we provide a study that employs competition, collaboration and other motivational techniques with debt environment. It is applied in a 6th semester eligible course of a bachelor degree. There were enrolled 36 students in the course and out of all 32 passed the class. Course syllabus is designed in a way that the students work both individually and within a team. They solve tasks, assignments, projects and tests, which rewards them imaginary dollars (iDollars) instead of points. The final grading supports competition, the final grade is based on the total amount of iDollars in comparison with other students. Furthermore, each student has to earn at least 15 iDollars each month throughout the semester in order to cover bank service charges; this amount is deduced from his/her account. At the beginning of the semester students are given tasks that they can submit throughout the semester based on their strategy. Course tasks are described in Table 1. The only dates that are fixed are Tests (h), which take place in the middle and at the end of the course. Small programming challenges (f) applying current topics are sent once in two weeks and the first correct solution is rewarded with 3 iDollars, second with 2 and the third with 1. Discussion on research paper (g) follows every other week. Other tasks (a-e) can be submitted at any time during the semester.

In this study teamwork supports both collaboration and competition. Teams consist only of two members, which apply a team student-to-student interaction. Teams compete with each other in order to reach the highest amount of iDollars and to solve the most of the SW challenges (f), these are similar to the problems assigned in ICPC contests. Furthermore, it is up to a team to split or cooperate on tasks (a-f), to divide the load. Students take part of the education process as they present a design pattern (a) to the entire class and the others can influence their iDollars reward by direct feedback. This presentation broadens to student-to-many-student interaction with the feedback. Teams also learn to write a research article about the topic presented in the class (b) or to apply it in their project (e). A proceeding of submitted research papers (b) is published after the class that gives the students both reference and the first experience with publishing. The coursework contains a demonstration of architectural decisions with documentation (c-d), where teacher-student interaction plays the primary role. Individual student approach is applied to the tests and discussions of the research paper.

5. STUDY EVALUATION

After our study was applied and the grades distributed, we received students' feedback on given techniques and tasks. Students who passed the course, and received a grade (Figure 1), filled out a questionnaire, asking about their feelings toward specific techniques applied in the course, and particular tasks. Students could answer if a particular technique was useful, motivating for them or whether they liked a given task. Answers were in form of options and in form of short survey. Survey options were: “Certainly yes, rather yes, do not know, rather no, certainly no”.

Figure 2 to Figure 6 show the response from the students on given aspects of the class. Graphs provide the amount of students divided by answer and grouped by the final grade. From responses can be seen that students rather like competition for the final grade, although they did not think that competition would increase their motivation level. Few students felt motivated by competition mostly because they were winners of SW challenges. Collaboration within a team was very well accepted, some students pointed out that it is better when the team is small rather than large, based on their experience from other classes. On the other hand, two students did not like team collaboration; the reason was an irresponsible team member. Students compared two motivation factors (Figure 5 and Figure 6). They felt that the monthly charge and the possible debt motivated them to work more intensely throughout the semester rather than the competition. Not all students liked debt environment; these were the students who experienced a loan. Students accepted well the imaginary dollars in comparison to regular points. In comparison with the results of the competition and collaboration, the collaboration was better accepted.

Particular tasks (a-h) from Table 1 are evaluated in Figure 7 to Figure 12. All tasks were well accepted. The most accepted was the contribution to student proceedings on SW design patterns. It could be understood that the students like to work on things that have a long-time value rather than shortly lived. There is almost no difference between a student attendance at practices and optional lectures, as shown in Figure 13 and Figure 14.

It can be seen from the attendance that there is no significant impact on the final grade. What is the difference in motivation and task evaluation among students with different grades? Better-graded students like competition, and some of the best students also felt motivated by the competition. Better-graded students also like the imaginary dollars more than the others. Regards the tasks; worse-graded students seemed to like SW projects. Better-graded students like to influence the education

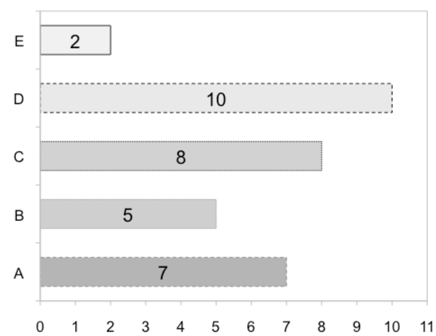


Figure 1. Final grades distributed in the course

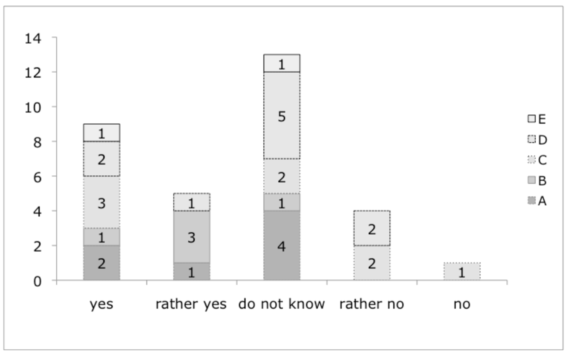


Figure 2. Do you think that iDollars are better than points?

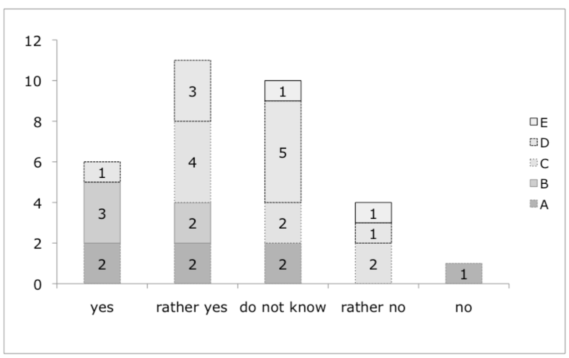


Figure 3. Did you like competition for the final grade?

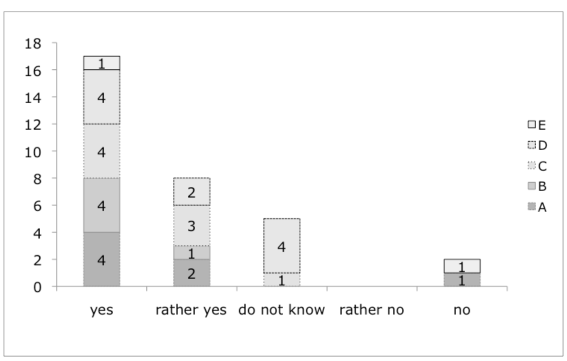


Figure 4. Did you like team collaboration?

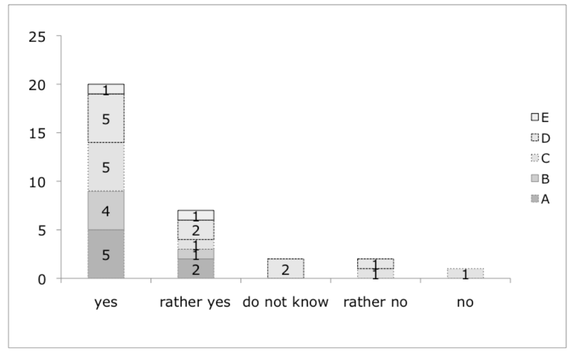


Figure 5. Did debt and monthly charges increase your motivation throughout the semester?

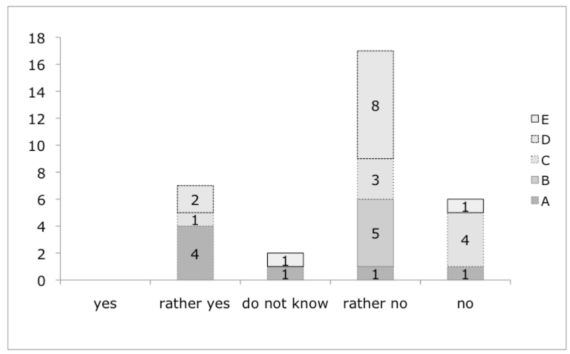


Figure 6. Did competition increase your motivation throughout the semester?

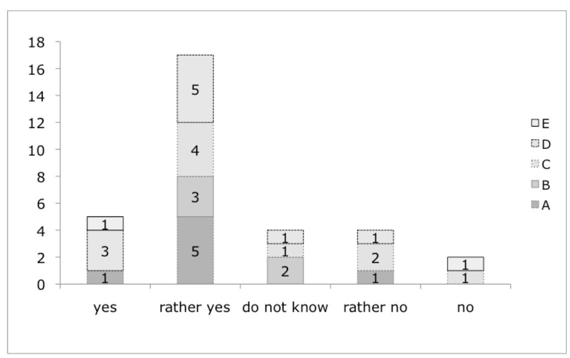


Figure 7. Did you like the option to influence education? (a)

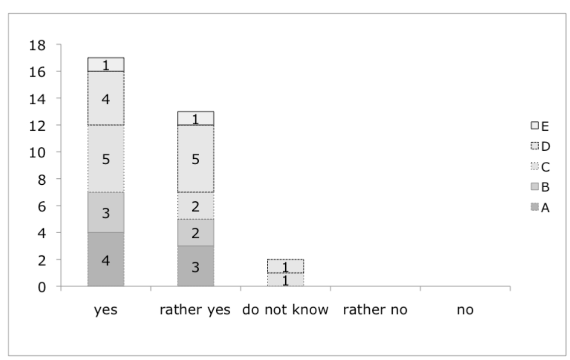


Figure 8. Did you like to contribute to student proceedings? (b)

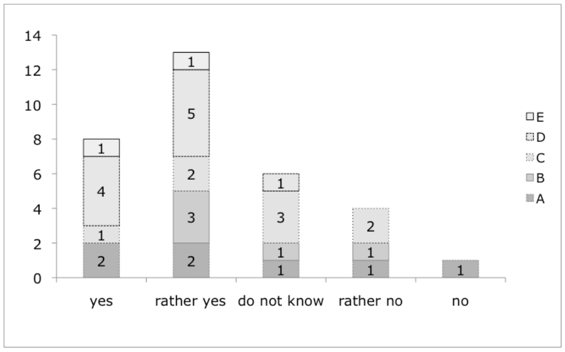


Figure 9. Did you like team SW projects? (c,d,e)

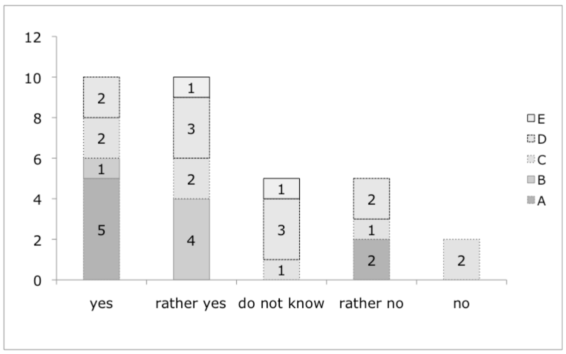


Figure 10. Did you like small SW challenge competition? (f)

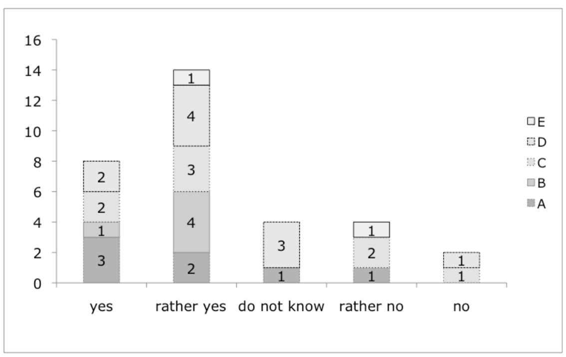


Figure 11. Did you like to read research papers? (g)

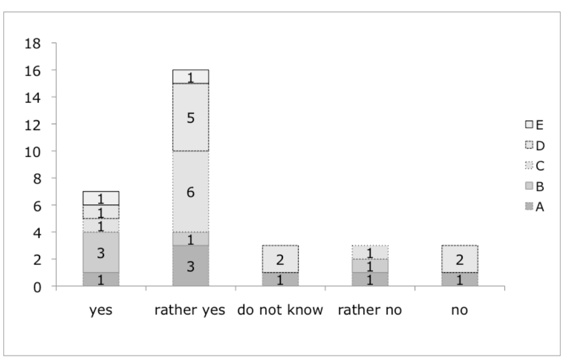


Figure 12. Did you like tests? (h)

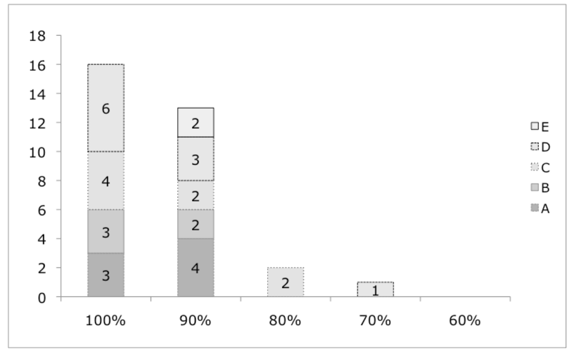


Figure 13. Participation at practices

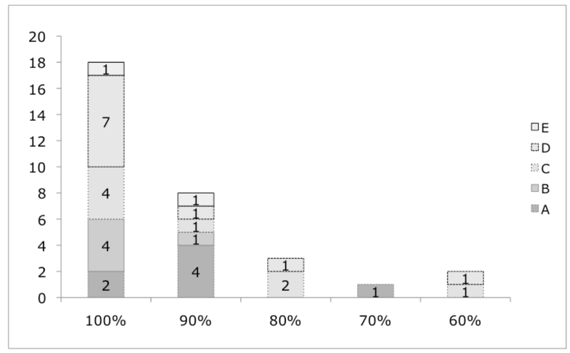


Figure 14. Participation at lectures

and contribute with a presentation; they also like research papers and SW challenges. Interesting is the feedback on tests, these are often not liked by students, but we must assume that a test in our environment is not understood as a verification of student knowledge, but as an opportunity to receive additional iDollars, which seems to be taken positively by students.

6. CONCLUSION

This paper suggests a new method to increase motivation for students throughout the semester. It can complement competitive and collaborative learning and other approaches. The method suggests the use of imaginary dollars instead of points. These dollars are deposited in a bank that monthly deduces service charges. It pushes students to deliver work in fixed time frames in order to prevent low-balance and possible debt causing increased bank charges the forthcoming month. The method was applied in a 6th semester course in a bachelor program. The results from the students' feedback show that this method motivates students more than competition. In our results competition seemed to be motivating small amount of students.

Our future work will focus on the comparison of other motivational techniques with debt environment. As next, we may reduce the time interval between bank charges. The method should also be applied in different disciplines in order to receive its impact on different fields. It is possible that the method may not be successful for some disciplines, other cultures and on different goals. From our results, we believe that the method fits well to the area of problem solving, which applies to computer science.

7. REFERENCES

[1] R. O. Anido, & R. M. Menderico, Brazilian Olympiad in Informatics. **Olympiads in Informatics**, Vol. 1, 2007. pp. 5-14.

[2] S. Attle, & B. Baker, Cooperative Learning in a Competitive Environment: Classroom, **International Journal of Teaching and Learning in Higher Education Applications**, Vol. 19, num 1, 2007, pp. 77-83

[3] T. Cerny, & B. Mannova, Competitive and Collaborative Approach Towards a More Effective Education in Computer Science, **Contemporary Education Technology**, Vol. 2, 2011, pp. 163-173.

[4] J. J. Coakley, **Sport in society: Issues and controversies** 6th ed., Irwin/McGraw-Hill, 1997.

[5] B. G. Davis, **Tools for Teaching**, San Francisco, Jossey-Bass, 1993. http://teaching.berkeley.edu/bgd/collaborative.html

[6] C. M. Hsiung, Identification of dysfunctional cooperative learning teams based on students' academic achievement, **Journal of Engineering Education**, Vol. 89, 2010, pp. 45-53.

[7] D.W. Johnson, & R.T. Johnson, An Overview of Cooperative Learning. In: THOUSAND, J., VILLA, A., & NEVIN, A. (Eds) **Creativity and Collaborative Learning : A Practical Guide to Empowering Students and Teachers**, 1994, pp. 31-43, Baltimore, MD: Paul H. Brookes Publishing Co.

[8] E. B. Kolawole, Effects of competitive and cooperative learning strategies on academic Performance of Nigerian students in mathematics, **Educational Research and Review**, Vol. 3 (1), 2008, pp. 33-37.

[9] L. K. Michaelson, B. K. Arletta, & L. D. Fink (Eds.) **Team-based learning: A transformative use of small groups in college teaching**, Stylus publishing, 2004.

[10] B. Morgan, Cooperative learning in higher education: Undergraduate student reflections on group examinations for group grades, **College Student Journal**, 37(1), 2003, pp. 40-50.

[11] B.A. Oakley, D. M. Hanna, A. Kuzmyn, & R.M. Felder, Best Practices Involving Teamwork in the Classroom: Results From a Survey of 6435 Engineering Student Respondents, **IEEE Transactions on Education**, Vol. 50 (3), 2007, pp. 266-272.

[12] M. Vansteenkiste, & E. L. Deci, Competitively contingent rewards and intrinsic motivation: Can losers remain motivated?, **Motivation and Emotion**, Vol. 27, num 4, 2003, pp. 273-299.

[13] E. Verdu, & R.M. Lorenzo, Effective strategies for learning collaboration and competition, **A New Learning Paradigm: Competition Supported by Technology**, 2010, pp. 11-40.

[14] ACM International Collegiate Programming Contest Competition for computer scientists, 2011, http://cm.baylor.edu/

[15] Indian Computing Olympiad, 2011, http://www.iarcs.org.in/inoi

[16] The British Informatics Olympiad, 2011, http://www.olympiad.org.uk

Student-voted Multiple-round Elimination Team Projects

Andrew Chen
Computer Science and Information Systems
Minnesota State University Moorhead
Moorhead, Minnesota 56563, USA

ABSTRACT

We introduce a novel form of classroom environment which offers the advantages of project-based classes, which can scale to large sizes, while reducing instructor grading load dramatically. This form of classroom environment functions well when students of different disciplines are able to enroll in the same section of this class, thus allowing the project teams to leverage off of the unique contributions that the different students can offer, both as individuals and as representatives of different disciplines. As such, this form of classroom environment is ideal for testing theories regarding how interdisciplinary interaction between students should be taught and practiced within an educational project context.

Keywords: Interdisciplinary, Team, Project, Education, Collaborative Learning

1. INTRODUCTION

The approach described in this paper was used to teach a game development class in which students from various disciplines, including computing, graphics, music, film, psychology, history, art, and others, were encouraged to enroll. This paper will consist mainly of a description of how the class was structured, along with some description of how the class turned out, and what might be done differently the next time this model of class is employed.

2. HOW THE CLASS WAS STRUCTURED

Students each pitched a game concept to the class. The class as a whole voted. Those with higher scores had their concept survive, and those with lower scores became available for selection as partners to work with on the surviving game concepts. This process repeated through several rounds of voting, elimination, and additional partner selection. The number of rounds that students were on a game concept that survived to the next round was the single largest part (although not a majority) of their grade in the class. Teams could voluntarily give up on their current project effort, and let all of their members become available for selection on the next round. Teams could eliminate ("fire") someone from their team, if everyone else on the team agreed (or at least didn't object). Smaller teams got to select new members ("hire") before larger teams, and higher-scoring teams got to hire before lower-scoring teams of the same size.

This model addressed many contemporary issues. It scales to larger class sizes (and certain phenomena, such as small voting alliances, are rendered less significant the larger the class size is). It does not suffer the same scaling issues as large lecture classes do. This model leverages student interest in game development and reality television shows such as "Survivor", as well as the students' knowledge and ability to evaluate and critique games (which they likely already practice, as consumers of games, given their interest in game development), but could be extended to other classes involving production of media that students typically consume (such as mobile or web applications, social media content, etc.).

3. CLASS ORGANIZATION DETAILS

The grade breakdown was as follows:

- 30% Survival Rate (curved)
- 10% Class participation, as measured through both quantity and quality of online contributions
- 10% Instructor Evaluation of Presentations
- 10% Evaluation of Presentations by surprise guests
- 10% Peer Feedback
- 10% Peer Evaluation of Final Game Product
- 10% Instructor Evaluation of Final Game Product
- 10% Presentations on Design Decisions

The survival rate was curved according to an exponential curve. Specifically, the number of percent lost was $(2^n - 1)$, up to the limit of 30%, where n was the number of rounds that a student was fired or on a losing team.

For non-computer games, provisioning of the physical materials was the responsibility of the students. For computer games using platforms other than what is available in the university laboratory environments (which consisted mainly of PCs of various ages depending on which lab), provisioning of the platforms was the responsibility of the students.

The class met once a week for 2.5 hours. Here is a rough description of the weeks:

1. Overview of the class
2. Voting round on initial game concepts
3. Hiring round
4. Voting round
5. Hiring round
6. Voting round
7. Hiring round
8. Presentations about what they will accomplish in the next three weeks
9. Guest critique
10. Spring break
11. Voting round - and what they said they would accomplish is displayed while voting to compare with what they are able to present
12. Hiring round
13. Guest critique
14. Lecture/workshop 1
15. Lecture/workshop 2
16. Playing the final games 1
17. Playing the final games 2
18. Playing the final games 3

Semesters at this institution are generally regarded as 14 week semesters, since most class days will occur 14 times, however, classes that meet on Wednesday meet 15 times (since many holidays are extended weekends, this explains why Wednesday class days meet more than days that are closer to the weekend).

The above crude schedule looks like the class ran for 18 weeks. This is because of the combination of a variety of different factors: (1) two of those weeks occur during the final exam time period, (2) one week occurs during spring break, (3) the class met on Wednesdays, thus providing an extra day.

During a “hiring round” class day, class typically had three parts:

- the students that were “available for hire” all gave short “why you should hire me” job pitches
- an open period of time that resembled a “job fair” whereby teams that were hiring could have one-on-one interviews with those that were up-for-hire
- the selection of who to hire by those teams that were hiring, in order by size first (smallest to largest), and then within each size category, by the score that was obtained during the voting round (highest score to lowest)

During the “playing the final games” time periods, class was broken into three different segments, and each segment was dedicated to playing one game. There were 9 surviving games after all of the elimination/voting rounds.

Given what was observed during the presentations, the lecture/workshop classes were tailored to what that particular class appeared to have been lacking, and were presented in such a way that the resulting knowledge could be immediately applied to make the students’ final product better.

There were tedious rules in place (which will not be repeated here) that covered what would happen if no member of a team that was hiring was present to do hiring - a default hire was assigned to them, one which was calculated to be the worst choice based on past performance.

4. RESULTS

The general concept seems to have worked. The instructor was pleased with the quality of the final products, and most of the students seemed to get the concept and contribute well, despite having their egos bruised in the process of the elimination rounds.

The default hiring provision came into place regularly during the first few weeks as students came to learn the hard way that attendance was important. While it was expected, it occurred far more than originally anticipated.

All of the provisions anticipated occurred: there were hirings, there were firings, and there were teams that voluntarily dissolved themselves.

Unexpected within the experience of teaching the class, was the desire of some students to “quit” or otherwise “leave” a team. No inquiry was done into the reasons for this, but reasonable speculation would involve lack of belief in the ability of the rest of the team to deliver a completed product, since ten percent of the grade was the entire class’s evaluation of the final game product, and ten percent of the grade was the instructor’s evaluation of the final game product.

The chief impediment that the instructor noticed within the formation and maintenance of the teams within the class was that of labor imbalance. Specifically, each discipline seemed to think that its own had the most work, and so programmers tended to hire programmers, artists tended to hire artists, and so on (note that programmers weren’t technically necessary since the types of games that could be developed were broad enough to include board games and card games and such).

5. FUTURE VARIATIONS

To address the issue of labor imbalance, the following plan has been devised:

1. Identify the typical roles within these teams (which may vary from real-world roles due to the fact that these are students).
2. At the beginning of class, students must declare which roles they believe they can satisfy.
3. Prior to hiring, teams must specify which role they want to satisfy, and it must be a role that they have not yet satisfied.
4. In hiring, teams can only hire those that would satisfy the role that they are hiring.

To address the issue of students wishing to quit a team, the following protocol has been devised:

Between a class period with voting and a class period with hiring, quitting may occur. Between a class period with hiring and a class period with voting, firing may occur. This precludes any confusion regarding which is which, and ensures that the presentations given during the voting round do not have the demoralizing effect of someone having just quit just prior to the giving of the presentation.

6. GUIDELINES FOR APPLICATION

This only works for classes that involve long, ongoing, team projects.

The project should be structured so as to allow teams to be able to absorb a new member effectively in a way that increases productivity - this works for something like a game in which there are various roles and one can always do more testing, but for writing a paper, this may be less appropriate due to clashing abilities to write/proofread/edit/critique.

This only works for classes in which the class itself as a whole can be held to be a good grader of the presentations through a cultural understanding that would equate the student demographic of the class with being the primary consumers of the type of product produced by the teams. The larger the class the better, for these purposes.

While it isn't necessary that this be an interdisciplinary class, such an environment helps reinforce the idea that every student has a role that they can contribute and helps with the structuring of teams so as to be able to incorporate new members as the class progresses. Keep in mind the general principle is to honor the following goals:

- ensure a quality product is produced
- channel the students' egos, both individual and collective, into working on that product.
- set up hiring/elimination in such a way that the instructor designs the mechanism, and the students carry it out, so as to avoid having the students place any blame for decisions on the instructor

7. RESEARCH DIRECTIONS

In attempting to think of the structure of this class in a larger pedagogical framework, a reasonable analogue that comes to mind, from the realm of critical game theory, is that of the narratology versus ludology debate. In narratological terms, games are to be seen as consisting primarily of a story, which makes them amenable to the same sort of analysis as other media that tell a story, such as film, literature, and so on. In ludological terms, games are to be seen as consisting primarily of play and the elements of play (ludic elements), and as such, a focus on the elements that contribute to the interactive feedback loop that facilitates play becomes central.

This student-voted multiple-round elimination team projects approach can be seen as modeling that ludological approach, since there is no predefined story for the students to progress through, and they could potentially have such wildly divergent stories as having their original concept survive until the end versus consistently being hired by losing teams throughout the semester, and thus is clearly not narratological in nature. The feedback loop of present, vote, eliminate teams, hire from the eliminated teams, and repeat, provides a clearly ludic function throughout the duration of the class.

In the future, I expect to research and further develop a more thoroughly articulated ludologically informed pedagogical viewpoint that should be amenable to application in ways other than the specific form that can be found in the central focus of this paper.

I also expect to gain additional experience on the application of the approach presented here, and to report on that in future research.

On Understanding Statistical Data Analysis in Higher Education

Vera MONTALBANO
Department of Physics, University of Siena
Siena, Italy

ABSTRACT

Data analysis is a powerful tool in all experimental sciences. Statistical methods, such as sampling theory, computer technologies necessary for handling large amounts of data, skill in analysing information contained in different types of graphs are all competences necessary for achieving an in-depth data analysis. In higher education, these topics are usually fragmentized in different courses, the interdisciplinary integration can lack, some caution in the use of these topics can missing or be misunderstood. Students are often obliged to acquire these skills by themselves during the preparation of the final experimental thesis. A proposal for a learning path on nuclear phenomena is presented in order to develop these scientific competences in physics courses. An introduction to radioactivity and nuclear phenomenology is followed by measurements of natural radioactivity. Background and weak sources can be monitored for long time in a physics laboratory. The data are collected and analyzed in a computer lab in order to understand the importance of statistical analysis in a not trivial case.

Keywords: Data Analysis, Statistics, Experimental Data, Radioactivity, Improvement of Scientific Competences, Interdisciplinary Integration, Experiential Inquiry Learning.

1. INTRODUCTION

Data analysis is an important scientific task which is required whenever information needs to be extracted from raw data. It combines several methodologies from statistics, computer science and experimental sciences in which it is widely utilized and developed. Very sophisticated tools are commonly applied in different disciplines, such as high energy physics, medicine, agriculture, biology, economics, environmental sciences, biophysics, and so on. It is possible to use extremely sophisticated data analytic techniques in an attempt to squeeze more and more information out of a given sample of data, or to achieve even more predictive power. But this should not be done out of context. Data analysis must be directed towards solving the real problem, not towards solving a mathematical abstraction of the problem which ignores the fundamental uncertainties intrinsic to the real problem [1].

Many sorts of dangers have become more important as time has passed and computer software has become more powerful. It is all too easy for researchers, no doubt extremely knowledgeable and skilful within their own discipline, to apply very sophisticated data analytic techniques in which they have no expertise whatsoever. The thrust in the development of data analytic software has, quite properly, been in this direction of making software easier to use. In physics, software is often written and utilized by researchers with data analytic expertise joint with physical specific expertise, but this is not always true, especially in other disciplines, so that users are not protected from rushing in and making silly mistakes [1].

In higher education teaching practice it is possible to notice a fragmentation of learning contents related to these topics into different courses. This can be seen in the separation into individual learning subjects, among which exist rare and weak links and what follows is a further separation to individual lessons. Due to the high fragmentation of knowledge, students experience many difficulties in linking information into an understandable, useful and meaningful tool.

Furthermore, introduction to statistics is performed in courses with no link with experimental data and related troubles, thus statistical tools are applied to data sets generated *ad hoc* or very far from student experience. On the other side, statistical tools are often used in laboratory courses without the intent of improving the understanding of their usage. Undergraduate students usually deal with undestanding data analysis by themselves during the preparation of the final experimental thesis, sometimes missing some important aspect. But if they follow a theorical curriculum, it is possible that they never develop this scientific competence. It is not obvious that this lack of knowledge could be recovered in further courses.

In paragraph 2, an early learning path on related topics is presented. The activities were planned and realized in the last class of high school. Remarks and considerations about that experience convinced me to consider a similar path in higher education. Such a proposal is presented in the next paragraph, in order to develop learning paths on statistical data analysis in physics laboratory realized in introductory physics courses for undergraduate students of disciplines different from physics. Finally, a brief discussion for physics students is performed.

2. PREVIOUS EXPERIENCES

In the last years, many Italian Universities are involved in National Plan for Science Degree (i.e. Piano nazionale Lauree Scientifiche or PLS), in order to enhance the interest of high school students towards scientific degrees, in particular in Physics, Mathematics and Chemistry.

The PLS guidelines are the following:

1) orienting to Science Degree by means of training
2) laboratory as a method not as a place
3) student must become the main character of learning
4) joint planning by teachers and university
5) definition and focus on PLS laboratories.

There are several types of PLS laboratories. Laboratories which approach the discipline and develop vocations, self-assessment laboratories for improving the standard required by undergraduate courses and deepening laboratory for motivated and talented students.

All activities described in the following were realized in this context, like a PLS laboratory [2] or a part of it [3].

An early experience

Since 2006, the Pigelleto's Summer School of Physics is an important appointment for orienting students toward physics. It is organized as a full immersion school on actual topics in physics or in fields rarely pursued in high school, i.e. quantum mechanics, new materials, energy resources. The students, usually forty, are engaged in many activities in laboratory and forced to become active participants [3]. During the 2009 edition, titled *The Achievements of Modern Physics*, a learning path on radioactivity was proposed by me to small groups of students. After a brief introduction to the nuclear phenomenology, they were involved in measures of radioactivity from a weak source of Uranium in order to characterize the emitted ionizing radiation by using an educational device provided by a school.

A teacher was impressed by student's involvement and suggested of elaborating a learning path on nuclear physics in order to propose this activity in her school.

A learning path on nuclear phenomena in high school

The activity was supported by the National Plan for Science Degree and the school obtained financial support from a regional project of Tuscany (Progetto Ponte, i.e. Bridge Project) whose purpose was to promote collaboration between Universities and High Schools for orienting students.

The learning path was planned for the last year of high school (5^a Liceo Tecnologico) after the lessons on electromagnetism and about 16 hours were scheduled for the presented activities.

The path consisted in four different but correlated activities [2]: a brief introduction to nuclear phenomena in classroom, measurement of radioactivity in physics laboratory, an introduction to statistical data analysis and some selected elaboration of experimental data performed in computer lab.

Introduction to nuclear phenomena: A brief history of discovery of radioactivity was given. The neutron discovery and atom structure were described with a particular attention in student's understanding on dimensions of atom and its constituents. Nuclear radioactivity phenomenology was presented and more details were given for α, β and γ emission. Isotopes, forces into atoms and nuclei were outlined. Transmutations and law of radioactive decay were established. Natural radioactive chains were showed. A special care was made in explaining binding energy and mass of a bound state, equivalence mass-energy and binding energy per nucleon. Finally, fission and fusion were introduced. Topics related to fission and fusion can be very stimulating for students, specially in our society in these times of discussion about usage of nuclear energy and health security.

Figure 1: the experimental set-up

Measures of radioactivity: The measures were realized by means of a commercial Geiger counter usually used for dosimetry. It can separate different contribution from α, β and γ rays and data were collected without any unit conversion and downloaded into a computer. The detector and the sources were aligned on an optical bench as shown in Fig. 1. The sources were small Uranium sheets (m_1 = 0.152 g), another Uranium sheets (m_2 = 0.888 g) and a very weak Uranium ore source (fluorescent marble). Students realized background measures, measurement of Uranium sources at different distances, α, β and γ rays measurement, the previous ones with different shields, measures with all three sources.

Introduction to statistical data analysis: A question was proposed: there are limits to the precision of a measure? By answering to this question, one can fix the cases in which the use of statistical data analysis make sense. The first step was to recall all previous knowledge of students on uncertainties and statistics [4] , which are many but almost never used in physics by teachers. Starting from the beginning, main topics were remembered with examples, such as representation of a casual phenomena, mean, median, variance and standard deviation, Bernoulli, Poisson and Gauss distributions At this point, it was possible to introduce some elements of sampling theory [5], such as existence of a parent distribution, characterization of a good statistical sample, its mean and variance. Many links to experimental data collected in physics laboratory were proposed and examples were constructed by using them, when it is possible.

Elaboration in computer lab: Many elaborations were proposed in order to improve student's understanding of statistical data analysis. Let me give some example: verify that radioactive decay follows Poisson's distributions, test different samples by using 10, 20, 30, 50 data from the background, compare means and standard deviations of these samples with the mean and standard deviation of very long measure, find for which values of mean the Poisson's distribution becomes indistinguishable from the normal one. All possible informations are extracted from experimental data. Finally, slightly different measures were compared. The raw data are showed in Fig, 2.

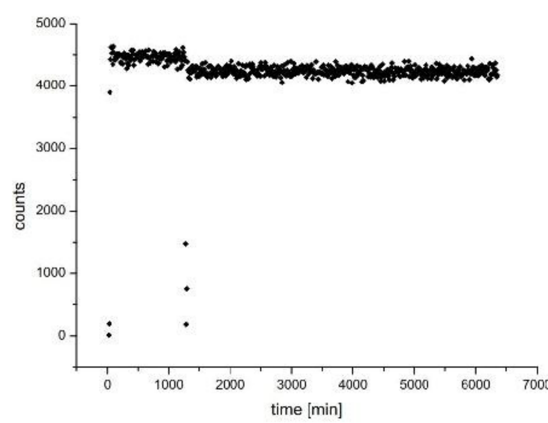


Figure 2: measures of ionizing radiation emitted from a source of Uranium at two different distances from the Geiger detector, counts are summed and registered every ten minutes

It is possible to recognize two different set of data (N_1 = 121, mean 4472, standard deviation 71, sample standard deviation 6 and N_2 = 506, with mean 4238, standard deviation 67, sample standard deviation 3). It is easy to compute the difference between means but it is overwhelmed by the uncertainty if one uses standard deviation, on the contrary the use of sample standard deviation allows to estimate quantitatively the difference $N_1 - N_2 = 226 \pm 18$.

Remarks: The learning path described above was performed on three last classes in 2010 and 2011. All classes made the introduction to nuclear phenomena in class and in laboratory, only one class made some activity in computer lab. In my opinion, this path is very interesting for students but hardly feasible in last class due to graduation exam. It can be an excellent optional proposal for interested students but if a teacher want to implement it in his classes, the presentation of statistical topics in previous years needs to be revised.

3. SOME PROPOSAL IN HIGHER EDUCATION

The above considerations led me to think about the possibility of proposing similar paths for undergraduate students. Nuclear phenomena remains one of more exciting topics rarely treated in introductory courses of physics. Some informal interview with our PhD students and some physics faculty convinced me that almost nothing is changed in Italian higher education in teaching this interdisciplinary topic, despite all curricula are revised many times in the last ten years.

A learning path on nuclear phenomena in introductory physics courses

In many undergraduate courses of study, such as Medicine, Biology, Environmental or Geological Sciences, physics remains a basic issue and one or more introductory physics courses are compulsory in the curricula.

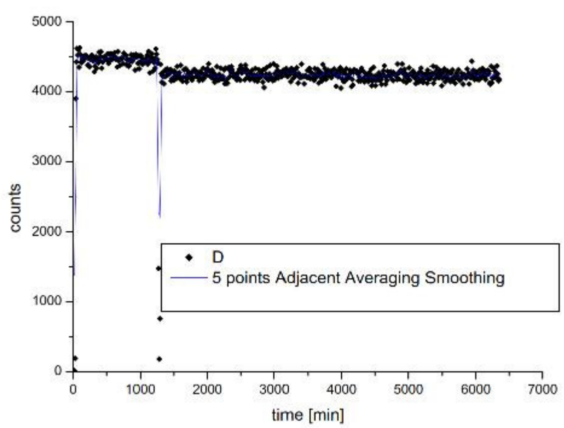


Figure 3: raw data and smoothing curve

Nevertheless, in these courses there is not enough time for offering a complete and earnest learning path on classical physics. Furthermore, vary other topics are requested such as some elements of applied or modern physics, often with the possibility of using physics laboratory. Statistics is usually taught by a mathematician and no link to experimental data is made.

Another problem that can arise is teaching in very large classes (greater then 100). In this situation, it is not easy to choose which topics are more effective and useful for these students. In my opinion, it could be a good choice to dedicate part of a physics course to a learning path similar to that described in the previous paragraph. Nuclear phenomena can be chosen as experimental topic, but other choices may be equally valid, or perhaps more suited to certain curricula. For instance, for student in environmental sciences a noise pollution measurement or nuclear measures on radon pollution can be more interesting. What matters is that experimental data must be referred to measures of a meaningful physical quantity.

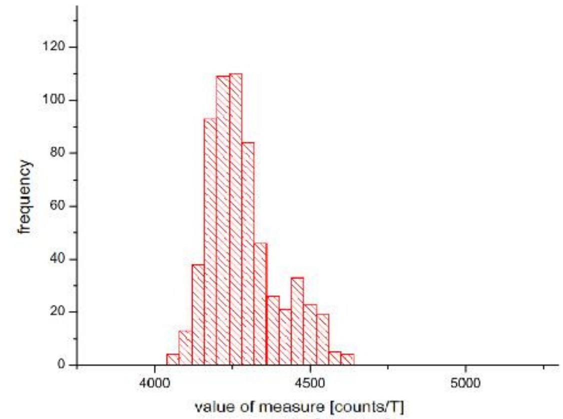


Figure 4: histogram of frequencies (T = 10 min)

Large classes can make this activity too, because it is possible to show the functioning of measuring devices to all students which can perform measures divided in small groups at different times. Since these students have more mathematical and scientific competences compared to students in high school, one can utilizes more advanced instrumentation for measuring physic al quantities, collecting and elaborating experimental data. For instance, the data showed in Fig.2, can be analyzed by using a common software for fitting data (e. g. Origin) and more sophisticated statistical tools can be used. In Fig. 3, a smoothing of data is performed. The histogram of frequencies of total data is showed in Fig. 4. It is easy to recognize that it is not a good statistical sample. Nevertheless, if a fit with two normal curves is performed (Fig. 5), one can understand that it is the convolution of two normal distributions.

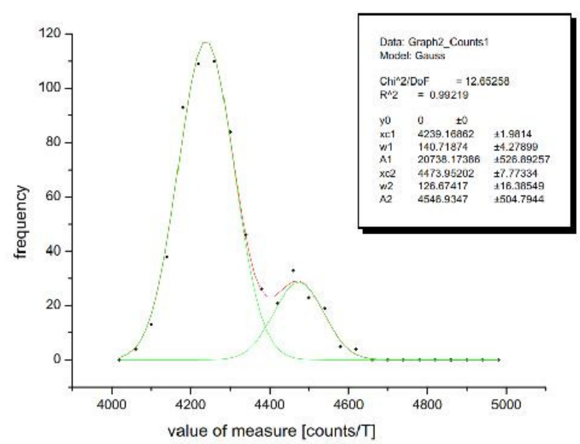


Figure 5: fit of histogram of frequencies with two normal curves

Thus, data are obviously the union of two good statistical samples with different means and standard deviations which are different from that obtained at the end of previous paragraph. Discussions with students about the origin of this behavior can be very instructive for a depth understanding of statistical data analysis.

A different learning path on data analysis for physics laboratory in course of study in basic sciences

In the case of undergraduate courses of study for basic sciences, such as Physics, Astronomy or Chemistry, the situation is more complex. Physics laboratories are very important courses, together with introductory, and sometimes advanced, courses in statistical data analysis. Different curricula can prepare in a very different way students. Usage of statistical tools can be really advanced for researchers in some discipline, like for High Energy Physics. An instance in this case is shown in Fig. 6 [6].

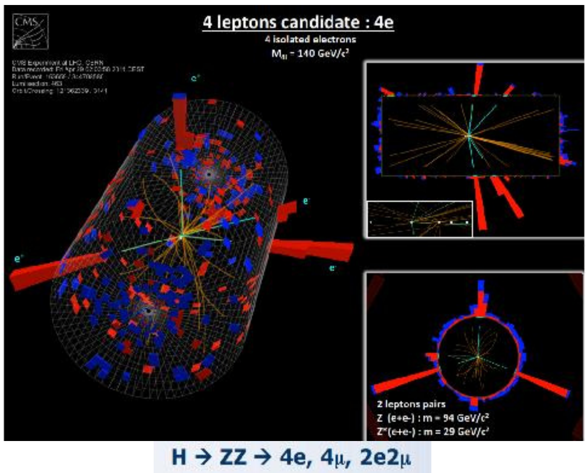


Figure 6: a four leptons candidate for Higgs Boson decay from Compact Muon Solenoid Experiment at Large Hadron Collider, CERN

In Fig. 6, different graphs from statistical data analysis of the event merged with pictorial schemes of the detectors set-up are shown .

How many data have been collected and analysed for obtaining a candidate on this hot research topic? The obvious answer is a huge amount. A raw estimate of this magnitude is given by integrated luminosity of LHC at the energy of 7 TeV times the inelastic cross section for proton-proton collision measured from TOTEM Experiment (Total Cross Section, elastic scattering and diffraction dissociation measurement) at LHC. The result is that, in order to obtaining about 25 candidates for Higgs boson decay in four leptons, as the one showed in figure, it was necessary to produce and consider $350x10^{12}$ events, i.e. less than one useful event every 10000 billions of produced events [7].

The researchers in this discipline must have, of course, a real solid formation in data analysis that usually develop by themselves in large international teams. It is singular that their training in this field is fragmented during their course of study and almost completely postponed to further research experience. Work is in progress with the purpose of conceiving a meaningful learning path on data analysis both in undergraduate and graduate courses for these kind of students.

4. CONCLUSIONS

Some proposal in higher education for realizing a learning path on statistical data analysis strongly linked to physics laboratory has been displayed. The next step will be to test at least one of these learning paths in undergraduate courses in order to proof their effectiveness. More care must be posed in elaborating new learning path in data analysis for basic sciences. The case of physics was outlined in previous paragraph but they exist other delicate and peculiar contexts, such as Mathematics or Pharmacy, that remain unexplored.

5. ACKNOLEDGEMENTS

This work is based on previous activities and experiences which were performed within the National Plan for Science Degree supported by Italian Ministry of Education, University and Research. The author would like to thank Angelo Scribano for the active support and useful discussions.

6. REFERENCES

[1] D. J. Hand, “Intelligent Data Analysis: Issues and Opportunities”, **Intelligent Data Analysis**, 2, 1998, pp. 67-79.

[2] V. Montalbano, S. Quattrini, “Measures of radioactivity: a tool for understanding statistical data analysis”, Proceedings 12th Symposium Frontiers of Fundamental Physics, Udine 21-23 November 2011, arXiv:1201.0553v1 [physics.ed- ph], 2012.

[3] R. Benedetti, E. Mariotti, V. Montalbano, A. Porri, “Active and cooperative learning paths in the Pigelleto’s Summer School of Physics”, Proceedings 12th Symposium Frontiers of Fundamental Physics, Udine 21 - 23 November 2011, arXiv:1201.1333v1 [physics.ed-ph], 2012.

[4] P. R. Bevington, **Data Reduction and Errors Analysis for Physical Science**, New York, McGraw-Hill Book Company, 1969.

[5] Stat Trek, http://stattrek.com/Lesson3/SamplingTheory.aspx, accessed 2011 November

[6] CMS Public Website, http://cms.web.cern.ch/, accessed 2012 January

[7] A. Scribano, private communication.

Perceptions of Teachers on UFABC Interdisciplinarity and its Practice

Fernanda Carla PANTALEAO
Centro de Ciências Naturais e Humanas, Universidade Federal do ABC - UFABC
Santo André, 09210-190, Brasil

and

Adelaide FALJONI-ALARIO
Centro de Ciências Naturais e Humanas, Universidade Federal do ABC - UFABC
Santo André, 09210-190, Brasil

ABSTRACT

The Federal University of ABC (UFABC) was established in 2005 with an innovative educational project, founded on principles such as interdisciplinarity, autonomy in the learning process, student mobility and technology development, among others. In August 2006, approximately one hundred PhD researches were hired with a wide range of disciplines and course projects to be developed and to be started in a month. This study intends to investigate the experience of these researches professors during the first five (05) years of activity in UFABC, how they interpret the guiding principles and the concepts outlined in the theoretical and structural design of the institution. The verification of these points is very pertinent at this moment because it can collaborate with the decision about the future courses to be held and/or modified in the institution, as well as contribute to the discussion and suggestion of points that can be adopted by other institutions all over the country wishing to adopt similar educational projects.

Keywords: Interdisciplinary, UFABC, Higher Education, Teaching

INTRODUCTION

Modern science and all knowledge produced in the twentieth century were marked by the belief that reducing the complex problems into simple objects allows a more objective and practice knowledge. Its influence can be easily noticed in the main concept on the history of science, which is the conviction that a finite set of elements and the analysis of each one of these enables the understanding of the whole (POMBO, 2004).

This knowledge process has a great historical relationship with the way that universities have been organized during the last centuries (MULHOLLAND, 2007). Almost all of them have been attended by departments focused on different disciplines and specialities. This practice has served largely to scientific progress observed in sec XX, mostly because the separation in areas allowed more precise delimitation of the research object and their further and rigorous analysis, reducing the number of methodologies and techniques required to be a specialist researcher in whichever field (POMBO, 2004).

In this view, it is truthful that specialization has possible many important scientific discoveries, like the cure of a great number of diseases, amazing technologies, medicine machines and others.

However, the sec XXI is marked by increasingly complex issues and challenges. The social scene is characterized by the advance of information technologies, the growth of corporations, and a life grounded in consumption, what promotes the massive use of natural resource and the increase of poverty and social inequality.

The emergence of nanotechnology and genetic engineering are also important examples, as they bring new and challenging questions, whose answers will require an equally complex approach. The ability to create new substances and materials through nanotechnology and/or genetic engineering brings us new realities, enabling actions on plants, animals and humans to modify features in scale never expected before (REYNAULT, 2011).

Ready to act to these new needs, there is a large movement worldwide to revisit models in their higher education. Different solutions are being attempted, but all of them with the same purpose, to adapt urgently the researches and education to new times (BEVILACQUA, 2011).

In Europe, for example, university education system has been widely discussed by the political authorities of the European Union. In 1999, European ministers of education signed one international treaty, known as the Bologna Declaration, where they bring the proposal to form a "European Higher Education system". The document is the basis of formation of a single system that equalizes university education among all signatory countries. The main objectives are promoting mobility to teachers and student between institutions and countries and encourage interdisciplinary research and high-level training to their citizens.

In 2006, the Faculty of Arts and Sciences at Harvard published a document with the studies and proposals for a curriculum review of his "College”. As part of this review, was launched in 2007 a new General Education program for the institution that aims to meet the educational needs of the students with the new needs of the twenty-first century.

The university proposed a re-viewed of the curriculum around eight themes, avoiding the classical disciplines used in antique knowledge partition. These axes are called: Aesthetic and interpretive understanding, Culture and faith, Empirical reasoning, Ethical reasoning, Science of living systems, Science of the physical universe, Societies of the world and, the United States in the world.

The purpose of the new division is to develop in students the ability to interpret cultural artifacts, to participate in political processes, to deal with the implications of technological change, interacting with people from diverse backgrounds to evaluate the various claims made in scientific discourse, deal with ethical dilemmas in their personal and professional lives, and make connections between the material covered in the classroom and everyday life (HARVARD UNIVERSITY, 2006).

In the U.S., it is possible to found other examples of university programs that are aware with the twenty-first century demands. This is the case at UC-MERCED, which started in 2005 in California. His academic structure has a very innovative profile, being organized in three major centers: the School of Engineering, College of Natural Science and the Faculty of Social Sciences, Humanities and Arts. The university focuses on research for the twenty-first century and seeks to interdisciplinarity as a mean to promote cutting-edge thinking and knowledge creation.

Regarding the Brazilian example, it is worth mentioning the creation of the Federal University of ABC - UFABC in 2005. The university was designed with the challenge of developing science and technology appropriating the knowledge in a more constructive, humane, and interdisciplinarity as the linchpin of their pedagogical project (UNIVERSITY OF ABC, 2005). According to Bevilacqua (2011), an author of the UFABC pedagogical project, emphasizes on the encouragement of creativity, scientific discovery, technological invention and well-founded criticism are the important points to be achieved.

Like Harvard, the proposal sought to develop an interdisciplinary academic structure organized into topics, such as the Structure of Matter, Energy, Communication and Information, Process Transformation, Representation and Simulation, Humanities and Social Sciences, not axes classics such as Physics, Chemistry, Biology and Mathematics that promote the continuity of a division that no longer exists (BEVILACQUA, 2011).

The project proposes to work with a flexible and open structure, unlike the departmental structure that characterizes the traditional institutions. Like no other university in Brazil, instead of departments the university was structured in three major centers: one for the Engineering and Applied Social Sciences, another for the Natural Sciences and Humanities and one for Mathematics, Computing and Cognition. The organization of academic and administrative centers, somehow, "portrays the actions to discover, systematize and make up that comprises, in overview, the present set of activities in university life" (UNIVERSITY OF ABC, 2005).

Being the first university in the country to adopt a completely new teaching model, the UFABC cannot rely on any previous experience, and therefore with no methodology traveled or tested. Therefore, experience becomes even more challenging in its two major source responsibilities. The first is the trying of something new, different and unknown. The second responsibility lies in the influence that this model may have to the reform of the Brazilian university. In other words, the UFABC may serve as north to other higher education institutions in Brazil.

Interdisciplinarity and the case of UFABC

According to Bevilacqua (2011), UFABC is new and free from defects of classical structures, and the construction of their academic project was hindered to accommodate new initiatives.

We can see in their pedagogical project a significant concern in translating some of the main obstacles, as discussed by renowned authors on the subject, for the concretization of interdisciplinarity.

Among the work entitled "Integration and Interdisciplinary Teaching in Brazil: Effectiveness and Ideology," Fazenda (2002) makes an important survey of the major obstacles founded to the realization of interdisciplinary teaching. Among them, the author points out the epistemological, institutional, socio-psychological and cultural obstacles.

With respect to epistemological obstacles, Japiassu (1976) highlights the fact that during the PhD formation, the research professor becomes trapped in his own specialty, getting hooked by the details, losing the notion of unity. To illustrate what this means, Japiassu argue that Binet admits "intelligence is what I measure with my tests" (p.94). Given this placement, Japiassu warns that the only detail was overlooked by the psychologist to ask what intelligence really means. Complementing this vision, Fazenda (2002) notes that "the very fact of reality have many and varied facets, is no longer possible to look at it under a single angle, through a single discipline" (p.52). Thus, the interdisciplinary approach emerges as complementary and necessary in dealing with this reality.

According to this perspective, the committee who designed the structure of academic UFABC was convinced that one of the ways to overcome the discontinuities of the accelerated advance of scientific knowledge was to change the traditional paradigm of separation of knowledge (BEVILACQUA, 2011). So was decided to design a new curriculum with a thematic division instead of the traditional division, according to the axes of classical physics, biology, chemistry, mathematics, etc. This new division, according to Bevilacqua (2011), brings with it the induction of interdisciplinary cooperation of professionals from different fields, since knowledge was reorganized into thematic fields rather than disciplinary axes, as is usually done.

Alongside the epistemological obstacles are institutional obstacles. These are characterized, in turn, by the usual rigidity with which the institutions are organized. In this regard, Fazenda (2002) points out that each course, hoping to preserve their "status", seeks to maintain its power against the other. Then, each course creates its own language and thus end up cutting off communication with everything else. Japiassu (1976) names this phenomenon as a kind of epistemological capitalism, in which there is a hoarding of knowledge. This fact can be easily seen in departmental structures universities. In there, they form veritable islands of pooling, catching each other real wars for space, funding and power.

Attentive to this issue, the departmental barriers were ignored by the committee who structured the UFABC. There are no departments in the institution, but three major centers: Center for Nature and Humanity Sciences, Center for Mathematics, Computing and Cognition and Centre for

Engineering and Social Sciences. These centers are distributed over different areas of teaching, thus the chemical shares the same space with the biologist, physicist, philosopher, engineer, mathematician, social scientist, etc. "The centers portray the actions to discover, systematize and make up that comprises, in overview, the present set of activities in university life" (PEDAGOGICAL UFABC PROJECT, 2005).

The committee responsible for the project notes that the elimination of departments is a breakthrough that allows a constant interaction between students and teachers to perform a work directed to interdisciplinary training. Although they know that the "institutional structure alone does not ensure the desired integration of knowledge, but the idea is that it induces and facilitates interdisciplinary, promoting systemic and, through them, the appropriation of knowledge by society, without weakening of strict disciplinary culture "(PEDAGOGICAL UFABC PROJECT, 2005).

Although structural change is not sufficient for effective interdisciplinary, one must recognize that the measure increases to a great extent the opportunities of meeting and dialogue between these professionals from different areas with teachers and students, since it breaks with the enclosure of knowledge and the professional isolation of a single area, characteristic of the departmental structure.

More challenging than changing the physical structures of an institution is to transform the mental structures. Thus, the socio-psychological and cultural obstacles become major barriers to be overcome for the realization of interdisciplinarity. According to Fazenda (2002), is generally a bias to adhere to interdisciplinarity, which is always regarded as something superficial. For the author, there are several causes that can grounds this attitude, among them the lack of the real meaning of the interdisciplinary project, the lack of specific training for this type of work and accommodation and conference staff.

In addition, participants must learn to deal with a new requirement of interdisciplinary work, the question of anonymity. Treasury warns that requires an interdisciplinary project participants, as with any change, a new overload of work and dedication. However, the personal prestige cancel each other out in front of a larger goal. Thus, the professional is canceled before a project is not intended to give "status", space, power or prestige to a single area of knowledge or a single department. The commitment of an interdisciplinary project is to guide the human intelligence against a particular issue, regardless of which area is. Thus, Fazenda (2002) emphasizes an interdisciplinary project requires the participants a certain amount of humility, openness and curiosity as fundamental characteristics. It also emphasizes that the aim is to join a team of experts, but they depart in search of a common language in addressing the problems of society.

Regarding the socio-psychological barriers, Japiassu (1976) states that the model of fragmentation of knowledge is encouraged, as it serves to divide and rule expert. Insofar as the specialty grows, the will to power and domination of the expert, thus forming a barrier to interdisciplinary cooperation. Likewise, Japiassu occurs with culture obstacles. "Each specialist becomes an ivory tower, whose entrance is forbidden to the uninitiated" (p.96).

Associated with this obstacle, we also have difficulty with training. The professionals of the twentieth century, as well as teachers, were formed in a structure marked by the paradigm of modern science, where the construction of knowledge occurred through the fragmentation of objects. Thus, the introduction of interdisciplinarity can be performed by persons who received a non interdisciplinarity training. This means a great pedagogical challenge for an institution that wishes to effect the interaction between the different areas.

Concerning the pedagogical relationship in a multidisciplinary project, Fazenda (2002) emphasizes the need for change in the relationship between teachers and students. For the author, it is necessary to overcome the pedagogical relationship based upon transmission of knowledge in a particular discipline, to a teaching relationship based on dialogue, where a position is the position of all. “Accordingly, the teacher becomes active, the critic, the entertainer par excellence” (p.56).

Faces to those aspects, some questions are quite relevant for the assessment of the project. Among them are: how the adoption of interdisciplinarity is occurring in the practice of teachers working in the UFABC? It is indeed found in practice or is just in the speech? The organization in centers has fostered interdisciplinary works?

The search for these answers led to this research, still in progress, whose objective is to investigate how teachers perceive the UFABC interdisciplinarity in teaching and in research.

THE RESEARCH METHODOLOGY

This research is a qualitative research because it focused in investigate the researches perceptions about the interdisciplinary and practice within a particular context, the UFABC, trying to understand how the phenomenon occurs in this context. It is an interpretative analysis.

The main activity involves carrying out semi-structured interviews with faculty members of the first group of teachers hired by the university. The selection of respondents was randomly selected among three different centers of the university, regardless of the training area.

Interview scripts were defined with a set of issues common to all respondents. The questions referred to three main focuses: personal process of adaptation at the university, vision about teaching and learning process and about researches projects. The material obtained was transcribed and used for analysis seeking to produce inferences that allow discussing and understanding the design of these teachers.

PRELIMINARY RESULTS

Despite the work still in progress, the material already obtained allows some observations. The first is the complexity that represents the implementation of a university based on an innovative structure, especially considering the aspect of conducting interdisciplinary work. This difficulty is by several factors. Initially, can be notice that the main obstacle is the lack of a common comprehension about an interdisciplinary teaching and research work. This aspect can observed in the statements below:

"Well, I knew about the UFABC proposal before making the call, I just did not know I had no idea it was interdisciplinary. I thought I knew what it was; I came by the proposal, found it interesting. When we actually started here in August 2006, we realized that in fact we did not know what it was interdisciplinary".

"The negative aspects or difficulties I would say are basically the following. First we do not have nowhere near a minimal understanding of what is interdisciplinarity, I'd go a little further, I'm not really sure what I'll say, but I would venture to say that UFABC is not yet a university interdisciplinary today it is multidisciplinary"

With respect to research, there are teachers who were able to establish collaborations with teachers from different areas and are very happy for that. However, there are researches who have failed to establish an effective dialogue for this occurrence:

"The centers are interdisciplinary, I'm here, I'm an engineering center, modeling and applied social sciences, ok. Now you ask me in my research that I have the presence of content engineering, modeling or applied social sciences, none".

"In relation to research what we call interdisciplinarity has worked very well in terms of infrastructure, laboratories, equipment, and the contact with other researchers is excellent".

Regarding the structure of centers and no departments, some teachers argue that this division actually improves communication and collaboration between them. However, some interviewed reports a lack of interaction due to the difficulty that someone have to transpose their area of knowledge and not believe in the need for interdisciplinarity.

Other difficulties mentioned by most respondents are directed to issue administrative centers. The center has a number of researches and more than one department, and several different areas compete in a single center. This fact has some implications as the formation of blocks of teachers in the same area to claim money, space allocation and teaching.

FINAL THOUGHTS

The UFABC is the first university in Brazil created with an innovative educational project willing to overcome the traditional model of higher education in Brazil. The preliminary results seem to indicate that interdisciplinary work is still limited. Yet, there is a design for a common understanding of what this practice.

With the results in this work we hope to contribute to a understanding of the UFABC experience and contribute for future reforms in higher education in Brazilian system.

REFERENCES

BEVILACQUA, L. Primórdios da Ärea Multidisciplinar da Capes e suas influencias na Pós-Graduação e na Graduação. In: In: PHILIPPI JR., A.. NETO, A. J. S (Ed) **Interdisciplinaridade em Ciência, Tecnologia & Inovação**. 1ª ed. Barueri, SP: Manole, 2011. Chapter 28, p.785-802.

FAZENDA, I. C. A, **Integração e interdisciplinaridade no ensino brasileiro: efetividade ou ideologia**. São Paulo: Edições Loyola, 2002.

HARVARD UNIVERSITY – **Faculty Of Arts And Sciences. Curricular Renewal in Harvard Coleege**, Janary 2006. Avalaible in: http://generaleducation.fas.harvard.edu/icb/icb.do

JAPIASSU, H., **Interdisciplinaridade e Patologia do saber**. Rio de Janeiro: Imago, 1976

MULHOLLAND, T., Multi, Inter e Transdiciplinaridade na concepção acadêmica das universidade brasileiras In: AUDY, J. L. N, MOROSINI, M. C., (Orgs) **Inovação e Interdisciplinaridade na Universidade**. Innovation and Interdisciplinarity in the University. Porto Alegre: PUCRS, 2007.

POMBO O. , **Interdisciplinaridade: ambições e limites.** Lisboa: Relógio D'água, 2004

RAYNAUT, C. Interdisciplinaridade: mundo contemporâneo, complexidade e desafios à produção e à aplicação de conhecimentos. In: PHILIPPI JR., A.. NETO, A. J. S (Ed) **Interdisciplinaridade em Ciência, Tecnologia & Inovação**. 1ª ed. Barueri, SP: Manole, 2011. Chapter 2, p.69-105.

UNIVERSITY FEDERAL OF ABC-UFABC. Pedagogical project, 2005. Available in: http://www.ufabc.edu.br. Acesses in 20 of february 2012.

UNIVERSITY OF CALIFORNIA – MERCED. Available in: http://www.ucmerced.edu/. Acesses in 20 of february 2012.

ZABALBA, M. A., **O ensino universitário: seu cenário e seus protagonistas.** Porto Alegre: Artmed, 2004.

Fostering the Self-Renewal of Teachers: An Underutilized Approach to Innovating Interdisciplinary Education

David J. Waters
Dept. Veterinary Clinical Sciences and Center on Aging and the Life Course, Purdue University; Center for Exceptional Longevity Studies, Gerald P. Murphy Cancer Foundation
West Lafayette, Indiana 47906, USA

and

Emily C. Chiang
Center for Exceptional Longevity Studies, Gerald P. Murphy Cancer Foundation
West Lafayette, Indiana 47906, USA

ABSTRACT

Our goal is to call teachers' attention to the need for self-renewal, challenging them to consider it a necessary approach to innovating interdisciplinary education. Our prescription for sustained self-renewal: *Each teacher assembles a gallery of intellectual heroes — gifted and articulate thinkers — to serve as their own life-long teachers.* In this paper, we share our experience teaching a "skills course" to interdisciplinary graduate students in Purdue University's Center on Aging and the Life Course. The course, entitled "To See and To Seize Opportunities", exposes scientists-in-training to an array of skills and attitudes that foster peak performance and self-renewal. Leading educators must work hard to create better opportunities for self-renewal. By envisioning even our best teachers as unfinished and under construction, we open up a new dialogue situating the self-renewal of teachers at the very core of educational excellence across a broad range of disciplines. To innovate interdisciplinary education, we believe it is time for a curricular re-think, emphasizing the importance of a transdisciplinary skills course in which students and their teachers can explore transformative ideas on personal development and self-renewal — in the classroom together.

Keywords: personal skills, learning attitudes, peak performance, leadership, heroes, self-renewal, interdisciplinary education.

Good teachers instruct students on how to function in the world. Great teachers show students how they can transform it. By seeing education as a dynamic and transformative process, great teachers grow to recognize their own unfinishedness. Their philosophies and methods are in a state of both permanence and flux. The nameplate outside their door reads "Under Construction." They are experts, yet open to change.

Disciplinary education journals publish perspectives on great teaching written by award-winning educators who share wide-ranging insights on achieving excellence in the classroom, offering innovative ways to inspire active learning. To these thoughtful recipes for educational success, we would add another key ingredient — *the teacher's process of self-renewal*. Where does the teacher turn for self-renewal? What skills and attitudes prime the teacher for continued high performance, to navigate the ever-changing territory that is the teaching-learning space? What keeps teachers from becoming closed-minded, from slipping into the ruts of their own expertise? Both teachers and administrators recognize the immense importance of self-renewal [1]. Yet few opportunities for the self-renewal of teachers are built into the educational system.

This paper is a call for teachers to take action — to innovate the education process by reaching their highest potential through self-renewal. We seek to call teachers' attention to the need for self-renewal, challenging them to consider it a necessary approach to innovating interdisciplinary education. We have proposed a schema for self-renewal that we believe can assist even the most accomplished teachers with their unfinishedness [2]. Our prescription for sustained self-renewal: *Each teacher assembles their own personalized, hand-picked gallery of intellectual heroes — gifted and articulate thinkers — who in turn serve as their life-long teachers.* The approach has an appealing rationale. If we as teachers can spend a bit more time *thinking about our own thinking*, then we will begin to see our own teaching philosophy from new angles, both analytical and creative. By investing in our own self-renewal, we are putting students first — harnessing new energy, gaining fresh insights into structuring the kinds of educational experiences that will nurture the skills and attitudes that can enable each student to go beyond knowledge to expertise.

In Purdue University's Center on Aging and the Life Course, an interdisciplinary unit fostering both research and education, we are promoting the value of tethering oneself to intellectual heroes. By assembling your own personalized gallery of intellectual heroes, you gain greatly by becoming both teacher and student. You direct your attention toward rapidly expanding your skills of reading and listening — the art of being taught [3]. This activity fosters personal growth, shaping new understandings that enable teachers to perform at the highest level.

We posit that progress, whether in the public domain of scientific knowledge or the private domain of personal thought, is benchmarked not by the results, but rather by the questions we ask [4]. A commitment to question making is a prized product of the "skills course" offered to our interdisciplinary graduate students in Purdue's Center on Aging and the Life Course. The course, entitled "To See and To Seize Opportunities", exposes students to an assortment of skills and attitudes that encourage peak performance and self-renewal. Fresh insights provoked by our intellectual heroes have sparked in us a deeper reflection, coming to see our teaching and the learning process from new angles of vision. The product of this effort has been a steadily evolving collection of questions that invite thoughtful exploration, providing teachers with a framework that is well-suited for across-discipline inquiry:

- *Are your students trained in the art of problem finding, not just problem solving?*
- *Could the quality of your thinking be enhanced by increasing your ability to be precise with language? Have you ever considered taking a course in general semantics to transform your language behavior?*
- *Do you teach the history of your discipline, enabling your students to witness the uneven spits and spurts of progress in that discipline?*
- *Have you reflected on the important role that provocation, rather than instruction, plays in your own learning?*
- *How often do we mislead students with a false impression of how much we know versus what we believe? Are you training students in the art of making judgments under uncertainty?*
- *If mastering the art of seeing the similar as different, the different as similar is an enviable achievement, shouldn't you be teaching a course on comparative something?*
- *Have you dedicated yourself to acquiring the opposing skills (analytical vs. creative) necessary for developing a disciplined imagination?*
- *Have you ever considered how your writing influences your reading? Is it time to re-think the way you write?*

CONCLUSIONS

The intent of this paper was to call teachers' attention to the need for self-renewal, challenging them to consider it a necessary approach to innovating education. Leading educators must work hard to create better opportunities for the self-renewal of teachers [2]. By envisioning even our best teachers as unfinished and under construction, we open up a new dialogue situating the self-renewal of teachers at the very core of educational excellence. To innovate interdisciplinary education, we believe it is time for a curricular re-think, emphasizing the importance of a transdisciplinary "skills course" in which students and their teachers can explore transformative ideas on personal development and self-renewal — *in the classroom together*.

REFERENCES

[1] Gardner J. **Self Renewal: The Individual and the Innovative Society**. New York: Harper and Row, 1963.

[2] Waters DJ, Waters LS. On the Self-Renewal of Teachers**. Journal of Veterinary Medical Education** 2011; 38: 235-41.

[3] Adler MJ. **How to Read a Book**. New York, NY: Simon and Schuster, 1940.

[4] Pirie NW. Selecting facts and avoiding assumptions. In: Berthoff AE, ed**. Reclaiming the Imagination.** Upper Montclair (NJ): Boynton/Cook Publishers Inc., 1984, p. 1249-61.

How the Deep-Rooted Metaphor "More is Better" Compromises Interdisciplinary Communication and Jeopardizes Public Health

Emily C. Chiang
Center for Exceptional Longevity Studies, Gerald P. Murphy Cancer Foundation
West Lafayette, Indiana 47906, USA

and

David J. Waters
Dept. Veterinary Clinical Sciences and Center on Aging and the Life Course, Purdue University;
Center for Exceptional Longevity Studies, Gerald P. Murphy Cancer Foundation
West Lafayette, Indiana 47906, USA

ABSTRACT

The notion that "more is better" is one of the most deeply rooted metaphors in our culture. What we come to value – as scientists, as citizens – is closely tied to the metaphorical concepts we live by. We propose that the conceptual system of MORE IS UP – GOOD IS UP – MORE IS BETTER compromises our ability to accurately communicate disciplinary research findings across disciplines because it fails to adequately emphasize that context determines meaning. Further, based upon our own research experience in the field of cancer-fighting nutrients, we believe that this failure seriously jeopardizes public health. The perception that is pervasive among the public is that, when it comes to taking dietary supplements, more is better. A growing body of scientific evidence, however, suggests that the dose-response between cancer-associated DNA damage and the dietary intake of nutrients, such as the antioxidant trace mineral selenium, is in fact U-shaped. Therefore more of "good things" may not necessarily be a good thing. This concept is especially relevant to health-conscious men and women, who are ironically at highest risk for the ill-effects of oversupplementation because they are already consuming high-quality diets rich in vitamins and minerals. In conclusion, the deep-rooted metaphor "more is better" is a significant obstacle to our efforts to successfully communicate life style decisions that really do promote health. Because a single recommendation for everyone is plainly inadequate, high priority should be placed on developing young scientists who have interdisciplinary communication skills and who see communicating health research to the public as vital to society's progress.

Keywords: U-shaped, dose-response, health research, context, communication skills, personalized health promotion.

Thirty years ago in Mind and Nature [1], Gregory Bateson wrote, "There are no monotone values in biology". Bateson explained further: "Desired substances, things, patterns, or sequence of experiences that are in some sense "good" for the organism … are never such that more of the something is always better than less of something. Rather, for all objects and experiences, there is a quantity that has optimal value." We contend that one of the real villains in public health today is the ineffective cross-disciplinary communication which paints an overly simplistic view of health promotion. This attitude is captured in the expression "Just show me the good things and I'll grab as much of them as I can." Based upon our own research experience in the field of cancer-fighting nutrients, we see how seriously this failure jeopardizes public health [2].

Selenium is an essential trace nutrient required to protect our bodies against oxidative stress. Even though selenium abundantly exists in many of the foods we eat, its anti-cancer activity reported in more than 400 animal studies allures the public to get even more selenium from supplements to be "fully protected". Why was the idea of selenium supplementation so popular even before scientists could pin down the optimal dose of selenium for cancer prevention? Apparently, when it comes to taking dietary supplements, the perception among the majority of people is that more is better.

Almost a decade ago, we began focusing our efforts on determining what is the best dose of selenium for prostate cancer risk reduction [3]. To investigate the optimal dose of an intervention, whether it be a dietary supplement or other lifestyle choice such as exercise, researchers generate what are called dose-response curves. So, we set out to construct the dose-response curve between selenium status and DNA damage within the prostate. We conducted a randomized selenium supplementation trial in 49 elderly male dogs that were physiologically equivalent to 65 year-old men. After 7 months supplementation, we found the relationship between selenium status and prostatic DNA damage was U-shaped — indicating that selenium status either below or exceeding the optimal range was not useful in reducing DNA damage in the prostate. Moreover, our U-shaped dose response curve remarkably paralleled the data from human studies of selenium and prostate cancer risk and anticipated the null results of the SELECT prostate cancer prevention trial in which most participants were already in the optimal range prior to supplementation [4]. The take-home message: When it comes to selenium and cancer prevention, more is not necessarily better.

The U-shaped curve is not peculiar to selenium and prostate cancer. Other examples include the association between different cancer risks and the status of vitamin E, zinc, and beta-carotene. When it comes to developing an effective cancer prevention regimen, we need to know who will benefit, at what dose, for how long, and when to intervene. In other words, a single recommendation for everyone is inadequate.

Now, more than ever, we need a new approach to cancer prevention – one that is personalized. We define personalized cancer prevention as a strategy that will enable each person to reduce his or her risk for lethal cancer by matching the dose, duration, and timing of an intervention with their own cancer risk profile [5]. Defining the U-shaped relationship between DNA damage and cancer-protective nutrients moves us one step closer to developing personalized, cancer-reducing interventions. It follows from this understanding that not all individuals will necessarily benefit from increasing their nutrient intake.

The notion that "more is better" is one of the most deeply rooted metaphors in our culture. What we come to value – as scientists, as citizens – is closely tied to the metaphorical concepts we live by [6]. Context determines meaning and a "more is better" mindset seriously compromises our ability to communicate our research findings across disciplines amid shifting contexts.

CONCLUSION

Promoting health in a U-shaped world – a world where more is not necessarily better – poses distinctive challenges. To have the greatest impact, scientists need to effectively communicate their research across disciplines. More importantly, scientists must help the public become more savvy about health promotion. An over-simplified message such as "Selenium is good for you" is meaningless and potentially dangerous. The public is hearing the results of health research out of context – a problem of paramount importance because context determines meaning. A high priority should be placed on developing young scientists who have interdisciplinary communication skills and who see communicating health research to the public as vital to society's progress.

REFERENCES

[1] G. Bateson, **Mind and Nature: A Necessary Unity**, London: Wildwood House. Pub., 1979.

[2] D.J. Waters, E.C. Chiang, "It's a U-shaped world: a Batesonian prescription for promoting public health", **ETC: A Review in General Semantics**, Vol 67, 2010, pp. 218-226.

[3] D.J. Waters, S. Shen, L.T. Glickman, D.M. Cooley, D.G. Bostwick, J. Qian et al., "Prostate cancer risk and DNA damage: translational significance of selenium supplementation in a canine model", **Carcinogenesis**, Vol. 26, No. 7, 2005, pp. 1256-1262.

[4] S.M. Lippman, E.A. Klein, P.J. Goodman, M.S. Lucia, I.M. Thompson, L.G. Ford, et al., "Effect of selenium and vitamin E on risk of prostate cancer and other cancers: the Selenium and Vitamin E Cancer Prevention Trial", **Journal of the American Medical Association**, Vol. 301, No. 1, 2009, pp. 39-51.

[5] D.J. Waters, E.C. Chiang, D. G. Bostwick, "The art of casting nets: fishing for the prize of personalized cancer prevention", **Nutrition and Cancer**, Vol. 60, No. 1, 2008, pp. 1-6.

[6] G. Lakoff, M. Johnson, **Metaphors We Live By**, Chicago: University of Chicago Press. Pub.,1980.

An Emergent Model: A Vehicle for Research Methods and Research Activities

Drs. C. Pierre, G. Chung, F. Abebe, T. Kebede

Clark Atlanta University
Department of Mathematical Sciences
Atlanta, Georgia 30314
Email contact: cpierre@cau.edu

ABSTRACT

The emergent model is born out of the field of complexity theory, a discipline studied by mathematicians, physicists, biologist, engineers, chemists, as well as political scientists and economists. One of the many models covered in complexity theory is the ant model which concerns individual, self-contained bodies, or ants, that strike out on undetermined paths to find food. These ants have no central contact, or central manager, to direct their paths. Once food is found by an ant, other ants are made aware via feedback in the form of pheromones left along a path. Within a very short period, all ants in the field will have converged on the food area. The emergent model gives a fresh approach to research design, and it stresses small research groups operating as individual, self-contained entities which answer to no central manager or control center. The ideas of both feedback and interdisciplinary research will undergird this model.

Keywords: Chaos, Dynamical Systems, Mathematical Biology, Complexity Theory, Emergent Model, Interdisciplinary Research

1. INTRODUCTION

This initiative is an outgrowth of a continued research effort by the Four Faculty Research Group (FFRG) in the Department of Mathematical Sciences at Clark Atlanta University to transform undergraduate and graduate education in science technology, engineering and mathematics. Over the last nine months, the FFRG has been active in research into the areas of dynamical systems, mathematical biology, and complexity theory. The FFRG has also been actively recruiting and mentoring student research participants to enhance student learning via research activities. This research group has been meeting weekly over the past year to scour the depths of the areas. To date, we have a wonderful grasp of these emergent and cutting-edge research areas of chaotic dynamical systems, mathematical biology and complexity theory. The interdisciplinary nature of these fields makes the pursuit very ideal for broader interdepartmental collaborations because the emerging market needs trained STEM graduates that are adept at interdisciplinary thinking and problem solving. It is an established pedagogic fact that interdisciplinary learning fosters interest in subject matters, and enhances the learning and retention process. [1, p. 62.] Therefore, the training and the improvement of student learning by the research oriented instruction for the undergraduate STEM majors must be done in a timely and urgent manner to bring about lasting improvement in undergraduate STEM education.

2. BACKGROUND AND INITIAL IDEAS

We formed the research group in our department at Clark Atlanta University in Atlanta, GA last fall to study what we believe to be some of the most recent cutting-edge research being done in the area of mathematical sciences. These topics are

1. Chaotic Dynamical Systems
2. Mathematical Biology
3. Complexity Theory, and Chaos and Complexity Theory

The research group FFRG (Four Faculty Research Group) formed in the Department of Mathematical Sciences at Clark Atlanta University prior to the summer of 2011was intended to do the most recent cutting-edge research in the areas of

1. Chaotic Dynamical Systems
2. Mathematical Biology/Mathematical Ecology
3. Chaos and Complexity Theories

There are many commonalities that exist between these topics. [2] [3]. FFRG has been doing intensive research in these areas for approximately nine months. One faculty member in the department has been doing research in this area, off-and-on, for more than 10 years including intensive summer work at the Oak Ridge National Lab in Oak Ridge, Tenn. Another faculty member in the department has been doing research in this area, and applying it to both mathematics and physics for years. All faculty members have been doing work in similar areas for at least 10 years. As a collective research group, the FFRG has been working together on research, which includes weekly debriefings, brainstorming sessions, and Q & A sessions on every Tuesday for about two hours. This has been lasting consistently for nine months. We have had teleconferences and mini-meetings between the weekly meetings. We are working to leave no stone unturned in these areas—if that is possible. The sessions are intense and there is no weak link.

Specifically, the charge of the FFRG is five-fold:

1. To seek funding to support the research we are doing;
2. To publish broadly in these areas;
3. To create new undergraduate (as well as graduate) courses, and have these courses approved to be included in the department's curriculum;
4. To be a leading research group and authority, in the above mentioned areas, among the HBCUs and the major universities;
5. To train upper-classmen (as well as graduate students) to do research in these areas and to delve into these topics.

Currently, three members of our FFRG are directing research for four undergraduate mathematics majors in our department. These students are doing research in the area of chaotic dynamical systems and they are making decent strides. The students' research efforts are being funded by the HBCU-UP - a federally funded (NSF) program. The Co-PIs of the HBCU-UP are the University President Carlton Brown and Dr. Cass Parker, chair of the Department of Chemistry at Clark Atlanta University (CAU). We are currently seeking funding for future efforts of this kind.

One of the FFRG members is teaching a chaotic dynamical systems course to graduate students in the graduate Topics in Mathematics (CMAT 651) course offered in the Department of Mathematics. Each student taking this graduate course has the opportunity to seek advice from any member of the FFRG at any time.

The FFRG has asked a biology professor of the Department of Biological Sciences and a computer scientist/physicist from the Department of Computer Science to join the group for research efforts. The biology professor has worked as a research biologist at the Oak Ridge National Lab, and his knowledge of, and experience in, the field of biology will help us tremendously. We are excited about the current collaboration with both professors. The biologist will participate in our future grant writing efforts.

Our research has carried us to many places in the mathematical sciences and in physics (we have a mathematical physicist in our group) and we are excited about the research we are doing. Moreover, we are working towards a point when we will be able to collaborate with other disciplines as well:

- We will work with research biologists in our Research Cancer Center (mathematical biology);
- We will work with research biologists, public policy experts and political scientists (complexity theory);
- we will collaborate with mathematicians from AIM, ORNL, Georgia Tech (dynamical systems);
- The list goes on.

3. RESEARCH DRIVES TEACHING AND IT PROMPTS STUDENT LEARNING

To meet the demands of tomorrow's workforce students must be better prepared. They must possess the right tools and skills. With improved student learning come improved skills and improved confidence to compete in the job market. Research skills and student learning are complementary. As the students' research skills improve, the students' inquisitiveness peaks, the students' interest increases and ultimately, student learning improves. The FFRG believes that:

- Research prompts student interest, student success, student learning, etc.
- Guided student research prompts student inquisitiveness, student learning, student interest, and student success.
- Research prompts students to pose the right questions in an effort to solve a problem or resolve an issue.
- One-to-one, two-to-one, three-to-one, or four-to-one interactions prompt student interest and student learning.

4. AN OVERVIEW OF THE EMERGENT MODEL

One method the FFRG will use is the *Emergent Model*. We will use the Emergent Model as one of the key components to facilitate improvement in undergraduate, and graduate, student learning at Clark Atlanta University (CAU). We believe the *Emergent Model* will prompt student interest, high student performance, student learning, etc.

The *Emergent Model* is an excellent research vehicle within which contemporary research methods may be used and, hopefully, exploited to seek out emerging schemes to resolve issues and/or to solve problems. [4] This model, based on concepts found in complexity theory (ant model, bee model, genetic algorithm, etc.) gives a fresh approach to research design, and it stresses groups operating as individual, self-contained entities to which no central manager, or central center, is answered. [5] [6] The ideas of both feedback and interdisciplinary research will undergird this model. By using this model, we believe the resulting research results will be superior results. We are trying to produce a superior research model using positive feedback, observation, etc. This is equivalent to ants leaving pheromone along a path. [5, pp.38-40] "…the resulting process will usually provide an efficient and fair solution that optimizes…" [5, p. 41] FFRG are always trying to get the best solutions, or the best methods when using the *Emergent Model*. "…getting the best solutions, however, requires very carefully setting conditions…" [5, p. 41]

This model is in its theoretical infancy. More and more activities and experiments will be designed to test this model. Interdisciplinary research will be stressed. The *Emergent Model* should be great for interdisciplinary groups consisting of a mathematician, a statistician, a physicist, and a generalist. It should be great for interdisciplinary groups consisting of a public policy expert, an economist, a decision sciences expert, a physicist, a mathematician and a computer scientist. Moreover, it should be great for interdisciplinary groups consisting of a biologist, an ecologist, a mathematician, a sociologist and an engineer. The list goes on.

5. TOOLS NEEDED BEFORE UTILIZING THE EMERGENT MODEL

Both graduate and undergraduate researchers/research trainees will be equipped with both basic and pure research tools before they embark on their research assignments. They will be exposed to basic tools of differential equations and dynamical systems, they will be trained to interpret large sets of data and they will be exposed to some of the designs of experiments, along with hypothesis testing, etc.

Research Planning and Research Methodology:
Young researchers will be exposed to the research methodology employed by separate academic disciplines in collecting and processing data within the framework of the research process [7, p.83], since they will be heavily involved in interdisciplinary research.

The young researchers will be asked to plan and design their strategies before they begin. They will ask to consider the principal area of investigation (the problem) by dividing it into more manageable subareas. They will be asked to ponder over the main issues and seek direction through appropriate hypotheses based upon obvious assumptions.

In essence, the young researchers will be exposed to the characteristics of both basic and pure research before they begin their field work. As an example, the characteristics of pure research will be promulgated into the researchers before they embark on the research mission.

Characteristics of Pure Research [7, p. 5]
(in Research Methodology)

1. Research originates with a question.
2. Research demands a clear articulation of a goal
3. Research requires a specific plan of procedure.
4. Research usually divides the principal problem into more manageable subproblems.
5. Research is tentatively guided by constructs called hypotheses.
6. Research will countenance only hard, measurable data in attempting to resolve the problem.
7. Research is, by its nature, circular; or more exactly, helical (research begets research).

The Emergent Model is given below.

6. THE EMERGENT MODEL

Ways to create superior groups, projects, products and schemes:

1. Create groups of small sizes, say 3-4.
2. The number of groups should be no more than 10.
3. Work within groups and across groups.
4. Assign a complex task, the same task, to all groups.
5. Send out a directive that all groups, initially, should work independently for about two weeks.
6. The task assignment should take about one to two months.
7. After the first two weeks, , pair the independent groups to review, brainstorm, and strategize about the assignment for a day (two to four hours).

   Example – Week 1
   {Group 1 with Group 10}
   {Group 4 with Group 2}
   {Group 8 with Group 3}
   {Group 5 with Group 9}
   {Group 6 with Group 7}

   Note: The pairings may change every week.

8. During the week or bi-week, individual groups will research and collect data, review, brainstorm, evaluate and strategize about the assignment as well as the groups' plans and progress.
9. By pairing groups each week, or bi-weekly, it is hoped that the feedback and collaborations will make each group better. We are counting on the fact that the feedback knowledge will make each group better, improve each group's process, strengthen the communication skills of all involved, prompt greater responsibility in all groups, prompt all groups to work towards a common goal and improve the ties and relations between groups.

10. There will be no central figure or manager controlling or managing the groups.
11. The groups, as self-contained and self-supporting entities, will work independently.
12. By the end of the task assignment, members of each of the manager-less or self-directed groups will have worked together to complete the task assignment.
13. The end result should be the emergence of a superior project, product or scheme.
14. Communication within and across groups must be stressed and done regularly.
15. In all, it is hoped that a deeper learning and engagement will be prompted by the processes and that the passion for persistence and performance will undergird the overall operation.

The members of the FFRG in the Department of Mathematical Sciences are:

Dr. Fisseha Abebe
Dr. Gary Chung
Dr. Temesgen Kebede
Dr. Charles Pierre

Undergraduate and graduate student researchers will be given short-term and long-term research assignments. After each assignment, the student researchers will attend both lecture and Q&A sessions hosted by the faculty researchers. During these sessions the faculty researchers will also conduct debriefings with the student researchers to monitor their progress and to improve the students' understanding of the concepts and ideas they researched.

Based on the idea that research drives teaching, the Q&A and debriefing sessions will facilitate the improvement of the students' understanding of the topics and ideas covered; hence, this process should improve student research skills and student learning.

7. PROPOSED MECHANISMS FOR INTERACTION AMONG UNDERGRADUATES AND FACULTY

It is believed that the *Emergent Model* will work very well with undergraduate students as well. Undergraduate students will be mentored by the Faculty Researchers. All undergraduate student researchers should be grouped, monitored, and given exceptional care. They will be treated as apprentices and given 1-to-1, 2-to-1, 3-to-1 or 4-to-1 attention throughout the process, whenever needed. Retention of students is a top priority and the FFRG will be focusing on this issue. The FFRG will endeavor to make sure that at least 80% of all student research candidates involved in the process graduate within four years and go on to be gainfully employed in the workforce, or go on to graduate school.

8. REFERENCES

[1] N.C. Anderson, et. al., **Facilitating Interdisciplinary Research,** The National Academies, Washington, DC: The National Academies Press, 2005.

[2] S.H. Strogatz, **Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering, Cambridge**, MA**:** Perseus Books Publishing, LLC, 1994.

[3] R.L. Devaney, **A First Course in Chaotic Dynamical Systems: Theory and Experiment,** Reading, MA: Addison-Wesley Publishing Company, Inc., 1992.

[4] M.M. Waldrop, **Complexity: The Emerging Science at the Edge of Order and Chaos**, New York: Simon & Schuster Paperbacks. 1992.

[5] L. Fisher, **The Perfect Swarm: The Science of Complexity in Everyday Life, New** York: Basic Books, 2009.

[6] M. Mitchell, **Complexity: A Guided Tour,** New York: Oxford University Press, Inc., 2009.

[7] P.D. Leedy, **Practical Research: Planning and Design (4th ed,),** NY: MacMillan Publishing Company, 1989.

Auto segmentation for Malay Speech Corpus

Tan Tian Swee, Ting Chee Ming, Chin Wee Lip, Lau Chee Yong, Sh-Hussain Salleh

Centre of Biomedical Engineering (CBE), Transportation Research Alliance (TRA), Universiti Teknologi Malaysia, 81310, Skudai, Johor, Malaysia

***Abstract*—Abstract-This paper deals with the automatic segmentation of Malay continuous speech database. Auto segmentation is a process of producing a sequence of discrete utterance with particular characteristics remaining constant within each one. In terms of quality, hand crafted segmentation would be the best method. However, due to the large database size, manual speech segmentation and labeling become tremendous. It is time consuming and error prone. Besides, even if the database is segmented by an expert, the segmentation rule may become subjective and not reproducible. Inconsistency result may occur from different linguistic experts. Thus, an automated segmentation rule was drawn to consistently segment the large scale database with satisfactory level of quality. Automated segmentation of Malay Language syllable is not a tough task because all syllables in Malay Language are pronounced almost equally and moreover it is not a tonal language like English. The manipulation and identification of the segment boundaries of Malay Language is straight forward and easy to understand. For the segmentation, the HMM based approach with adapted Viterbi force alignment technique is used. Composite HMM with Baum Welch reestimation was utilized to ease the process of phonetic segmentation. All the data from the database was fed into the segmentation tool directly without prior trained sample for pre-training purpose. For the design of the sentence coverage of the database, the scripts are consisting of 1000 sentences. 620 sentences are selected from primary school Malay Language text book and 380 sentences were computed using the 70% highest frequency words that appear in the 10 million words online digital text. This configuration of Malay Language script already promises a phonetically balanced database which covers all the vowels and consonants. The objective evaluation method is used to identify the performance. The result from the auto-segmentation was verified to obtain the accuracy degree and overall quality. The result was tested perceptually and it is proven to have satisfactory high quality.**

I. INTRODUCTION

In the speech technology, auto segmentation is a crucial step to process the database. Before the database can be feed into the system for another processing, the database should be segmented according to the needs of the system. Auto segmentation is a process of producing a sequence of discrete utterance with particular characteristics remaining constant within each one [3]. It is the first step in any tools or systems like morphological analyser, POS tagger, syntactic parser, etc., and applications, like machine translation, information extraction, information retrieval, etc. Finding the boundary of either a sentence or a token is nontrivial [1]. To obtain the best result of segmentation, manually craft the database is the best solution. But hand crafted process is time consuming and expensive [4]. Moreover, due to the large database size, manual speech segmentation and labeling become tremendous and error prone. Besides, even if the database is segmented by an expert, the segmentation rule may become subjective and not reproducible [2]. Inconsistency result may occur from different linguistic experts. Thus, an automated segmentation rule was drawn to consistently segment the large scale database with satisfactory level of quality. Automated segmentation of Malay Language syllable is not a tough task because all syllables in Malay Language are pronounced almost equally [5]. The segmentation of Malay Language has some similarities if compared to English. Firstly, Malay language is a phonetic language and it is also written in Roman characters like English. Secondly, it is not a tonal language because all the syllables in Malay are pronounced almost equally. As a review of Malay Language, there are six (6) main vowels and 29 consonants in standard Malay (SM). SM have a total of nineteen of the consonants, where /m/, /n/, /f/, /l/, /s/ and /y/ are pronounced almost the same way as in English. In Malay language, the syllabic structure is well-defined and can be unambiguously derived from a phone string. The basic syllable structure of the Malay language is generated by an ordered series of three syllabication rules. The linguists claimed that Malay is a Type III language, namely Consonant-Vowel (CV) and Consonant-Vowel- Consonant (CVC) are the most common and they can be found almost in every Malay primary word (Noraini et al. 2008). To perform the segmentation process, hidden Markov model (HMM) is used. HMM is a statistical model which is able to model patterns and sequence. It is widely used in speech technology like recognition and speech synthesis. The basic idea of HMM is the transition of states and observations. All the states inside the HMM consist of a mean and variance which is able to release the desire output with high probability after trained. The model can be illustrated as the figure below.

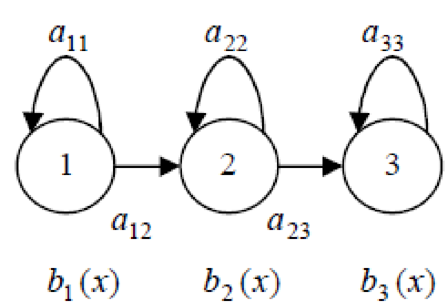


Figure 1. Block diagram of hidden Markov model

Previously proposed idea was adding post refinement to increase the quality of the segmentation result. The post refinement method was called implicit boundary refinement with using Viterbi forced alignment. The method mechanism is to extend the start point and end point of the training data to the point next to it or its adjacent point. So, each of the training data was embedded with wider range of boundary. That means the training process can be conducted in an easier way and the HMM can be better trained and better model the phonetic boundaries. This post refinement would enable the Viterbi alignment process conducted with better accuracy and less errors. The block diagram of the process can be illustrated in the figure below.

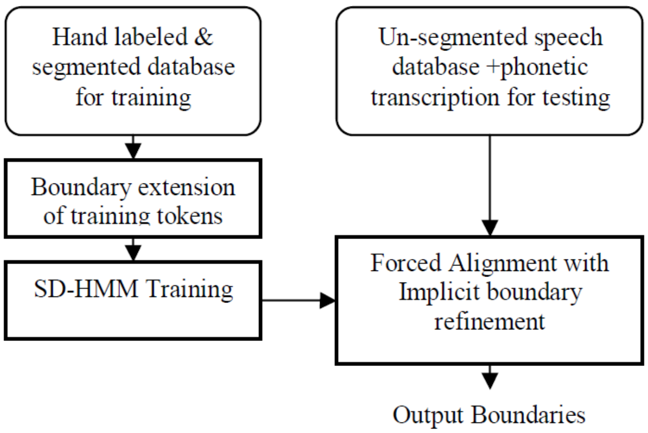


Figure 2. Block diagram of automatic segmentation with implicit boundary refinement

For the implicit boundary refinement method, the starting point and end point of the phoneme were extended until the adjacent point. The solid line represent the original boundary while the dotted line is the extended boundary of the phoneme. The result is shown in the figure below. The problem of this method is the decision to point the boundary. Manual refinement would result in the inconsistency and inaccurate. All the manual refinement is just based on approximation. But the key idea of the refinement is to enable the training process to become easier and reduce error.

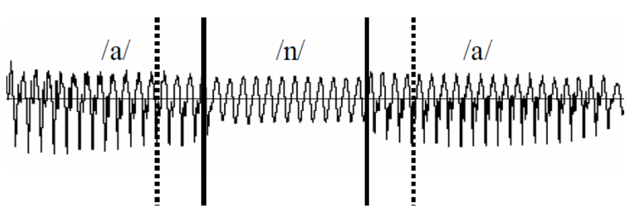


Figure 3. Starting point and end point extension from the original boundary

However, this kind of automatic segmentation requires a pre-manual hand seeded segmented data for the training process. If the language is changed the hand seeded process also has to change. Moreover, manual refinement could result in inconsistency and error prone. So, a fully automated segmentation tool is anticipated to ease the problem. In this project, a fully automated phonetic segmentation is proposed rather than using post refinement method.

II. Methodology

This paper introduces the study of fully automated segmentation of a Malay language database. The theory used to perform the auto segmentation is the composite HMM trained by Baum Welch reestimation and implement the Viterbi alignment on all samples. The advantage of this method is fully automated and also able to obtain satisfactory level of quality. The segmentation system used in this study consists of two phases: Initialization and Iteration. The block diagram of the process is shown below.

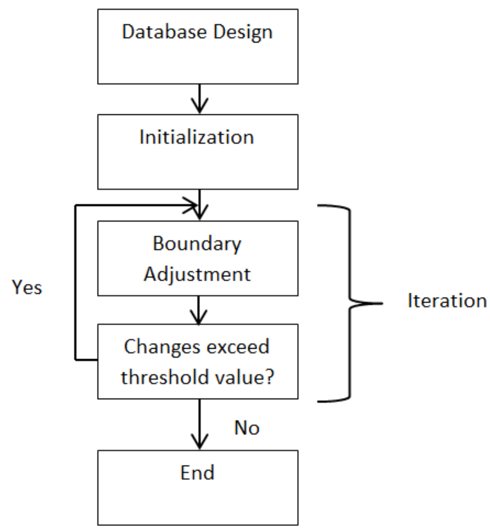


Figure 4. Block diagram of auto segmentation

For the initialization phase, all the data from the Malay Language database will go through the feature extraction process. A normal speech can be model using 39 element feature vector model. The 39 element consist of 13 of static features, 13 delta vector and 13 delta-delta vector. The 13 static features contain 12 MFCC computed from 24 filter banks and log energy. The features extraction can be illustrated as the figure below.

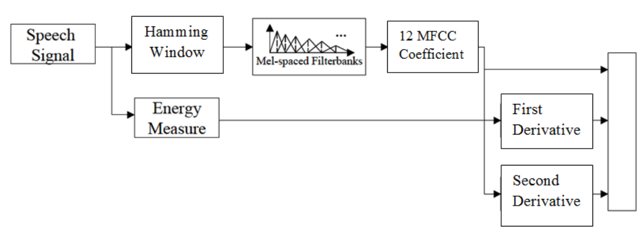


Figure 5. Features extraction process

After that, the HMM was trained using the features extracted from the process above, then the HMM now knows how to segment the rest of the data. All the data were fed into the system to perform the initialization. The initialization means to equally distribute the frame number for all the phoneme by dividing the total frame to the total number of phoneme in the sentence. This is the first entry step for every data so that the iteration can be conducted based on this initialization result. After that, the iteration takes place to adjust the boundary of the phoneme according to the clusters. The first iteration is based on the initialization result and slowly move the boundary according to the frame characteristic. The iterations after that were proceed if the changes of the boundary exceed the stopping threshold of the segmentation process. The process can be illustrated in figure below.

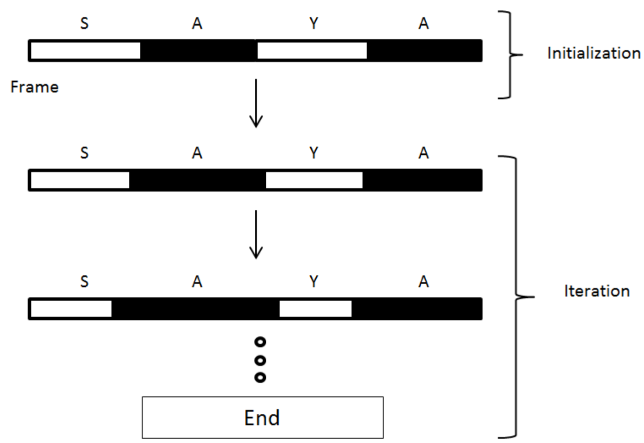


Figure 6. Iteration process of auto segmentation

After the iteration stops, that means the auto segmentation process is done and the result is ready.

A. Database Design

For the design of the sentence coverage of the database, the scripts are consisting of 1000 sentences. 620 sentences are all the words in primary school Malay Language text book and 380 sentences are the 70% highest frequency words that appear in the 10 million words online digital text. This configuration of Malay Language script already promises a phonetically balanced database which covers all the vowels and consonants.In Malay language, there are 24 pure phonemes and 6 borrowed phonemes, divided into 8 categories. Among the pure phonemes, there are 18 consonants and 6 vowels. The borrowed consonantal phonemes are /f, z, sy, kh, gh, v/. Five diphthongs can be found in Malay language which are /ai/, /au/, /oi/, /ua/, /ia/. This Malay phone set cover all Malay phoneme unit in Malay language. The segmentation experiment is based on the 35 phones above and a silent model for pausing /pau/. The /gh/ was folded to /g/ due to limited training tokens. All the sentences have been hand labeled and segmented according to the chosen Malay phone set in Tabel 1.

Category	Malay Phones
Vowels	/a/, /e/, /eh/, /i/, /o/, /u/
Plosives	/b/, /d/, /g/, /p/, /t/, /k/
Affricates	/j/, /c/
Fricatives	/s/, /h/, /f/, /z/, /sy/, /kh/, /gh/, /v/.
Nasal	/m/, /n/, /ng/, /ny/
Trill	/r/
Lateral	/l/
Semi-vowel	/w/, /y/

Table I
LIST OF MALAY PHONES ACCORDING TO CATEGORIES

III. RESULT

The auto segmentation process was taken place to segment all the Malay sentences. The result shows the consistency of the segmentation output which can avoid the inconsistency due to manual segmentation. The result was verified according to its waveform and the result is shown in the figure below.

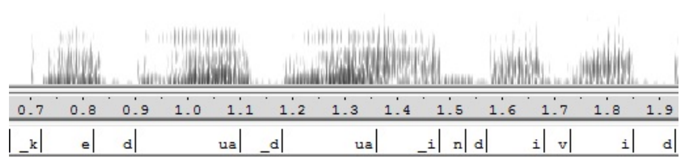


Figure 7. Result of auto segmentation

This method is purely automated and no manually hand crafted input needed. Once the HMM was trained using the features extracted from the same database, then the HMM knows how to segment the waveform. The result shows the waveform was segmented according to its phoneme. The perceptual evaluation is carried through by selecting the phoneme segment and playback to verify whether the segment is correct or not. The data was randomly picked and listen as an evaluation. The result shows that the segment of all the data was correct according to its phoneme. So can be concluded that the segmentation tool is reliable. The result was also compared with the auto segmentation with implicit boundary refinement perceptually. For the segmentation with implicit boundary refinement, the phonetic segmentation was first carried out without the implicit boundary refinement and the result obtained was regard as the standard segmentation result to be compared to another method. The phonetic segmentation with implicit boundary refinement also carried out and the result was compared to the previous result without implicit boundary refinement. The segmentation result was improved by implicit boundary refinement and the post refinement is more accurate in the zone of small tolerances. It can be said that the implicit boundary method is capable to increase the precision of segmentation. However, the comparison between the method proposed in this paper and the method with post refinement, the quality and precision of both of the result shows no much difference and changes using the subjective evaluation. So can be concluded that the method proposed in this paper is capable to obtain the same result with the auto segmentation with implicit boundary refinement. However the main advantages of this method is it is purely automated without hand crafted data as training data.

IV. Conclusion

In this study, automatic Malay Language segmentation is described. This provides the basis for preparing segmented speech database for Malay TTS. Composite HMM based with Baum Welch reestimation approach using Viterbi alignment is used for the segmentation. From the obtained result, it can be proven that the auto segmentation tool is able to get the satisfactory level of result. The syllable based segmentation can perform as good as the segmentation with implicit boundary refinement. For the future work, the testing can be conducted with more feature set and different number of Gaussian number to verify the reliability of the system.

V. Acknowledgement

This research project is supported by CBE (Central of Biomedical Engineering) at Universiti Teknologi Malaysia and funded by Minister of Higher Education (MOHE), Malaysia under GUP grant Tier 1 vot number 01H49 with the project title Development of Malay Speech Technology for Preschool Vocabulary Learning System.

References

[1] Bali Ranaivo-Malancon (2011). Building a Rule-based Malay Text Segmentation Tool. International Conference on Asian Language Processing.

[2] F. Brugnara, D. Falavigna and M. Omologo (1993). Automatic segmentation and labeling of speech based on Hidden Markov Models. Speech Communication 12 (1993) 357-370

[3] Matthew A. Siegler, Uday Jain, Bhiksha Raj, Richard M. (1997). Automatic Segmentation, Classification and Clustering of Broadcast News Audio. Proc. DARPA Speech Recognition Workshop page 97-99

[4] Noraini Seman, Kamaruzaman Jusoff (2008). Automatic Segmentation and Labeling for Spontaneous Standard Malay Speech Recognition. International Conference on Advanced Computer Theory and Engineering.

[5] Zaharah Othman and Sutanto Atmosumarto (1995). Colloquial Malay: a complete language course for beginners.

A Glance at Educational Governance for IDREC 2012

Dr. Laura THOMPSON
Graduate Programs
Keiser University
Fort Lauderdale, FL 33309 USA

ABSTRACT

The purpose of this study is to look at global governance in varying aspects of higher educational environments. Governance is defined as the "control and oversight over policies, procedures and standards for application development…., design, implementation, testing and monitoring of deployed services." As an essential part of this process, the researcher used information systems evaluation to derive the components that make up educational systems today. These include but are not limited to the instructor, the student, which was part of a traditional educational model; with the content and technology, which were viewed as part of a fixed model of instruction [15].

Keywords: Governance, Information Systems Evaluation, Music Stakeholder Alignment Model, Cloud-based Networking, Information Sharing

INTRODUCTION

University network systems are continuously charged with the task of allocating sufficient resources for its faculty and its ever-changing population of students. In these environments, the student and or faculty member are often referred to as the system users (consumers), and receive a system identification (ID) number to access the computer network. The ID number helps the information systems department keep track of who's on the network at any given time.

Although sharing is not recommended by Information System management, there are no precautions limiting clients from sharing ID numbers. In most cases, the consumer will not access the same computer each time he or she logs into the network. Yet in other cases, Information Technology (IT) management decides what to "police/govern" and determines what information to make accessible in general. A typical example of this is seen in environments where a generic password gets students and faculty on the network. However, individual passwords with provide access to student person records. In addition, student, faculty and administrative networks are separated, with no information access between the networks. Further, to access previously downloaded files, the system user will have to find the computer he/she worked on and hope to find it vacant of other users.

To rid the network of unwanted files, one popular university technique is to take individual computer systems go through a process called "re-imagining". Re-imaging is where the files that do not match the "software requests" are erased from the system. Because of the length of time needed to complete the process can take up to 3 hours per computer, technology managers deem the best time to re-image a computer is during the semester break. Of course for the computer hacker, this could pose as the ideal time to access and compromise a network prior to the start of the next semester. While this cleaning process is essential for the individual function of the computer and the overall functionality of the network, it serves as an inconvenience for consumers trying to find their file downloads, forcing them to seek alternative storage methods.

Consequently other university systems provide their users with laptops. Each time the user logs into the network the laptop is updated with the user's "roaming profile", which contains up to 6 megabytes of files he/she added to his/her computer system. This profile can be accessed from any computer on the network.

In this scenario, but both campus locations rely on some sort of system evaluation to dictate how the network will be set up each term. However, at the end of each semester, computers go through the endless cycle of re-imagining again, which is an indication that users are not part of the decision-making process.

According to Thompson (2008) the fixed and traditional models can come together to create strategic alignment [18]. Educational models like e-Learning (Guri-Rosenblit, 2009) are distributed using technology across multiple channels. According to presenters at the Axmedis Conference on Content and Media Distribution (2006), future models will incorporate web based, local and mobile based distribution of metadata content for multi-channel distribution [1].

REVIEW OF LITERATURE

The concept of information systems evaluation is seen in an article by Vassilis Serafeimidis and Steve Smithson (July, 2003), which discusses the principles surrounding information systems evaluation. In the article, the authors state that cognitive psychology attempts to provide the necessary structures to reveal and understand the individual's knowledge, views, emotions and perceptions [26]. Although challenging, it is important for IT managers to closely simulate the potential work environments of its system users. This along with the principles taught in the classroom prepares students for the real work world.

By definition, the word schema is used to determine an individual's reality. 'Interpretative schemes' then are seen as a 'modality' of shared cognitive structures for understanding and as an initiative for actions [26]. When considering the interpretative schemes, the authors recommend that the following people/groups should be considered in the initial information systems decision-making process:

- Senior Executives
- Department Managers
- End users
- IS Managers & Support Staff

In addition, Serafeimidis and Smithson have identified that the personality of the group, the interests of the individual users and the organizational role are factors that will eventually affect the evaluation process [26].

Consequently, system users are not included in the decision making process. Thus the re-imaging of computers becomes part of the information system post-evaluation process each semester. In an article reprinted by Hiltz and Johnson, authors Ives, Hamilton and Davis (1990) state "The behavior of one subsystem is dependent on its environmental relationship to other subsystems so that different system characteristics may be optimal for users and groups with different characteristics" [13]. In order to see the ramifications of this statement, it's important to investigate the types of files system users may be adding to the network and or their individual roaming profiles.

Distribution across multiple channels is a tool that will minimize inefficiency among stakeholders, reduce overall costs to do business and furthermore provide opportunities for consumers to be direct distributors using the forms of electronic devices they have at their disposal. In certain countries like Algeria, for example, the distribution of e-learning can be standardized on a national level [2].

A redesign of the IT infrastructure is required to promote the information access (for student access and e-learning verification process), the capture of information and content (so that students have the freedom to select and send files from a variety of formats), and the sharing of information and content (between stakeholders such as students and instructor or student peer groups – where applicable). According to Benchicou, Aichouni and Nehari (2010), websites can serve as portals providing access to both students and faculty, and thus setting up the school to be accessible globally. The platform engages people to interact with each other (forming communities) and with the university [2].

With the trends in campus networks shifting more to wireless technology, information system managers will ultimately determine if students and faculty stay or leave a university. According to Roach, despite the economic challenges that exist in IT budgets, there has been growth in wireless networking (based on the "green reports"). This enables users to access campus Internet connections "from their laptops and other computing devices" [25]. In addition to this trend, approximately a sixty-six percent of campus networks have adopted some policy prohibiting the illegal download of digital content [25]. Further Roach states that college students are easy prey for "media industry officials" to target when trying to locate violators of these policies.

Both information content stakeholders and consumers base the redesign process on strategic balance between realistic system capabilities, limitations and technology accessible. From this, the researcher derived a definition of technological preparedness as the freedom to promote the access, the capture and the sharing of information and content through technology. However like the music stakeholder alignment model [28], misalignment may occur when stakeholders do not have access to the technology or can not make adequate decisions based on the technology available. Benchicou, Aichouni and Nehari, (2010) state that this is seen more with countries with low economic status or those whose infrastructure faces challenges including power allocation, lack or resources and technology, or even lack of skilled persons accessing the technology [2].

Challenges also exist in countries like Nigeria, where the population size for youth attending school has outgrown the nation's educational supply. In this case, the lack of academic standard make it difficult adequately distribute education. With this comes the need to establish policies on secure and regulatory. One of the main areas to regulate would be information access and information sharing.

According to Jansen and Grance (2011), the motivation toward cloud computing should increase efficiency, not decrease measures to provide network security (p. 5). Unfortunately in some cases, measures to secure network resources in the cloud have failed. The authors concur that the growth of cloud computing stems from the culmination of networks accessing and sharing resources and information. It can also be concluded that cloud computing is the force behind the expansion of network boundaries in organizations. Because of this, the organizational infrastructure must now implement control and oversight of the 'potential' access of other networks that are not part of the organizations initial perimeter of stakeholders [15].

Because of its technologically savvy end user base, university networks are geared to be the most sophisticated in nature. Further the interconnected nature of university networks and the simplicity with which, network platforms can be accessed create, and in some cases, mandate the need for some form of global governance in education. Finally, the cost cutting measures used to reduce capital investment in IT resources increases the need to utilize cloud computing options. But with it, also amplifies the need to impose standards in information access, information capture and information sharing. Some form of alignment between shared networks can address issues of illegal file sharing/network sharing.

Technology companies invest lots of resources in the protection of intellectual property [23]. In higher education, this cost gets transferred to the student in what's commonly known as a "technology fee". Software and DRM providers like Microsoft ™ have the ability to set and enforce standards for intellectual property access. This is seen especially in cases where the Internet is used as a means to access digital content through hyperlinks or via direct file downloads.

Authors Boucqueau, Delaigle, and Goray (1999) describe a variety of case scenarios in the project and the procedures governing secure transactions in open network environments. Several network distribution models are presented as they relate to the secure web transaction process, which requires the registration of web content through Certification Authorities (CA) or Trusted Third Parties (TTP) (e.g. Equifax) [5].

HYPOTHESIS

The literature review contains a good basis for examination of variables and constructs that determine online behavior among college students. For the proposed study the following hypotheses was examined:
H1. Students' music download behavior will be influenced by available storage capacity.
HO1. Student's music download behavior will not be influenced available storage capacity.
Methodology

Participants in this study were college students from a university in Georgia. The National Center for Education Statistics shows that in Fall, 2003 the university's Decatur location had a total enrollment of 5,151 students with a total of 3,974 undergraduates. At that time, the undergraduate enrollment was made up of 59.3% men and 40.7% women. This is relevant to the study because Goldsmith (2001) found that music downloads are 75% more prevalent with men than with women (Goldsmith, 2001).

The dependent variable was Student downloads and the independent variable for this study was: "Student downloads with the influence of available storage capacity" (Active Independent Variable (IV)). The active IV was manipulated through survey questions that will determine student computer skill/usage, behavior characteristics, and ethnic origin. The attribute variable measured the general demographics and motivation factors of students that download music.

From this, the study investigated the motivational factors relating to downloads by academic major, ethnic origin, sex and age, as well as the influence of storage and available capacity via thin client/cloud-based networks on student's download decisions. The dependent variable is download behavior.

One third of the questions were consistent with surveys conducted in the literature reviews for the purpose of comparison analysis and validation. Approximately six questions were designed to measure motivational factors/characteristics that comprise a thin client/cloud-based network. These questions were measured using a Likert scale. The remaining questions were nominal scale based (e.g. race, sex, college major, age range, etc.).

FINDINGS

The more significant of the two clusters contained General Demographics (Age, College major and Sex) and questions relating to Motivational factors, as seen below.
Although 77.6% of the participants have experience downloading music, only 71% see illegal music downloads as theft. Approximately 69% of the participants were male. 46% of the participants were 25 years of age and younger. Finally, study showed that the ethnicity of the student was not significant to become part of the group membership. The overall results of this initial study show that Behavior, Computer skill/Usage, and Ethnic Origin are significant in determining group membership for online behavior, but are insignificant in determining the motivational factors for downloading music.

To test the null hypothesis, measures of central tendency were conducted as the analysis moved from a three step to a two step cluster analysis. The results indicate that there is a change as the centroids merge clusters 1 and 2 from the 3 step cluster to the two step cluster. However, there is no change in mean as the clusters centroids migrate in the second cluster.

Figure 1:Three Step Cluster Centroids

Three Step		ResponseID Mean	ResponseID Std. Deviation
Cluster	1	182.40	106.656
	2	187.80	106.718
	3	183.50	106.570
	Combined	184.50	106.377

Two Step		ResponseID Mean	ResponseID Std. Deviation
Cluster	1	185.10	106.488
	2	183.50	106.570
	Combined	184.50	106.377

Figure 2: Two Step Cluster Centroids

Thus the null hypothesis was rejected as the analysis moves from a three step to a two step cluster analysis.

SUMMARY

Many in information systems would argue that system users do not directly pay the salary of information systems officers, and therefore should not be part of the decision making or evaluation process. A viewpoint like that is considered positivistic. In the positivist view, data is objective & independent of the researcher. Within this viewpoint, there is a single scientific method. However, in an article written by Bichler and Loebbecke, the concept of "user modeling and personalization of information systems" is expressed "as important components of software engineering and information systems development" [3].

An example of this positivistic view in information systems is seen in an article written by Walsham (1995). Here, it shows that information systems (IS) evaluation has been dominated by positivistic scientific paradigm (Walsham, 1995). According to Walsham, information systems IS evaluation is referred to as a complex decision making process because of the various social and organizational processes it is embedded in.

Cognitive psychology attempts to provide the necessary structures to reveal and understand the individual's knowledge, views, emotions and perceptions. On the other hand, the relativist viewpoint recognizes that there's no purely observational language. Further there is not one single scientific method & and no method is superior to the other. Moreover, theories about consumer research are not universal and all data are theory-laden.

In a comparison analysis of the social interaction with ants, Stengal, Bleimann, and Stynes discover new ways to create a "virtual, university network architecture" [27]. The purpose of this study was to further break down the digital divide between members of the community, the working public, and the handicap. The authors break down the system needs of the university which include "mentoring, assignment delivery and evaluation" [27].

According to the authors, the behavioral patterns of ants were analyzed in order to enhance network routing. Similarities were identified in territory marking, in the exchange of food and data. The results of the study show that a university network of this nature requires both an asynchronous and a synchronous network to be implemented due to the complex nature of the ants' patterns.

Finally, McGrath (2002) writes an article, which discusses how librarians organize material, make it available and assert their power through consortia. Meanwhile, aggregators and intermediaries look for new ways to mediate; technologists invent new devices and we all have to grapple with the consequences.

The process of unauthorized data access and measures used to protect databases is presented in an article by Nance (1998). Here, the implementation of IDs is used to access databases enabled servers in order to identify and validate users on the network. Many of the mechanisms used to protect database servers, (such as firewalls, proxy servers, private networks, private links, the reliance on secure internet addresses/host names, and confining database server traffic to specific node address), provide a false sense of security to database administrators [21].

Lee (2003) defines a document as "any record of information and any other material data stored or recorded by mechanical or electronic means" [17]. In a study on university networks, Lee presents a university server containing illegal documents uploaded by members of the math department. Evidence of the server activity is captured on activity logs and was admissible by the courts. According to Lee, "control could be destructive to university life because it would create a chilling environment, when universities generally strive for an environment of intellectual freedom and creative exploitation" [17].

None of the methods mentioned provide guarantees of secured systems [21]. The overall goal of the models described is to provide protection of intellectual property rights, especially in the case of piracy. Based on the distribution network content providers contend with issues such as quality of service, transmission at the TCP/IP layer of the OSI model, public keys, and authentication [21].

The Internet has made the way for many technological and logistical advances, making it more efficient to transfer resources from business to business (B2B) and to a large variety of customers. Even today, many of the traditional models for distribution have failed. This demise has caused brick and mortar stores to share resources through networks online where the Internet (electronic media) is used as the means to transfer/distribute products and services globally [30].

iTunes is described as a sophisticated collaboration of network and database relationships and dependencies enabling consumers to react with copyright and contract law and DRM technologies with one click [9]. Industry experts predict drastic increases with this interaction as it's anticipated that nearly all entertainment and music media will take on a digital format by 2009. This has manifested itself by way of Podcasting Technology, which is used in what's called the Opencast Mobile Learning environment.

According to authors Boyinbode, Bagula and Ngambi, (2011), students have proven to increase their level of class participation with the advent and implementation of podcasts in the e-learning environment. The main challenge with Opencast Mobile instruction is found in the preparedness of instruction. Along with podcasts mobile learning (called M-learning) also has architectural challenges ("intermittent connections and lack of cross-platform solutions") that enable the users to have adequate access to content [6]. However it, like other options in the Opencast Framework provides the most flexibility for instruction.

CONCLUSION

Through this study students' intentions were examined as they go through the process of downloading music. The study indicates that students want are motivated to save music after they've downloaded it. Based on the answers to the motivational factor/characteristics that comprise a thin client/cloud-based network, younger students under 25 years of age desire to be have more mobile with the music they download than students over 40 years.

It has been said that one of the best ways to keep track of unauthorized data access is through centralization. In the case of the EINSTEIN project, the U.S. Government has created an intrusion detection and prevention system "on all incoming communication" (Bellovin, Bradner, Diffie, Landau and Rexford, 2011). According to government reports, the increase in intrusion based attacks created the need for the EINSTEIN project. The intent of the project is to check all internet traffic coming into the federal government. The nature of EINSTEIN is one that through Correlation, inspects "network traffic in real time as they appear" (Bellovin, Bradner, Diffie, Landau and Rexford, 2011).

Authors like Goldsmith (2002), LaRose and Eastin (2002) say that most online consumers are college students. In fact a high percentage of the studies dealing with online buying behavior are conducted with college students. According to LaRose and Eastin (2002), "college-aged consumers spend 12% of their total income online, and those purchases are concentrated in categories (clothes, music, and books) frequently associated with compulsive buying" [16]. The authors "Interpret research on impulsive, compulsive, and addictive online buying within a theory of human behavior well known to media effects researchers: Albert Bandura's (1986) social cognitive theory" [16].

REFERENCES

(1) AXMEDIS.org (2006, December). Automating production of cross media content for multi-channel distribution. *Axmedis 2nd Annual Conference on Content Distribution.*

(2) Benchicou, S., Aichouni, M., & Nehari, D. (2010). E-learning in engineering education: a theoretical and empirical study of the Algerian higher education institution. *European Journal Of Engineering Education*, *35*(3), 325-343. doi:10.1080/03043797.2010.483610

(3) Bichler, M., & Loebbecke, C. (2000, April – June). Pricing strategies and techniques for online delivered content, *Journal of End User Computing*, 12 (2), 4 – 10.

(4) Book, M. (2000). Where commerce, culture and the first amendment collide: Questions on ethics in software engineering raised by the Napster case. Retrieved August 3, 2005 from http://www.matthiasbook.de/papers/napster/files/napster-paper.pdf

(5) Boucqueau, J. M., Delaigle, J. F., Goray, E. (1999, June 8). *How to secure IPR overexisting open networks. On the feasibility to protect IPR in existing open networks*. ISO/TC 46/SC 9/Working Group 1 for ISO Project 15706: International Standard Audiovisual Number (ISAN), Europe. Retrieved from http://www.nlcbnc. ca/iso/tc46sc9/isan.htm

(6) Boyinbode, O., Bagula, A., & Ngambi, D. (2011). An Opencast Mobile learning Framework for Enhancing Learning in Higher Education. *International Journal Of U- & E-Service, Science & Technology*, *4*(3), 11-18.

(7) Coyle, K. (2001). Stakeholders and standards in the e-book ecology: Or, it's the economics, stupid. *Library Hi Tech*, Bradford: 19(4), 314-324. Retrieved on July 29, 2003 from http://proquest.umi.com

(8) Deeter-Schlmelz, D., R., Bizzari, A., Graham, R., & Howdyshell, C. (2001). Business-to-business online purchasing: Suppliers' impact on buyers' adoption and usage intent, *Journal of Supply Chain Management*, Tempe: 37(1), 4-10.

(9) Fisher, III, W. W. (2004, June). iTunes how copyright, contract and technology shape the business of digital media – a case study, Berkman Publication Series No. 2004 – 07, 1 - 103. Retrieved December 29, 2005 from http://scholar.Google.com.

(10) Gaither, C. (2003, September 24). Recording industry withdraws suit: mistaken identity raises questions on legal strategy [Electronic Version]. *The Boston Globe*.

(11) Goldsmith, R., (2001). Using the domain specific innovativeness scale to identify innovative Internet consumers, *Internet Research*, Bradford: 11(2), 149-158.

(12) Hendry, C., Brown, J., & Defillippi, R., (2000, April). Regional clustering of high technology-based firms: Opto-electronics in three countries. *Regional Studies.* Cambridge: 34(2), 129-144.

(13) Hiltz, S. & Johnson, K. (June, 1990). User Satisfaction with Computer-Mediated Communication Systems, *Management Science*, Newark: v16, i6, pp 739 – 764.

(14) Intracom S. A. (2004, August 8). Multimedia streaming of interactive content across mobile networks (MUSICAL), *eContent*: EDC-22131-Musical/27193: 1(3) 1 –50. Retrieved May 5, 2005 from www.google.com

(15) Jansen, W., & Grance, T. (2011, January). Draft NIST Special Publication 800 – 144: guidelines on security and privacy in public cloud computing, *National Institute of Standards and Technology*, *U.S. Department of Commerce*, Gathersburg: 800 – 144, 1 – 60

(16) LaRose, R., & Eastin, M., S. (2002). Is online buying out of control? Electronic commerce and consumer self-regulation, *Journal of Broadcasting & Electronic Media*, Washington: 46(4), 549 – 564.

(17) Lee, C. (October, 2003). Music Companies on Trail of Copyright Offenders, *Managing Intellectual Property*, i133, p 53

(18) Majchrzak, A., Rice, R. E., Malhotra, A., King, N, & Ba, S. (2000). Technology adaptation: The case of a computer-supported inter-organizational virtual team. *MIS Quarterly,* 24(4), 569-600.

(19) Malhotra, N. K. (2004). *Marketing research: An applied orientation,* (4th ed.). Upper Saddle River: Prentice Hall, Inc.

(20) McGrath, M. (2002). Assumptions versus reality: User behaviour in sourcing scholarly information, Interlending & Document Supply, Bradford: v30, i3, pp 120 – 125.

(21) Nance, B. (1998, October 1). Improving data access security. *Network Computing, 9(19),* 122-124.

(22) Patel, A. (2003, March). Generating ROI is a top priority for data-capable networks. *Wireless Systems Design*. , 45 – 50.

(23) Peitz, M. & Waelbroeck, P. (2005). An economist's guide to digital music. *CESinfo Economic Studies*, *51*(2-3), 359 – 428. Retrieved December 30, 2005 from http://scholar.Google.com.

(24) Privacy and security as simple as possible - but not more so. (2011). *Communications of the ACM*, *54*(8), 30-33. doi:10.1145/1978542.1978553

(25) Roach, R. (January 1, 2004). "Wireless Movement Dominating Campus IT Change", *Black Issues in Higher Education*. Reston: v20, i23, p 38.

(26) Serafeimidis, V. & Smithson, S. (July 2003). Information systems Evaluation as an Organizational Institution Experience from a case study. Information Systems Journal. v13, i3, pp 251 – 275.

(27) Stengal, I., Bleimann, U., & Stynes, J., (2003). Social Insects and Mobile Agents in a Virtual University. Campus – Wide Information Systems. Bradford: v20, i3, pp 84 – 90.

(28) Thompson, L. A., (2008). Globalization and its impact on music marketing and distribution, Argosy University, Sarasota: pp 5 – 90.

(29) Thompson, L. A., (2005). Download Behavior in Thin Client/Cloud-based Environments, Argosy University, Sarasota: pp 7 – 13.

(30) Vaccaro, V. L., & Cohn, D. (2004). The evolution of business models and marketing strategy in the music industry. *The International Journal of Media Management,* 6 (*1 & 2*), 46 – 58.

Analysis of the Largest Normalized Residual Test Robustness for Measurements Gross Errors Processing in the WLS State Estimator

Breno CARVALHO
Electrical Engineering Department, University of São Paulo
São Carlos, São Paulo, Brazil

and

Newton BRETAS
Electrical Engineering Department, University of São Paulo
São Carlos, São Paulo, Brazil

ABSTRACT

This paper purpose is to implement a computational program to estimate the states (complex nodal voltages) of a power system and showing that the largest normalized residual (LNR) test fails many times. The chosen solution method was the Weighted Least Squares (WLS). Once the states are estimated a gross error analysis is made with the purpose to detect and identify the measurements that may contain gross errors (GEs), which can interfere in the estimated states, leading the process to an erroneous state estimation. If a measure is identified as having error, it is discarded of the measurement set and the whole process is remade until all measures are within an acceptable error threshold. To validate the implemented software there have been done several computer simulations in the IEEE´s systems of 6 and 14 buses, where satisfactory results were obtained.

Another purpose is to show that even a widespread method as the LNR test is subjected to serious conceptual flaws, probably due to a lack of mathematical foundation attendance in the methodology. The paper highlights the need for continuous improvement of the employed techniques and a critical view, on the part of the researchers, to see those types of failures.

Keywords: Power Systems, State Estimation, Gross Errors, Normalized Residual.

1. INTRODUCTION

The real time power systems operation has the main objective to keep the electrical system operating. To achieve this goal it is necessary that the voltage, frequency, lines power flow and lines and equipment load levels be kept within safety thresholds.

The state estimation process plays an essential role for the monitoring and analysis of an electrical system, because it handles analog redundant information measurements contaminated by noise, in order to better estimate the complex voltages in the buses belonging to the supervised system [1].

The ability to detect and identify GEs is one of the important attributes of the state estimation process in power systems. Some GEs are obvious and can be initially identified and eliminated from the estimation process, through a simple verification of the measures input data. Such errors can be: absurd values of effective voltage, values far beyond those expected for measures of power and /or electrical current, etc. [2]. However, not all types of GEs are easily detectable and identifiable in this way, requiring the use of other methodologies.

The WLS state estimator works well when the noise in the measurements are Gaussian, but fails in the occurrence of one or more GEs [2]. To overcome this limitation, methods were developed for detection and identification of GEs, among which the most widely used, are based on the analysis of the measurement residual, because they provide information on possible violations of assumptions concerning the measurement model (the residue is the difference between the measured and estimated value of the measures).

The WLS estimator, associated with GEs processing techniques based on the analysis of the measures residues, give satisfactory performance in the occurrence of simple GE, or when there are multiple non-interactive GEs [3], but can fail in the following situations:

i) GEs associated with measures with low redundancy (critical measures or pertaining to critical sets of measures);

ii) Interactive multiple GEs;

iii) GEs that have the characteristic of being highly influential, i.e., to attract the convergence of state estimation process, called leverage point measures [3].

Because of the simplicity of its formulation, as well as the ease of its computer implementation, the WLS estimator associated with the largest normalized residual test is the most used in operation centers.

2. PROBLEM FORMULATION

This section will present a general formulation for the WLS estimator associated with the largest normalized residual test.

Power System Mathematical Modeling

In this work we adopted as a representation of an electrical system branch a generalization of the equivalent model of transmission lines, in phase and lagged transformers, from the model presented in [4]. From the application of Kirchhoff's laws, on the general model π, we obtain the following expressions for the active (P_{kl}) and reactive (Q_{kl}) power flow in the branch that connects the buses k e l [4]:

i) From bus k to bus l

$$P_{kl} = a_{kl}^2.V_k^2.g_{kl} - a_{kl}.V_k.V_l.g_{kl}.\cos(\theta_{kl}+\phi) + - a_{kl}.V_k.V_l.b_{kl}.\sin(\theta_{kl}+\phi)$$

$$Q_{kl} = -a_{kl}^2.V_k^2.(b_{kl}+b_{kl}^{sh}) + a_{kl}.V_k.V_l.b_{kl}.\cos(\theta_{kl}+\phi) + -a_{kl}.V_k.V_l.g_{kl}.\sin(\theta_{kl}+\phi) \quad (2.1)$$

ii) From bus l to bus k

$$P_{lk} = V_l^2.g_{kl} - a_{kl}.V_k.V_l.g_{kl}.\cos(\theta_{lk} - \phi) + - a_{kl}.V_k.V_l.b_{kl}.\sin(\theta_{lk} - \phi)$$
$$Q_{lk} = -V_l^2.(b_{kl} + b_{kl}^{sh}) + a_{kl}.V_k.V_l.b_{kl}.\cos(\theta_{lk} - \phi) + - a_{kl}.V_k.V_l.g_{kl}.\sin(\theta_{lk} - \phi) \quad (2.2)$$

Being:
b_{kl} - the component series susceptance;
b_{kl}^{sh} - the shunt susceptance of the transmission line;
$t_{kl} = a_{kl}.e^{j\phi}$ - the transformer turns ratio;
V_k and V_l - the voltage magnitudes at buses k and l;
θ_k and θ_l - the voltage phase angles at buses k and l.

In the complex power injection equating, at any power system bus, you must consider the possible existence of shunt elements connected to it. Thus, the expressions for the active and reactive power injections, at a generic bus k, can be written as:

$$P_k = \sum_{l \in \Omega k} P_{kl}$$
$$Q_k = -V_k^2.b_k^{sh} + \sum_{l \in \Omega k} Q_{kl} \quad (2.3)$$

Being:
b_k^{sh} - the shunt susceptance of a capacitor connected to the bus k;
Ωk - the set of adjacent buses to the bus k.

For more details of the complete equating of flow and power injection expressions see [4]. The next section will address the solution method chosen for the WLS state estimator.

Normal Equation Method

The state estimation is the calculation of unknown state variables through a set of inaccurate measures; therefore, the estimation obtained will not be exact. Thus, the estimation problem is to find a way to achieve the best estimation and, for this, from many existing statistical criteria, the one which has been used in most of the power systems is the weighted least squares. In this paper it is assumed that there are no errors in the model parameters. With respect to this representation, the nonlinear equations for state estimation in power systems are represented as:

$$\underline{z} = h(\underline{x}) + \underline{w} \quad (2.4)$$

Being:
$\underline{z}$ - the measures vector (m x 1);
$h(.)$ - the vector of nonlinear functions, which lists the measures with the state variables (m x 1);
$\underline{x}$ - the vector of state variables to be estimated (n x 1);
$\underline{w}$ - the vector of measures errors (m x 1);
m - the measures number;
n - the state variables to be estimated number.

The measures errors are considered as independent random variables, with zero Gaussian mean [5]. Calling R the covariance matrix of the measures error vector, with size m x m, we have:

$$R = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_m^2 \end{bmatrix} \quad (2.5)$$

Where σ_i^2 is the variance of measurement error i. Therefore $\underline{w_i} \sim N(0,R_{ii})$ for every "i".

By applying the methodology of weighted least squares, the best estimation of the state variables vector $\underline{x}$, called $\hat{\underline{x}}$, can be obtained by calculating the value of $\underline{x}$ that makes minimum the $J(\underline{x})$ index [5], given by:

$$J(\underline{x}) = \frac{1}{2}.\underline{w}^t.R^{-1}.\underline{w} \quad (2.6)$$

or

$$J(\underline{x}) = \frac{1}{2}.[\underline{z} - h(\underline{x})]^t.R^{-1}.[\underline{z} - h(\underline{x})] \quad (2.7)$$

Being R^{-1} the inverse of the covariance matrix of the measures error vector, used here as a weight matrix for the measurements.

The $J(\underline{x})$ index becomes a minimum when:

$$\frac{\partial J(\underline{x})}{\partial \underline{x}} = 0 \quad (2.8)$$

or

$$H^t(\hat{\underline{x}}).R^{-1}.[\underline{z} - h(\hat{\underline{x}})] = 0 \quad (2.9)$$

being $H(\hat{\underline{x}})$ the matrix of first derivatives of the nonlinear functions of vector $h(\underline{x})$, known as the Jacobian, calculated at the point represented by the vector of estimated state variables $\hat{\underline{x}}$, and represented by:

$$H(\hat{\underline{x}}) = \left.\frac{\partial h(\underline{x})}{\partial \underline{x}}\right|_{x=\hat{x}} \quad (2.10)$$

Because the $J(\underline{x})$ index is a quadratic nonlinear function, to obtain $\hat{\underline{x}}$ we apply an iterative method to solve a linear equation at each iteration k, in order to calculate the current estimation of the state variables vector, through successive corrections [5], given by:

$$\underline{x}^{k+1} = \underline{x}^k + \Delta\underline{x}^k \quad (2.11)$$

However, to determine the correction $\Delta\underline{x}^k$, we perform the linearization of the equations $h(\underline{x})$ around the point $\underline{x}^k$, represented by the expression:

$$h(\underline{x}^{k+1}) \cong h(\underline{x}^k) + H(\underline{x}^k).\Delta\underline{x}^k \quad (2.12)$$

Rewriting equation (2.4), in relation to the approximations made in $h(\underline{x})$, we obtain the measurement model which has become linear:

$$\underline{z} = h(\underline{x}^k) + H(\underline{x}^k).\Delta\underline{x}^k + \underline{w} \quad (2.13)$$

Or:

$$\Delta\underline{z}(\underline{x}^k) = \underline{z} - h(\underline{x}^k) = H(\underline{x}^k).\Delta\underline{x}^k + \underline{w} \quad (2.14)$$

Being $\Delta\underline{z}(\underline{x}^k)$ defined as vector of the measurements residues.

From the model of linear measurement, the objective function $J(\Delta\underline{x})$ becomes:

$$J(\Delta\underline{x}) = \frac{1}{2}.[\Delta\underline{z}(\underline{x}^k) - H(\underline{x}^k).\Delta\underline{x}^k]^t.R^{-1}.[\Delta\underline{z}(\underline{x}^k) + -H(\underline{x}^k).\Delta\underline{x}^k] \quad (2.15)$$

Whose minimum is calculated from:

$$\frac{\partial J(\Delta\underline{x})}{\partial \Delta\underline{x}} = H(\underline{x}^k)^t.R^{-1}.[\Delta\underline{z}(\underline{x}^k) - H(\underline{x}^k).\Delta\underline{x}^k] = 0 \quad (2.16)$$

Therefore, the solution can be obtained by the following equation:

$$\Delta\underline{x}^k = [H(\underline{x}^k)^t.R^{-1}.H(\underline{x}^k)]^{-1}.H(\underline{x}^k)^t.R^{-1}.\Delta\underline{z}(\underline{x}^k) \quad (2.17)$$

Which is called the normal equation, where:

$$H(\underline{x}^k)^t.R^{-1}.H(\underline{x}^k) = G(\underline{x}^k) \quad (2.18)$$

is the gain matrix (G).

The iterative process starts from an initial value $\underline{x}^0$ and, at each iteration k; the corrections in the state variables $\Delta\underline{x}$ are obtained using equation (2.17). The vector of state variables update is obtained using equation (2.11) until a stopping criterion is satisfied, such as:

$$max|\Delta\underline{x}^k| \leq \varepsilon \quad (2.19)$$

where ε denotes a predetermined error tolerance.

Thus, this criterion indicates that the iterative process will be terminated when the magnitude of adjustments in state variables is negligible. The algorithm of the WLS state estimator can be summarized by the following steps:

Step 1: Set $k = 0$ and choose an initial solution $\underline{x}^k = \underline{x}^0$;
Step 2: Calculate the matrices $H(\underline{x}^k)$ and $G(\underline{x}^k)$ at the point $\underline{x} = \underline{x}^k$;
Step 3: Get the state variables correction through the normal equation and update the variables:

$$\Delta\underline{x}^k = G(\underline{x}^k)^{-1}.H(\underline{x}^k)^t.R^{-1}.\Delta\underline{z}(\underline{x}^k)$$
$$\underline{x}^{k+1} = \underline{x}^k + \Delta\underline{x}^k$$

Step 4: Test the stopping criterion: if $max|\Delta\underline{x}^k| \leq \varepsilon$, the process converged. Otherwise, make $k = k + 1$ and return to Step 2.

Largest Normalized Residual Test

The method used in this paper to detection and identification of measurement GEs is through the normalized residues vector ($\underline{r}^N$) analysis. The residues vector is defined by:

$$\underline{r}(\underline{\hat{x}}) = \underline{Z} - h(\underline{\hat{x}}) \tag{2.20}$$

To normalize the residue is necessary to calculate the residues covariance matrix, defined by the equation:

$$\Omega(\underline{\hat{x}}) = R - H(\underline{\hat{x}}).G^{-1}(\underline{\hat{x}}).H^t(\underline{\hat{x}}) \tag{2.21}$$

Thus, the normalized residue is calculated by:

$$\underline{r}_i^N(\underline{\hat{x}}) = \frac{r_i(\underline{\hat{x}})}{\sqrt{\Omega_{ii}(\underline{\hat{x}})}} \tag{2.22}$$

where Ω_{ii} is the i diagonal element of the residues covariance matrix.

The importance of residues standardization can be understood if it is taken into account that different types of meters have generally different variances, so that a discrepant measurement's residue value can be perfectly acceptable to another. The residues standardization places them in a single reference, thus allowing a fair comparison of their absolute values.

Admitting the hypothesis that the measurements errors ($\underline{w}_i$) are independent random variables with normal distribution with zero mean and known variance it is proved, in [6], that the elements of the normalized residues vector presents standard normal distribution, i. e.:

$$\underline{r}_i^N \sim N(0,1)$$

Thus, the existence of GEs can be verified by the following test:

- If any $|\underline{r}_i^N| > \beta$, with i = 1,...,m, there is suspicion of GE;
- If all $|\underline{r}_i^N| \leq \beta$, with i = 1,...,m, supports the hypothesis that there is no GE.

Usually it is assumed $\beta = 3$ [2].

Considering the hypothesis of a single measurement containing a GE and all other measurements as perfect, in [7] and [2], it is shown that for a measurement system, free of critical measurements and critical sets, the measurement with GE attend the largest normalized residual ($\underline{r}_{max}^N$). Thus, we can perform both detection and identification of the measurement with a GE, at the same time, by testing:

$$\underline{r}_i^N > \beta \text{ (Threshold)} \tag{2.23}$$

In the presence of single GE, the method does not identify GE on critical measurements, or on measurements pertaining to critical sets of measures. This is due to the fact that the critical measurements have zero residues [8] and measurements of critical sets have normalized residues equal in magnitude [9].

After identifying the measurement with a GE, some special treatment should be given to this measurement, in order to minimize its effect. Traditionally, the effect of the measurement with GE can be suppressed in two ways [7]:

i) withdrawal of the measurement with a GE from the measurement set and re-estimates the states;

ii) recovery of the measurement with GE through the value of estimated error and performs again an estimation of the states.

In this paper the measurement identified as having a GE is removed from the measurement set.

OBS. 1: In all the gross error detection test procedure, it was used the residual as a measure of the measurement gross error, without any proof that this assumption is correct.

OBS. 2: No proof at all is presented that the measurement with gross error is the one with the largest normalized residual. Again they are mixing measurement error with measurement residual, and they are completely different quantities [10].

OBS. 3: The conventional methodology is not considering that the residual space is of dimension equal to the measurements´ number minus the system state variables, that is, a correlated space. Otherwise the measurement error is a not correlated space; that is the measurement errors are not correlated.

OBS. 4: The consequence of the correlated space for the measurements is that instead of using a hyper-sphere in order to identify the measurement with error one should use instead a hyper-ellipsoid.

3. PROPOSED METHODOLOGY

The implemented program has the following flowchart:

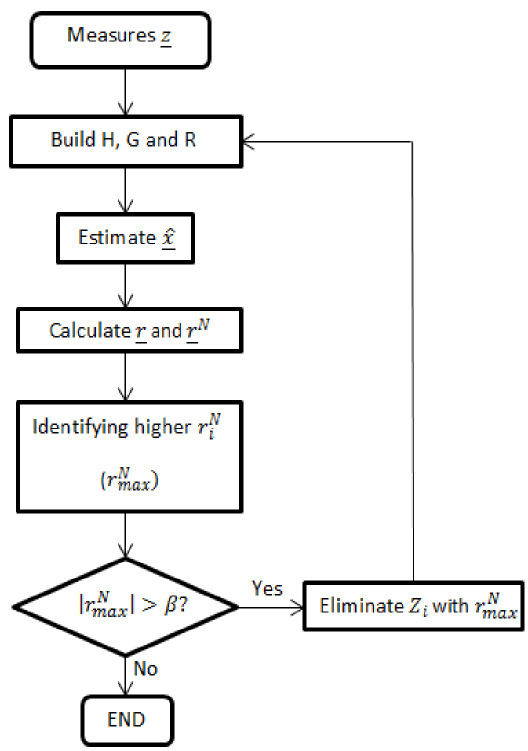


Fig.1: WLS estimator flowchart.

Obs.: β is chosen (how many standard deviations are accepted). In this paper it was considered $\beta = 3$.

The systems chosen for computer simulations are the IEEE's 6-bus and the IEEE-14 bus, where the program reads automatically the database in *.txt* format.

4. RESULTS

In this section we present the computer simulations results for IEEE's 6 bus and 14 bus systems.

IEEE´s 6 Buses System

The system has the following topology:

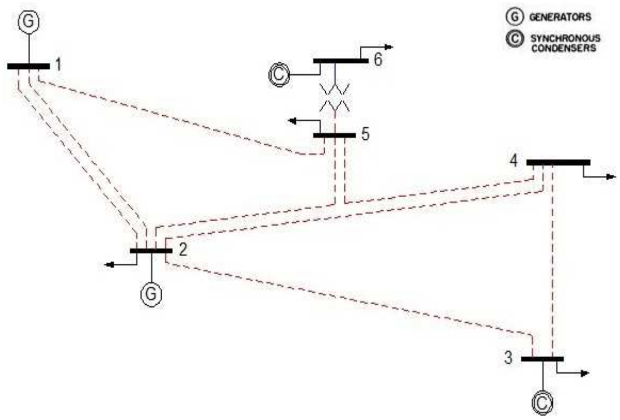


Fig. 2: IEEE's 6 buses system topology.

From the results of a load flow program, the measurement plan was built with a measuring overall redundancy index equal to three times the number of state variables to be estimated, therefore, consists of 33 measurements, without adding in the initial case, random noise and without the presence of critical measurements or critical sets of measurements. The measurements are shown in the following tables:

Table 1. Power injection measurement values

Active Measurements		Reactive Measurements	
Measurement	Value (MW)	Measurement	Value (MVAr)
AI 1	149.0	RI 1	-3.7
AI 2	18.3	RI 3	2.2
AI 4	-47.8	RI 4	3.9
AI 5	-7.6	RI 6	-14.0

Table 2. Power flow measurement values

Active Measurements		Reactive Measurements	
Measurement	Value (MW)	Measurement	Value (MVAr)
AF 1-2	103.1	RF 1-5	3.2
AF 1-5	45.9	RF 2-3	4.7
AF 2-3	62.5	RF 2-4	-0.6
AF 2-5	22.8	RF 3-4	4.4
AF 3-4	-33.4	RF 4-5	8.9
AF 4-5	-48.3	RF 5-6	14.7
AF 2-1	-101.3	RF 2-1	6.7
AF 3-2	-60.8	RF 5-1	-4.4
AF 4-2	-33.7	RF 4-2	-1.2
AF 5-2	-22.5	RF 5-2	-4.0
AF 5-4	48.6	RF 4-3	-3.8
AF 6-5	-11.2	RF 6-5	-14.0

Table 3. Voltage measurements

Measurement	Value (V)
V1	1.060

To weigh the measurements used by the WLS state estimator it was assumed that all meters have standard deviation calculated by the following equation:

$$\sigma_i = \frac{pr.\left|z_i^{lf}\right|}{3} \tag{4.1}$$

Where *pr* is the meter precision (considered 3% in this work by author´s choice) and z^{lf} is the measurement value obtained from a load flow simulation. After running the implemented software the following results were obtained for the estimated state variables:

Table 4. Estimated state variables

Bus	Magnitude (pu)	Angle (rad)
1	1.0602	0
2	1.0452	-0.0558
3	1.0102	-0.1700
4	1.0255	-0.1115
5	1.0284	-0.0911
6	1.0701	-0.1150

After the estimation process the largest normalized residual test is performed, to detect possible measurements containing GEs. In this case, the calculated largest normalized residual was: $\underline{r}_{max}^N = 0.0150$ on the flow measure RF 1-5. As the $\underline{r}_{max}^N \leq 3$, we accept the hypothesis that there is no measurement with a GE.

Now let's add a 5σ error on the measurement AF 1-2 (chosen randomly). Repeating the estimation process the following results are obtained:

Table 5. Estimated state variables

Bus	Magnitude (pu)	Angle (rad)
1	1.0587	0
2	1.0431	-0.0580
3	1.0074	-0.1743
4	1.0232	-0.1143
5	1.0261	-0.0936
6	1.0676	-0.1182

In this case: $\underline{r}_{max}^N = 4.3806$ on the measurement AF 1-2. As expected, the test detected and identified the measurement carrying the GE. By eliminating this measurement from the measurement plan, the results were the same as the original case without the existence of measurement error. With the results obtained it can be observed the effect a measurement with GE can cause on the state estimation process, leading it to erroneous values for the estimated state variables.

Now, let´s admit that the measurements are no more perfect, but having an associated random noise, so that they vary from $\pm 3\sigma$ of its original values, so, not characterizing a measure with GE. This was done to test the robustness of the largest normalized residual test, since, in real systems; the set of measurements is subject to noise. In this case: $\underline{r}_{max}^N = 3.6882$ on the measurement AF 6-5. We conclude that the test failed, since the added noise in the measurements was less than 3σ (in magnitude), not characterizing GE, as the test found. Now, adding a GE of 5σ on AF 2-3, the test resulted in: $\underline{r}_{max}^N = 4.5204$ on the measurement AI 1. It appears that the test detected the presence of GE, but not on the measurement that GE was inserted, failing again.

IEEE´s 14 Buses System

The system has the following topology:

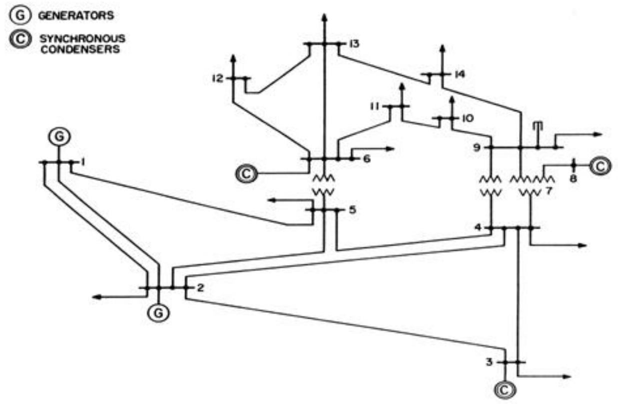


Fig. 3: IEEE's 14 bus system topology.

For this system the measurement plan consists of 81 measurements, leading the overall redundancy level equal to 3.0.

Table 6. Power injection measurement values

Active Measurements		Reactive Measurements	
Measurement	**Value (MW)**	**Measurement**	**Value (MVAr)**
AI 1	232.4	RI 1	-16.5
AI 2	18.3	RI 3	6.1
AI 4	-47.8	RI 4	3.9
AI 5	-7.6	RI 6	5.2
AI 7	0	RI 7	0
AI 8	0	RI 9	-16.6
AI 10	-9.0	RI 10	-5.8
AI 11	-3.5	RI 12	-1.6
AI 13	-13.5	RI 13	-5.8
AI 14	-14.9	RI 14	-5.0

Table 7. Power flow measurement values

Active Measurements		Reactive Measurements	
Measurement	**Value (MW)**	**Measurement**	**Value (MVAr)**
AF 1-2	156.9	RF 1-2	-20.4
AF 2-3	73.2	RF 1-5	3.9
AF 2-4	56.1	RF 2-4	-1.55
AF 3-4	-23.3	RF 2-5	1.2
AF 4-5	-61.2	RF 4-5	15.8
AF 4-9	16.1	RF 4-7	-9.7
AF 5-6	44.1	RF 5-6	12.5
AF 6-12	7.8	RF 6-11	3.6
AF 6-13	17.7	RF 6-13	7.2
AF 7-9	28.1	RF 7-8	-17.2
AF 9-10	5.2	RF 9-10	4.2
AF 10-11	-3.8	RF 9-14	3.6
AF 12-13	1.6	RF 13-14	1.7
AF 2-1	-152.6	RF 2-1	27.7
AF 5-1	-72.7	RF 5-1	2.2
AF 3-2	-70.9	RF 3-2	1.6
AF 4-2	-54.5	RF 4-2	3.0
AF 4-3	23.7	RF 5-2	-2.1
AF 5-4	61.7	RF 4-3	-4.8
AF 7-4	-28.1	RF 5-4	-14.2
AF 9-4	-16.1	RF 7-4	11.4
AF 11-6	-7.3	RF 6-5	-8.1
AF 12-6	-7.7	RF 11-6	-3.4
AF 13-6	-17.5	RF 12-6	-2.4
AF 8-7	0	RF 13-6	-6.8
AF 9-7	-28.1	RF 9-7	-5.0
AF 14-9	-9.3	RF 10-9	-4.2
AF 11-10	3.8	RF 14-9	-3.4
AF 13-12	-1.6	RF 11-10	1.6
AF 14-13	-5.6	RF 14-13	-1.6

Table 8. Voltage Measurements

Measurement	Value (V)
V1	1.060

After running the implemented software the following results were obtained for the estimated state variables:

Table 9. Estimated state variables

Bus	Magnitude (pu)	Angle (rad)
1	1.0599	0
2	1.0449	-0.0870
3	1.0098	-0.2222
4	1.0175	-0.1800
5	1.0194	-0.1532
6	1.0697	-0.2483
7	1.0614	-0.2333
8	1.0899	-0.2333
9	1.0558	-0.2609
10	1.0508	-0.2636
11	1.0567	-0.2582
12	1.0549	-0.2631
13	1.0501	-0.2646
14	1.0353	-0.2799

After performing the state estimation the obtained largest normalized residual with perfect measurements was: $\underline{r}_{max}^{N} = 0.0177$ on the measurement RF 6-11. Note that $\underline{r}_{max}^{N} \leq 3$ then the hypothesis that there is no measurement containing a GE is accepted. Now let's add a -6σ noise on the measurement FA 5-6 (chosen randomly) in the same perfect measurements set. Remaking the process we obtained the following results:

Table 10. Estimated state variables

Bus	Magnitude (pu)	Angle (rad)
1	1.0627	0
2	1.0477	-0.0863
3	1.0125	-0.2214
4	1.0210	-0.1769
5	1.0230	-0.1498
6	1.0743	-0.2265
7	1.0646	-0.2258
8	1.0932	-0.2254
9	1.0586	-0.2518
10	1.0539	-0.2528
11	1.0594	-0.2426
12	1.0592	-0.2410
13	1.0544	-0.2436
14	1.0390	-0.2656

For this case: $\underline{r}_{max}^{N} = 4.5849$ on the measure AF 5-6. As expected, the test detected and identified the measurement containing a GE. Eliminating it from the measurement plan the results were the same as the original case without GE, validating the test for this case.

Similarly to the case of the 6 buses system, we added random noise in the set of measures. For this case: $\underline{r}_{max}^{N} = 4.1846$ and, the test failed again when the measures have noise, even though they were lower than 3σ. Finally, adding a GE of

6σ on the measure AF 13-6 we obtained $r^N_{max} = 5.5143$ on this same measure, thus the test was effective in this case. However, simulating the system again, we obtained $r^N_{max} = 3.6290$ on the measure AI 10, thus the test detected the GE, but was not able to correctly identify the measurement containing a GE.

As stated initially, it is shown that the largest normalized residual test fails many times. Through the geometric interpretation, the author [11] proves mathematically that the measurement error is composed of components detectable and undetectable, also shows that the detectable component of the error is exactly the noise of the measurement error. The methods previously used for the processing of gross errors (GEs), consider only the detectable component of the error, then as a consequence, may fail.

Through orthogonal projections defined by the equation of the projection matrix, [11] also showed that errors in measurements that are very close to the range space of the Jacobian matrix, relative to other measures, are difficult to detect when using the largest normalized residual test, thus, depending on the amplitude of the components of the error, this method may fail. So is being studied and proposed a new methodology to process the measures with GE. This proposition is obtained by decomposing the measurement error in two components: the first is orthogonal to the range space of the Jacobian matrix, whose amplitude is equal to the residue of the measure; the other belongs to the range space of the Jacobian matrix and therefore does not contribute to the residue of the measure.

5. CONCLUSION

By the tests presented, it is verified the effectiveness of the proposed algorithm, applied to power systems state estimation, by using the WLS state estimator method.

The implemented software allows the operator to directly read the database from the solution of a load flow problem in a *.txt* format. After reading the database it solves the problem of state estimation by the weighted least squares method, taking into consideration the measurement quality, given by their respective variances.

To validate the results obtained by the WLS, the largest normalized residual test is performed in an attempt to detect and identify possible measurement containing GEs, which interfere negatively in the estimation process. If such measurement is detected, through a threshold test for r^N_{max}, it is discarded from the measurement set and the process of estimation is remade.

The results obtained using the software shows that the theoretically expected results of the classic state estimation analysis fails some times and the reasons for that is the lack of theoretical consistency used in the classical state estimation proposition. The simulations results showed that the largest normalized residual test fails and again the reason are the inconsistencies of the used theoretical background. For example even in the case of not having gross error in the measurement set, the test detected GE and in other cases, correctly detected the presence of GE, but erroneously identified the measurement containing a GE. These facts make clear the lack of robustness of the largest normalized residual test when using measurements sets with random noise.

6. REFERENCES

[1] R.A.S. Benedito, Undetectability index of gross errors in power systems state estimation, Ph.D. dissertation, Dept. Elect. Eng., São Paulo Univ., São Carlos, SP, Brazil, 2011. (In Portuguese)

[2] A. Abur and A. G. Expósito, *Power systems state estimation: theory and implementation,* Marcel & Dekker Publishers, Nova York, USA, 2004.

[3] A. J. Monticelli, Electric power system state estimation, *Proceedings of the IEEE*, *88*(2), 2000, 262-282.

[4] A. J. Monticelli, *Load flow in electric power networks*, Edgard Blucher, São Paulo, Brazil, 1983. (In Portuguese)

[5] F. C. Schweppe and E. J. Handschin, Static state estimation in electric power systems, *IEEE Transactions on Power Apparatus and Systems*, *62*(7), 1974, 972-982.

[6] A. J. Monticelli, *State estimation in electric power systems: a generalized approach,* Kluwer Academic Publishers, Massachusetts, USA, 1999.

[7] A. J. Monticelli and A. Garcia, Reliable bad data processing for real-time state estimation, *IEEE Transactions on Power Apparatus and Systems*, *102*(5), 1983, 1126-1139.

[8] K. A. Clements, G. R. Krumpholz and P. W. Davis, Power system state estimation residual analysis: an algorithm using network topology, *IEEE Transactions on Power Apparatus and Systems*, *100*(4), 1981, 1779-1787.

[9] L. Mili, T. Van Cutsem and M. Ribbens-Pavella, Hypothesis testing identification: a new method for bad data analysis in power system state estimation, *IEEE Transactions on Power Apparatus and Systems*, *103*(11), 1984, 3239-3252.

[10] N. G. Bretas, A. S. Bretas, S. A. R. Piereti, Innovation concept for measurement gross error detection and identification in power system state estimation, *IET Generation, Transmission & Distribution*, *5*(6), 2011, 603-608.

[11] N.G. Bretas, J.B.A. London Jr., L.F.C. Alberto, R.A.S. Benedito, Geometrical approaches for gross errors analysis in power system state estimation. *IEEE Power Tech*, Bucharest, Romania, 2009, 1-7.

Real time analysis of spectrum of supply current with utilization of full digital system

Peter DRGONA*, Michal FRIVALDSKY*, Pavol SPANIK*, Jozef SEDO*

and

Jozef KUCHTA**

* University of Zilina, Department of mechatronics and electronics,
Zilina, 01026, Slovakia

** EVPU (Electronic research and testing institute),
Nova Dubnica, 01851, Slovakia

ABSTRACT

This paper deals with construction of device for online, real-time analysis of recuperative current with the use of microprocessor and thus will be representing full digital system for real time analysis. The proposed system should be utilized for railway transport for the measurement of recuperative currents which is being supplied from power converter of traction vehicle. In the first part of paper, the current sensor and digital filter are described. Then the computation algorithm is being presented. Last part of paper shows implementation in microprocessor and experimental verification of proposed solution.

Keywords: real-time measurement, spectrum, full digital, FFT.

1.INTRODUCTION

Nowadays, a high frequency power semiconductor converters are used in all areas including industry and transportation, for control the energy flow. Their main advantage is high efficient power conversion, high frequency and small size. On the other side, converters uses semiconductor devices in high frequency switching mode, so they represent nonlinear load for distribution network and supply current for converters is not harmonic. As a result, the higher harmonics of supply current are generated, which significantly affects the EMC of converter and other electronic devices. Emissions of higher harmonics are strictly set by certain standards. For railway applications, these limits are set by standard UIC 550-3 (European Union).

Without utilization of power factor corrector (PFC) the supply current of converter acts as current with high portion of non-harmonic elements. Even PFC is being used the given amount of these non-harmonic elements are still being taken away from distribution network. This problem can be also registered during recuperative braking of traction vehicles. As a results of described behavior of power semiconductor systems, the higher harmonics of supplying current are significantly affecting EMC of the semiconductor converter and other electronic devices.

This paper is presenting novel full digital control methodology for higher order frequency spectrum identification provided online, in the real time. Core of proposed system is based on modern microprocessor unit from TMS320F Family from Texas Instruments. Other additional measurement equipment and selection of most suitable additional features will be consequently described. Last part of paper is given for implementation and verification through various experiments.

2. CURRENT SENSING

This chapter describes selection of the current sensing device/method which is suited in final system. In principle there exist three general methods for current sensing:

a) utilization of shunt resistor

b) utilization of current transformer

c) measurement through Hall sensor

Shunt resistor is able to be used for various amplitudes and shape of current waveforms. The output signal for further processing is voltage which is proportional to flowing current. The main criterions for shunt resistors are very low resistance value (in order to meet acceptable values of voltage which has to be processed) and power dimensioning according to flowing current. Base on worldwide standards the most of shunts are designed to have voltage drop (sensing voltage) 50mV, 75 mV or 100 mV at the nominal current. Disadvantages related to usage are temperature stresses which in extreme condition can expressively change value of resistance and thus value of sensing voltage, what can lead in measurement failure and thus collapse of total system. Also at the measurement of time variable waveforms of current, the proposed shunt resistor must be made from material which acts as low inductance due to suppression of measurement distortion.

Therefore more suitable methodology is utilization of current transformer. It transforms flowing current to corresponding voltage, which is being measured by sensing device. Requirement according current transformer is that it cannot operate in no load condition. Also there are other disadvantages according magnetic saturation and power dimensioning, what makes this solution for high power application (what traction vehicle introduces) non attractive due to large volume.

The last mentioned proposal according current measurement is base on well - know methodology, which is based on utilization of Hall sensor. In principle there exist two possible ways:

- measurement in open loop
- measurement in closed loop

In our proposed measuring device we have focused on second way, measurement in closed loop, during which importance has to be given mainly for operation of hall sensor in linear region. Closed loop operation is characterized by total compensation of magnetic flux around magnetic circuit. Each deviation from zero offset leads to generation of Hall effect. Electrical circuit then supplies secondary current, which is necessary for compensation of magnetic field. Advantages of closed loop operation are high accuracy, large available bandwidth, fast response, perfect linearity, possibility of overloading.

Digital filters

Due to fact, that hall sensor is equipped with current output, the transformation from current to acceptable voltage level for A/D converter (microprocessor) is necessary. A/D converter is operating with input voltage level between 0 VDC - 3 VDC.

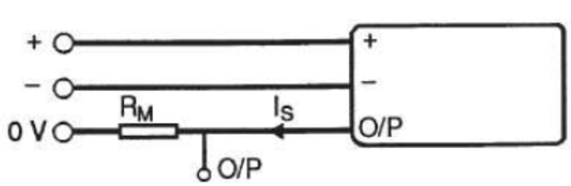


Fig. 1 Principle schematics of Hall sensor connection with measuring resistor

Fig. 2. shows schematics for Hall sensor connection with measuring resistor which is suited for current transformation into voltage. Determination of measuring resistor R_M was done using next formula:

$$R_M = \frac{U_{AD\max}}{I_{pp\_o\max}} = \frac{U_{AD\max}}{\frac{1}{n} I_{pp\_i\max}} = \frac{3}{\frac{1}{1000}300} = 10\Omega \qquad (1)$$

Hall sensor has maximum measuring range ± 150 A, so the maximum amplitude of measured current is 300A. Due to this value the voltage drop on measuring resistor will have to be in the range of ± 1.5 V, but for A/D converter the conversion of negative voltages is not possible. For the purpose of the voltage level adjustment on the measuring resistor without loss of information of measured signal the utilization of next circuit (fig. 3.) was necessary.

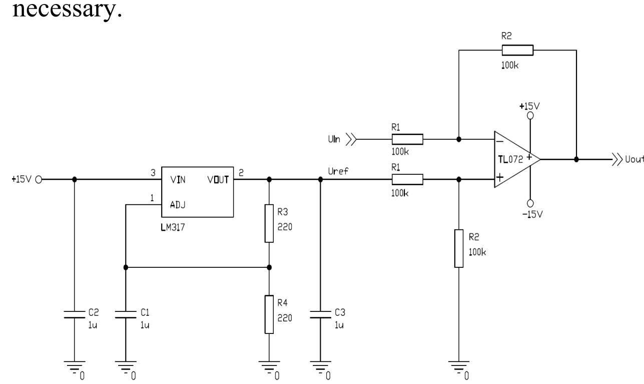


Fig. 2 Block scheme of circuit for voltage adjustment

This circuit shifts input voltage ± 1.5 V to range from 0 to 3V, without modification of measured signal. This is very important, because even small deviation in frequency of the signal can lead to inaccurate analysis. For the output voltage, next formula is valid:

$$U_{out} = \frac{R_2}{R_1}(U_{ref} - U_{in}) \qquad (2)$$

, where U_{ref} is voltage at the output of voltage stabilisator LM317, and U_{in} is voltage across sensing resistor R_M. If $R_1 = R_2$ then:

$$U_{out} = U_{ref} - U_{in} \qquad (3)$$

In such way the output voltage U_{out} is modified voltage in the range 0 V - 3 Vdc which is supplied into A/D converter.

3. FILTERING OF INPUT SINGAL

The Hall sensor with closed loop was used in proposed system. This sensor is capable to sensing AC, DC or impulse currents. Due to use of AD converter in microprocessor, the voltage level from current sensor needs to be shifted in range 0 to 3V.

Digital filters

Digital filter as representation of digital system, process digital input and after processing, at the output of the system, digital signal appears. The role of the digital filter is to affect the spectrum of the input signal in desired form. This means either select a part of the spectrum which will stay unchanged and suppress the rest. (low pass filter, high pass filter, band pass or band stop, or shaping the frequency response. [1] [2]

For given application, a digital IIR filter was chosen, due to its small time delay during processing of input signal (measured current). Another advantage of this type of filter is small memory usage. Unlike FIR filters, stability of IIR filters is not always guaranteed. After the proposal, check whether all poles of transfers function lies inside unit circle in z-domain. Another disadvantage of IIR filters is sensitivity to saturation of processor's arithmetic and sensitivity to quantization of values. Advantages of IIR filters are lower degree of transfer function at the same requirements. Consequently, memory demands for accumulation of coefficients and state variables are lower and also, time for processing of input variable is lower also. General transfer function of IIR filter is in form:

$$H(z)=\frac{Y(z)}{X(z)}=\frac{\sum_{i=0}^{M} a_i \cdot z^i}{\sum_{i=0}^{M} b_i \cdot z^i} \tag{4}$$

In final it is possible to use design tables of analog filters, because each digital filter can be described by its analog equivalent.

Main criterion for filter design is its stability, that means that all poles of transfer function must belong to internal part of unity circle line (see fig. 3).

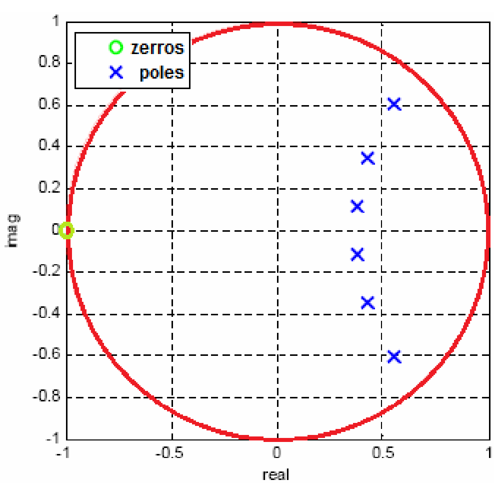


Fig.3. Distribution of poles and zeros of proposed system

For the analysis of spectrum of supply current we have decided to assign antialiasing function for proposed digital filter, what means that it will behave as low-pass filter. For the sampling frequency of f_s = 192307,69 Hz we design filter with cutoff frequency f_c = 26 kHz. Digital filter in such way won't change amplitude frequency spectrum in the range from 0 up to 26 kHz. During filter design we have utilized methodology of analog to digital transformation. First we have researched analog low-pass which was consequently through the use of bilinear transformation transformed into p and z plain. Next, with the use of tables for design of Butterworth filter we have selected analog filter whose low-pass was 6th grade polynomial function. This proposed filter is giving compromise between fastness and slope of transition zone of modulated frequency characteristic. Analog form of proposed filter is given by next equation:

$$H(p)=\frac{1}{p^6+3{,}863703\,p^5+7{,}464102\,p^4+9{,}141620\,p^3+7{,}464102\,p^2+3{,}863703\,p+1} \tag{5}$$

Using mentioned tables of Butterworth filter we get equation (5) in digital form what is given by next equation:

$$H(z)=\frac{b_0+b_1z^{-1}+b_2z^{-2}+b_3z^{-3}+b_4z^{-4}+b_5z^{-5}+b_6z^{-6}}{1+a_1z^{-1}+a_2z^{-2}+a_3z^{-3}+a_4z^{-4}+a_5z^{-5}+a_6z^{-6}} \tag{6}$$

where the coefficients of polynoms are:

b_0 = 0,001552195356958 b_1 = 0,009313172141747
b_2 = 0,023282930354368 b_3 = 0,031043907139157
b_4 = 0,023282930354368 b_5 = 0,009313172141747
b_6 = 0,001552195356958 a_1 = -2,733967214763194
a_2 = 3,600856207068566 a_3 = -2,715090353346278
a_4 = 1,219380058434311 a_5 = -0,304731364108709
a_6 = 0,0032893169560607

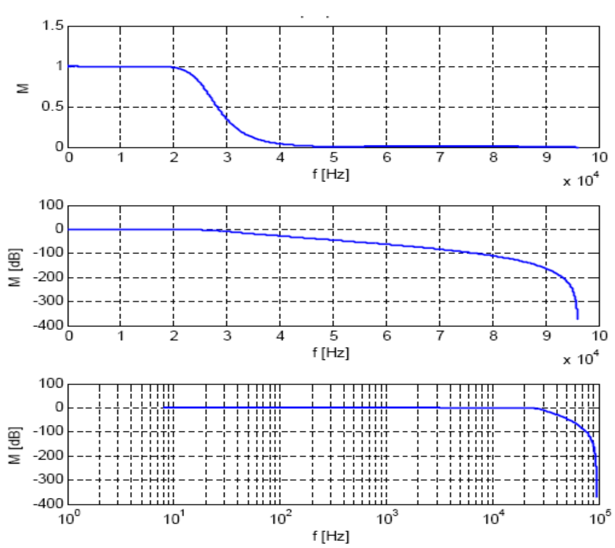


Fig.4 Frequency characteristic of proposed filter

Fig.4 is showing modulated frequency characteristic of proposed digital filter. At the cutoff frequency the attenuation is around -3dB. This confirms proper design.

4. COMPUTATION OF DISCRETE FOURIER TRANSFORM AND AMPLITUDE SPECTRUM

For computing of discrete Fourier transform the FFT algorithm with "in place computing" was used. In program, two fields each with 512 samples were used. Both fields were in 16b floating point form. One of the field was for real part and second was used for imaginary part of complex number. For realization of FFT function, standard function-void fft(float32*p_real, float32*p_imag) was used. Input to the function is array of real numbers, but computing of imaginary numbers is also possible. Generally, output from the discrete Fourier transform is sequence of complex numbers. From above mentioned function, the output is array of real parts of numbers and array of imaginary numbers. Hence, the input values are overwritten with output values. Advantage of this algorithm is low memory usage [3] [4].

For amplitude spectrum computing, function „void Spectrum_Amp(float32 *p_real, float32 *p_imag)" was used. this function reconstruct the frequency axis and compute amplitudes of each harmonic component by equation:

$$A[n] = \frac{2}{N}\sqrt{X_{real}^2[n] + X_{imag}^2[n]} \qquad (7)$$

where:

n - is n-th component of Fourrier transform

X_{real} - is real part

X_{imag} - is imaginary part

Function, which resolution is determined by number of samples N, assign the value of amplitude to each frequency from frequency range, whereby identify the spectrum of analyzed signal, in our case the current from converter.

5. Implementation into microprocessor

Algorithms for filtering, computing and analysis of measured signal are realized in software form in microprocessor. For this purpose a 32bit float-point microprocessor TMS 320F28335 Delfino. This 150MHz microprocessor contains float-point C2000 core and powerful peripherals, from which a 12bit A/D converter was used for measurement of signal from Hall sensor. Software for computing and analysis was written in standard C-language. Maximum frequency of analyzed signal is up to 20kHz, so the sampling frequency of 192kHz was chosen, which means, that Shannon criteria was observed

Implementation of digital filters

Digital filter in device for real time measurement and analysis of recuperative currents from electric locomotive acts like antialiasing filter, which is basically low pass filter. For sampling frequency f_s=192307.69 Hz, a low pass filter with cut-off frequency f_c=26kHz was designed. Digital filter do not affect amplitude frequency spectrum from 0 to 26kHz, but all higher frequencies will be damped.

Filter was designed using analog to digital transformation method or design by emulation method. First, analog low pass filter was designed and consequently, this filter was transformed to discrete z-domain by bilinear transform. By use of tabs for design of Butterworth filter, an analog low-pass filter of 6-th degree was chosen. Degree of this filter gives good compromise between computing speed and rate of Bode characteristics. Transfer function of analog filter is in form:

$$H(s) = \frac{1}{s^6 + a_1 \cdot s^5 + a_2 \cdot s^4 + a_3 \cdot s^3 + a_4 \cdot s^2 + a_5 \cdot s + a_6} \qquad (5)$$

This function was transformed by bilinear transform to z-domain. This two steps can also be performed in reverse. For implementation in microprocessor, transfer function in z-domain must be transferred into time domain. After transformation we get two differential equations, state and output. These two differential equations can be easily implemented into microprocessor.

Realization of window function

Because of analyzed signal consists of combination of sinusoidal waveforms and duration of transient effects is longer than window, a Hanning window was chosen. This window has good frequency resolution and limited spectral leakage. Realization of the window function consists in multiplication of the window value, with the value of the sample. This is performed by function „void Okno_Hann(float32 *p_pole)“.

Computing of discrete Fourier transform

For computing of discrete Fourier transform the FFT algorithm with "in place computing" was used. In program, two fields each with 512 samples were used. Both fields were in 16b floating point form. One of the field was for real part and second was used for imaginary part of complex number. For realization of FFT function, standard function - void fft(float32*p_real, float32*p_imag) was used. Input to the function is array of real numbers, but computing of imaginary numbers is also possible. Generally, output from the discrete Fourier transform is sequence of complex numbers. From above mentioned function, the output is array of real parts of numbers and array of imaginary numbers. Hence, the input values are overwritten with output values. Advantage of this algorithm is in low memory usage and relatively short time necessary for computing of this function. For application in real time, computing time of algorithms has big impact on performance of whole device. Time necessary for computing of all algorithms was 6.8μs, whereby single precision 32b floating point operands were used for variables.

5. SYSTEM FOR REAL TIME ANALYSIS OF CURRENT'S SPECTRUM

Our proposal of system for real time analysis of current's spectrum is primarily indented for use in railway applications. But this system can be normally used in every application where exact measurement of recuperative current is necessity. As we previously mentioned, according to standard UIC 550-3, the recuperative current which is flowing through rail back into power network cannot affect signal devices and other equipment which could be potentially safety risk for railway system.

Whole analysis of recuperative current must be done in real time. It means that if analyzed current contains specific harmonic content, or if content of higher harmonics is above allowable limit, then the proposed system has to perform certain measurements. Control system and control algorithm will prevent distorted current from flowing back into the distribution point and power network.

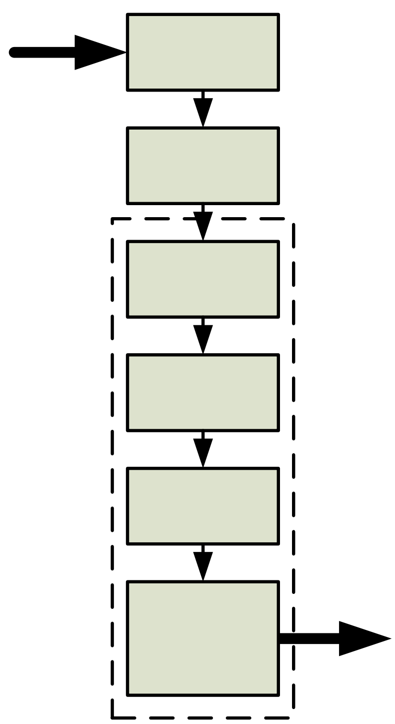


Fig. 5. Block scheme of system for analysis of recuperative currents

Fig. 5 shows block scheme of proposed system. It consists from current sensing device, and from microprocessor (in our case we have been working with TMS320F28335). Analog to digital conversion and all functions of filters and computational algorithms which are necessary for measurement and analysis of harmonic content in recuperative currents are in digital - in software form.

6. SIMULATION AND EXPERIMENTAL VERIFCATION

For verification of proposed algorithm - digital filter and its accuracy, the simulation in MATLAB environment was made. Input to the simulation was signal, which contains 20 harmonic functions. Each function has amplitude equals to 1 and frequencies were in range from 5kHz to 100kHz. Sampling frequency was 192kHz

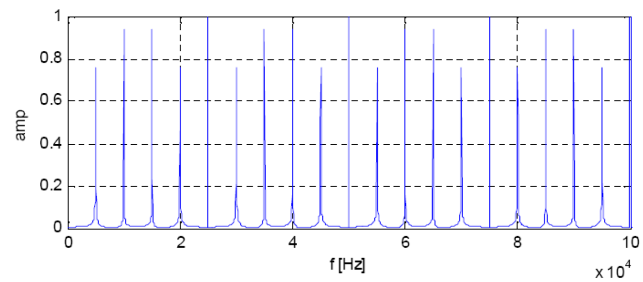


Fig.6. Amplitude spectrum of original input signal

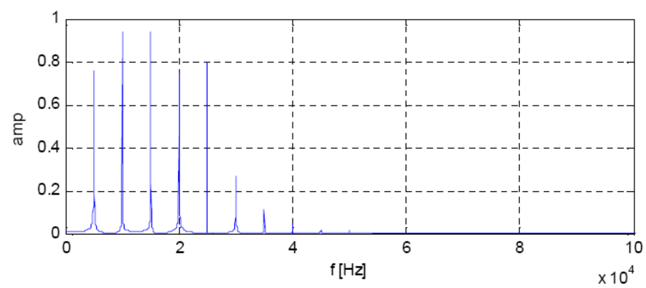


Fig.7. Amplitude spectrum of filtered signal

Fig. 7. shows amplitude spectrum of filtered signal. From figure it is clear to say, that frequencies of each harmonic component are clearly identified, but due to spectral leakage, some frequencies are still present also above range of 26kHz. But this problem has only minor impact on application of real time measuring device. Results from simulations also show good accordance with standard MATLAB function fft().

Experimental verification

For verification of whole system, measurements and tests of sensor and digital filter were made. On fig.8 the waveforms of analyzed signal are shown, where upper wave is related to current and bottom wave is measured voltage. On fig.9 the spectrum of this signal is being showed, whereby it was computed through the use of MATH function which is given by measuring oscilloscope. From this figure, the basic harmonic content (frequency around 50 Hz) of measured signal is able to be clearly identified. It serves for comparison with our proposed system, whose real time spectrum analysis is shown on fig.10.

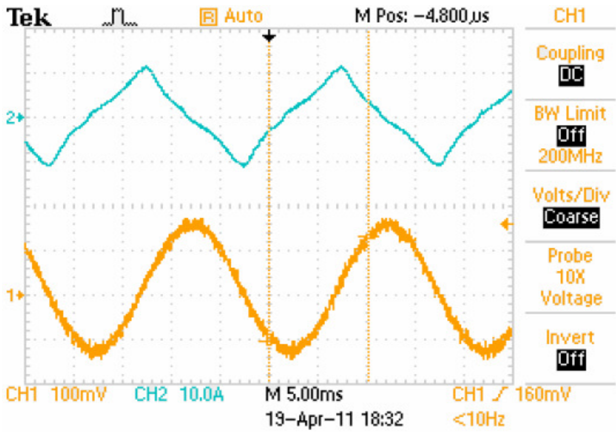


Fig.8. Signal waveforms used for verification of proposed system (up-current, down-voltage)

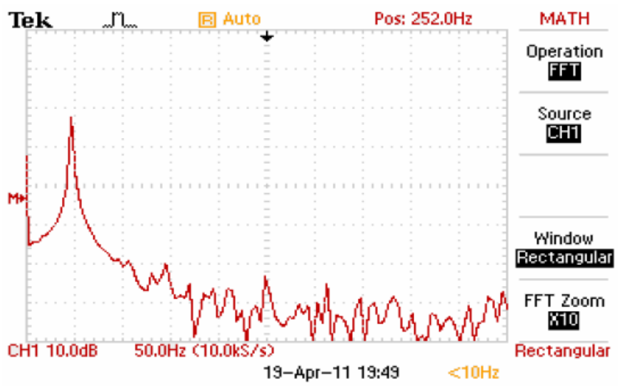


Fig.9. Amplitude spectrum measured on sensing resistor (computed by oscilloscope)

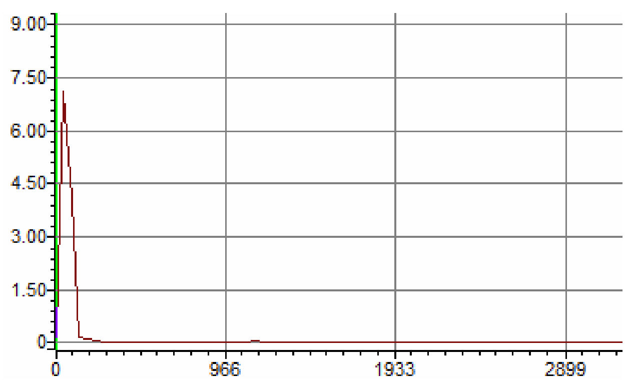


Fig.10. Amplitude spectrum computed by proposed system

From fig. 10 can be seen that proposed measuring system is acting correctly and that it analyzed measured current waveform in expected manner. There exist some problems according accuracy, what was caused by utilization of low-performance sensing resistor and microprocessor unit.

5.CONCLUSION

In this paper the system for analyzing of spectrum of supply current is proposed. All system, except the current sensor is in digital form. Proposed filter and computing algorithm shows good results for application as system for monitoring the recuperative currents from converter in electric locomotive. Performance and resolution is limited by used microprocessor, so for other application the more powerful processor should by used.

10. ACKNOWLEDGMENTS

The authors wish to thank to Slovak grant agency APVV for project no. APVV-0138-10 - Research and development of the small power drives with two-phase motors, and for project no. LPP-0366-09 - High frequency power electronic converter. Special thanks is also given to Slovak grant agency VEGA for project no. 1/0943/11.

REFERENCES

[1] Moudgalya, M.: **Digital Control**. Chichester: John Wiley & Sons, Ltd., 2007. 545 ISBN 978-0-470-03143-8

[2] Oppenheim, A. V., Schafer, R. W., – Buck, J. R.: **Discrete – Time Signal Processing**. 2nd edition, New Jersey: Prentice Hall, 1999. 878 s. ISBN 0-13-754920-2

[3] Dobrucky, B., Drgona, P., Spanik, P., Bobek, V.: **Using Power PC-MPC for real-time VHFIM sensorless drive control**, In: The 2nd International multiconference on complexity , informatics and cybernetics, 27-30 march, 2011, Orlando, FL-USA, pp-167-173, ISBN-13:978-1-936338-20-7

[4] Tiefenbacher, M., Jakubek, S., Kozek, M.: **Modeling and identification of a chain pendulum**, In: The 2nd International multiconference on complexity , informatics and cybernetics, 27-30 march, 2011, Orlando, FL-USA, pp-306-312, ISBN-13:978-1-936338-20-7

[5] Sitar, J., Racek, V., Havrila, R., Gerek, I.: **Input Active Filter for the DC/DC Converter Supplied from the Traction Troley Net**, 11th European Conference on Power Electronics and Aplycations, EPE 2005, 11-14.9.2005, Dresden, ISBN: 90-75815-08-5, IEEE Catalog Number: 05EX1132C

[6] Hargas, L., Koniar, D., Stofan, S. **Sophisticated biomedical tissue measurement using image analysis and virtual instrumentation**, LabVIEW - practical applications and solutions, INTECH, ISBN 978-953-307-650-8, 2011 pp.155 - 181

Database Performance Issues

Mihyar Hesson, College of Engineering and Information Technology,
Al Ain University of Science and Technology, UAE
mihyar.hesson@aau.ac.ae
Hayder Al-Ameed, College of Information Technology,
UAE University, UAE
Hader@uaeu.ac.ae
Mohammed Semaka, College of Science, Qatar University, Qatar
mohammed.semaka@yahoo.com

ABSTRACT

One of the major issues to address in regards to ORACLE database performance is the size of the database. The bulk of the information consists of a large number of records, contained in many tables, each ranging from thousands of rows. There are many factors that can have direct effects on the performance of the database. CPU, Memory, Network, Disk I/O are among other factors. In order to make a database up and run efficiently, each factor must be addresses carefully, and the best tuning strategy must be applied for optimum performance. ORACLE performance issues are complex, and for a DBA, there is a large number of values to monitor and examine in order to decide on best tuning strategy.

The aim of the work behind this paper was to design and implement a Database Performance Tuning Measurement Toolkit for ORACLE Database Servers working on MS-Windows platforms. This system is called ORACLE Performance Monitoring Toolkit (OPMT for short). It is used for testing, analysis and reporting of the database performance.

Keywords: ORACLE Database, Database Performance, Tuning, Availability, Windows operating system

INTRODUCTION

Database vendors are becoming aware that the human cost of operating large database systems is growing dramatically. As the scope of relational database functions has expanded in recent years, the complexity of database systems has also grown. The added complexity and the increase in data size (now frequently into tens of terabytes) have increased the burden on database administrators. The combination of increased data volumes, larger systems, and increased function, has motivated the need for autonomic capability within database management systems in order to reduce cost of ownership and to enable databases to operate in environments with limited access to skilled administration personnel [1].

A new and interesting approach to this management problem is an *autonomic DBMS* that is able to automatically manage its resources to maintain acceptable performance in the face of changing conditions [2]. Other terms used in the literature for an autonomic DBMS are *self-tuning DBMSs* [3] and *no knobs operation* [4]. An autonomic DBMS must be able to perform the configuration tasks currently carried out by DBAs to initially set up a system. Typical configuration tasks include determining appropriate allocations for main memory areas such as the buffer pools and the sort heap, mapping database objects (user defined tables and indexes) to disk storage and mapping database objects to buffer pools. Performing configuration tasks requires knowledge of the available system resources and of the characteristics of the workload.

Many system performance evaluation models have been proposed, varying in degree of accuracy, cost to build and run, and the reliance of the method on configuration-specific parameters. Many of these models are concerned with evaluating the appropriateness of a particular hardware configuration on the mix of programs that will run on the system. These methods, generally used for capacity planning, include queuing models, simulation, and monitors of contrived or real workloads. Detailed simulation models, such as IBM's CSS, FIVE [5], and ANCICSVS [6], QSIM [7], and APLOMB [8], are complicated, requiring large initial investments of time and money. These investments are not warranted by a single use; thus, these simulators are best suited for development by specific manufacturers and may not be publicly available. General packages, such as SCERT and CASE, are more useful as flexible evaluation tools [9].

On the other hand, and as businesses across all industries are getting more dependent on IT infrastructure, the problem of availability of IT systems becomes an increasingly strategic business concern. This is because the interdependencies of modern business are often linked through software. On one hand, employees, customers, and partners communicate and conduct commerce through networked systems, at a more intense level today than ever before. On the other hand, when one system in the network fails, it can break dependencies and impact business processes across the enterprise and beyond to customers and partners.

To avoid the ramifications of database downtime, many corporations have taken a renewed interest in database availability. For some, the goal is continuous availability, where a database server never fails, which would result in so-called unplanned downtime. Most companies do not need such a stringent level of database availability; they are satisfied with high availability, which allows for a small room of planned downtime allocated for database maintenance. Most strategies for continuous availability and high availability assume that a slight repair to hardware and/or software is all that is required for recovery.

High availability is often associated with fault-tolerant systems. The term *fault-tolerant* means a system can operate in the presence of hardware component failures A single component failure in a fault-tolerant system will not cause a system interruption because the alternate component will take over the task transparently. As the cost of components continues to drop, and the demand for system availability increases, many *non-fault-tolerant* systems have redundancy built-in at the subsystem level. As a result, many non-fault-tolerant systems can tolerate hardware faults— consequently, the line between a fault-tolerant system and a non-fault-tolerant system becomes increasingly blurred [10, 11, 12, 13]

Another key point in bringing database systems users' satisfaction is the system performance. Performance can be defined as the ability of a system to deliver the results based on the request of the users, while keeping them satisfied. The key point to remember is satisfaction. If current database key indicators reveal that the database is running optimally, but users are not satisfied with the response times, then there will be a need for tuning [14].

One of the biggest responsibilities of a DBA is to ensure that the database is tuned properly. The ORACLE RDBMS is usually tunable and allows the database to be monitored and adjusted to increase its performance. One should do performance tuning for the following reasons [15]:

- The speed of computing might be wasting valuable users' time (users waiting for response);
- Enable your system to keep-up with the speed business is conducted; and
- Optimize hardware usage to save money (companies are spending millions on hardware).

Tuning the database performance is not a simple task and it depends on the specific requirements, the operating system and the target hardware. There is no "one fits it all" approach. The goal is to avoid obvious slowdowns and balance the available resources (I/O bandwidth, memory and CPU) [16].

DATABASE PERFORMANCE TUNING

Every DBA has experienced a situation in which an application slows down after it has been in production for a while. But why this happens is not always evident. Perhaps the number of transactions issued has increased or maybe the volume of data has increased [17].

The overall performance of a system can be generally measured according to transaction response times, that is, the time it takes to complete a query or task, which is also the time the user must wait for the task to finish and possibly return results. Slow (long) response times translate into bad performance and frustrated users, whereas quick response times mean better performance and happy users. Database system may show different levels of performance (good or bad) at different times of the day, according to heavy or light user activity. If a user query takes a relatively long time to finish, based on previous tuning, it is an indication that system may have a performance problem that may be resolved, and this should be investigated.

Another way to know if there is a performance problem is by simply monitoring the system on a regular basis, and this is the theme of the work behind this paper. Before we describe this work, we will summarize some of the existing tools, services and some other techniques available in the market that address or deal with the tuning problems.

U.S. Computer Software

Performance improvement process conducted by U.S. Computer Software Inc. is iterative. Generally, it consists of step-by-step detection, analysis, and elimination of performance bottlenecks. At each step, different ways to remove a given bottleneck are considered, and taking into account all important project's constraints such as timeline and budget, to choose the optimum solution [18].

The solutions may include:

- Tuning operating system performance on computers running different operating systems.
- Tuning performance of database servers (including MS SQL Server, Oracle, and DB2). Optimization of database structures and schemas. Optimization of SQL queries sent to the SQL-based database servers. Optimization of the source code responsible for utilization of database servers (including the code using such APIs as JDBC, ODBC, DAO, ADO, OLE DB, and others).
- General optimization of Java and C++ code. Optimization of input/output using asynchronous I/O techniques. Optimization of multi-threaded code.
- Optimization of hardware configurations (including RAID systems and networking devices). Hardware upgrade planning.

SQL-Optimizer

Once a connection to an Oracle database is established, SQL-Optimizer opens a number of windows that display current user connection information [19]. An administrator can view all current threads that are operating on the RDBMS as well as drill down to information that is quite valuable in identifying inefficient SQL. Users of SQL-Optimizer can select sessions to see the query start and end times, the actual SQL that was issued to the database, and I/O statistics including a query’s "hit ratio," which is a measure of how much data the query obtained from memory vs. physical I/O. A high hit ratio (normally defined as 80 percent or greater) is good, while lower hit ratios indicate table scan problems and/or inefficient SQL.

Although the information displayed by SQL-Optimizer is very useful, the GUI itself is difficult to view for long periods due to the constant refreshing the tool must perform to keep up with database activity. The tool’s visual presentation suffers from jerky and choppy behavior in the statistical windows displaying the connection data.

SQL-Optimizer offers much in the way of quickly analyzing currently running SQL. By selecting the correct options within the tool, an administrator can instantly see the actual SQL a user has submitted along with a formatted explain analysis. A number of other tools are capable of doing this (such as Platinum’s Plan Analyzer for Oracle), but SQL-Optimizer interrogates the current SQL and, using a built-in SQL "Expert" engine, would determine if the SQL could be written differently to improve performance. Any suggestions the tool has, are displayed to the user in the form of rewritten code and a new EXPLAIN analysis, which details the latest code’s access paths.

RDM Services (Remote Database Monitoring)

This is a rather on-line service than a tool. A Progress Professional Services consultant initiates the RDM deployment by conducting an on-site database tuning service, tailored to the individual needs. The objective is to provide users with the insurance that the database will achieve top performance in its operating environment and workload [20].

During this engagement, the experienced consultants will:

- Identify performance bottlenecks
- Review and adjust standard database parameter settings and configurations
- Make environmental and OS specific adjustments
- Conduct a review of procedures, such as backup, dump and reload, and disaster recovery
- Make additional recommendations for ongoing monitoring and tuning
- Provide a detailed report outlining adjustments made and summarizing recommendations

Database aggregation to improve database performance

The development of the database aggregation and calculation engine has created a vastly more cost-effective alternative for performance improvement. Like the turbo supercharger on a high-performance racing engine, this new form of software works by addressing the critical factor that limits performance. In very large databases, the limiting factor is usually the data aggregation function. For some years, critics like Ralph Kimball have argued that "the single most dramatic way to affect performance in a large data warehouse is to provide a proper set of aggregates" [21, 22, 23, 24]. In some cases, Kimball argues, this strategy can speed queries "by a factor of 100 or even 1000. No other means exists to harvest such spectacular gains."

One data aggregation engine product in particular combines patented aggregation and caching algorithms with advanced load-balancing administrative facilities to exploit the full potential of modern microprocessor-based servers. The cache resides on the existing database server and is accessible as an existing database table. Once installed, the aggregation and caching engine can be referenced by standard SQL queries, existing applications and OLAP tools, and is completely invisible to business intelligence users and their applications. In operation, the aggregation and caching engine replaces the summary tables that traditional RDBMSs must generate to aggregate data. The caching engine emulates multidimensional as well as traditional relational summary tables and can be accessed by many different kinds of tools, including OLE-DB, ODBC or JDBC.

Perhaps the most extraordinary feature of database aggregation engines is their connectivity. The database aggregation engine is designed for fast, economical integration into an existing database infrastructure. It can operate with a wide variety of currently popular operating systems such as MS Windows, NT, and UNIX, with DBMS systems such as Oracle, Microsoft, Oracle Express and SQL to name only a few and can interface directly with existing reporting tools and database applications.

THE OPMT SYSTEM

The structure of the system include buffer cache, database blocks tuning, input/output tuning, operating system tuning, shared pool and others. It is out of the scope of this paper to include the flowcharts of these.

OPMT consists of two functional parts, each is interfaced through a number of screens, the first part is dedicated for testing the system, and it is interfaced by four screens that the user can move between by selecting the appropriate tab. Each screen is related to different category.

The second part is for displaying the analysis results and reports. There are three interface screens designed for that purpose.

System Tests

The user interface screens are available to perform system tests, these screens are used for:

- Shared pool. The first screen "Shared Pool" is divided into three database blocks. These are:
  - Library cache—in which three records were retrieved, these are: Gethitratio, Pins and Reloads. When the button “Calculate reloads–to–Pins Ratio” is pressed the system will display the value of “sum(reloads)/sum(pins)”. This value is analyzed and the result is displayed in a separate window as either "good" or "bad” alongside some recommendation that will appeare if the analysis is bad. This is to inform the user about the necessary actions that must be taken to improve the system. In addition to the recommendation, some brief comments that are related to the monitored database system status, will be displayed also as shown in figure 1 below.
  - Shared pool reserved area—in which two records are retrieved; these are: "Request Misses" and "Request Failures". These values are displayed when the “V$SHARED_POOL_RESERVED” button is pressed.
  - Data dictionary cache—in which three values are retrieved; these are: "Parameter", "Gets" and "Getmisses". When the user clicks on “sum of GETMISSES to sum of GETS” button, OPMT calculates the percentage of the sum of "GETMISSES to the sum of GETS" and checks if this value is less than 15 per cent, then it will show some analysis regarding this calculated value.
- Redo Log buffer. The second screen “Redo log buffer” is used for two system testing functionalities; these are:
  - Session Wait. This is used to check the length of system waiting period in seconds. Pressing on "Ckeck Seconds in Wait" button, as shown in the upper part of figure 2 below, four pieces of information will be retrieved and displayed. These are SID “Session ID”, the Event, the waiting period in Seconds, and State of the system. If the waiting period is relatively too long, the system will display "Bad" in the "Analysis" window in the left bottom corner of the screen.

It also gives the necessary recommendations to deal with this problem and some comments too.

  - System Status "Sysstat". This to check for the value of redo buffer allocation retries. Again, if it is relatively large the system will show the results of the analysis, recommendation and comments too.

- Input/Output. The third user interface is the “Input/Output” screen and it is about scan statistics. This screen is mainly used for retrieving and displaying two values, namely: the name and the value of the scanned table. Again, if the value is undesirably high, the system will display "Bad" in the Analysis window and recommendations and comments will be given too. This screen is shown in figure 3 below.

- Undo Segments. This is the forth screen which is used for checking the number of deleted values. The user can get the number of deleted values by clicking on the “Check if number of waits > 1% number of requests” button. The desirable value must be less than 1% of the total number of requests. If this is the case, the system will display "Good", otherwise, the user will get "Bad" result as described above. This is shown in figure 4 below.

System Analysis and Reports

The operations performed in the previous part are available for analysis and reporting in this part as shown in figure 5. In this screen, some of operations that have been previously performed are summarized according to Category, Type, Analysis Results, Recommendations and Comments.

The user can save these results if he/she likes by pressing on the "Save" button. A summary report can be displayed by pressing on the "Summary Report" button. The summary report will display all the previous operations in more details in a form of pages as shown in figure 6 below. The last screen combines the analysis summary with test details in two separate windows. The common link between the two is the test id as shown in figure 7 below

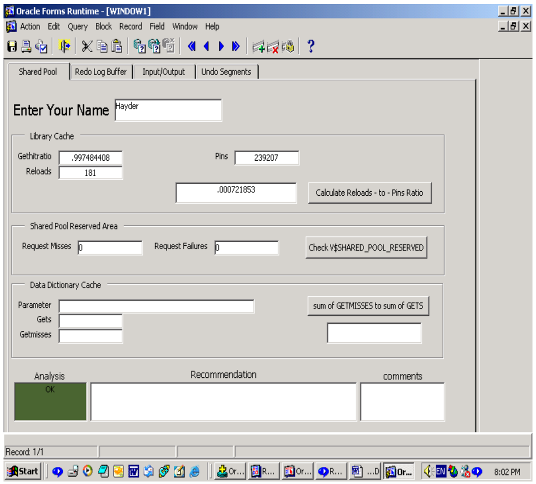


Figure 1 The Shared Pool Screen

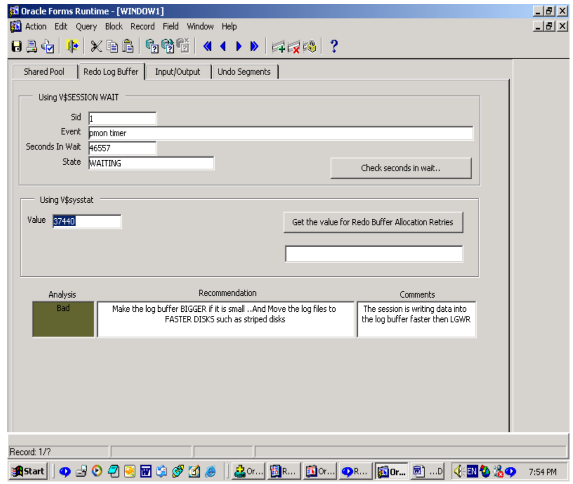


Figure 2 The Redo Log Buffer Screen

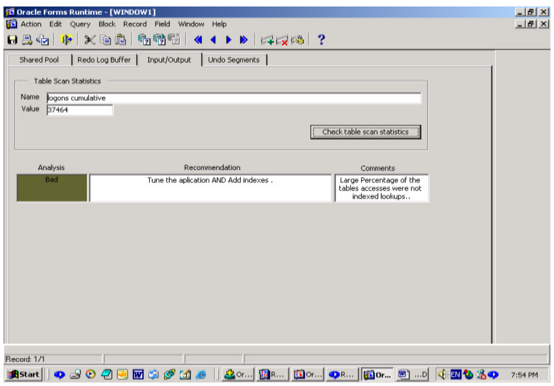


Figure 3 The Input/Output Screen

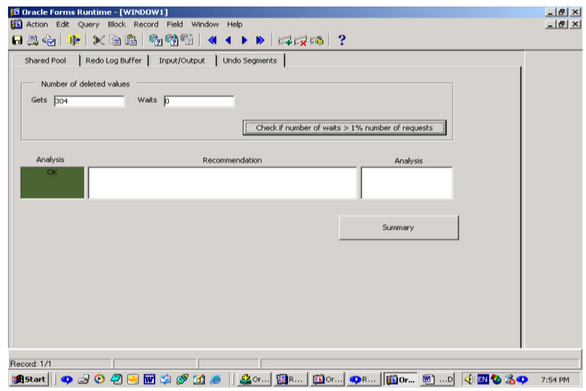


Figure 4 The Undo Segments Screen

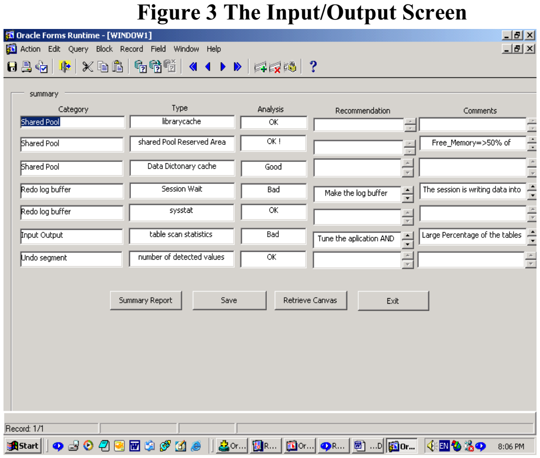


Figure 5 The Operations Summary Screen

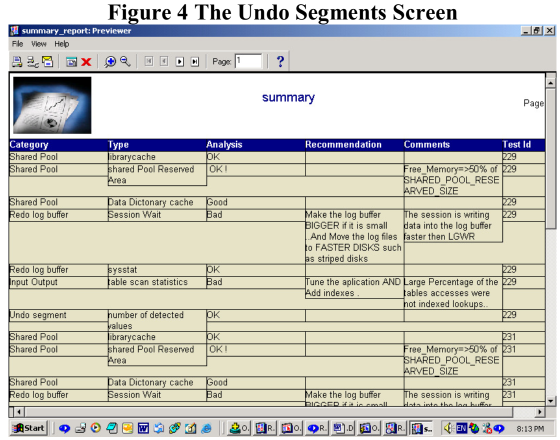


Figure 6 The Report Screen

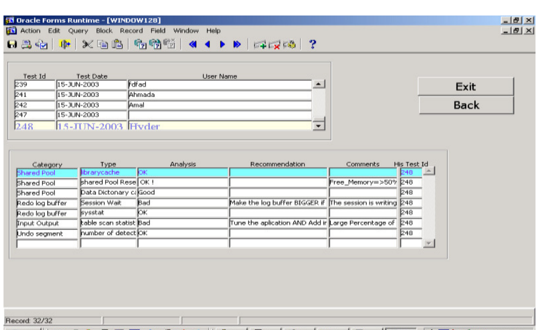


Figure 7 More Test Details Screen

CONCLUSION

OPMT is a database performance monitoring and tuning system that was developed to be used with ORACLE Database system run on Microsoft Windows operating systems. The main objective of speeding up database systems performance tuning processes and simplifying the DBA duties in a reliable and flexible manner was achieved. The system has been tested and proved efficient and reliable.

OPMT provides a flexible structure that can be further developed with minor changes. It provides friendly interfaces that can be easily used by the DBA and developers to monitor database performance. All test information and produced history files for each performance test are stored and can displayed and examined at any time.

The system design is flexible and can be easily expanded. Future expansions and changes may include:

- Make the system multi-lingual.
- Adding graphics and charts.
- Expand it to work on other database management systems like MS-SQL Server, DB2, and Sybase.
- Make it to run on other operating systems, like UNIX and Linux.

ACKNOWLEDGEMENT

We would like to thank Amal Khalifa Al-Ghufli, Alia Ali Al-Amimi**,** Salamah Abdullah Al-Qaydi, Enas Saleh Shehatah, and Hiba Nawfal Yonis for their valuable contributions in this work.

REFERENCES

[1] Lightstone, S. *et al.*, "Toward Autonomic Computing with DB2 Universal Database", SIGMOD Record, Vol. 31, No. 3, September 2002.

[2] Xu, X., Martin, P. and Powley, W., "Configuring Buffer Pools in DB2 UDB", IBM Canada Ltd., the National Science and Engineering Research Council (NSERC) and Communication and Information Technology Ontario (CITO), 2002.

[3] Chaudhuri, S. (ed). Special Issue on, "Self-tuning Databases and Application Tuning", IEEE Data Engineering, *Bulletin 22(2)*, June 1999.

[4] Bernstein, P. *et al*., "The Asilomar Report on Database Research", ACM SIGMOD Record 27(4), December 1998, pp. 74 - 80.

[5] NGUYEN, H. C., OCKENE, A., REVELL, R., AND SKWISH, W. J., "The role of detailed simulation in capacity planning". IBM Syst. J. 19, 1 (1980), 81-101.

[6] SEAMAN, P. H., "Modeling considerations for predicting performance of CICS/VS systems", IBM Syst. J. 19, 1 (1980), 68-80.

[7] FOSTER, D. V., MCGEHEARTY, P. F., SAUER, C. H., AND WAGGONER, C. N., "A language for analysis of queuing models", Proceedings of the 5th Annual Pittsburgh Modeling and Simulation Conference (Univ. of Pittsburgh, Pittsburgh, Pa., Apr. 24-26). 1974, pp. 381-386.

[8] REISER, M., AND SAUER, C. H., "Queuing network models: Methods of solution and their program implementation", Current Trends in Programming Methodology. Vol. 3, Software Modeling and Its Impact on Performance, K. M. Chandy and R. T. Yeb, Eds. Prentice-Hall, Englewood Cliffs, N. J., 1978, pp. 115-167.

[9] BOROVITS, I., AND NEUMANN, S., "Computer Systems Performance Evaluation", D.C. Heath and Co., Lexington, Mass., 1979.

[10] Enrique Vargas, High Availability Fundamentals, Sun BluePrints™ OnLine - November 2000, http://www.sun.com/blueprints

[11] Harry Singh, Distributed Fault-Tolerant/High-Availability Systems, Trillium Digital Systems, a division of Intel Corporation, 12100 Wilshire Boulevard, Suite 1800 Los Angeles, CA 90025-7118 U.S.A. Document Number 8761019.12.

[12] David McKinley, High availability system High availability system platforms, Dedicated Systems Magazine - 2000 Q4 (http://www.dedicated-systems.com)

[13] Sasidhar Pendyala, Oracle's Technologies for High Availability, Oracle Software India Ltd., India Development Centre.

[14] James Koopmann, Database Performance and some Christmas Cheer, an article in the Database Journal, January 2, 2003.

[15] Frank Naudé, Oracle Monitoring and Performance Tuning, http://www.orafaq.com/faqdbapf.htm

[16] Michael Marxmeier, Database Performance Tuning, an article, http://www.hp-eloquence.com/support/misc/dbtuning.html

[17] Craig S. Mullins, eDBA: Online Database Reorganization, an article published in the dbazine.com magazine May 2002, http://www.dbazine.com/edba10.html

[18] Oracle® Express**,** Performance and Database Design Guide: Basic**,** Release 6.3**,** January 2000**,** Part No. A65169-01

[19] Robin Schumacher, SQL-Optimizer is a performance management SQL monitor for Oracle servers, DBMS, January 1998, http://www.dbmsmag.com/9801d09.html

[20] PROGRESS SOFTWARE, Progress Software Professional Services, http://www.progress.com/za/services/index.ssp

[21] Ralph Kimball, http://www.informatik.uni-trier.de/~ley/db/indices/a-tree/k/Kimball:Ralph.html

[22] Information Builders, OLAP–Online Analytical Processing–Tools, http://www.informationbuilders.com/olap-online-analytical-processing-tools.html

[23] Ralph Kimball, Intelligent Enterprise Magazine articles, http://www.rkimball.com/html/articles.html

[24] OLAP Tool Functionalities, Data Warehousing - OLAP Tool Selection, http://www.1keydata.com/datawarehousing/toololap.html

Analysis of Performance Parameters of Just In Time and Non- Just In Time Based Manufacturing Industries in India

Sunand Kumar
Department of Mechanical Engineering, National Institute of Technology,
Hamirpur (Himachal Pradesh) PIN-177 005, India
(Email:sunandkumar_gupta@yahoomail.com)
and

Sarbjeet Singh
Department of Mechanical Engineering, Government College of Engineering & Technology,
Jammu (Jammu &Kashmir) PIN-180 001, India

ABSTRACT

The present paper discusses about identifying critical elements of Just in Time and comparison of performance analysis of Just in Time and non-Just in Time based Indian manufacturing industries viz: automobile component manufacturing, steel manufacturing, electronic component manufacturing industries etc. Questionnaire based on various performance variables were sent to nearly 89 industries however about 45 responses were received. Based on the data statistical tests such as t-test, Cronbach α-analysis and β-regression model have been implemented to find the role of performance parameters in these industries. Based on T-test analysis it was reported that various performance parameters such as effectiveness of production planning (3.02), accuracy of demand forecasting (2.77), accuracy of completing production plan (3.01), zero deviation schedule (3.36), level of inventory control (2.76), played vital role in case of JIT based industries. Whereas non JIT based industries showed poor performance in terms of continuous improvement (0.105), level of WIP reduction (0.14), level of inventory control (0.55) and operation cost (0.315). Similarly, in regression analysis viz: setup time reduction (0.130), effectiveness of production planning (0.231), accuracy of demand forecasting (0.19), level of work in process reduction (0.075), reduction in manufacturing lead time (0.127), level of inventory control (0.185) also showed great improvement in terms of JIT based industries. However, non JIT based industries showed poor performance in terms of zero defect manufacturing (0.082), level of inventory control (0.124), reduced manufacturing lead time (0.096) and degree of ontime delivery (0.004). The paper concludes that Just in time based industries have competitive edge over non Just in time based industries on some selected performance parameters.

Key words: Just in Time, Statistical Technique, Regression Analysis, Performance Parameter.

1. INTRODUCTION

Just in Time has been widely implemented in manufacturing industries as a survival strategy against global market competition and offers various benefits such as greater though put high productivity and better quality. The JIT has more impact on production control performance than production planning performance [1]. The similarities between just in time manufacturing and simultaneous new product development is of great concern [2]. JIT was not complicated, but implementation stage is not easy to face at all. However this could bring high level of competitiveness to big companies that in general suffer from excessive stocks at different stages of production chain [3-4]. **S**everal types of production systems had been studied and researched in the past using either simulation or optimization method [5-6]. JIT culture in India can bring significant improvements in work culture and JIT success is not possible without the involvement of employees and others who are associated with the industries directly or indirectly [7-8]. There is need to improve the work environment before the JIT can be implemented in India [9]. There are about 29 elements which are responsible for implementing JIT in Indian manufacturing industries [10]. The methodological and implementation issue in JIT system are proving to be of vital importance for viability and advancement of manufacturing enterprises in a continually changing competitive business environment [11]. In the era of liberalization and privatization, Indian industries needed instant support of meaning full and effective manufacturing system like JIT to follow the trend of globalization in order to become more competitive [12]. JIT is an over all organizational phenomena and a very strong relationship between JIT practices, infrastructure practices, combination of JIT management and infrastructure practice was related to manufacturing performance [13]. Just in Time is not complicated, but implementation stage is not easy to face at all. However this could bring high level of competitiveness to big companies that in general suffer from excessive stocks at different stages of production chain [6]. There are various problems of developing countries like India where stress is on maintaining efficiency rather then reducing cost in contrast to developing countries better infrastructure and practices automatically efficiency. Due to wide differences in the operating environment of Indian industries, there is a need to improve the work environment before the JIT can be implemented in India. The current JIT index is 23.38 on a 40 point scale employing that though it is quite difficult, JIT implementation is possible. In today's competitive market the performance in manufacturing industries depend on a superiority of manufacturing and design technology. Just in time is a measure of performance in manufacturing. JIT is a concept of manufacturing and supplying goods that are needed, when they are needed in exact quantity and at right time. This results in many different strategies adopted by the industries to bench mark there performance. So JIT is one of the manufacturing strategies for comparing the market. Little research has been carried out to improve the quality at optimum cost in spite of appreciable industrialization in India, Due to limited availability of resources; it is highly difficult for India to promote advanced manufacturing technology. Indian industries need new technology, which is characterized by low investment, high short-term pay off, high rate of productivity and high quality. At present, Indian industries run with traditional manufacturing systems like job shop and continuous flow manufacturing systems, whereas Japanese and European industries have developed and employed new manufacturing systems. Indian industries require major effort to convert the traditional manufacturing system into new manufacturing systems. In the present work the performance of different Indian industries in terms of selected performance variables has been analyzed and comparative analysis of JIT and non JIT industries has been done.

2. METHODOLOGY

Based on five point Likert Scale [10] the questionnaire has been prepared to collect the relevant data regarding the quantum of importance, difficulties and constraints in the implementation of JIT in Indian manufacturing industries. The questionnaires were sent to various manufacturing industries and were mainly selected from Jammu & Kashmir, Himachal Pradesh, Punjab, Haryana, Delhi and Utter Pradesh in India. Out of 89 manufacturing industries only 45 responses were received. The questionnaire has been broadly divided into Section A and section B. The objective of the questionnaire was to collect general information regarding annual turnover, company ownership, production strategy, implementing JIT or not, type of product manufactured, degree of importance in the case of implementation of JIT and also on evaluation of performance parameters degree. The collected data was arranged and analyzed by checking it at significant level of 0.05% through various statistical tests such as t-test, Cronbach α-analysis and β-regression model to find the role of performance parameters in case of JIT and non JIT based industries Data analysis was carried out on the performance parameters like effectiveness of production planning, accuracy of demand forecasting, level of WIP reduction, degree of on time delivery, degree of quality, setup time reduction, continuous improvement, process flexibility, zero defects, zero deviation schedule, accuracy of completing production plan, flexibility of production planning, reduced manufacturing lead, level of inventory control ,operation cost.

3. RESULTS AND DISCUSSIONS

In the study, Cronbach's α- analysis was used to test the reliability of fifteen response variables identified for measuring the production performance of just in time and non just in time based industries. It has been observed that Cronbach's alpha lie between 0.75 and 0.85 which indicates that the selected performance parameters are reliable for comparative analysis of performance of just in time and non just in time based industries. On the basis of comparative analysis as per T-values (Table 1) for various production performance variables of JIT and non JIT based industries, it is concluded that the performance parameters such as effectiveness of production planning, accuracy of demand forecasting, zero deviation schedule, accuracy of completing production plan has positive impact on JIT based industries whereas accuracy of demand forecasting, degree of quality, process flexibility has positive impact on non JIT based industries. It is further concluded that non JIT based industries perform poor in terms of various production performance variables such as zero defect manufacturing, level of inventory control, continuous improvement, on time delivery. In case of high impact production variables such as PP01(effectiveness of production planning), PP010 (zero deviation schedule) and PP011(accuracy of completing producing plan) PP03 (level of WIP reduction) and PP014 (level of inventory control), the percentage variation is more than double in case of JIT based industries as compared to non JIT industries. That is in case of JIT based industries effectiveness of production planning (3.02), zero deviation schedule (3.36), accuracy of completing production (3.01), level of WIP reduction (0.90), level of inventory control(2.76) plays a very vital role when compared with non JIT industries. Similarly incase of some production variables such as PP06 (setup time reduction), PP08 (process flexibility), PP013 (reduced manufacturing lead time), PP09 (zero defects) the percentage variation is more then 100% and incase of production variables such as PP07 (continuous improvement), PP05 (degree of quality), the percentage variation is around 50% and these variables have little impact when comparing JIT industries and non JIT industries. The T-values for different production variables for comparative analysis of JIT base and non JIT based industries has been reflected in Figure 1. On the basis of comparative analysis as per regression analysis, it is concluded that various performance parameters such as effectiveness of production planning, set up time reduction, degree of on time delivery, zero deviation schedule, level of inventory control have positive impact on JIT based industry as compared to non JIT based industries. Performance parameters such as operation cost, flexibility of production planning have positive impact on non JIT based industries. However, non JIT based industries suffer a lot on various production performances variables such as zero defect manufacturing, level of inventory control, reduce manufacturing time, degree of on time delivery. In case of high impact production variables such as PP01 (effectiveness of production planning), and PP010 (zero deviation schedule), PP02 (accuracy of demand forecasting), PP07 (continuous improvement), the percentage variation is more than 99%. That is in case of JIT based industries effectiveness of production planning and zero deviation schedule play a vital role when compared with non JIT industries. Similarly in case of production variables such as PP012 (flexibility of production planning), PP09 (zero defects), PP06 (setup time reduction), the percentage variation lies between 50% and 99% that is these variables play moderate role when comparing JIT industries with non JIT industries (Table 2). Also production variables PP014 (level of inventory control), PP013 (reduced manufacturing lead time) PP05 (degree of quality), the percentage variation is less than 50% that is above variables have very little role to play when comparing JIT industries and non JIT industries. Figure 2 shows that plot of β values for different production variables for comparative analysis of JIT and non JIT based industries. Figure clearly depicts the production variables having high impact and low impact on production performance of JIT based industries as compared to non JIT based industries.

4. CONCLUSIONS

It is concluded that the factors having positive influence on the performance of industries are, setup time reduction, effectiveness of production planning, accuracy of demand forecasting, level of work in process reduction, degree of on time delivery, reduced manufacturing lead time, level of inventory control with respect to Just in time component. Further, Non Just in time industries suffer a lot on zero defect manufacturing, level of inventory control, reduced manufacturing lead time, continuous improvement, degree of on time delivery. On the basis of T-test analysis it can be concluded that in case of some high impact production variables such as effectiveness of production planning (3.02), zero deviation schedule (3.36), accuracy of completing production plan (3.01), accuracy of demand fore casting (2.76), the percentage variation is more then double in case of JIT based industries when comparing with non JIT based industries. In case of production variables such as setup time reduction (1.96), process flexibility (0.24), reduced manufacturing lead time (1.82), zero defect (1.72), the percentage variation is nearly 100% when comparing JIT based industries with non JIT based industries on the basis of T-test analysis. Based on the T-test values, non JIT based industries showed poor performance in terms of production variables such as level of WIP reduction (0.14), operation cost (0.315) and continuous improvement (0.015).

Table 1 T-value for JIT and Non JIT industries

Production performance variable	T-value		Description of variable
	JIT industries	Non JIT industries	
PP01	3.020	1.080	Effectiveness of production planning
PP02	2.770	1.610	Accuracy of demand forecasting
PP03	0.900	0.140	Level of WIP reduction
PP04	0.200	1.050	Degree of on time delivery
PP05	1.800	1.001	Degree of quality
PP06	1.900	0.742	Setup time reduction
PP07	0.140	0.105	Continuous improvement
PP08	0.240	1.122	Process flexibility
PP09	1.720	0.722	Zero defects
PP010	3.360	1.072	Zero deviation schedule
PP011	3.010	1.142	Accuracy of completing production plan
PP012	1.920	0.645	Flexibility of production planning
PP013	1.820	0.742	Reduced manufacturing lead time
PP014	2.760	0.550	Level of inventory control
PP015	0.210	0.315	Operation cost

Table 2 β -value for JIT and Non JIT industries

Production performance variable	β-value		Description of variable
	JIT industries	Non JIT industries	
PP01	0.231	0.103	Effectiveness of production planning
PP02	0.190	0.092	Accuracy of demand forecasting
PP03	0.075	0.006	Level of WIP reduction
PP04	0.002	0.004	Degree of on time delivery
PP05	0.140	0.097	Degree of quality
PP06	0.130	0.079	Setup time reduction
PP07	0.126	0.008	Continuous improvement
PP08	0.094	0.009	Process flexibility
PP09	0.135	0.082	Zero defects
PP010	0.230	0.116	Zero deviation schedule
PP011	0.207	0.122	Accuracy of completing production plan
PP012	0.184	0.107	Flexibility of production planning
PP013	0.127	0.096	Reduced manufacturing lead time
PP014	0.191	0.124	Level of inventory control
PP015	0.003	0.014	Operation cost

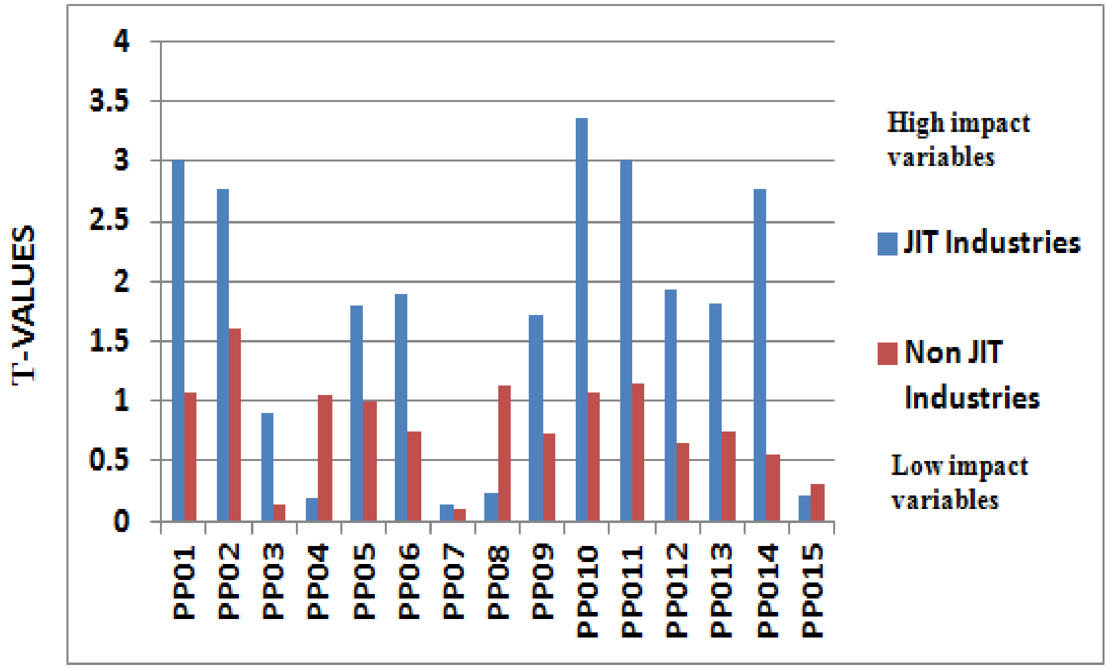


Figure 1 **Production Variables versus T-values**

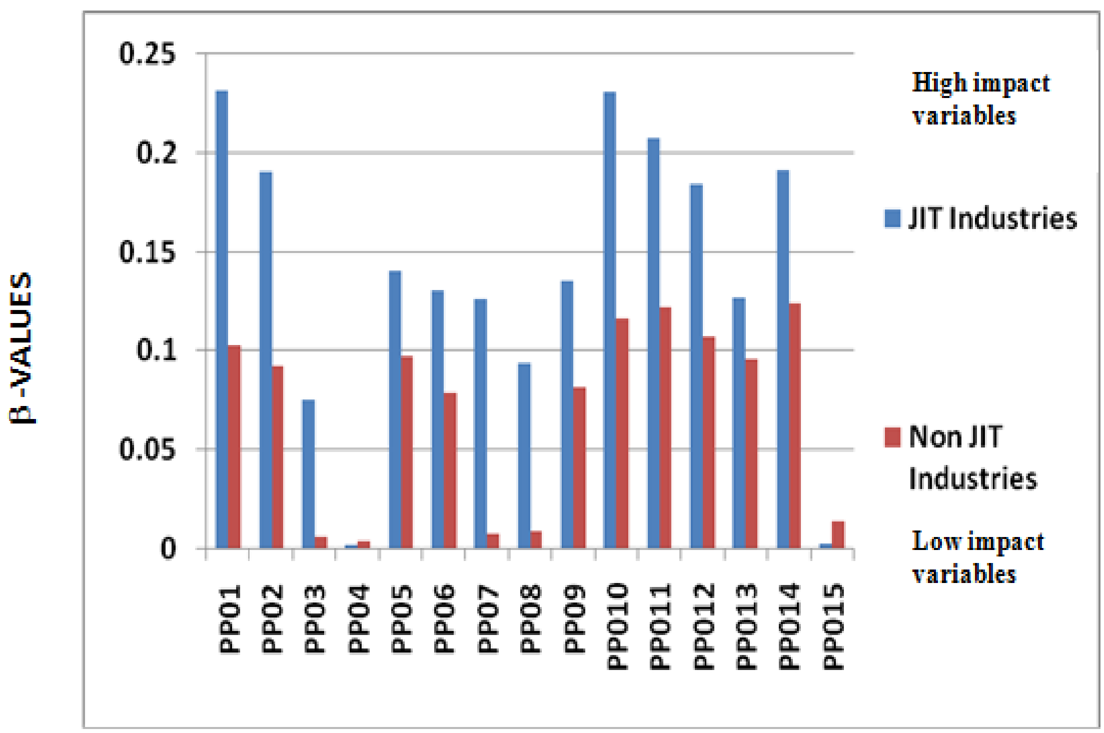


Figure 2 **Production Variables versus β --values**

Further, on the basis of comparative analysis of JIT industries and non JIT industries it has been concluded that JIT component that has noticeable positive influence on performance of industries are, setup time reduction, effectiveness of production planning, accuracy of demand forecasting, level of work in process reduction, degree of on time delivery, reduced manufacturing lead time, level of inventory control. Non JIT industries suffer lot on zero defect manufacturing, level of inventory control, reduced manufacturing lead time, continuous improvement, degree of on time delivery. From the regression analysis it is clear that some high impact production performance variables such as effectiveness of production planning (0.021), zero deviation schedule (0.230), accuracy of demand forecasting (0.19), continuous improvement (0.126), differs more than 99% when comparing JIT based industries with non JIT based industries. In case of non JIT based industries, regression analysis shows poor performance in terms of degree of on time delivery (0.004), level of WIP reduction (0.006), process flexibility (0.009) and continuous improvement (0.008). The regression analysis also points out that in case of production variables such as reduced manufacturing lead time (0.127), degree of quality (0.14), the percentage variation is less than 50%, that is these variables have little role to play when comparing JIT industries with non JIT based industries. At last it can be concluded that comparative analysis of JIT based industries have competitive edge over non JIT based industries on selected performance parameters.

REFERENCES

[1] Aniko. H. Prajogo and Robert B Johnston; "Barriers to just in time implementation in small manufacturing enterprises", proceedings of APICS conference, Las Vegas, Nevada, 1988, pp. 486 – 489

[2] A. Kumaravadivel and R. Sathiyamoorthy; "Determination of optimum base stock level (JIT Complementary), Industrial Engineering, Journal, Vol. 36, No. 4, April 2007, pp. 16 - 23.

[3] Wilhelm Dangelmaier and Kiran R. Mahajan; "simulation assisted optimization and real time control aspects of flexible production system subjected to disturbances", IEEE, Vol. 1, 4244, No. 0501, July 2006, pp. 1785 - 1794.

[4] Kumar, Vikas and Garg, Dixit; "JIT elements in Indian context: An analysis" productivity, Vol. 41, No. 2, July - Sept 2006, pp. 217 - 222.

[5] Chandra, Pankaj and Sastri Trilochan; "Competitiveness of Indian manufacturing" Vikalpa, Vol. 23, No. 3, June 2005, pp. 25 - 36.

[6] Kibara, Barbara, B and Flynn; " Impact of Just in time manufacturing as infrastructure in manufacturing performance"; JSTOR, management science, Vol. 43, No. 9, Sept 1997, pp. 1246 - 1257.

[7] Ansari, A., 1996, "Survey identifying critical analysis in successful implementation of just in time pulling techniques," Industrial Engineering Journal, 12, pp. 46 - 50.

[8] Singvi, S., 1996, "Employee involvement in JIT success: The Eicher Experience", Productivity, 33(2), pp. 15 - 25.

[9] Vrat, P., and Mittal S., 1998, "Implementation of JIT in Indian environment: A Delphi study", Productivity, 14(3), pp. 215 - 221.

[10] Garg, D., 1998, "Critical analysis in JIT purchasing in Indian context" Productivity, 37, pp. 271 - 277.

[11] Satir, A., 2001, "Just in time manufacturing system", Proceeding of international conference, Canada, pp. 111 - 115.

[12] Garg, D., 2000, "JIT quality techniques in Indian Industries", Productivity, 42(4), pp. 582 - 589.

[13] Ebrahimpur, M., and Schonberger, R.J., 2004, "The Japanese Just In Time/Total quality control production system:Potential for developing countries", International Journal of Production research, 22(3), pp. 421 – 430.

Uniqueness of Optimal Finite Dimensional Flows for Mixing Enhancement

Weijiu Liu
Department of Mathematics, University of Central Arkansas
Conway, AR 72035, USA

ABSTRACT

The problem of mixing enhancement is considered by tuning finite dimensional flows in an optimal way, which is defined as the linear combination of linearly independent steady flows. Optimal mixing control problem is formulated by using the flow as control variable. Using variational principles, we prove the existence of an optimal flow and derive an optimality system that consist of nonlinear convection-diffusion equations and ordinary differential equations. If the initial concentrations are sufficiently small or the diffusivity is sufficiently large, we prove that the optimal flow is unique, and then synthesize an optimal dynamical state feedback controller following the dynamical programming procedure.

Keywords: Mixing Enhancement, Optimal Control, Convection-Diffusion Equations, Variational Principle and Finite-Dimensional Flows.

I. INTRODUCTION

Mixing enhancement is central to the vast majority of processes in the chemical, pharmaceutical, aeronautical, and hydrocarbon processing industries. Well-mixed chemical reactions can yield substantial product benefits and enhanced mixing of fuel can optimize combustion chamber.

Mixing can be enhanced by destabilizing a flow [4], [12], [13], [14], [15], [17], [27], [30]. The flow can be destabilized using passive control devices such as the backward facing step [32] and lobed nozzles [8], open-loop active excitations through flaps and wall-jets [16], and active feedback controls [1], [5], [33], [31].

The objective of this paper is to continue the first author's work [23] by deriving mathematical criteria for an optimal finite dimensional flow for mixing enhancement and proving the uniqueness of the optimal flow under certain conditions, which was left as an open problem in [23].

A simple mathematical mixing model is the convection-diffusion equation

$$\begin{aligned} \frac{\partial c}{\partial t} + (\mathbf{v}\cdot\nabla)c &= \kappa\nabla^2 c \quad \text{in } \Omega, \\ c(\mathbf{x}, t_0) &= c^0(\mathbf{x}) \quad \text{in } \Omega, \\ \frac{\partial c}{\partial \mathbf{n}} &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{1}$$

Supported by the University Research Council Fund of the University of Central Arkansas.

in the absence of a source or sink. In the above equation, $c = c(\mathbf{x}, t)$ denotes the concentration of a physical quantity called a passive scalar, $c^0(\mathbf{x})$ is an initial concentration, $\kappa > 0$ denotes the molecular diffusivity of the scalar, Ω is a bounded domain in $\mathbb{R}^n$, $\frac{\partial}{\partial \mathbf{n}}$ denotes the normal derivative along the boundary $\partial\Omega$ ($\mathbf{n}$ denoting the unit normal on the boundary), $\mathbf{v} = \mathbf{v}(\mathbf{x},t)$ denotes an incompressible velocity field ($\nabla\cdot\mathbf{v} = 0$), $\nabla = \left(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n}\right)$, and $\nabla^2 = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$. We assume that $\mathbf{v}$ satisfies no-slip boundary conditions on the boundary $\partial\Omega$ ($\mathbf{v} = 0$).

In the preliminary study [23], we assumed that an arbitrary unsteady flow can be generated. This may not be realistic. Thus we consider finite dimensional flows given by

$$\mathbf{v} = \sum_{i=1}^{m} \bar{\mathbf{v}}_i(\mathbf{x}) u_i(t), \tag{2}$$

where $\bar{\mathbf{v}}_i(\mathbf{x})$ ($i = 1, \cdots m$) are given steady flows and $u_i(t)$ ($i = 1, \cdots m$) are weight controls. These steady flows prescribe how the control action is distributed in the flow field. Such finite dimensional flows were suggested in the preliminary study [23] and were studied in [24].

Following our preliminary study [23], we define mixing efficiency functionals by penalizing the average of variance of a diffusive scalar and the average of the flow weights. Using variational principles, we prove the existence of an optimal flow weight and derive an optimality system that consists of nonlinear convection-diffusion equations and ordinary differential equations. Furthermore, we show that if the initial concentrations are sufficiently small or the diffusivity is sufficiently large, then the optimal weight is unique. This uniqueness result enables us to synthesize an optimal dynamical state feedback controller following the dynamical programming procedure.

The optimal mixing problem has been studied in the literature. Using the entropy of automorphisms of dynamical systems as the measure of mixing efficiency, D'Alessandro, Dahleh, and Mezic [2] formulated an optimal mixing problem by maximizing the entropy among all permissible periodic sequences composed of two shear flows orthogonal to each other. They derived the form of the protocol which maximizes the entropy by developing appropriate ergodic-theoretic tools. Another optimal mixing problem was defined by Noack, Mezic, Tadmor, and Banaszuk [26], who used the flux across a recirculation region as the measure of mixing efficiency and then maximized the flux among all permissible controlled vortex motions.

The paper is organized as follows. We define a mixing efficiency functional in Section II. We then derive an optimality system for the optimal flow weight in Section III. Using the Banach fixed point theorem of contraction mapping, in Section IV, we prove the uniqueness of the optimal flow weight if initial concentrations are sufficiently small or the diffusivity is sufficiently large.

II. Mixing Efficiency Functionals

We need a number of function spaces for our discussions. We denote by $H^s(\Omega)$ the usual Sobolev space [11] for any $s \in \mathbb{R}$. For $s \geq 0$, $H^s_0(\Omega)$ denotes the completion of $C_0^\infty(\Omega)$ in $H^s(\Omega)$, where $C_0^\infty(\Omega)$ denotes the space of all infinitely differentiable functions on Ω with a compact support in Ω. We set

$$\begin{aligned}
\mathbf{L}^2(\Omega) &= \{L^2(\Omega)\}^n,\\
\mathbf{H}^1_0(\Omega) &= \{H^1_0(\Omega)\}^n,\\
\mathbf{H}^2(\Omega) &= \{H^2(\Omega)\}^n,\\
\mathbf{H}^1_{0,div}(\Omega) &= \{\mathbf{v} \in \mathbf{H}^1_0(\Omega) \ : \ \mathrm{div}(\mathbf{v}) = 0 \text{ in } \Omega\},\\
\mathbf{L}^2_{div}(\Omega) &= \text{ the closure of } \mathbf{H}^1_{0,div}(\Omega) \text{ in } \mathbf{L}^2(\Omega).
\end{aligned}$$

The $\mathbf{L}^2$ norm of a function $\mathbf{f}(\mathbf{x}) \in \mathbf{L}^2(\Omega)$ is denoted by

$$\|\mathbf{f}\| = \left(\int_\Omega |\mathbf{f}(\mathbf{x})|^2 dV\right)^{1/2}.$$

The strain tensor of the velocity $\mathbf{v} = (v_1, v_2, v_3)$ is denoted by

$$\nabla \mathbf{v} = \begin{pmatrix} \frac{\partial v_1}{\partial x_1} & \frac{\partial v_1}{\partial x_2} & \frac{\partial v_1}{\partial x_3} \\ \frac{\partial v_2}{\partial x_1} & \frac{\partial v_2}{\partial x_2} & \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_3}{\partial x_1} & \frac{\partial v_3}{\partial x_2} & \frac{\partial v_3}{\partial x_3} \end{pmatrix}.$$

The mean concentration of $c(\mathbf{x}, t; \mathbf{v})$ is defined by

$$\langle c(t; \mathbf{v})\rangle = \frac{1}{\mathrm{mes}(\Omega)} \int_\Omega c(\mathbf{x}, t; \mathbf{v}) dV.$$

Let $\bar{\mathbf{v}}_1(\mathbf{x}), \cdots, \bar{\mathbf{v}}_m(\mathbf{x}) \in \mathbf{H}^1_{0,div}(\Omega)$ be a set of linearly independent velocities and $u_1, \cdots, u_m \in H^1_0(0,T)$. As in [23], we define the following mixing efficiency functional

$$\begin{aligned}
&J(u_1, \cdots, u_m)\\
= &\int_0^T \Big(\rho\|c(t; \mathbf{v}) - \langle c(t; \mathbf{v})\rangle\|^2 + \alpha\|\mathbf{v}(t)\|^2\\
&+\beta\|\nabla \mathbf{v}(t)\|^2 + \gamma\left\|\frac{\partial \mathbf{v}(t)}{\partial t}\right\|^2\Big) dt\\
&+\mu\|c(T; \mathbf{v}) - \langle c(T; \mathbf{v})\rangle\|^2, \qquad (3)
\end{aligned}$$

where $\mathbf{v}$ is given by (2), $T > 0$ is some desired time, and $\alpha > 0, \beta, \gamma, \mu, \rho \geq 0$ are weight constants. For the physical motivation of this functional, we refer to [23].

The weight constants in (3) play an important role in determining the control strength. For small values of α, β, γ, the functional will result in an optimal solution with a small variance of the scalar, but with big magnitudes of the velocity $\mathbf{v}$, of the strain tensor $\nabla \mathbf{v}$, and of the acceleration $\frac{\partial \mathbf{v}}{\partial t}$. This implies that the smaller the weights, the more turbulent the optimal flow, and then the better the mixing enhancement.

There are different measures for mixing efficiency such as Lagrangian and Eulerian time-averages of a flow [3], the mixing variance coefficient [7], and the Mix-Norm defined by Mathew, Mezić, and L. Petzold [25]. For the convenience of treatment of our optimal control problem, we use the L_2 norm of a scalar variance as the mixing efficiency measurement.

We note that the mean is conserved. In fact, integrating equation (1) over Ω gives

$$\frac{d}{dt}\langle c\rangle = \frac{\kappa}{\mathrm{mes}(\Omega)} \int_\Omega \nabla^2 c \, dV = 0,$$

where we have used the boundary conditions on $\mathbf{v}$ and c. Therefore we can assume zero mean without loss of generality. With the zero-mean assumption, the cost functional reduces to

$$\begin{aligned}
&J(u_1, \cdots, u_m)\\
= &\int_0^T \Big(\rho\|c(t; \mathbf{v})\|^2 + \alpha\|\mathbf{v}(t)\|^2 + \beta\|\nabla \mathbf{v}(t)\|^2\\
&+\gamma\left\|\frac{\partial \mathbf{v}(t)}{\partial t}\right\|^2\Big) dt + \mu\|c(T; \mathbf{v})\|^2. \qquad (4)
\end{aligned}$$

Then the optimal control problem is to minimize J in an admissible weight space $\mathcal{U} = (H^1_0(0,T))^m$

$$J(u_1^*, \cdots, u_m^*) = \min_{(u_1, \cdots, u_m) \in \mathcal{U}} J(u_1, \cdots, u_m). \qquad (5)$$

The minimizer $(u_1^*, \cdots, u_m^*)$ is called an *optimal weight*.

III. Optimality Systems

The existence of an optimal weight was proved in [23] and the optimality system for the optimal weight can be derived as in [23].

Theorem 3.1: Let $\bar{\mathbf{v}}_1(\mathbf{x}), \cdots, \bar{\mathbf{v}}_m(\mathbf{x}) \in \mathbf{H}^1_{0,div}(\Omega) \cap \mathbf{C}(\Omega)$ be a set of linearly independent velocities. If $(u_1^*, \cdots, u_m^*)$ is an optimal weight under the efficiency functional J defined by (4), then it satisfies the following system

$$\frac{\partial c}{\partial t} + (\mathbf{v}^* \cdot \nabla)c = \kappa \nabla^2 c, \qquad (6)$$

$$\frac{\partial g}{\partial t} + (\mathbf{v}^* \cdot \nabla)g = -\kappa \nabla^2 g + \rho c, \qquad (7)$$

$$\begin{aligned}
&\sum_{i=1}^m \frac{d^2 u_i^*}{dt^2} \int_\Omega \gamma \bar{\mathbf{v}}_j \cdot \bar{\mathbf{v}}_i dV = \int_\Omega g \bar{\mathbf{v}}_j \cdot \nabla c \, dV\\
&+\sum_{i=1}^m u_i^* \int_\Omega \bar{\mathbf{v}}_j \cdot (\alpha \bar{\mathbf{v}}_i - \beta \nabla^2 \bar{\mathbf{v}}_i) dV, \qquad (8)
\end{aligned}$$

$$\mathbf{v}^* = \sum_{i=1}^m \bar{\mathbf{v}}_i(\mathbf{x}) u_i^*(t), \quad j = 1, \cdots, m, \qquad (9)$$

$$\frac{\partial c}{\partial \mathbf{n}} = \frac{\partial g}{\partial \mathbf{n}} = 0 \quad \text{on } \partial\Omega, \qquad (10)$$

$$u_i^*(0) = u_i^*(T) = 0, \quad i = 1, \cdots, m, \qquad (11)$$

$$c(\mathbf{x}, 0) = c^0(\mathbf{x}), \quad g(\mathbf{x}, T) = -\mu c(\mathbf{x}, T). \qquad (12)$$

IV. UNIQUENESS OF OPTIMAL WEIGHT

To prove that the optimal weight is unique, it suffices to prove that the optimality system (6)-(12) has a unique solution (note that the system has at least one solution since an optimal weight exists). We can achieve this by developing a number of estimates about the solutions of (6)-(12).

For convenience, we state a well-known estimate [23] about the solution of (1) as follows.

Lemma 4.1: Let $\mathbf{v} \in L^2(0,T;\mathbf{L}^2_{div}(\Omega))$. Then the solution c of (1) satisfies the following estimate

$$\|c(t)\|^2 + 2\kappa \int_0^t \|\nabla c(s)\|^2\, ds = \|c^0\|^2. \qquad (13)$$

Lemma 4.2: Let $\bar{\mathbf{v}}_1(\mathbf{x}), \cdots, \bar{\mathbf{v}}_m(\mathbf{x}) \in \mathbf{H}^1_{0,div}(\Omega) \cap \mathbf{C}(\Omega)$ be a set of linearly independent velocities. Let c_1, c_2; g_1, g_2; $(u_1, \cdots, u_m)$, $(w_1, \cdots, w_m)$ be the solutions of (6)-(12) corresponding the initial conditions c_1^0, c_2^0, respectively. Set $\mathbf{u} = \sum_{i=1}^m \bar{\mathbf{v}}_i u_i, \mathbf{w} = \sum_{i=1}^m \bar{\mathbf{v}}_i w_i$. Then the solutions satisfy the following estimates:

$$\begin{aligned}
&\max_{0\le s\le t} \|c_1(s) - c_2(s)\|^2 \\
\le\ & 2\|c_1^0 - c_2^0\|^2 + \frac{2}{\kappa}\|c_2^0\|^2 \int_0^t \|\mathbf{w}(s) - \mathbf{u}(s)\|^2_\infty ds, \qquad (14)\\
&\kappa \int_0^t \|\nabla(c_1 - c_2)(s)\|^2\, ds \\
\le\ & \|c_1^0 - c_2^0\|^2 + \frac{1}{4\kappa}\|c^0\|^2 \int_0^t \|\mathbf{w}(s) - \mathbf{u}(s)\|^2_\infty ds, \qquad (15)\\
&\max_{0\le s\le t} \|g_1(s) - g_2(s)\|^2 \\
\le\ & \left(\frac{6M^2\rho^2(T-t)}{\kappa^2}\|c_2^0\|^2 + \frac{4\mu^2}{\kappa}\|c_2(T)\|^2\right) \\
&\times \int_0^T \|\mathbf{w}(s) - \mathbf{u}(s)\|^2_\infty ds + 2\mu^2\|c_1(T) - c_2(T)\|^2 \\
&+ \frac{2M^2\rho^2(T-t)}{\kappa}\|c_1^0 - c_2^0\|^2, \qquad (16)
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^T \left\| \frac{d}{ds}(u_1(s), \cdots, u_m(s)) \right.\\
&\left. -\frac{d}{ds}(w_1(s), \cdots, w_m(s)) \right\|^2 ds \\
\le\ & \frac{K\|c^0\|^2}{2\kappa} \int_0^T \|g_1(t) - g_2(t)\|^2 dt \\
&+ K\left(\frac{M^2\rho^2 T\|c^0\|^2}{\kappa} + \mu^2\|c_2(T)\|^2\right) \\
&\times \int_0^T \|\nabla c_1(t) - \nabla c_2(t)\|^2 dt, \qquad (17)
\end{aligned}$$

where M is the Poincaré's constant in the Poincaré's inequality and K is a positive constant.

Proof. Set $e_c = c_1(\mathbf{u}) - c_2(\mathbf{w})$. A direct calculation shows that e_c satisfies

$$\begin{aligned}
\frac{\partial e_c}{\partial t} + (\mathbf{u}\cdot\nabla)e_c &= \kappa\nabla^2 e_c + ((\mathbf{w} - \mathbf{u})\cdot\nabla)c_2(\mathbf{w}), \qquad (18)\\
e_c(\mathbf{x}, 0) &= c_1^0 - c_2^0, \quad \frac{\partial e_c}{\partial \mathbf{n}} = 0 \quad \text{on } \partial\Omega.
\end{aligned}$$

Multiplying (18) by e_c and using the boundary conditions, we obtain the equation

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|e_c(t)\|^2 \\
=\ & -\kappa\|\nabla e_c\|^2 + \int_\Omega e_c((\mathbf{w} - \mathbf{u})\cdot\nabla)c_2(\mathbf{w})dV \\
\le\ & \|\mathbf{w}(\mathbf{t}) - \mathbf{u}(\mathbf{t})\|_\infty \|e_c(t)\| \|\nabla c_2(t;\mathbf{w})\|. \qquad (19)
\end{aligned}$$

Integrating over $[0, t]$ gives

$$\begin{aligned}
&\|e_c(t)\|^2 \le \|c_1^0 - c_2^0\|^2 \\
&+2 \max_{0\le s\le t} \|e_c(s)\| \\
&\times \int_0^t \|\mathbf{w}(s) - \mathbf{u}(s)\|_\infty \|\nabla c_2(s, \mathbf{w})\| ds \\
\le\ & \|c_1^0 - c_2^0\|^2 \\
&+2 \max_{0\le s\le t} \|e_c(s)\| \left(\int_0^t \|\mathbf{w}(s) - \mathbf{u}(s)\|^2_\infty ds\right)^{1/2} \\
&\times \left(\int_0^t \|\nabla c_2(s, \mathbf{w})\|^2 ds\right)^{1/2},
\end{aligned}$$

which implies that

$$\begin{aligned}
&\max_{0\le s\le t} \|e_c(s)\|^2 \le \|c_1^0 - c_2^0\|^2 \\
&+2 \max_{0\le s\le t} \|e_c(s)\| \left(\int_0^t \|\mathbf{w}(s) - \mathbf{u}(s)\|^2_\infty ds\right)^{1/2} \\
&\times \left(\int_0^t \|\nabla c_2(s, \mathbf{w})\|^2 ds\right)^{1/2} \\
\le\ & \|c_1^0 - c_2^0\|^2 \\
&+ \frac{1}{2}\left(\max_{0\le s\le t} \|e_c(s)\|\right)^2 + 2\int_0^t \|\mathbf{w}(s) - \mathbf{u}(s)\|^2_\infty ds \\
&\times \int_0^t \|\nabla c_2(s, \mathbf{w})\|^2 ds.
\end{aligned}$$

It then follows from (13) that

$$\begin{aligned}
&\max_{0\le s\le t} \|e_c(s)\|^2 \le 2\|c_1^0 - c_2^0\|^2 \qquad (20)\\
&+4 \int_0^t \|\mathbf{w}(s) - \mathbf{u}(s)\|^2_\infty ds \int_0^t \|\nabla c(s, \mathbf{w})\|^2 ds \\
\le\ & 2\|c_1^0 - c_2^0\|^2 + \frac{2}{\kappa}\|c_2^0\|^2 \int_0^t \|\mathbf{w}(s) - \mathbf{u}(s)\|^2_\infty ds.
\end{aligned}$$

This proves (14).

Using the first equation of (19), we derive that

$$\begin{aligned}
&\frac{1}{2}\|e_c(t)\|^2 + \kappa\int_0^t \|\nabla e_c(s)\|^2\, ds \\
\leq\ & \|c_1^0 - c_2^0\|^2 \\
&+ \max_{0\leq s\leq t}\|e_c(s)\|\int_0^t \|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty \|\nabla c_2(s,\mathbf{w})\| ds \\
\leq\ & \|c_1^0 - c_2^0\|^2 \\
&+ \max_{0\leq s\leq t}\|e_c(s)\|\left(\int_0^t \|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds\right)^{1/2} \\
&\times\left(\int_0^t \|\nabla c_2(s,\mathbf{w})\|^2 ds\right)^{1/2}.
\end{aligned}$$

This implies that

$$\begin{aligned}
&\frac{1}{2}\max_{0\leq s\leq t}\|e_c(t)\|^2 + \kappa\int_0^t \|\nabla e_c(s)\|^2\, ds \\
\leq\ & \|c_1^0 - c_2^0\|^2 \\
&+ \max_{0\leq s\leq t}\|e_c(s)\|\left(\int_0^t \|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds\right)^{1/2} \\
&\times\left(\int_0^t \|\nabla c_2(s,\mathbf{w})\|^2 ds\right)^{1/2} \\
\leq\ & \|c_1^0 - c_2^0\|^2 \\
&+\frac{1}{2}\max_{0\leq s\leq t}\|e_c(s)\|^2 + \frac{1}{2}\int_0^t \|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds \\
&\times\int_0^t \|\nabla c_2(s,\mathbf{w})\|^2 ds.
\end{aligned}$$

It then follows from (13) that

$$\begin{aligned}
&\kappa\int_0^t \|\nabla e_c(s)\|^2\, ds \leq \|c_1^0 - c_2^0\|^2 \\
&+\frac{1}{4\kappa}\|c^0\|^2\int_0^t \|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds. \qquad (21)
\end{aligned}$$

This proves (15).

Set $e_g = g_1(\mathbf{u}) - g_2(\mathbf{w})$. A direction calculation shows that

$$\begin{aligned}
&\frac{\partial e_g}{\partial t} + (\mathbf{u}\cdot\nabla)e_g = -\kappa\nabla^2 e_g \\
&+((\mathbf{w}-\mathbf{u})\cdot\nabla)g_2(\mathbf{w}) + \rho e_c \quad \text{in } \Omega, \qquad (22) \\
&e_g(\mathbf{x},T) = \mu[c_2(T)-c_1(T)],\ \frac{\partial e_g}{\partial \mathbf{n}} = 0 \quad \text{on } \partial\Omega.
\end{aligned}$$

Multiplying (22) by e_g and using the boundary conditions, we obtain the equation

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|e_g(t)\|^2 = \kappa\|\nabla e_g\|^2 \\
&+\int_\Omega e_g((\mathbf{w}-\mathbf{u})\cdot\nabla)g_2(\mathbf{w})dV + \rho\int_\Omega e_g e_c dV \\
\geq\ & \kappa\|\nabla e_g\|^2 - \|\mathbf{w}(\mathbf{t})-\mathbf{u}(\mathbf{t})\|_\infty\|e_g(t)\|\|\nabla g_2(t;\mathbf{w})\| \\
&-\rho\|e_g(t)\|\|e_c(t)\|. \qquad (23)
\end{aligned}$$

Since $\langle e_g\rangle = 0$, we have the following Poincarés inequality [10], [11]

$$\|e_g(t)\| \leq M\|\nabla e_g(t)\|.$$

where M is a positive constant. Using the Young's inequality, it therefore follows from (23) that

$$\begin{aligned}
&\frac{d}{dt}\|e_g(t)\|^2 \geq 2\kappa\|\nabla e_g\|^2 \\
&-2\|\mathbf{w}(\mathbf{t})-\mathbf{u}(\mathbf{t})\|_\infty\|e_g(t)\|\|\nabla g_2(t;\mathbf{w})\| \\
&-2M\rho\|\nabla e_g(t)\|\|e_c(t)\| \\
\geq\ & 2\kappa\|\nabla e_g\|^2 - 2\|\mathbf{w}(\mathbf{t})-\mathbf{u}(\mathbf{t})\|_\infty\|e_g(t)\|\|\nabla g_2(t;\mathbf{w})\| \\
&-2\kappa\|\nabla e_g(t)\|^2 + \frac{M^2}{2\kappa}\|e_c(t)\|^2 \\
=\ & -2\|\mathbf{w}(\mathbf{t})-\mathbf{u}(\mathbf{t})\|_\infty\|e_g(t)\|\|\nabla g_2(t;\mathbf{w})\| \\
&-\frac{M^2\rho^2}{2\kappa}\|e_c(t)\|^2. \qquad (24)
\end{aligned}$$

Integrating over $[t,T]$ gives

$$\begin{aligned}
&\|e_g(t)\|^2 \leq 2\max_{0\leq s\leq t}\|e_g(s)\| \\
&\times\int_t^T \|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty\|\nabla g_2(s,\mathbf{w})\| ds \\
&+\mu^2\|c_1(T)-c_2(T)\|^2 + \int_t^T \frac{M^2\rho^2}{2\kappa}\|e_c(s)\|^2 ds \\
\leq\ & 2\max_{0\leq s\leq t}\|e_g(s)\|\left(\int_t^T \|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds\right)^{1/2} \\
&\times\left(\int_t^T \|\nabla g_2(s,\mathbf{w})\|^2 ds\right)^{1/2} \\
&+\mu^2\|c_1(T)-c_2(T)\|^2 + \int_t^T \frac{M^2\rho^2}{2\kappa}\|e_c(s)\|^2 ds.
\end{aligned}$$

It therefore follows from (20) that

$$\begin{aligned}
&\max_{0\leq s\leq t}\|e_g(t)\|^2 \leq 2\max_{0\leq s\leq t}\|e_g(s)\| \\
&\times\left(\int_t^T \|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds\right)^{1/2} \\
&\times\left(\int_t^T \|\nabla g_2(s,\mathbf{w})\|^2 ds\right)^{1/2} \\
&+\mu^2\|c_1(T)-c_2(T)\|^2 + \int_t^T \frac{M^2\rho^2}{2\kappa}\|e_c(s)\|^2 ds \\
\leq\ & \frac{1}{2}\max_{0\leq s\leq t}\|e_g(s)\|^2 + 2\int_t^T \|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds \\
&\times\int_t^T \|\nabla g_2(s,\mathbf{w})\|^2 ds \\
&+\mu^2\|c_1(T)-c_2(T)\|^2 + \frac{M^2\rho^2(T-t)}{\kappa}\|c_1^0 - c_2^0\|^2 \\
&+\frac{M^2\rho^2(T-t)}{\kappa^2}\|c_2^0\|^2\int_0^T \|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds, \qquad (25)
\end{aligned}$$

which gives

$$\begin{aligned}
&\max_{0\le s\le t}\|e_g(t)\|^2\\
\le\;& 4\int_t^T\|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds\int_t^T\|\nabla g_2(s,\mathbf{w})\|^2 ds\\
&+2\mu^2\|c_1(T)-c_2(T)\|^2+\frac{2M^2\rho^2(T-t)}{\kappa}\|c_1^0-c_2^0\|^2\\
&+\frac{2M^2\rho^2(T-t)}{\kappa^2}\|c_2^0\|^2\int_0^T\|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds. \quad (26)
\end{aligned}$$

To estimate $\int_t^T\|\nabla g_2(s,\mathbf{w})\|^2 ds$, we multiply (7) (change g to g_2) by g_2 and integrate over $\Omega\times(t,T)$. Using (13) and the Poincaré's inequality, we can readily derive that

$$\begin{aligned}
&\frac{1}{2}\|g_2(t)\|^2+\frac{\kappa}{2}\int_t^T\|\nabla g_2(s,\mathbf{w})\|^2 ds\\
\le\;&\frac{M^2\rho^2(T-t)}{2\kappa}\|c_2^0\|^2+\frac{\mu^2}{2}\|c_2(T)\|^2. \quad (27)
\end{aligned}$$

It therefore follows from (26) that

$$\begin{aligned}
&\max_{0\le s\le t}\|e_g(t)\|^2\\
\le\;&\left(\frac{6M^2\rho^2(T-t)}{\kappa^2}\|c_2^0\|^2+\frac{4\mu^2}{\kappa}\|c_2(T)\|^2\right)\\
&\times\int_0^T\|\mathbf{w}(s)-\mathbf{u}(s)\|_\infty^2 ds\\
&+2\mu^2\|c_1(T)-c_2(T)\|^2\\
&+\frac{2M^2\rho^2(T-t)}{\kappa}\|c_1^0-c_2^0\|^2. \quad (28)
\end{aligned}$$

This proves (16).

Using the well known estimate [11, Chapter 6] on the boundary value problem (8), it follows from (13) and (27) that there exists a constant K, independent of κ, c^0 such that

$$\begin{aligned}
&\int_0^T\left\|\frac{d}{ds}(u_1(s),\cdots,u_m(s))\right.\\
&\left.-\frac{d}{ds}(w_1(s),\cdots,w_m(s))\right\|^2 ds\\
\le\;& K\max_{0\le t\le T}\|g_1(t)-g_2(t)\|^2\int_0^T\|\nabla c_1(t)\|^2 dt\\
&+K\max_{0\le t\le T}\|g_2(t)\|^2\int_0^T\|\nabla c_1(t)-\nabla c_2(t)\|^2 dt\\
\le\;&\frac{K\|c^0\|^2}{2\kappa}\int_0^T\|g_1(t)-g_2(t)\|^2 dt\\
&+K\left(\frac{M^2\rho^2T\|c^0\|^2}{\kappa}+\mu^2\|c_2(T)\|^2\right)\\
&\times\int_0^T\|\nabla c_1(t)-\nabla c_2(t)\|^2 dt. \quad (29)
\end{aligned}$$

This proves (17).

Theorem 4.1: Let $\bar{\mathbf{v}}_1(\mathbf{x}),\cdots,\bar{\mathbf{v}}_m(\mathbf{x})\in\mathbf{H}^1_{0,div}(\Omega)\cap\mathbf{C}(\Omega)$ be a set of linearly independent velocities. If the initial condition c^0 is sufficiently small or the diffusivity κ is sufficiently large, the optimality system (6)-(12) has a unique solution, and then the optimal weight is unique.

Proof. Suppose the solutions of the optimality system (6)-(12) is not unique. Noting that $c_1^0=c_2^0=c^0$, we deduce from (14), (15), (16), (17), and the Poincaré's inequality that there exists a constant K, independent of κ, c^0 such that

$$\begin{aligned}
&\int_0^T\|(u_1(s),\cdots,u_m(s))-(w_1(s),\cdots,w_m(s))\|^2 ds\\
&\le\tfrac{K\|c^0\|^4}{4\kappa^3}(12M^2\rho^2T^2+M^2\rho^2T+17\kappa\mu^2)\\
&\times\int_0^T\|(u_1(s),\cdots,u_m(s))-(w_1(s),\cdots,w_m(s))\|^2 ds.
\end{aligned}$$

If the initial condition c^0 is sufficiently small or the diffusivity κ is sufficiently large such that

$$\frac{K\|c^0\|^4}{4\kappa^3}(12M^2\rho^2T^2+M^2\rho^2T+17\kappa\mu^2)<1,$$

then we must have

$$\int_0^T\|(u_1(s),\cdots,u_m(s))-(w_1(s),\cdots,w_m(s))\|^2 ds=0.$$

This is a contradiction.

REFERENCES

[1] O. M. Aamo, M. Krstic, T. R. Bewley, Control of mixing by boundary feedback in 2D channel flow, Automatica, 39 (2003), pp. 1597-1606.

[2] D. D'Alessandro, M. Dahleh and I. Mezic, Control of mixing in fluid flow: a maximum entropy approach, IEEE Trans. Automat. Control, 44 (1999), pp. 1852–1863.

[3] D. D'Alessandro, I. Mezic and M. Dahleh, Statistical properties of controlled fluid flows with applications to control of mixing, Systems Control Lett., 45 (2002), pp. 249–256.

[4] T. M. Antonsen, Z. Fan, E. Ott, and E. Garcia-Lopez, The role of passive scalars in the determination of power spectra of passive scalars, Phys. Fluids, 8(1996), pp. 3094-3104.

[5] A. Balogh, O. M. Aamo and M. Krstic, Optimal mixing enhancement in 3D pipe flow, IEEE Trans. on Control Systems Technology, 13(2005), pp. 27-41.

[6] Alain Bensoussan, Giuseppe Da Prato, Michel C. Delfour, Sanjoy K. Mitter, Representation and Control of Infinite Dimensional Systems. Birkhauser, Boston, 2006.

[7] R. Bottausci, I. Mezić, C. D. Meinhart, and C. Cardonne, Mixing in the shear superposition micromixer: three-dimensional analysis, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 362 (2004), pp. 1001–1018.

[8] B. V. N. Charyulu, J. Kurian, P. Venugopalan, and V. Sriramulu, Experimental study on mixing enhancement in two dimensional supersonic flow, Experiments in Fluids, 24 (1998), pp. 340-346.

[9] R. F. Curtain and H. Zwart, An introduction to Infinite-dimensional Linear Systems Theory, Springer-Verlag, New York, 1995.

[10] R. Dautray and J.L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology, Vol.2, Functional and Variational Methods*, Springer-Verlag, Berlin, 1990.

[11] L. C. Evans, Partial Differential Equations, American Mathematical Society, Providence, Rhode Island, 1998.

[12] A. Fannjiang and G. Papanicolaou, Convection enhanced diffusion for periodic flows, SIAM J. Appl. Math., 54 (1994), pp. 333-408.

[13] D. R. Fereday, P. H. Haynes, and A. Wonhas, Scalar variance decay in chaotic advection and the Batchelor–regime turbulence, Phys. Rev. E, 65(2002), pp. 035301(R).

[14] M. Giona, S. Cerbelli, and V.Vitacolonna, Universality and imaginary potentials in advection-diffusion equations in closed flows, J. Fluid Mech., 513 (2004), pp. 221-237.

[15] M. Giona, V. Vitacolonna, S. Cerbelli, and A. Adrover, Advection diffusion in nonchaotic closed flows: Non-Hermitian operators, universality, and localization, Phys. Rev. E, 70 (2004), pp. 046224.

[16] D. Greenblatt and I. J. Wygnanski, The control of flow separation by periodic excitation, Progress in Aerospace Sci., 36(2000), pp. 487.

[17] P. H. Haynes and J. Vanneste, What controls the decay of passive scalars in smooth flows? Phys. Fluids, 17 (2005), pp. 097103.

[18] I. Lasiecka and R. Triggiani, Control Theory for Partial Differential Equations, Volume 1: Abstract Parabolic Systems, Cambridge University Press, 2000.

[19] X. Li and J. Yong, Optimal Control Theory for Infinite Dimensional Systems, Birkhäuser, Boston, 1995.

[20] J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations. Springer-Verlag, Berlin, 1971.

[21] W. Liu and G. Haller, Strange eigenmodes and decay of variance in the mixing of diffusive tracers, Phys. D 188 (2004), pp. 1-39.

[22] W. Liu, Does a fast mixer really exist? Phys. Rev. E, 72 (2005), pp. 016312.

[23] W. Liu, Mixing enhancement by optimal flow advection, SIAM J. Control Optim., 47, no. 2, (2008), 624-638.

[24] G. Mathew, I. Mezić, S. Grivopoulos, U. Vaidya, and L. Petzold, Optimal control of mixing in Stokes fluid flows, J. Fluid Mech., 50, (2007), pp. 261281.

[25] G. Mathew, I. Mezić, and L. Petzold, A multiscale measure for mixing, Phys. D, 211 (2005), pp. 23–46.

[26] B. R. Noack, I. Mezić, G. Tadmor, and A. Banaszuk, Optimal mixing in recirculation zones, Phys. Fluids, 16 (2004), 867–888.

[27] A. Pikovsky and O. Popovych, Persistent patterns in deterministic mixing flows, Europhys. Lett., 61 (2003), pp. 625-631.

[28] I. H. Shames, Mechanics of Fluids, Fourth Edition, McGraw-Hill, New York, 2003

[29] T.H. Solomon and J.P. Gollub, Chaotic particle transport in time-dependent Rayleigh-Bénard convection, Phys. Rev. A 38 (1988), pp. 6280-6286.

[30] Yue-Kin Tsang, Thomas M. Antonsen, Jr. and Edward Ott, Exponential decay of chaotically advected passive scalars in the zero diffusivity limit, Phys. Rev. E, 71 (2005), pp. 066301.

[31] Y. Wang, G. Haller, A. Banaszuk, and G. Tadmor, Closed-loop Lagrangian separation control in a bluff body shear flow model, Phys. Fluids, 15 (2003), pp. 2251.

[32] W.M. Yang, S.K. Chou, C. Shu, Z. W. Li and H. Xue, Combustion in micro-cylindrical combustors with and without a backward facing step, Applied Thermal Engineering, 22 (2002), pp. 1777-1787.

[33] C.C. Yuan, M. Krstic, and T. Bewley, Active control of jet mixing, IEE Proceedings: Control Theory and Applications, 151 (2004), pp. 763-772.

[34] E. Zeidler, Nonlinear functional analysis and its applications III, Variational Methods and OPtimization, Springer-Verlag, New York, 1985.

Plan-for-Gov[IT] - Planning for Governance of IT Method: use of the Techniques of "Text Retrieval" for mapping the expected support needs from IT Area to serve of the Corporation's Core-Business expectations

Altino José Mentzingen de MORAES
altino.moraes@gmail.com
Graduate School, IBMEC College (http://www.ibmec.br/df) - Brasília, DF 70012-200, Brazil

ABSTRACT

The IT - Information Technology, in accordance with the philosophy of the IT Governance (and also as defined by Authors listed as follows) requires its integration to the process of Strategic Planning of its Corporation, with the intention to align its actions with the Core-Business aiming at to reach the expected results by the IT Area.

The question is how the IT can, under a methodological and direct way, to know how to interpret the expectations expressed by the Strategic Planning (a component of the Corporative Governance), in actions that are addressed to its Area in a practical manner and with an adequate tool kit related to the Frameworks (Models) focused to the implementation of the IT Governance, for posterior creation of the necessary Effectiveness Indicators for monitoring about success level of the actions of IT in alignment with the Business.

The result of this work is the proposal of the Text Retrieval and its subsequent validation (as a plausible resource for actual use to try to help the Governance of IT in its primary task of assisting the Corporation Core-Business), which was named as Plan-for-Gov[IT] - Planning for Governance of IT Method, which can be automated by the use of resources of "word finding" in Word Processors or in another software products with also this purpose.

Keywords: IT Governance; IT Management; Strategic Planning; Frameworks (Models) of IT Governance; Text Retrieval, Keyword Search, Survey.

1. INTRODUCTION

The Rau's [27] approach defines Governance as a way for that Organizations can guarantee that Strategic Goals are defined, monitored and achieved and, according to this same Author, when Governance is applied to Information Technology (IT) this is called as IT Governance and it means how Top Managers interact and communicate with IT Leaders to ensure that investments in this sector contribute to the achievement - in an efficient and effective way - the targets set as evaluation criteria for Business Strategy.

In reference to the aspects between Corporate Strategy and the adequacy of IT to meet it, Grosvenor & Brown [13, p.5] argue that "...the biggest challenge for IT Management is to maintain alignment between the connection points of the Business with its own IT...", in other words, to focus its work operations in practical results in accordance to the planning of the Corporation.

Under the vision of Bergstein & Sviokla [4], IT is the essence of an Organization and the analysis of how the Board of Directors manages such an asset, perhaps the most critical, may be useful to consider the type of administrative attitude that CIO - Chiefs of Information Office should have before the Board, in an Organization that intensively uses IT in its Business.

The necessity of alignment with the objectives of the Organization is described by these Authors when they stated that IT can become viable as an integrant part of the productive machine, as much in its own activity as well as in its interaction with the Business, mainly serving the determinations and achievement of the objectives related with the Corporative Strategic Planning (which are based on quantification metrics to monitor results). This is due to the crucial importance for conducting Businesses, achieved by IT, in a scenario of strong competition faced by Organizations today.

2. PROBLEM STATEMENT

As a paradigm of IT Governance, presented as a "Puzzle" to be built and also to be resolved, Altino Moraes [20] establishes the following questions (graphically exposed in Annex A in the Appendix 1 with green color and italic format font for English terms/words translation).

According to this argument, by supplementing also that of other Authors, the basis work of the IT Governance is the Indicators Management (which, can be, to control Performance, Results, Quality and Effectiveness).

As an example, about the theory of Construction of Indicators, according Goethert & Hayes [11] on the goal of proposing a result measurement via Indicators, the main question is not what metrics should be used, but what information you want to retrieve and what data you need to know. To answer this questioning, the Authors conduct Workshops with Executives of Organizations through the GQ(I)M Methodology - Goal, Question, Indicator and Measurement.

In Annex A in the Appendix 2 is shown graphically the difference between the application of Types of indicators to assess compliance with targets set for the results that may add value to the business.

For answer the question rise in the Topic "**ABSTRACT**", this work has the intention of assisting this process proposing (and using the "Text Retrieval" Techniques) a relationship of Keywords, that can be retrieved in resultant texts of the planning activity, with the identified Fields of Action (that are also defined in this work) of the Frameworks (Models), what, would assist the IT Governance to understand how to put in practice its activities (beyond the activities of monitoring its own Effectiveness Indicators) in order to support the Core-Business' Organization needs. This proposal has the name of Plan-for-Gov[IT] - Planning for Governance of IT Method.

3. RESEARCH METHODOLOGY

After the exposition of the concepts of IT Governance, of Construction of Indicators and Techniques of Text Retrieval, was held a Bibliographical Research based in the literature of the Frameworks (Models) that can be applied by the IT Governance as support for its Management, and then selected the most disclosed (but others are also mentioned) for each one of the ten (10) Fields of Action (which are also defined in this article as one of its results) and that could be identified this way by the direction given to its use by the Institutes, Organizations and Associations that have created and still maintain them.

In the sequence, a Survey was conducted (during the entire year of 2011) with 320 (three hundred and twenty) Graduate Students - in the IT Area - in order to evaluate this work statements and its proposal, and also, to validate the applicability (or not) of the Keywords proposed as initial idea (between Verbs and Nouns) to implement the Plan-for-Gov [IT] - Planning for Governance of IT Method in their Companies.

4. LITERATURE REVISION

4.1. Text Retrieval

The Text Retrieval Technique describes the use of Keywords (also called "reserved-words") to search for terms in free-texts that can, by consolidation, providing a unique understanding to the context in which these are.

The most used technique, by IT to identify classes and elements of a database, is proposed by Russell Abbott [1] in his article "Program Design by Informal English Descriptions" published in 1983 in the "Communication of the ACM" periodical. This paper argues that the objects of the database can be identified by parsing the grammatical text that describes the problem. Another Literature sample is the Attar article [3].

This technique will be applied, in the proposed Model presented in this work, to link the documents generated by the Strategic Planning with the actions that IT must implement to meet them. This will be done by the Proposal of the Keywords that should make this link of liaison with the Frameworks (Models) for IT Governance. The connection point will be the Fields of Action of IT those will be identified in the Frameworks (Models).

4.2. Frameworks (Models) for IT Governance

To put into practice their control activities of the IT environment, in order to exercise its governance, IT uses Frameworks (Models) that point to procedures and propose controls in various fields of its actions.

Following are presented (in Alphabetical Order) the Frameworks (Models) selected, from the various among those which were studied, and which were chosen among the most publicized and recognized by the IT Community. Others, also researched, are mentioned and referenced in the Bibliography of this work.

During this analysis, were identified in which fields of IT these would be applied, since have been identified 10 (ten) Fields of Action, according to direction given to its use by Institutes, Organizations and Associations that have created and still maintain them. These 10 (ten) Fields will be defined in the Topic "**4.3. Definition of the Fields of Action of IT**".

4.2.1. CMMI©

This Framework (Model) was created and is still being maintained by SEI - Software Engineering Institute [29]. By the direction given to its use by this Entity, this Framework (Model) can be defined as DEVELOPMENT for the Field of Action of IT Governance. The Figure 1, presented after the Topic "**8. REFERENCES**", shows its architecture. Others options, for this same Field of Action, are Brazilian MPS br [28] and ISO/IEC 15504 [15].

4.2.2. CobiT©

This Framework (Model) was created ITGI - IT Governance Institute [16] and is still being maintained by ISACA [14]. By the direction given to its use by this Entity, this Framework (Model) can be defined as MANAGEMENT for the Field of Action of IT Governance. It is referenced by Gartner Group in its "IT Governance Report" [10]. The Figure 2, presented after the Topic "**8. REFERENCES**", shows its architecture. Other option, for this same Field of Action, is TOGAF [31].

4.2.3. ISO 17799©

This Framework (Model) was created and is still being maintained by British Standards Institute [6] under BS15000 scope. By the direction given to its use by this Entity, this Framework (Model) can be defined as SECURITY for the Field of Action of IT Governance. Other option, for this same Field of Action, is NIST 800-14 [22].

4.2.4. IT BSC© - Balance Score Card

This Framework (Model) was created and is still being maintained by Grembergen [12]. This work is based in another original work written by Kaplan & Norton [17], but also, including concerns about IT controls. By the direction given to its use by this Author, this Framework (Model) can be defined as PLANNING for the Field of Action of IT Governance. The Figure 3, presented after the Topic "**8. REFERENCES**", shows its architecture.

4.2.5. ITIL©

This Framework (Model) was created and is still being maintained by ITGI - IT Governance Institute a branch of Office of Government Commerce [23]. It has 7 (seven) Books. By the direction given to its use by this Entity, this Framework (Model) can be defined as PRODUCTION for the Field of Action of IT Governance. The Figure 4, presented after the Topic "**8. REFERENCES**", shows its architecture.

4.2.6. PMBoK©

This Framework (Model) was created and is still being maintained by PMI [26]. By the direction given to its use by this Entity, this Framework (Model) can be defined as DESIGN for the Field of Action of IT Governance. Other option, for this same Field of Action, is PRINCE II [25].

4.2.7. SAS 70©

This Framework (Model) was created and is still being maintained by AICPA [2]. By the direction given to its use by this Entity, this Framework (Model) can be defined as AUDITING for the Field of Action of IT Governance. Other option, for this same Field of Action, is COCOMO [7].

4.2.8. Six Sigma©

This Framework (Model) was created by Motorola [21], based in Deming [9] studies, and was later improved by 3 (three) Authors, which are, Pande, Neuman and Cavanagh [24]. By the direction given to its use by these Authors, this Framework (Model) can be defined as QUALITY for the Field of Action of IT Governance.

4.2.9. SOX

This Framework (Model) was created by 2 (two) Senators from US Republican and Democrat Party, is also known as Sarbanes-Oxley or Sarbox and was normalized by COSO [8]. By the direction given to its use by these Authors, this Framework (Model) can be defined as COMPLIANCE for the Field of Action of IT Governance. The Figure 5, presented after the Topic "**8. REFERENCES**", shows its architecture. Others options, for this same Field of Action, are Basel II [5] and Solvência II [30].

4.2.10. TMMI©

This Framework (Model) was created and is still being maintained by 3 (three) Authors, which are, Liebman, Paes and Menezes [19]. By the direction given to its use by these Authors, this Framework (Model) can be defined as TESTING for the Field of Action of IT Governance. The Figure 6, presented after the Topic "**8. REFERENCES**", shows its architecture. Other option, for this same Field of Action, is Krause [18].

4.3. Definition of the Fields of Action of IT

The 10 (ten) Fields of Action identified according to direction given to its use by Institutes, Organizations and Associations that have created and still maintain them, were (in Alphabetical Order): AUDITING, COMPLIANCE, DESIGN, DEVELOPMENT, MANAGEMENT, PLANNING, PRODUCTION, QUALITY, SECURITY e TESTING.

In Table 1 the acronyms, for the 10 (ten) Fields of Action listed in the preceding paragraph, were also defined (along with colors that are added below in order to better understanding) to facilitate the classification of them under the Plan-for-Gov [IT] - Planning for Governance of IT Method proposed by this work.

Table 1 - Definition of the Fields of Action of IT

AD - AUDITING	
CN - COMPLIANCE	
DG - DESIGN	
DV - DEVELOPMENT	
MG - MANAGEMENT	
PL - PLANNING	
PR - PRODUCTION	
QL - QUALITY	
SC - SECURITY	
TS - TESTING	

5. PLAN-FOR-GOV[IT] PRESENTATION

The Plan-for-Gov[IT] - Planning for Governance of IT Method proposes the use of the "Text Retrieval" Techniques, in resultant texts of the planning activity, for the Keywords selection that can drive activities for the 10 (ten) Fields of Action identified in the Frameworks (Models), what means to say that, the 10 (ten) Fields of Action are the link (and are in the middle of the connection) among the Keywords and the Frameworks (Models) defined to be applied.

After this step, already with the identified Field of Action, the Framework (Model) that could be better adjust for support this activity, would be placed operational (with the creation, and posterior monitoring, of KPI - Key Performance Indicators) with the objective to help the auditing of the IT Governance tasks.

In Annex A in the Appendix 3, there is a Figure that presents graphically the relationship among the Keywords, the 10 (ten) Fields of Action of the Frameworks (Models) and also the 10 (ten) Frameworks (Models) select in this study.

6. SURVEY EXECUTION

This universe of 320 (three hundred and twenty) Graduate Students - in the IT Area - was asked whether the below Words (divided in Verbs and Nouns) could point to one (or more than one) of the Fields of Action that were identified in the Frameworks (Models) analysis done in the selected Frameworks (Models) in the Topic "LITERATURE REVISION" of this paper (those were also shown to them).

Verbs: To Administer; To Administrate; To Auditing; To Build; To Componentize; To Control; To Cost; To Define; To Design; To Development; To Enhance; To Estimate; To Evaluate; To Implement; To Improve; To Institute; To Manage; To Measure; To Normalize; To Plan; To Processing; To Product; To Prospect; To Raise; To Record; To Rule; To Run; To Schedule; To Support; To Systematize; To Test

Nouns: Adjusting; Administration; Application; Auditing; Component; Componentization; Compliance; Contingency; Control; Coordination; Cost; Deadline; Definition; Design; Development; Directives; Enhancement; Evaluation; Expectation; Goals; Indicators; Legislation; Management; Method; Methodology; Metrics; Mission; Performance; Planning; Procedure; Proceduring; Processes; Production; Productivity; Prospection; Quality; Recording; Rules; Regulation; Result; Schedule; Scope; Security; Solution; Strategic; Support; Survey; Systems; Tasks; Time; Tests; Values.

After the tabulation, the results are shown in a table in Annex A in the Appendix 4, where the Verbs and Nouns chosen by the Graduate Students in the Survey done are presented (Keywords). Also, in this Table, were included a new column to define which Framework (Model), suggested by this paper, can be used to implement the IT Governance auditing and controlling activities.

To put in practice the Plan-for-Gov[IT] - Planning for Governance of IT Method, this work proposes the Formulary in Annex B. Using this one and also the Keywords, which can be revised in future works, and Frameworks (Models) referenced in the Table above (in Annex A in the Appendix 4), it is possible to record the job done (in the Strategic Documents connected to an already existent Project or driving to an opening of new one), to support the tasks execution and future revisions about the procedures.

7. CONCLUSIONS

By the supervised manner with that this work was built (assembly and survey), it is possible to conclude that the proposal presented can be useful, by the more methodological way of treatment regarding to the integration of the IT Governance with the processes of Strategic Planning, assisting a first and initial approach, for posterior development of the tasks of this Area.

Despite of this Proposal points to some initial Frameworks (Models) suggestions, the application of the Plan-for-Gov[IT] - Planning for Governance of IT Method could be kept even that others Frameworks (Models) are defined or exist more than 1 (one) for each Field of Action, by the reason of the connection with the planning directives are done by these Fields of Action and its Keywords (which, in this Proposal, are kept constants).

8. REFERENCES

[1]ABBOTT, Russell J. **Program design by informal English descriptions.** *Communications of the ACM*. Volume 26, Issue 11, 1983.

[2]AICPA. **SAS 70**. 2007, available in: http://www.aicpa.org/, accessed in: 07/02/22.

[3]ATTAR, R.; FRAENKEL, **A. S. Local Feedback in Full-Text Retrieval Systems.** Journal of the ACM (JACM), Volume 24 (3) Association for Computing Machinery **-** Jul 1, 1977.

[4]BERGSTEIN, M.; SVIOKLA, J. **Board Oversight**. *Optimize*, ABI/INFORM Global, p.74, 6 p., Jan 2005.

[5]Bis Organization. **Basel II (Accord),** 2006, available in: http://www.bis.org/publ/bcbsca.htm/, accessed in: 10/09/05.

[6]British Standards Institute. **BS15000**. 2006, available in: http://www.bsi-global.com/, accessed in: 08/03/15.

[7]COCOMO. **Constructive Cost Model II**. 2006, available in: http://sunset.usc.edu/csse/research/COCOMOII/cocomo_main.html, accessed in: 09/03/11.

[8]COSO. **COSO Enterprise Risk Management (ERM) Framework**. 2006, available in: http://www.coso.org/, accessed in: 09/04/04.

[9]DEMING, W. E. **Qualidade: a revolução da administração** . São Paulo: Marques Saraiva, 1990.

[10]GARTNER, G. **IT Governance Report**. Publication Date: 26 May 2004/ID Number: TG-22-6285, p.3-5, Gartner, Inc, 2004.

[11]GOETHERT, Wolfhart; HAYES, Will. **Experiences in Implementing Measurement Programs**. Technical Note CMU/SEI-2001-TN-026, Software Engineering Measurement and Analysis Initiative, SEI - Software Engineering Institute, November 2001.

[12]GREMBERGEN, W. V. **Strategies for Information Technology Governance**. Hershey, PA: Idea Group Publishing, 406 p., 2004.

[13]GROSVENOR, J.; BROWN, P. **Presenting A Unified Face**. *Optimize*, Massachusetts: p. 35, 5 p., 2005.

[14]ISACA. **CobiT®- Control Objectives for Information and Related Technology**. 2005, available in: http://www.isaca.org/, accessed in: 06/12/08.

[15]ISO/IEC 15504 (SPICE) INTERNATIONAL ORGANIZATION FOR STANDARTZATION. **ISO/IEC TR 15504: Information technology - Software Process Assessment**, published as a set with 9 documents: TR 15504-1 to TR 15504-9, ISO/IEC Technical Report, 1998.

[16]IT Governance Institute. **CobiT® - Control Objectives for Information and Related Technology**. Rolling Meadows USA: Board Briefing on IT Governance, 4rd Edition, 2006.

[17]KAPLAN, R. S.; NORTON, D. P. **Organização orientada para a Estratégia: como as empresas que adotaram o BSC - *Balanced Scorecard* prosperam no novo ambiente de negócios**. Rio de Janeiro: Editora Campus, 2001.

[18]KRAUSE, H. Mitchel. **A Maturity Model for Automated Software Testing**. *Medical Device & Diagnostic Industry Magazine*, MDDI Article Index, 1994.

[19]LIEBMAN, A., PAES, F., MENEZES, H.. **Assessment of the Integration Test Execution Team According to Level II of the Testing Maturity Model - TMM**. 2006.

[20]MORAES, Altino J.M.. **Proposta do M2A3-GovTI - Modelo de Maturidade para Análise do Alinhamento das Atividades relacionadas à Governança da TI em conformidade com as expectativas de resultado planejadas pela Organização**. Production Engineering Doctoral Thesis. PPGEP - Engenharia de Produção da Universidade Paulista, 406 p., São Paulo, 2011.

[21]MOTOROLA University. **Six Sigma - An Introduction to Process Improvement Teams**. 2006, available in: http://www.motorola.com/motorolauniversity.jsp/, accessed in: 07/01/31.

[22]NIST-Federal Information Processing Standards. **NIST 800-14**, FIPS, 2006.

[23]Office of Government Commerce. **ITIL®**. 2006, available in http://www.ogc.gov.uk/, accessed in: 06/03/20.

[24]PANDE, P. S.; NEUMAN R. P.; CAVANAGH R. R. **The Six Sigma Way Team Fieldbook: An Implementation Guide for Process Improvement Teams**. McGraw-Hill, 2002.

[25]PRINCE. **Prince2**. 2002, available in: http://www.prince2.org.uk/, accessed in: 10/02/14.

[26]Project Management Institute. **PMBok®- A Guide to Project Management Body of Knowledge**. 2005, available in: http://www.pmi.org/, accessed in: 07/09/06.

[27]RAU, K. G. **Effective Governance of IT: Design Objectives, Roles, and Relationships**. *Information Systems Management*, Boston, v.21, n.4; p.35, 8 p., 2004.

[28]SOFTEX. **MR MPS br: Modelo de Referência para Melhoria do Processo de Software Brasileiro**. 2007, available in: http://www.softex.br/mpsbr/_guias/default.asp/, accessed in: 10/02/22.

[29]Software Engineering Institute. **CMMI®**. 2003, available in: http://www.sei.cmu.edu/, accessed in: 06/08/17.

[30]Solvência. **Solvência II**, 2006, available in: www.apseguradores.pt/solvencia/, accessed in: 11/09/03.

[31]TOGAF - The Open Group Architecture Framework. **TOGAF 8.1**, 2006, available in: http://www.opengroup.org/togaf/, accessed in: 10/11/10.

Figure 1 - Capability Maturity Model for Software Integrate [29]

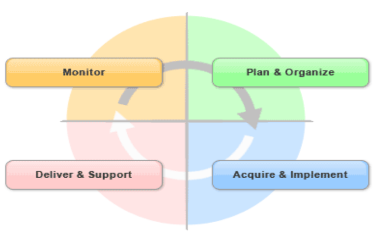


Figure 2 - Control Objectives for Information and Related Technology [16]

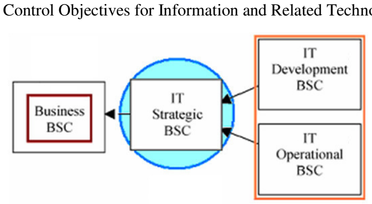


Figure 3 - Grembergen's IT BSC - Balance Score Card [12]

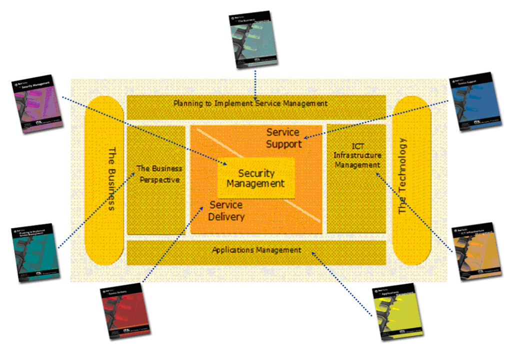


Figure 4 - ITIL: IT Infrastructure Library (adapted from [23])

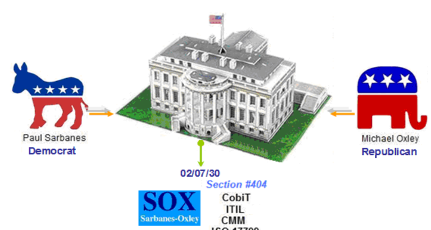


Figure 5 - Sarbanes-Oxley Law [8]

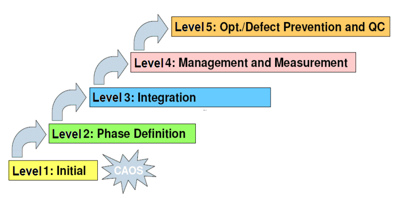
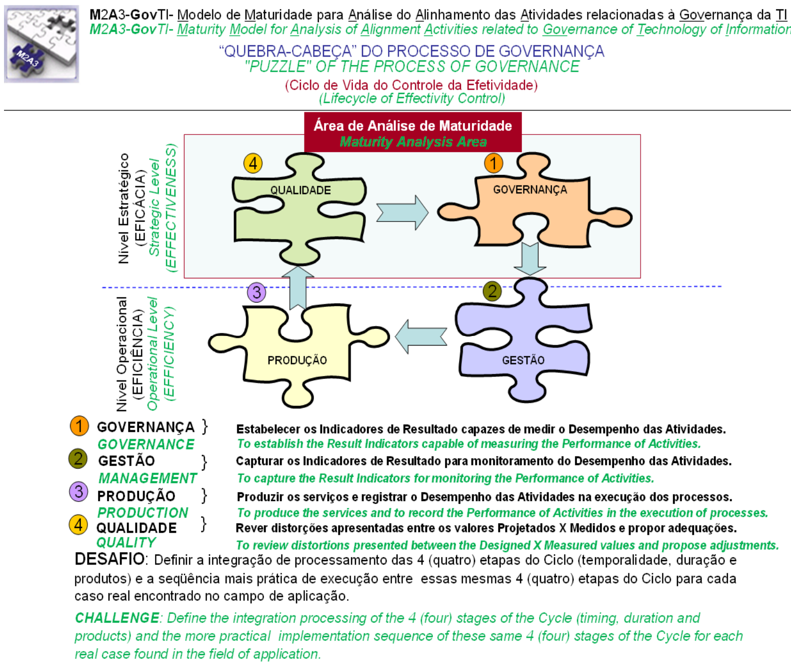
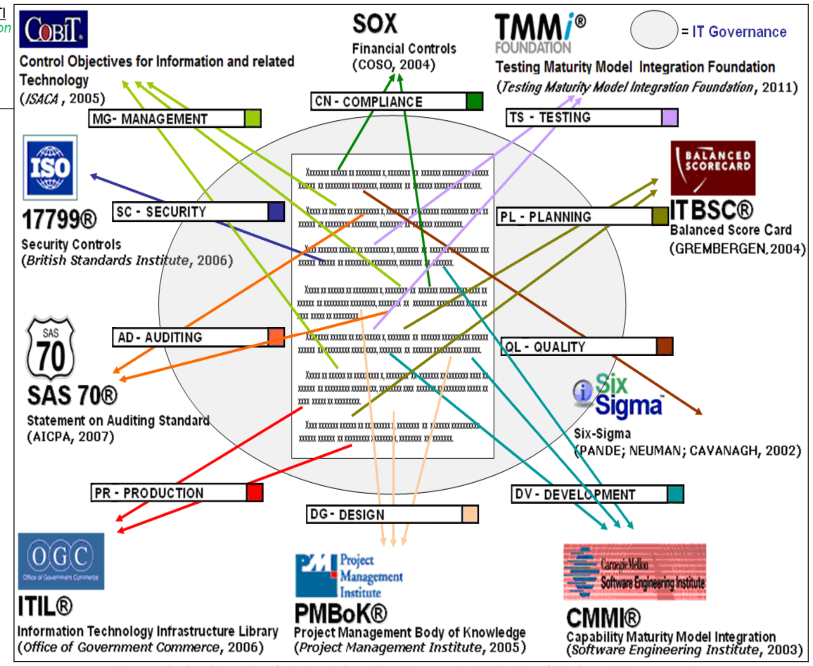


Figure 6 - TMM - Test Maturity Model (adapted from [19])

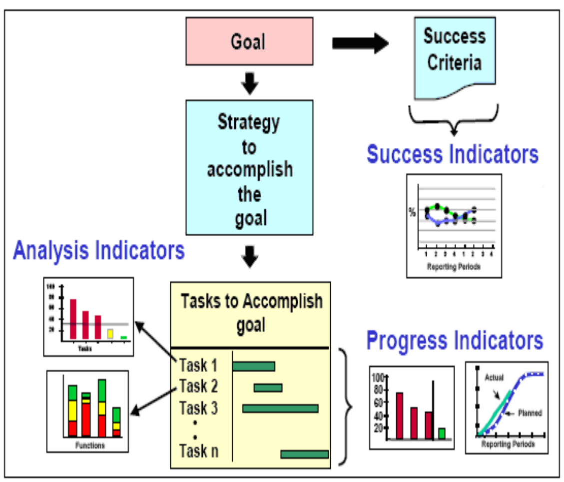
Annex A



Appendix 1 - IT Governance Puzzle of M2A3-GovTI Method [20]

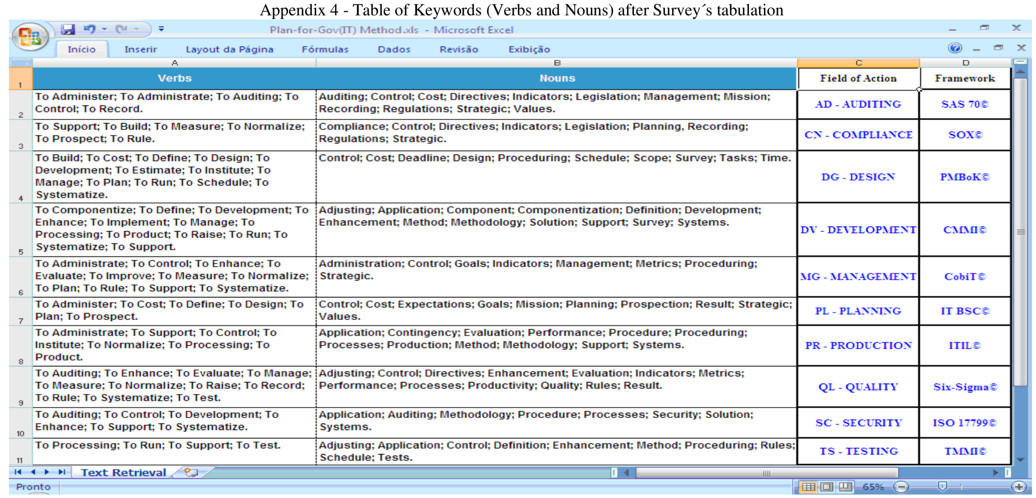


Appendix 3 - Graphic of the Relationship: Keywords X Fields of Action X Frameworks



Appendix 2 - Indicators Application Process Flow (adapted from [11])

Appendix 4 - Table of Keywords (Verbs and Nouns) after Survey´s tabulation



Verbs	Nouns	Field of Action	Framework
To Administer; To Administrate; To Auditing; To Control; To Record.	Auditing; Control; Cost; Directives; Indicators; Legislation; Management; Mission; Recording; Regulations; Strategic; Values.	AD - AUDITING	SAS 70©
To Support; To Build; To Measure; To Normalize; To Prospect; To Rule.	Compliance; Control; Directives; Indicators; Legislation; Planning, Recording; Regulations; Strategic.	CN - COMPLIANCE	SOX©
To Build; To Cost; To Define; To Design; To Development; To Estimate; To Institute; To Manage; To Plan; To Run; To Schedule; To Systematize.	Control; Cost; Deadline; Design; Proceduring; Schedule; Scope; Survey; Tasks; Time.	DG - DESIGN	PMBoK©
To Componentize; To Define; To Development; To Enhance; To Implement; To Manage; To Processing; To Product; To Raise; To Run; To Systematize; To Support.	Adjusting; Application; Component; Componentization; Definition; Development; Enhancement; Method; Methodology; Solution; Support; Survey; Systems.	DV - DEVELOPMENT	CMMI©
To Administrate; To Control; To Enhance; To Evaluate; To Improve; To Measure; To Normalize; To Plan; To Rule; To Support; To Systematize.	Administration; Control; Goals; Indicators; Management; Metrics; Proceduring; Strategic.	MG - MANAGEMENT	CobiT©
To Administer; To Cost; To Define; To Design; To Plan; To Prospect.	Control; Cost; Expectations; Goals; Mission; Planning; Prospection; Result; Strategic; Values.	PL - PLANNING	IT BSC©
To Administrate; To Support; To Control; To Institute; To Normalize; To Processing; To Product.	Application; Contingency; Evaluation; Performance; Procedure; Proceduring; Processes; Production; Method; Methodology; Support; Systems.	PR - PRODUCTION	ITIL©
To Auditing; To Enhance; To Evaluate; To Manage; To Measure; To Normalize; To Raise; To Record; To Rule; To Systematize; To Test.	Adjusting; Control; Directives; Enhancement; Evaluation; Indicators; Metrics; Performance; Processes; Productivity; Quality; Rules; Result.	QL - QUALITY	Six-Sigma©
To Auditing; To Control; To Development; To Enhance; To Support; To Systematize.	Application; Auditing; Methodology; Procedure; Processes; Security; Solution; Systems.	SC - SECURITY	ISO 17799©
To Processing; To Run; To Support; To Test.	Adjusting; Application; Control; Definition; Enhancement; Method; Proceduring; Rules; Schedule; Tests.	TS - TESTING	TMMI©

Annex B

icEME 2012

International Conference on Engineering and Meta-Engineering

March25th-28th,2012-Orlando,FL,USA

Plan-For-Gov[IT]

INSTRUCTIONS: In the text created in the Strategic Planning reports prepared by the Corporation, mark which of the following Keywords (Verbs and Nouns) were found and will define the IT Area of Action to be considered.

Source-Document: Target-Project:

VERBS		AD	CN	DG	DV	MG	PL	PR	QL	SC	TS
To Administer	01										
To Administrate	02										
To Auditing	03										
To Build	04										
To Componentize	05										
To Control	06										
To Cost	07										
To Define	08										
To Design	09										
To Development	10										
To Enhance	11										
To Estimate	12										
To Evaluate	13										
To Implement	14										
To Improve	15										
To Institute	16										
To Manage	17										
To Measure	18										
To Normalize	19										
To Plan	20										
To Processing	21										
To Product	22										
To Prospect	23										
To Raise	24										
To Record	25										
To Rule	26										
To Run	27										
To Schedule	28										
To Support	29										
To Systematize	30										
To Test	31										

Field of Action	Framework
AD - AUDITING	SAS 70
CN - COMPLIANCE	SOX
DG - DESIGN	PMBoK®
DV - DEVELOPMENT	CMMI®
MG - MANAGEMENT	CobiT®
PL - PLANNING	IT BSC®
PR - PRODUCTION	ITIL®
QL - QUALITY	Six Sigma®
SC - SECURITY	ISO 17799®
TS - TESTING	TMMi

NOUNS		AD	CN	DG	DV	MG	PL	PR	QL	SC	TS
Adjusting	01										
Administration	02										
Application	03										
Auditing	04										
Component	05										
Componentization	06										
Compliance	07										
Contingency	08										
Control	09										
Coordination	10										
Cost	11										
Deadline	12										
Definition	13										
Design	14										
Development	15										
Directives	16										
Enhancement	17										
Evaluation	18										
Expectations	19										
Goals	20										
Indicators	21										
Legislation	22										
Management	23										
Method	24										
Methodology	25										
Metrics	26										
Mission	27										
Performance	28										
Planning	29										
Procedure	30										
Proceduring	31										
Processes	32										
Production	33										
Productivity	34										
Prospection	35										
Quality	36										
Recording	37										
Rules	38										
Regulation	39										
Result	40										
Schedule	41										
Scope	42										
Security	43										
Solution	44										
Strategic	45										
Support	46										
Survey	47										
Systems	48										
Tasks	49										
Time	50										
Tests	51										
Values	52										

Remarks:

altino.moraes@gmail.com

Plan-for-Gov[IT] Formulary Proposal

Physical feasibility study system for projects with a view to the decision making process

Renata Maria Abrantes Baracho
School of Information Science
Federal University of Minas Gerais
Belo Horizonte, Brazil
renatabaracho@ufmg.br

Marcelo Franco Porto, Renato Vieira de Almeida
School of Engineering
Federal University of Minas Gerais
Belo Horizonte, Brazil
marceloporto@ufmg.br

***Abstract* – This research presents the modeling and implementation of an information system, called a "A Physical Feasibility Study - PFS." It considers information on the law for land use and occupancy in the city of Belo Horizonte, Minas Gerais, Brazil. The system simulates the PFS construction possibilities and profitability of a given site or region based on the standards and parameters of the land, a fact that directly involves the profitability of the project and the construction companies' decision-making process. It uses concepts and techniques of Building Information Modeling - BIM to model and simulate the construction possibilities. The PFS responds to questions (research questions) on how to define the short term profitability of a site in order to allow company managers to make the decision to invest or not in the purchase of a particular site. The first part considers the parameters and variables concerning the rules of urban works legislation in the municipality of Belo Horizonte. Next comes the analysis, design and development of the PFS system to simulate the possibility of the civil construction project. The system collects input data on the area, front, zoning, type of use and occupancy of the land and the road classification. These data are processed through consulting tables of parameters that form urban regularization and the law on land use and occupation in a region. The system presents output data on the utilization coefficient, occupancy rate, minimum permeable area, clearances and other parameters that dictate what can be built on the land and the topology of the buildings with all their internal areas and respective uses. It then sets out the possible construction options so that the user may choose the desired option. After this, with the definition of the construction costs, you have the profitability of the land. Cost definition is based on the definition of the standard cost amount per meter squared for construction, considering the standard finishing and real estate value in the region. In 58 seconds, the PFS system presents the parameters with information on what can be constructed and the cost of implanting the project. Finally, in order to validate the research and analysis of system functionality, a feasibility study and tridimensional simulation of 113 sites in the Belvedere neighborhood of Belo Horizonte were performed[1].**

Information retrieval, information system , physical feasibility study, urban legislation, Building Information Modeling,

I. INTRODUCTION

The theme of site/terrain physical feasibility study mobilizes professionals and companies in Brazil with the intention of achieving the best utilization of a given site, taking into consideration the multiple parameters that can directly influence construction. Mindful of the over-riding need for quick and precise information these days, it is important to strive for solutions that support the decision-making process.

This article involves the study, design and proposal for a physical feasibility study of building. The research considers concepts and techniques of three-dimensional simulation to evaluate and optimize the possibilities for construction in a given site, in accordance with the rules, master plans and legislation of a region's municipal government. A system that automates the physical feasibility study of sites was implemented. The result of the study generates project construction possibility reports that directly influence in the project decision-making process.

The object of this research is to develop an information system that, using the parameters of urban legislation, assists the manager in the decision-making process.

During research development, we implemented a prototype of a Physical Feasibility Study – PFS computer system that produces a report with the information that managers need to quickly answer questions such as: What are the project options that generate economic yield and better utilization of a particular site/region? What the profitability of the site/region that I am analyzing?; What can be built on this site/region?; How many floors is it possible to build? In this context, an analysis of urban indices and occupancy guidelines of a given region can offer the best opportunities for projects by working on the maximum utilization of land for construction, and consequently by extracting socio-economic and environmental impact data with respect to the enterprise project.

Once the analysis parameters have been defined, other questions that arise are the difficulty in interpreting and analyzing these parameters within the context of maximum utilization of the land/site and the construction possibility calculations that meet legislative norms. One manner of reducing the possibility of errors is the automation of the physical implantation feasibility study process of the civil works project, through the database system with access programmed according to site data; this automation makes calculations and carries out studies in accordance with parameters stored in the database and the site's parameters. The

[1] The authors thank FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) for their financial support.

prototype uses database and graphic computing concepts and was developed in the VB/VBA programming language on the .NET platform.

To evaluate the results, a prosperous region in growth and with new urban paradigms was chosen in the Belvedere neighborhood. A feasibility study and tridimensional simulation of 113 sites/terrains was carried out in this neighborhood, located in the southern region of Belo Horizonte, capital of Minas Gerais State/Brazil, a state with 853 municipalities.

II. LITERATURE REVIEW

This research developed a feasibility analysis model for works implantation in accordance with the real estate sector dynamic. According to Gonzalez [1], this sector shows differentiating characteristics, such as the strong influence of location and the diversity of assets, as well as operating under a regime of imperfect competition, which generates problems in the study of feasibility, subjecting the process to evaluation errors and a resultant fall off in the utilization of a site's economic potential.

To explain and evaluate the main theories connected to the study of urban politics, Kelvin [2] clarifies what the methodological approaches are in the urban filed.

Within this context, city expansion planning is necessary, in the sense of achieving the full development of social functions and ensuring the well being of city inhabitants. Based on this, one of the strategies defined to address this situation is through the adoption of specific legislation. With the accelerated growth of large cities, such as São Paulo, Rio de Janeiro and Belo Horizonte, the need arose to control the planning of this growth as a form of ensuring better housing development conditions. For Belo Horizonte, one of the instruments of urban control is the Urban Legislation Applied to Building Projects - laws on the use and occupancy of land.

For Weber [3], the study of typologies becomes important, since as well as indicating the project criteria adopted in a certain time and place, it can serve as parameter for the feasibility studies, in relation to decision making for the implantation of new buildings.

Building Information Modeling - BIM is a new approach that uses a 3D computer model, with semantic contextual intelligence for the spatial coordination of conception, construction and management of buildings. The model contains information related to physical and functional construction in order to facilitate the integration, inter-operability, exchange, and reutilization of information by different users during the life cycle of a building, besides allowing for different visions and data analysis according to the digital information available. [4, 5, 6,7]

To evaluate results, a prosperous region in rapid and large scale growth was chosen, questioning the urban parameters of Belo Horizonte: the Belvedere neighborhood.

Interface and usability techniques were analyzed and utilized, with the objective of making the system easier to use. According to Nielsen [8], usability is an important factor that aims to measure ease of use: Usability is the characteristic that determines if the handling of a product is easy and quickly learned, not easily forgotten, does not provoke operational errors, offers a high degree of satisfaction in its users and efficiently resolves the tasks for which it was designed.

In current methods, measuring the ease of use level of an application is very important both for the developers, who may evaluate the content developed, and for the end users, who have the opportunity to evaluate how easy the content is to use, check if the operations can be carried out promptly and supply feedback to developers so that there is continuous improvement.

If an application is easy to use and accessible, it means that the user can navigate and consult the content, learn quickly, memorize easily and make fewer errors during operations.

The municipality of Belo Horizonte is divided in five regional administrations: *Oeste, Venda Nova, Pampulha, Centro-Sul and Noroeste*. The *Centro-Sul* region is made up of 49 neighborhoods, has a total area of 32.49 Km² and a population of 260,524, according to City Council data. Managed by the *Centro-Sul* Municipal Regional Administration Department, the region was divided in 13 planning units (PUs). The PUs are spatial units that bring together one or more neighborhoods, with homogeneous land occupancy characteristics and respecting physical, natural or constructed barrier limits.

The Belvedere neighborhood is in the *Centro-Sul* region, in the Belvedere PU. Its privileged location allows direct access through the region's busiest roads: direct access to the highway ring roads, an access channel that surrounds the city with exits to the principal regions of Brazil, including Federal Highway BR-040, which links Brasília to Rio de Janeiro; the MG-030 State Highway, which connects Belo Horizonte to Nova Lima; and the *Avenidas Raja Gabaglia*, linking the neighborhood to the center, and *Nossa Senhora do Carmo*, linking the neighborhood to *Savassi*, an important point of commerce in the region. The neighborhood is located near the *Serra do Curral*, one of the most scenic landscapes of Belo Horizonte, considered one of its symbols. As such, Belvedere is implanted in an up-market area of the city with an infrastructure boasting complete services, leisure, public road connection, as well as natural beauty.

According to Amorim[9], another positive characteristic of the area is the services infrastructure, with large scale commercial ventures, such as the BH and Ponteiro Lar Shopping Malls, as well as super markets, car dealerships, building material outlets, specialized hospitals, among other services, such as banks, bakeries and fruit and veg shops.

To simulate urban growth, according to Baracho[10], the utilization of tridimensional parametric modeling generating the information system involves software utilization. Normally, the parameters should be adapted to local norms and conditions where the Project and construction will be executed. Implantation of the system makes it necessary to carry out interoperability tests in order to verify if the exported data can be used directly.

III. BH LAND USE AND OCCUPANCY LAW

The first part of the research project may be defined as an information survey phase in which information is gathered on the relevant urban land use and occupancy law, and the technical manual on urban works legislation for Belo Horizonte/MG. This information forms the consultation base for the calculations to be done by the information system. In this first part, a study of specific legislative terms was developed in order to create concepts for the interpretation of information contained in legislation for use in project development.

The legislation applied to building projects supplies the data in chart form. These charts are divided according to the zone, roadway classification and site use, whereby the data varies according to these parameters.

Legal consultation begins with the identification of a given site. Based on chart consultation, you have the zoning classification (Figure 1). Zoning is a fundamental unit of legislation and is a widely used instrument in master plans, where the city is divided into areas to which differentiated guidelines for land use and occupancy are applied, especially urban indices.

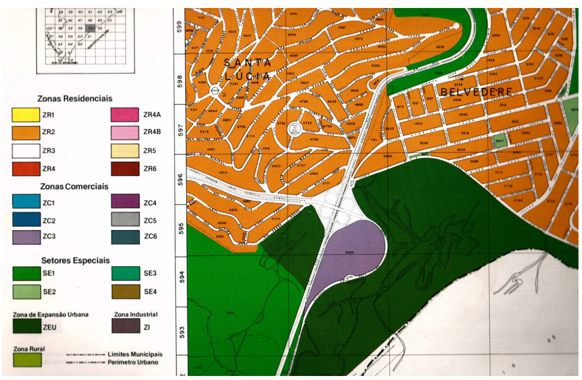


Figure 1. Zones defined by law.
Source: Urban land use and occupancy in Belo Horizonte

With the identification of zoning and use, through consulting the "Zones, Categories of use, Settlement Models". you arrive at the settlement model that may be utilized. The settlement model defines the parameters to be applied in relation to minimum number of car parking spaces, minimum permeabilization rate, minimum frontal distancing, maximum height of division, minimum side and rear distancing, maximum utilization coefficient and maximum occupancy rate. The parameter for car parking spaces indicates the minimum number of spaces the building should have according to the locality, site area and the use and type of building. The rate of permeabilization indicates the percentage of the site that must be reserved for water absorption by the soil, an area that may not be waterproofed or built on. Minimum frontal distancing indicates the minimum distance that a construction must maintain from the frontal limit of the site. Maximum height of the division defines the height of the building, preserving, for example, heights compatible with the 'Serra do Curral', the mountain that boarders the Belvedere neighborhood. Minimum side and rear distancing complete this. Maximum utilization coefficient is the number that, multiplies by the lot area, indicates the maximum quantity of squared meters that may be built on in a lot, adding up the area of all floors. The maximum occupancy rate is the percentage between the building design and the site area. In other words, it represents the percentage of the land that it is built on.

IV. PHYSICAL FEASIBILITY STUDY - PFS

The second part of the project entails the drafting of VB/VBA programming language concepts with the aim of the theoretical arsenal to develop the programming techniques necessary for the creation of a prototype. The VB/VBA language was chosen as it was more efficient in the development user graphic interfaces and had certain applicability with respect to the possibility for application connection and CAD software.

A database was implemented with the possible situations considering the municipality of Belo Horizonte/MG. The database was used to implement data and charts considered as a set of inter-related information, referring to one same subject and organized in a useful manner, with the proposition of serving as a base for the user to recover information, draw conclusions and make decisions.

The database in question has information such as: minimum number of car parking spaces, minimum permeabilization rate, minimum frontal distancing, maximum height of division, minimum side and rear distancing, maximum utilization coefficient and maximum occupancy rate.

The PFS starts on a welcome screen with the names of the researchers, the project name, and the name of the research financers (Figure 2). The system has an easy to use interface.

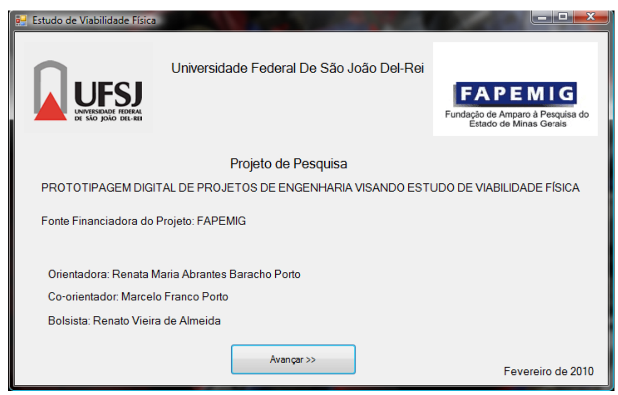


Figure 2. Opening screenshot of the Physical Feasibility Study System - PFS
Source: Prepared by author.

The PFS system is divided into three stages and a report module; the first stage is called 'Site', the second 'Building" and the third 'Profitability'. Each stage makes calculations related to their respective parameters. The three stages are available through the upper menu and its respective tags.

In accordance with the previous section on land use and occupancy law for Belo Horizonte, the necessary variables for data input in the system are defined, through consultation. The variables that need to be defined by the user are: site area, size of site front, type of use, zoning and roadway classification.

The PFS system shows the possible site uses through a consultation of legislation regarding possible uses for a given site. The user then chooses a type within the possibilities shown by the system.

The user starts to use the system through the Site tab, and then fills out the boxes in the order they appear. In the first tab, the program has compulsory data entry boxes: user name, site area, size of site front, site use, zoning and roadway classification of site – the other input data is optional (Figure 3). Figure 3 has the definition of variables on the left, and the frame with report of calculated parameters on the right hand side.

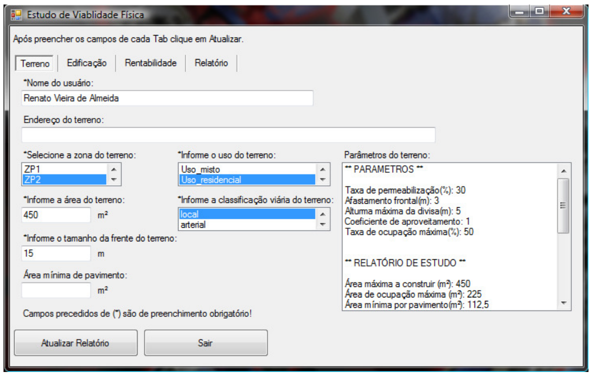


Figure 3. Physical Feasibility Study – EVF: Site tab. Source: Prepared by author

In this table the program provides output data. Parameters supplied by the database consultation: minimum permeabilization rate, frontal distancing, maximum height of division, utilization coefficient and maximum occupancy rate; an initial study report: maximum area to construct; area of maximum occupancy; minimum area per floor; permeable area; construction options: number of floors; area of each floor; construction height (H); minimum number of parking spaces. The second program tab, Building, uses the construction options calculated in the Site tab as input data. The user chooses the option of preference and a table automatically appears with distribution by compartment for that floor. The user can make simulations and change or compare the area occupied by the distribution generated with the maximum area per floor (Figure 4).

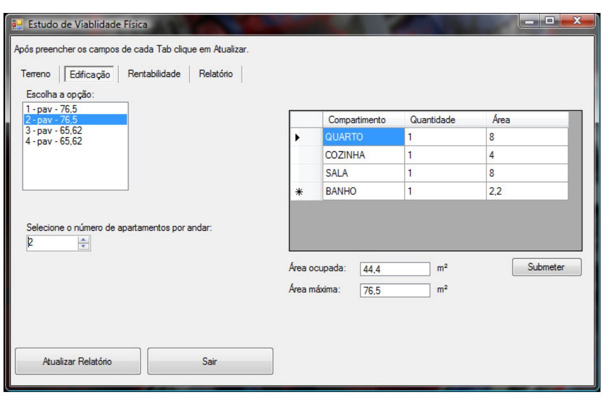


Figure 4. Physical Feasibility Study – PFS: Building Tab Source: Prepared by the author

In the third tab, Profitability, the users enters the meters squared construction amount in the site region, normally stipulated in a chart and released in specialized media on civil construction. The PFS system calculates the construction cost of each floor and the total cost of the works (Figure 5).

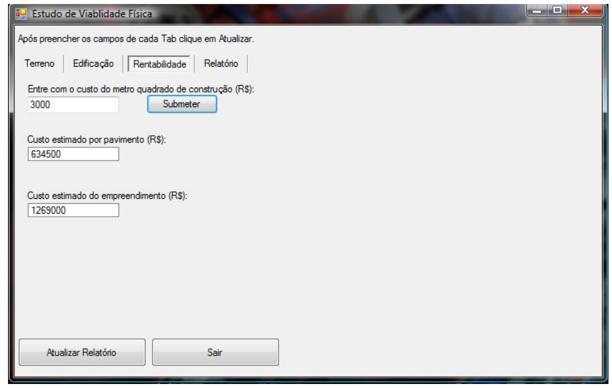


Figure 5. Physical Feasibility Study – PFS: Profitability Tab. Source: Prepared by author

The final report appears as the result of the study and the calculations executed by the PFS system. In this report, all the data generated by the PFS system are shown, as well as all the relevant information contained in the urban legislation database. Figure 6 shows the screen with report. The PFS system generates a report with all the parameters and calculations showing the result for site profitability.

Another result is the use of the report to simulate urban growth and generate supplemental data for studies on regional growth and social, economic and environmental impacts in a region. As research methodology strategy, a simulation of how to use the system was performed in a region of Belo Horizonte. The choice of the Belvedere neighborhood, in Belo Horizonte, as object of the study was made due to its rapid growth and the impact caused in the region. The speed of growth and real estate speculation provoked great debate between professionals involved in urban growth and buildings.

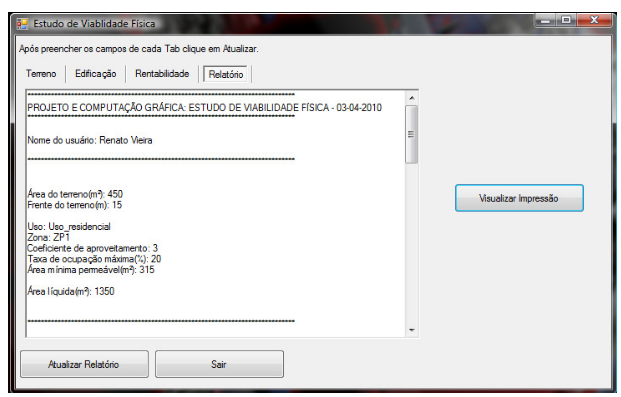


Figure 6. Physical Feasibility Study – EVF: Report Tab Source: Prepared by author.

Initially, a feasibility study of part of the Belvedere neighborhood was done; the neighborhood's zoning and the region's plan were researched. The plan was acquired through a consultation with the data processing company PRODABEL, Belo Horizonte, who are in charge of this information. The plan was transferred from pdf to CAD 2D format.

With the information generated by the PFS system, a simulation of the constructions within each lot was created, so as to modify the 2D plan and obtain a tridimensional - 3D simulation (Figure 7), in which you have a tridimensional virtual image of a part of the neighborhood; this view shows blocks that simulate the constructions with maximum possible utilization of the site.

Based on the definition of a set of lots in the Belvedere neighborhood, the PFS System was used to generate charts with information on the number, address, area and front of lots; the construction options possible were simulated, with the definition of numbers of floors, area of each floor and total construction area. The estimated project cost is related for each one of the options.

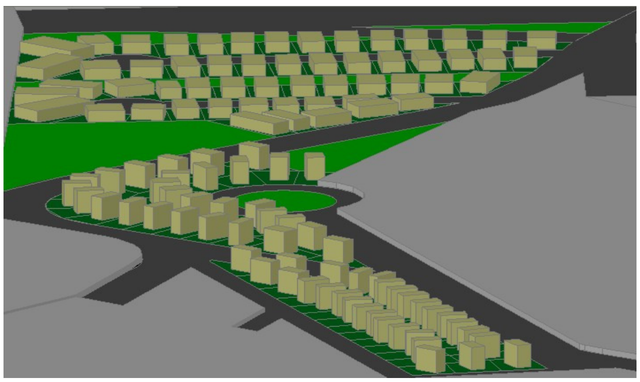


Figure 7. Tridimensional simulation, result of PFS system – View 1. Source: Prepared by author.

V. RESULTS

A step by step simulation carried out with the PFS System is shown below for the number 11 lot, located at Rua Juvenal de Melo Serra, Belvedere. The result of this simulation is described in the example that follows and the data summarized in Chart 1.

Input parameters – Site Tab: Area: 528m², Front: 16.5m, Zoning: ZP3, Use: Residential, Roadway classification: Local, Outut parameters – Site Tab: Utilization Coefficient: 1.5, Maximum occupancy rate (%): 50, Minimum permeable area: 158.4m². These parameters generate nine construction options as defined in floor numbers 1 to 9 (Chart 1)

Chart 1 – Physical Feasibility Study, study of 1 lot

Floor Num.	*Floor Area*	*Area to Construct*	*Height*	*Num. Spaces*
1	264.00	246.00	2.90	1
2	264.00	524.00	5.8	2
3	264.00	792.00	8.7	3
4	198.00	792.00	11.60	4
5	158.00	790.00	14.50	5
6	132.00	792.00	17.40	6
7	113.00	791.00	20.30	7
8	99.00	792.00	23.20	8
9	88.00	792.00	26.10	9

Input Parameters – Building Tab: Choice of option generated: 2; Number of apartments per floor: 1; Compartments: 1 BEDROOM - 8m²; 1 KITCHEN - 4m²; 1 LIVING ROOM - 8m²; 1 BATHROOM – 2.2m²; Observation: The values shown for each compartment are minimum default amounts obtained based on Urban Legislation Applied to Architecture Projects.

Input parameters – Profitability Tab: Cost per meter squared of construction in the region: R$ 1,200.00; Observation: the cost per meter squared of construction in the region was obtained from the Belo Horizonte Municipal City Council. Output parameters – Profitability Tab: Cost per floor: R$ 316,800.00; Total project cost: R$ 633,600.00.

In order to validate the PFS System, 113 simulations were carried out according to the example above. Lots in the ZP2 zone of Belvedere were considered. A second analysis was done in the ZP3 zone varying the number of lots between 1 and 4, these lots to be linked, creating a larger lot with an area equal to the sum of the areas of the adjacent lots. This second analysis was done so that the area of the lots would allow for larger buildings to be simulated.

Input parameters – Site Tab: Area: 1056m², Front: 33.0m, Zoning: ZP3, Use: Residential, Roadway classification: Local, Output parameters – Site Tab: Utilization co-efficient: 1.5, Maximum occupancy rate(%): 50, Minimum permeable area: 316.8m². These parameters generate 14 construction options according to the definition of floor numbers 1 to 14 (Chart 2)

Chart 2 – Physical Feasibility Study - PFS, study of 2 lots.

Floor num	*Floor Area*	*Area to Construct*	*Height*	*Num. Spaces*	*Cost R$*
1	528.00	528.00	2.90	1	633,600.00
2	528.00	1056.00	5.80	2	1,267,200.00
3	528.00	1584.00	8.70	3	1,900,800.00
4	396.00	1584.00	11.60	4	1,900,800.00
5	316.00	1580.00	14.50	5	1,896,000.00
6	264.00	1584.00	17.40	6	1,900,800.00
7	226.00	1582.00	20.30	7	1,898,400.00
8	194.00	1552.00	23.20	8	1,862,400.00
9	176.00	1584.00	26.10	9	1,900,800.00
10	158.00	1580.00	29.00	10	1,986,000.00
11	144.00	1584.00	31.90	11	1,900,800.00
12	132.00	1584.00	34.80	12	1,900,800.00
13	121.00	1573.00	37.70	13	1,887,600.00
14	113.00	1582.00	40.60	14	1,898,400.00

The program generates the profitability calculation of implanting such works. Along with this, the environmental impacts of the works may be calculated, since based on the parameters of the site and the zoning, we can define, for example, the rate of water absorption by the soil in the site analyzed; as a consequence of a study such as this, one could rethink the structural policy of works for sanitation and channeling of rain water, generating the consequent reduction of flooding points, the improvement of social conditions of the population and a reduction of public money being spent on sanitation and health.

A physical feasibility study digital prototype was constructed for part of Belvedere neighborhood, as shown in Figure 11, seeing as tridimensional visualization is a very useful tool in the context of data analysis when dealing with the works process and civil optimization.

VI. Conclusion

The study offers a new method for information retrieval based on visual content for architectural/engineering projects and starts an innovative line of research. With the Physical Feasibility System – PFS, in just 58 seconds, you can get the physical feasibility study result for a site in which you can find out all the construction and site parameters, simulate the construction options possible, as well as the project cost. A study of a block with 28 lots may be executed in 27 minutes and 4 seconds; that is, a study that would take days to complete without using this automated method. The expected research objective of using graphic computing techniques and a digital prototype to optimize and consistently improve the study process of civil works projects and physical feasibility studies of sites was achieved. In doing so, we built a physical feasibility study prototype for sites, with part of the Belvedere neighborhood, in Belo Horizonte/MG serving as digital prototype. Through this digital prototype you can simulate and calculate project parameters, creating a model and a general scheme for optimizations in project development. A physical feasibility study prototype was implemented with the development of a computational system in VB/VBA language; simultaneously, the architectural project of part of the Belvedere neighborhood was implemented using graphic computing software. The digital prototype was done with a tridimensional simulation of the architectural project; Reports on studies and final results for feasibility study possibilities were prepared, and through the simulations and calculations performed, information on the parameters and possible options that interfere, or could interfere, in the project concept was gathered.

References

[1] GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*, 3rd ed., Upper Saddle River, NJ: Prentice Hall, 2008.

[2] KEVIN, W., et al. Urban Politics: An Interdisciplinary Dialogue. International journal of urban and regional research 35.4 (2011):853-871.

[3] WEBER, I., SCHMITT, C. M. *Mapeamento de Variáveis Tipológicas para Edifícios Residencias Multifamiliares em Porto Alegre/RS, na Vigência do Plano Diretor de Desenvolvimento Urbano e Ambiental*. In: XI Encontro Nacional do Ambiente Construído, 2006, Florianópolis, 2006.

[4] GRILO, A.; JARDIM-GONÇALVES, R. *Value proposition on interoperability of BIM and collaborative working environments*. Automation in Construction, Vol.19, No.5, Elsevier Science, pp. 522–530, 2009.

[5] GOEDERT, J.D.; MEADATI, P. Integrating construction process documentation into building information modeling. *Journal of Construction Engineering and Management-Asce*, 2008. 134(7): p. 509-516.

[6] GUANGBIN, W.; YANG, Z. *The Ways of Building Information Modeling to Improve the Value Chain of Building Process*. in Management and Service Science, 2009. MASS '09. International Conference on. 2009.

[7] VOZZOLA, M., CANGIALOSI, G.; LO TURCO. *BIM Use in the Construction Process*. in Management and Service Science, 2009. MASS '09. International Conference on. 2009.

[8] HOWARD, R.; BJORK, B.C. *Building information modelling - Experts' views on standardisation and industry deployment*. Advanced Engineering Informatics, 2008. 22(2): p. 271-280.

[9] NIELSEN, Jakob. *Usability engineering*. Boston: Academic Press, 1993. 362 p.

[10] AMORIM, F. P. ; TANGARI, V. R. *Belvedere III analisys of the urban occupation of an environmental preservation area and its consequences on the landscape*. In: ISUF – International Seminar on Urban Form, 2007, Ouro Preto. Anais ISUF – International Seminar on Urban Form. 1 : 1, 2007. v. 1. p. 1-1.

[11] BARACHO, R.M.A.; PORTO, M.F.; SOARES, M.M.; CARNEIRO, V.M.; FELIX, F.R. *Information system for engineering projects aimed at building process optimization*. In: International Symposium on Information System and Software Engineering, 2011, Orlando. Information System and Software Engineering: ISSE 2011. Orlando : International Institute of Informatics and Systemics, 2011.

[12] BARACHO, R. A.; CENDÓN, B. V. *Information Retrieval for Engineering Projects: Using images to search for images*. In: I International Conference On Information Science And Engineering (ICISE2009), 2009, Nanjing. I International Conference On Information Science And Engineering (ICISE2009), 2009. v. 11. p. 1307-1310.

Experience from using the HYNI automated hydrostatic measuring system for monitoring movement at the 1000 MW Temelín nuclear power station in the Czech Republic.

Ing. Jiří Lechner, CSc., Michal Lodin, M.Sc., M.Sc.Eng.
The Czech Research Institute of Geodesy, Topography and Cartography
Ustecka 98, CZ-250 66 Zdiby, Czech Republic

Introduction

The issue of monitoring shifts (vertical deformation) of buildings is addressed by Czech technical standards CSN 73 0405 on "Measuring Shifts of Construction Objects" and "CSN 73 1001 Building Foundations, Foundation Soil under Areal Foundations". These standards describe measurements of location and shape (shift and deformation) of buildings and their parts relative to the position and shape in the initial stage of measurement caused by changes in the foundation soil material under the object, or in its vicinity due to construction or other activity, the effect of static, dynamic or seismic strain or other factors. The CSN 73 0405 standard deals with all building types, for which specific legislation does not apply. It does not apply to measurements done:

1. during load tests of building constructions,
2. for land subsidence and landslides,
3. for deviations of geometric parameters of actual completion of buildings during construction which are determined during the inspection accuracy by other CSN standards.

1. Purpose of Shift Measurements

The purpose of measuring the shifts is to:

1. Obtain documentation for the assessment of behavior of the foundation soil and he effect of construction of nearby objects,
2. comparison of actual shift magnitudes to predicted values, calculated in as part of the construction project,
3. monitoring of the state, functions and security of new constructions and existing buildings affected by construction activity in their vicinity.

Building shifts are measured during their construction and after its completion if:

1. shifts can be significant for safety and the use of the built object or facility
2. unusual or new structures, for example are used; e.g., foundations of the turbine plate for turbines of 100 MW output and above (CSN 73 1020),
3. high-rise objects are built with a height greater than 50 m (e.g., a dam wall),
4. built object are sensitive to the effects of shifts or are founded in adverse geological conditions,
5. objects are constructed on a previously undermined area (CSN 73 0039) and this measurement corresponds to the significance of the object.

It is necessary to establish a measurement project for each structure, whose shifts are to be measured, specifying the purpose and means of measurement, expected magnitudes of shifts, measurement accuracy with prior analysis, method of point stabilization and measurement time frame.

Construction of the Temelín nuclear power plant represents a unique technical undertaking with the use of new technological procedures, advanced construction methods and construction technologies. The main production block of the power plant consists of the reactor building (ground plan 68 x 68 m, height 66 m), engine house (128 x 49 x 49 m) distribution facility (98 x 22 x 23 m) and the exchange station (48 x 17 x 26 m). The engine house is a single-nave steel hall where the turbo engine aggregate and related service equipment of this non-nuclear part of the power plant. The building structure is divided into these parts:

1. base plate, - a reinforced concrete structure measuring 61,1 x 16,4 x 2,7 m,
2. reinforced concrete structure formed by the 2 x 8 pillars fixed in the base plate at -5.1 m connected in the upper parts by stringers and crosspieces. Stringers have 86 vibration isolators installed at the pillar locations at the +10.95m level and GERB 26 viscose absorbers (flexible mounting system enabling horizontal adjustment of the turbo engine aggregate).
3. Turbogenerator upper foundation plate at the +15m level. This reinforced concrete plate has built-in anchor elements and outer dimensions of 60 x 16 x 3.5m.

2. Measurement Accuracy

The accuracy of building shift measurements of shifts is defined by the deviation threshold value for determination of the length of the resulting shift vector, or its elements. The deviation threshold value, unless stated otherwise, is determined as:

$$\delta_1 \leq 2/15\ p. \qquad (1)$$

where p is the expected total shift or its component in mm.

The deviation threshold value for a device with increased requirement for stability, reliability, safety and operation economy, the δ_1 is reduced by 1/3, unless determined differently by the designer. The threshold deviation value of shift measurements of used building facilities, affected by construction activities in the vicinity, should not exceed the value of:

$$\delta_2 \leq 2/15\ p_k. \qquad (2)$$

where p_k is the critical shift value in mm. If reached, safety of the monitored object is threatened.

3. Stationary automated hydrostatic measuring system

Requirements on the stability of accuracy measurements (in height terms) of the technological facility and buildings hosting the engine aggregates, and turbogenerator of 500 MW and 1000 MW outputs and the foundation plate of the reactor room, were incorporated in the development of the HYNI and INVA sensors by the Research Institute of Geodesy, Topography and Cartography, in cooperation with other institutions. The measurement accuracy of these sensors is stated as standard deviation of elevation measurements a part of the system aggregate accuracy of the entire measuring system (up to the distance of 100 m) $\delta \leq \pm 0.05$ mm.

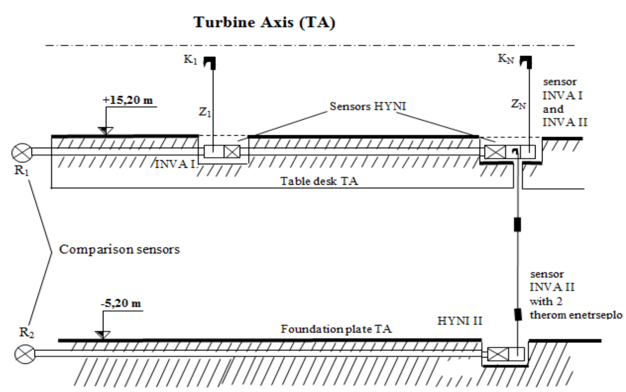
Fig. 1 – Compled HYNI sensors

The stationary hydrostatic measuring system consists of hydrostatic sensors HYNI which are connected by hoses filled with liquid and data transmission cables power supply. The number of sensors in a measuring system is determined

by the requirements of the case with up-to the maximum of 126 sensors used the measuring system. Due to possible measurement of vertical shifts at different height levels, it is possible to include the INVA sensors which enable connecting different levels and it is possible to continuously monitor distance changes of these interconnected levels. Evaluation of individual shifts is automated using the established reference points and on the basis analysis of the measured data, even height stability of reference points is assessed.
An important characteristic of the measuring system is its ability of acquiring continuous measurements without operator intervention. Data collection takes place in previously specified time intervals and measured values can be viewed at any time while the automated measurement programme is running.

Technical parameters of the sensor:
Sensor dimensions: width 205 mm,
height 275 mm, depth 275 mm
Screws for the input of fluid and air: 1/2

Weight:	13kg
Power supply:	18–30 Vss, 0,5 A
Measuring range:	10-90 mm
Communication:	RS 485
	double-wire circuitry



connections of height levels using the invar device
thermal sensor
console
sensor interconnections
sensors INVA I,II

Fig. 2 – Connection scheme of sensor HYNI and INVA in height terms at TA

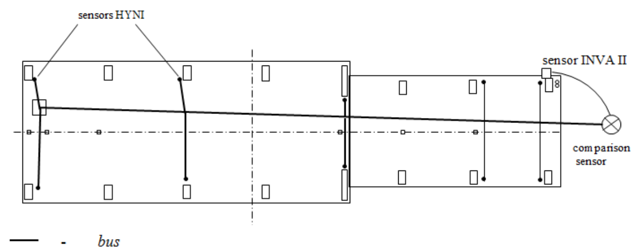


Fig. 3 – The HYNI system at the – 5,20 m level at the engine house (plan view)

Fig. 4 – Side view on the TA from the side of the exciter. Connection between sensors INVA to HYNI

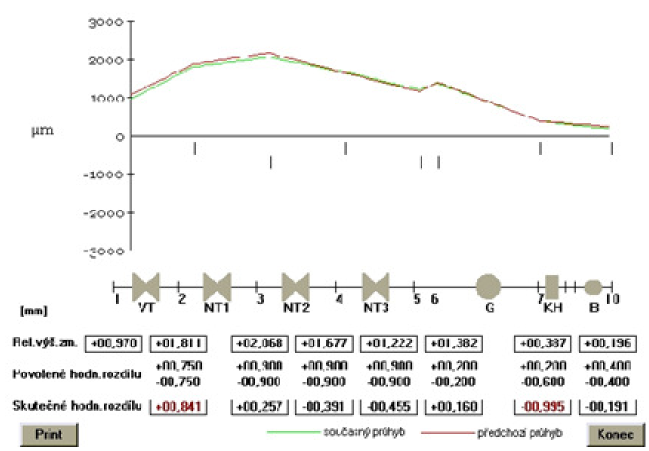


Fig. 5 – The sag curve of the TA

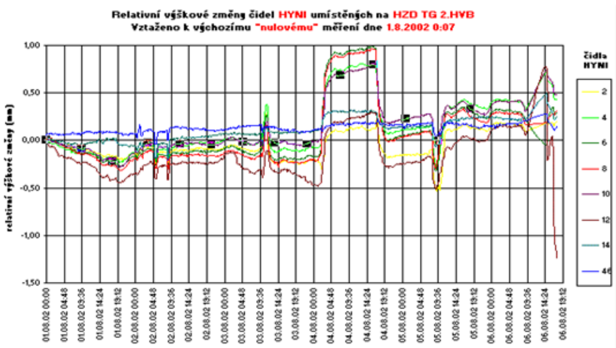


Fig. 6 – Recorded height deformations during discharge and recharge of condensers

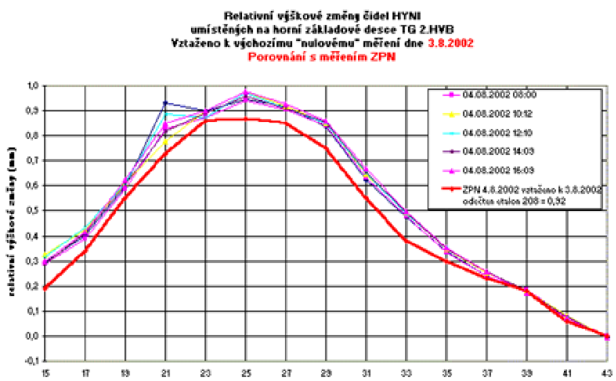


Fig. 7 – Comparing the results of measurement system HYNI and very accurate leveling

4. Conclusion

Results of measurements are used for a planned and efficient maintenance while providing evidence to enhance the security and reliability of the turbo engine and extension of its service life. Measured results of the development of the deformation of the bottom foundation plate shows that the magnitude of deformation does not exceed 20% of the criteria of the CSN 73 1020. Deformations of the upper foundation plate reach, in comparison with the relevant criteria, a maximum of 70% of their value, and a substantial part of the established magnitude of deformation of the upper foundation plate is mainly caused by the change of the temperature field base – TG.

Literature Cited

[1] ČSN 73 0405 Measurement of shift of building objects.
[2] ČSN 73 1001 Foundation of buildings. Base soil under areal foundations.
[3] ČSN 73 1020 Basics of design of torque machines.
[4] ČSN 73 0039 Designing of objects in an undermined area – Basic rules.

Adjusting the Efficiency Curves from Hydroelectric Plants

Ieda HIDALGO
FT, Unicamp
Limeira, Sao Paulo, Brazil
iedahidalgo@ft.unicamp.br

Darrell FONTANE
Civil Engineering, CSU
Fort Collins, Colorado, USA
fontane@engr.colostate.edu

João Eduardo LOPES
FEC, Unicamp
Campinas, Sao Paulo, Brazil
jelopes1@gmail.com

José Geraldo ANDRADE
FT, Unicamp
Limeira, Sao Paulo, Brazil
josegeraldo@ft.unicamp.br

ABSTRACT

The mathematic representation of the functions used to calculate penstock head loss, turbine efficiency and generator efficiency can be modified over time as the plant ages. These functions are important information to computational models used for optimization and simulation of the power generation systems operation. They depend on data quality to provide reliable results. Therefore, this paper presents a methodology to adjust an efficiency function for each generating unit based on measured data. It is applied using an optimization method available in Excel 2010 software, the *Evolutionary Algorithm*. A case study with data from a large Brazilian hydroelectric is shown. The benefits of the data adjusting are analyzed using a comparison metric. The results confirm that the optimal unit efficiency functions significantly improve the performance of simulation models to reproduce observed data and better describe the actual operation of the hydroelectric plant.

Keywords: Hydroelectric operation, Efficiency curves, Generating units, Optimization methods.

1 INTRODUCTION

A significant performance factor in the power generation from a hydroelectric plant is the efficiency of the units. Each generating unit experiences three types of losses. These losses occur in the turbine, the generator, and the penstock. In the turbine and the generator the losses happen due to mechanical friction and heat dissipation in the process to convert kinetic energy into mechanical energy and mechanical energy into electrical energy, respectively. The penstock head losses are associated with the friction of water in the pipeline [1].

Unfortunately the efficiency functions used by many operators are still those obtained from tests in reduced scale models when the company originally acquired the equipment [2]. It is important to have accurate estimates of the efficiency functions so that computational models used for optimization and simulation of the hydroelectric plants operation provide reliable results [3][4][5].

Diniz et al. (2007) [6] studied two mathematical models to represent the efficiency of turbine-generator sets. They used a linear regression technique on the points of the efficiency hill curve in order to find the coefficients of the models.

Colnago (2007) [2] adjusted turbine efficiency curves with the aid of methods for measuring of water discharge. According to him, an optimal dispatch using efficiency curves derived from a reduced model probably do not optimize the plant generation.

Resende (2007) [7] assessed the variation of the characteristics of pipelines infested by golden mussels. He states that using conventional tables for the adoption of coefficients of penstock head losses can lead to mistaken results if compared to reality.

This paper presents a methodology to adjust an efficiency matrix for each generating unit based on measured data. Each matrix considers the losses in the penstock, the turbine and the generator.

2 COMPONENTS OF THE PROBLEM

Production Function is the focus of this work. The goal of the production function, Eq. (1), is to quantify the power generation of a hydroelectric plant.

$$p = k.\eta_t.\eta_g.[h_{fb}(x) - h_{tr}(u) - h_{pl}].q \qquad (1)$$

where:

- p Is the power obtained in the conversion process of hydraulic potential energy into electric energy (MW);
- k Is the gravity constant, multiplied by the water specific weight and divided by 10^6. Its value is 0.00981 (MW/(m³/s)/m);
- η_t Is the turbine efficiency in the conversion process of kinetic energy into mechanical energy;
- η_g Is the generator efficiency in the conversion process of mechanical energy into electrical energy;
- $h_{fb}(x)$ Is the forebay elevation which is function of the water storage x (m);

$h_{tr}(u)$ Is the tailrace elevation which is function of the water release *u* (m);

h_{pl} Is the penstock head loss which is function of the water discharge (m);

q Is the water discharge by the turbines in the powerhouse (m³/s).

In order to apply the production function to short-term planning, turbine efficiency, generator efficiency and penstock head loss functions are required. The turbine converts kinetic energy into mechanical energy. This transformation does not occur completely because there are losses due to mechanical friction and heat dissipation. Turbine efficiency can be expressed as function of the power or water discharge by the gross or net head. It is also called a *Hill Curve* because of its shape.

The generator converts mechanical energy into electrical energy. This transformation does not also occur completely because there are losses in the armature winding known as copper loss, magnetic losses caused by hysteresis and eddy current and mechanical losses that consist of friction at bearings and commutator. Generator efficiency is function of the generated power.

Penstock head loss refers to the loss of water power, in meters, due to friction within the pipeline in the route from the reservoir to the turbine. This loss is calculated taking into account the length, diameter, curvature and roughness of the penstock's inner walls. The loss is usually represented as a quadratic function, Eq. (2):

$$h_{pl} = c_p \, . q^2 \qquad (2)$$

where:

c_p Is the constant that represents the characteristics of the pipeline (s^2/m^5).

3 METHODOLOGY

In order to consider the variations in each generating unit of the plant without needing to estimate the individual functions of turbine efficiency, generator efficiency and penstock head losses, an efficiency matrix is developed for each unit. In this case, Eq. (1) can be rewritten as follows:

$$p = k \, .\eta_t \, .\eta_g \, .[h_{fb}(x) - h_{tr}(u) - h_{pl}].q \qquad (1)$$

$$p = k \, .\eta^G .[h_{fb}(x) - h_{tr}(u)].q \qquad (3)$$

where:

η^G Is the unit efficiency considering the losses of the turbine-generator set and the penstock.

According to Eq. (3), the unit efficiency matrix depends on the power, the forebay and tailrace levels and the water discharge. In general, these variables are recorded in the database by the company that manages the plant operation. Therefore, they can be used to compute an efficiency matrix for each generating unit.

As shown in Eq. (4), the objective function for this problem is to optimize the values of the efficiency matrix in order to minimize the sum of squared errors between the measured water discharge and the water discharge calculated using the efficiency matrix.

Objective function:

$$Min \sum_{i=1}^{num} \left(q_i - \frac{p_i}{k \, . \, \eta^G{}_i \, . \, [h_{fb}(x)_i - h_{tr}(u)_i \,]} \right)^2 \qquad (4)$$

Subject to:

$min \leq \eta^G{}_{r,c} \leq max$

where:

num Is the number of operations recorded in the plant's database;
i Is the index of the operation recorded in the plant's database;
min Is the lowest acceptable value for the unit efficiency matrix;
max Is the highest acceptable value for the unit efficiency matrix;
r Is the row number of the unit efficiency matrix;
c Is the column number of the unit efficiency matrix.

The limits of the unit efficiency matrix should consider the estimated losses associated with the turbine, generator and penstock. According to Encina 2006 [8], the lowest loss related to turbine, generator and penstock together is 6% and the highest is 27%. Therefore, the value of the elements or cells in the unit efficiency matrix should range between 73% and 94%.

For application of the optimization methods the turbine efficiency matrix was used as an initial starting point. During the optimization process, this matrix is converted to a unit efficiency matrix.

4 CASE STUDY

The proposed methodology to optimize unit efficiency functions was applied to a large Brazilian hydroelectric plant. The hourly data of power, gross head and water discharge from 2006 to 2010 were transferred from the plant's database to a spreadsheet in Excel. This corresponds to 43,784 records to be used in the optimization process.

The optimization tool used was the *Evolutionary Algorithm* (EA) of the Solver package in Excel Version 2010. An EA applies the principles of evolution found in nature to the problem of finding an optimal solution. These algorithms maintain a population of candidate solutions, rather than a single best solution so far. From existing candidate solutions, they generate new solutions through either random mutation of single points or crossover or recombination of two or more existing points. The population is then subject to selection that tends to eliminate the worst candidate solutions and keep the best ones. This process is repeated, generating better and better solutions.

EA are designed to find "good" solutions to non-smooth, multi-modal optimization problems, but they can also be applied to smooth, non-linear problems to seek a globally optimal solution. The EA in the Solver tool is based upon a *Genetic Algorithm* (GA) approach [9].

Two personal computers were used for the tests, since personal computers are common in hydroelectric companies. Their characteristics are shown in Table 1. It was anticipated that the run-time could be significant for this problem and therefore it would be useful to determine how much the run-time might be reduced with an enhanced processor.

Table 1. Characteristics of the computers used for the tests

Item	Computer-1	Computer-2
Processor	Core 2 Duo	Core 2 Quad
Processor speed	2.00 GHz	2.83GHz
Memory (RAM)	3 GB	4 GB
System type	32 bits	64 bits

An EA is often overwhelmed by the dimensionality when problem size scales up, for example, from ten to a hundred or to a thousand decision variables [9]. Therefore, for all tests using this method the problem was sliced in pairs of gross head.

The tests were made using both “default options” and “customized options”. First, the optimization process was run using the default sets of population size, convergence and maximum time without improvement. Second, the convergence value was changed from 0.0001 to 0.00001 and the maximum time without improvement was altered from 30s to 60s. The population size was not changed from its’ default value of 100.

Table 2 presents the results of the optimization analysis for each computer. The table contains the test type, the sum of squared errors between recorded and calculated water discharge, the percentage that this sum represents in relation to the use of the initial efficiency matrix, and the total run-time.

The results for each computer are always different. An EA is designed for multi-model objective functions and is intended for exploration of the objective function surface. The EA algorithm begins from a population randomly distributed over the range of the decision values. The method may yield different solutions on different runs, even if the user has not changed the model at all.

Table 2. Results of the optimization analysis using EA

Comp.	Test Type	Sum of Squared Error	% Initial Matrix	Run-Time (hh:mm:ss)
1	Default	3,073,107.05	12.98	00:25:32
2	Default	1,166,047.67	4.92	00:25:22
1	Custom	1,230,683.52	5.20	00:41:54
2	Custom	411,886.48	1.74	00:37:21

The numbers of Table 2 show that in both types of options the sum of squared errors in Computer-2 is around 1/3 of the value of this variable in the Computer-1. Besides, the run-times are similar and longer with the customized option.

5 ANALYSIS OF RESULTS

In order to evaluate the model performance a comparison metric commonly used in hydroelectric operation simulators was analyzed; the water discharge. This analysis used a one year period (8,753 records) not included in the optimization process.

Two types of comparisons were made. In the first mode, the recorded water discharge values were compared to the calculated water discharge values using the current functions of turbine efficiency, generator efficiency and penstock head loss, as Eq. (1). In the second manner, the recorded water discharge values were compared to the calculated water discharge values using the best unit efficiency functions optimized with EA as proposed in this paper, according to Eq. (3).

Table 3 shows the values of Sum of Squared Errors (SSE), Mean Squared Errors (MSE), Sum of Absolute Errors (SAE) and Mean Absolute Errors (MAE) between the recorded and calculated water discharge for the two types. It also shows the percentage of the error reduction of these variables when using the optimized functions instead of the current functions.

As can be seen, comparing the second and the third columns, the use of the optimized functions reduces the error between recorded and calculated water discharge in relation to use of the current functions. For the analyzed period, the use of the optimized functions with EA reduces, on average, more than 50% the errors in the calculation of the water discharge.

Table 3. Statistical Summary

Statistics	Current Functions	Optimized Functions	% Error Reduction
SSE	313,853.30	122,219.22	61.06
MSE	35.86	13.96	61.07
SAE	44,215.41	24,501.61	44.59
MAE	5.05	2.80	44.55

A detailed analysis of the values that make up the SAE using current functions, 44,215.41, shows that 55% of this value consists of negative errors. This means that, in general, the calculated water discharge is lower than the recorded, probably, because the current efficiency functions of this plant are overestimated. This indicates the importance of updating the efficiency functions based upon observed data that reflects the changing condition of the plant.

6 CONCLUSIONS

This paper presents a methodology to adjust efficiency functions for hydroelectric generating units. It is based on measured data of power, gross head and water discharge. Its objective is to determine the actual performance characteristics of the set: penstock, turbine and generator.

The methodology was employed to the data from a Brazilian hydroelectric plant using the EA optimization method. The tests were also made in two ways: “default options” and “customized options”. The use of the “customized options” was around 2.66 times better; however, the run-time increased around 1.56 times. The best result using EA, sum of squared errors equal 411,886.48, was obtained in about 37 minutes.

After the optimization process of the unit efficiency functions, the benefits of the methodology were analyzed comparing values of water discharge. Recorded water discharge values were compared to calculated water discharge values using both current functions (based on manufacturer’s data at initial installation) and the unit efficiency functions (optimized with EA).

The use of the optimized functions reduces the error between observed and calculated water discharge when compared with the use of the functions currently in use. For the analyzed period, the percentage of the error reduction for the SSE using the optimized function with EA was 61.06%. This is a significant improvement and provides operators with a more accurate simulation result on which to base daily operating decisions.

7 ACKNOWLEDGEMENTS

The research reported herein was supported by the FAPESP e CNPq, Brazilian Government agencies dedicated to the development of science and technology, which have funded the post-doctoral studies of the first writer.

8 REFERENCES

[1] Encina, A.S.A., Ohishi, T., and Soares Filho, S. (2002). “Optimal dispatch of generating units of the Itaipu hydroelectric plant”. IEEE Transactions on Power Delivery,17(1), 154-158.

[2] Colnago, G.R. (2007). “Optimal dispatch model with individual treatment of generating units in hydroelectric plants”. MS thesis, State University of Campinas - Sao Paulo, Brazil.

[3] Hidalgo, I.G., Fontane, D.G., Soares Filho, S., Cicogna, M.A., and Lopes, J.E.G. (2010). “Data consolidation from hydroelectric plants”. ASCE Journal of Energy Engineering, 136(3), 87-94.

[4] Hidalgo, I. G., Fontane, D. G., Soares Filho, S., and Cicogna, M. A. (2009). “Computer aided system for managing, controlling and analyzing data from hydroelectric plants.” ASCE - World Environmental & Water Resources Congress: Great Rivers, p.4749-4757 - Kansas, United States.

[5] Hidalgo, I.G. (2004). “Search system for data analysis from hydroelectric plants”. MS thesis, State University. of Campinas - Sao Paulo, Brazil.

[6] Diniz, A.L., Esteves, P.P.I., and Sagastizabal, C.A. (2007). “A mathematical model for the efficiency curves of hydroelectric units”. IEEE - Power Engineering Society General Meeting, ID: 10.1109/PES.2007.385632, p.1-7 - Tampa, Florida, United States.

[7] Resende, M.F. (2007). “The variation of the hydraulic characteristics in penstock due to infestation by limnoperna fortunei”. MS thesis, Federal University of Minas Gerais – Minas Gerais, Brazil.

[8] Encina, A.S.A. (2006). “Optimal dispatch of generating units in hydroelectric systems by heuristic based on lagrangian relaxation and dynamic programming”. Ph.D. thesis, State University of Campinas - Sao Paulo, Brazil.

[9] Frontline Systems (2010). Developers of Excel Solver. Available at: www.solver.com. Access in: 05/11/2011.

Increasing Systems-Safety by Meliorating Policy-Processes under Conditions of Ambiguity

Analyzing Interdisciplinary Ascendancies of the German Traffic System by Using Cybernetic Hazard Analyzing Methodologies

René HOSSE
Institute for Traffic Safety and Automation Engineering, Technische Universität Braunschweig, Germany

and

Sven SIKATZKI
Chair of Governance and Public Policy, Technische Universität Braunschweig, Germany

and

Eckehard SCHNIEDER
Institute for Traffic Safety and Automation Engineering, Technische Universität Braunschweig, Germany

and

Nils BANDELOW
Chair of Governance and Public Policy, Technische Universität Braunschweig, Germany

ABSTRACT

This paper shows the results of incorporating the MSP (multiple streams perspective) into the cybernetic-based hazard analyzing methodology STAMP (systems-theoretic accident modeling and process) to improve the development-process of policies within socio-technical systems. Understanding policy processes is seen as crucial in improving safety-compliance of policies and thus increase system's safety. MSP illustrates several obstructions to safety-compliant policy making e.g. due to unclear preferences, time restrictions or manipulation. In contrast the application of cybernetic-based analyzing methodologies shows significant improvement in increasing systems' safety, because it generates a deeper understanding of hidden processes and feedback-loops within socio-technical systems. STAMP models control structures of systems' resources, showing the cascaded relationships between interdisciplinary control components, e.g. market actors and policy entrepreneurs. This paper initially focuses on the description of socio-technical systems, MSP and STAMP. Afterwards it focalizes the analogy between MSP and STAMP and subsequently generates a hybrid methodology incorporating MSP into the system's control structure modeled by STAMP. Furthermore the explanatory power of STAMP-MSP is exemplified by analyzing the German traffic system.

Keywords: STAMP, Policy Process, Hazard Analysis, Multiple Streams Perspective, Control Structures, Ambiguity

1. CHALLENGES OF COMPLEX SYSTEMS FOR POLICIES AND SAFETY

Policy within a socio-technical system is seen as a key to sustain system's safety and to control social components within safety-relevant constraints [1]. Recent research has shown that the application of system dynamic models improves the policy process and the intended impact. Some decades ago the adaptation of cybernetics to political science led to partial models of the policy process like "Information Flow in Foreign Policy Decisions " by Karl W. Deutsch which where renowned at due time [2]. Nevertheless political systems and policy processes have diversified and became more complex and those models do not seem to be state of the art anymore. Modern systems have similarly evolved from formerly simple mechanic systems to complex systems with a vast number of components, like hard-, and software, personnel, equipment, environment, functions, procedures, and policies [3,4]. Generally safety can be defined as the absence of undesired events [5], which means it is measurable by its complement, the accident [6].In contrast to this trend the so-called classic hazard analyzing methodologies have not been developed adequately to fulfill the exigencies of complex systems, especially the influence of policies. Methodologies like FMEA (failure mode and effect analysis) or FTA (fault tree analysis) are still based on the assumption that accidents result from chains of events leading to an undesired event, thence all events must be predictable. Breaking these chains consequently will hinder the accident to happen. The primary problem lies in the fact that complex systems cannot be understood thoroughly and that breaking event-chains is not sufficient in order to prevent accidents. [4]

There is a range of developments which abet the systems' complexity and strengthen the need for renewed policies and thus analyzing methodologies. Minding the fast pace of technological change while the range of a product's functions is increasing [7] the period of a product's life-cycle is decreasing [8]. Using software within systems generates new kinds of hazards, which do not stem from mechanical dysfunctions. Nowadays an incremental amount of information is necessary to control a social-technical system. [9] “The operation of some systems is so complex that it defies the understanding of all but a few experts, and sometimes even they have incomplete information about its potential behavior.“ [10] Consequently the less is known about accidents, the more often personnel are accused as the initiators of accidents. [5] Usually human behavior only is reported, if it generates an undesired event. Often humans need to intervene in a system’s operative process if an accident is already inevitable [4]. Ergo tolerance for simplified accident decreases and understanding of an accident must be based on the analysis of the system’s design and not on human failure [11]. Furthermore public opinion tends to force the policy giving institutions to take more responsibility in achieving safety and increase the control of human behavior to more safety compliance, because “[…] safety exists within a complex environment involving interactions between people, equipment, policies and operating conditions.” [1] “Effectively preventing accidents in complex systems requires using accident models that include that social system as well as the technology.” [11] Therefore analyzing safety requires an approach which is capable of identifying the meshed interactions between social and technical controllers. Every social-technical system is influenced by society, psychology, economy and politics. Following the ascendancies on shaping human behavior will be regarded more intensely [12].

Social learning postulates that individuals adopt habits of other role models. These habits may be contrary to policies but legitimated by others. Thus policies almost never get full compliance and the system drifts into an unsafe state [13]. Dulac also anticipates that social systems continuously shift into an unsafe state, because in some way unsafe states are higher rewarded to the social controllers than safe states [5]. This effect is supported by psychological ascendancies on humans within socio-technical systems. Walter [14] introduces a model, which is based on the Law of Effects by Thorndike [15] that can be used for describing the link between stimulus and response of humans. The model separates humans’ behavior into the state of rule-compliant and rule-non-compliant behavior. The non-compliant behavior is kept until it results in a negative response. If a stimulus results in a negative response humans switch back to the safety-compliant behavior [14]. Knowledge of this stimulus-response link is crucial because it forces politics to take this into account while creating safety-increasing policies. Furthermore it strengthens the need for combining interdisciplinary research approaches in order to create a better knowledge about systems and thus develop adequate policies.

In policy-processes numerous concepts and theories exist to explain the process and results of the policy process. These theories deliver conflicting perspectives though. The political sphere is a crucial variable for system-safety since its decisions affect the whole safety environment. In accordance [12] the cybernetic model describes the interaction between the political system and its environment mainly with the simple but established triad of input, throughput and output [12, 16, 17]. On the input side demand and support are generated and progress through the political sphere to meet certain ‘checkpoints’ “ occupied by gatekeepers” [17] who have a major influence on the systems agenda, thus its decisions or policies [16]. The underlying logic of this textbook perspective is the policy cycle that focuses separately on each stage. Input, throughput and output occur at an univocal moment in time involving specific actors and institutions [17].
MSP in contrast presumes that the policy process consists of changing constellations of different actors at different moments in time [18]. This assumption draws upon the “garbage can model of organizational choice”, which was introduced by Cohen, March and Olsen (1972) [19, 20]. Cohen, March and Olsen define choice in organizations as a garbage can into which an alternating set of actors discards solutions and problems [18]. Kingdon applies this concept to the policy process and presumes that those constellations emerge under conditions of ambiguity and with the possibility of political manipulation [20, 21]. MSP thereby assumes that the policy processes is shaped by "fluid participation, problematic preferences, and unclear technology" as Robinson and Eller (2010) underline by citing Kingdon´s key indicators for ambiguous policy processes [22]. The MSP model of the policy process consists of three rather independent streams which are coupled by a policy entrepreneur during open windows of opportunity [23]. The distribution of all necessary information for policy makers is hampered by systemic conditions which refer to a notion of a very broad spectrum of motifs, ideas, beliefs and other patterns of thinking that could be connected to the same phenomena [24]. Neither the epistemic and ontological foundation nor the perception of phenomena as problems or solutions is definite [20, 25].This affects the whole policy process. Problems and solutions are not determined and policy outcomes are utterly variable. Thence the conversion of different opportunities into decisions within a political system is not predictable [20, 26].
Against the background of the increasing challenges within complex socio-technical systems and the conflicting theories of policy-processes, the following chapter will present interdisciplinary perspectives on systems and thus generate a hybrid methodology for gaining a deeper understanding of policy-processes using cybernetics.

2. POLICY PROCESSES UNDER AMBIGUITY: THE MULTIPLE STREAMS PERSPECTIVE

The political system is capable of affecting system´s safety especially through legislation. It sets framework conditions for engineering, education, enforcement, and the economy. The assumption of ambiguity affects the model of the policy process as rationality cannot be improved by more information and policy choice is hardly predictable.

Zahariadis [20] names three assumptions of the MSP. Firstly the difference between serial and parallel attention, respectively processing: serial attention applies to individuals e.g. policy makers who are only able to regard one issue at a time due to the limits of human cognition and other constraints. Political systems in contrast have the possibility of parallel processing.. Secondly time restrictions are an important aspect: issues are non-permanent phenomena. They only remain salient for a certain time until another competing issue raises on the policy agenda. The third assumption focuses on the independence of the streams: The streams within a policy subsystem not only generate different contents but usually do so without considering each other. For example policy “solutions” might be produced even before an issue floats on top of the problem stream.

The implications for the policy process that result from these assumptions are quite unique. First of all it is necessary to consider which should be the level of analysis: Should this be the individual with serial attention and processing or the subsystem with parallel attention and processing. Increasing system’s safety by providing all necessary information might especially be hampered when decisions are made at an individual scale. Secondly policy makers’ shortness of time limits their rationality and problem orientation [26]. Furthermore data validity might not have an impact on policy output if data and policy makers meet at the wrong time. Thirdly to improve system’s safety one needs a deep understanding of how the streams of the policy process are structured.

- The *problem stream* consists of conditions and information that under conditions of ambiguity are neither obvious problems nor have public attention [22, 27]. The mechanisms of attention drawing include indicators, feedback and focusing events. Indicators like the number of deaths on the road or people without health insurance are either periodically published or emerge in single studies. Feedback from former programs highlights best practices as well as failed policies. Focusing events are able to disrupt the dominant pattern of thinking about a problem and thus the erupting policy process. [20] For instance multiple vehicle collisions or other unpredictable fatal traffic accidents might shift the actors' perception and increase the contingency of the underlying policy system. The conversion of such conditions into problems is a genuine political process. Kingdon alleges that problems have a "perceptual, interpretive element" [27]. In addition specialized actors draw policy makers´ attention to a limited spectrum of problems [22].
- The *policy stream* represents a "soup of ideas that compete to win acceptance in policy networks" [20]. These ideas might be "solutions" to one or a set of problems [22] but without an obvious nexus between problem and solution since the streams are independent and problems have an interpretative element. Zahariadis identifies "technical feasibility " and "value acceptability" as "selection criteria" of the policy stream. The selection effects on ideas reach from no change at all to their total disappearance. [20] Robinson and Eller point out that policy selection in MSP is dominated by an elitist set of actors and public opinion is mainly disregarded [22].
- The *politics stream* includes three elements, namely the "national mood, pressure group campaigns, and administrative or legislative turnover" [20]. National mood and pressure group campaigns are important to policy makers since they have to identify which policies are zeitgeisty and will be supported by interest groups. Turnover in administration or legislation change the political conditions in the way that a different set of actors with different beliefs and values might change the value acceptability of some policies. [20]

Policy Windows or "windows of opportunity"[1] are points in time when "advocates of proposals [are able] to push their pet solutions, or to push attention to their special problems" [27]. Policy windows do only open when each of the independent streams functions complementarily. Firstly at certain moments in time a phenomenon must be defined as a problem. Secondly a technical feasible and value acceptable solution has to be at hand and thirdly restrictions within the politics stream must not be too intense. Policy windows are only of short duration independently on whether they have opened from the politics stream e.g. through changes of individual actors or by new problem definitions e.g. shaped by focusing events [27]. Policy windows might be predictable e.g. through elections but can also emerge out of the sudden e.g. epidemics, nuclear or natural disasters [28]. Although those more or less predictable events might provide the opportunity for decisions to be made, MSP assumes another venue for policy change which involves active coupling or joining of the three streams [20, 22, 27]. Individual or corporate actors who are trying to use these windows by coupling of the streams are labeled "policy entrepreneurs". In order enforce policy change policy entrepreneurs try to manipulate policy makers and develop a decision context within which the political framework, problem definition and their pet solution work complimentary [20,27]. Zahariadis 2007 identifies three factors that have an impact on coupling. Firstly access through value compatibility is helpful to convert an entrepreneur´s solution into actual policy. Policy entrepreneurs who have general access to policy makers due to similar values are more successful than others. Secondly the more resources they have to promote their solutions the more successful they tend to be in achieving their goals. Thirdly policy entrepreneurs who are skilled at using manipulating strategies, e.g. salami tactics, framing or affect priming have greater chances of success [20]. In summary policy entrepreneurs that couple streams and open policy windows can use everything that either improves the perceived feasibility of a solution that increases the normative acceptance of a policy and/or that lets a problem appear more urgent.

The multiple streams perspective helps analyzing policy processes on the qualitative level, but it lacks capabilities in quantifying policy processes. Thence the aim of this research is to combine the multiple streams perspective with cybernetics-based hazard analyzing methodology and thus quantify policy processes to certain amount.

3. SYSTEMS-THEORETIC ACCIDENT MODEL AND PROCESS

An appropriate hazard analysis using cybernetics is represented by STAMP (systems-theoretic accident modeling and process), which has been introduced by Leveson. "[STAMP] is a new approach to hazard analysis that enables model-based simulation and analysis of risk throughout the system life cycle, including complex human decision-making, software errors, system accidents (versus component failure accidents), and organizational risk factors." [11] The primary aim of STAMP is to identify adequate safety constraints in systems, which are capable of sustaining safety and hinder accidents to happen. This can be achieved by analyzing all relevant control components, technical and social, within a system and ascertain their direct and indirect control loops. Within socio-technical systems most interactions between human and technical control components are cascaded control loops. The controlled process can be measured by sensors which deliver their information to the automated controller. The controlled process, also named operative process, transforms inputs into process outputs influenced by disturbances. The automated controller adhere a model of the process and a model of the interfaces. In order to conduct the adequate control actions, the automated controller controls the controlled process by the actuator. The human supervisor is located parallel to the automated controller. Thus the supervisor is able to influence the automated controller. The here presented control loop can be translated to the policy process by recoupling the human supervisor and automated controller by the actors of policy processes. In doing so, STAMP offers a further tool, which helps visualizing the relevant controllers. In our model the policy making process itself can be described as the controlled process, inputs from different streams are transformed into legislation. Policy makers function like automated controllers since focusing events or feedback as well as national mood meet their value based "sensors" and their decisions are comparable to the actuators. The policy entrepreneur's influence may lack of a predictable "if-then" function. Nevertheless their more or less successful attempts to frame information and manipulate have an influence on the policy maker´s sensors thus their decisions for instance. [4]

At the center of the STAMP analysis is the safety control structure. It represents all relevant control components involved in a

[1]Windows of opportunity and policy windows are often used interchangeably by MSP scholars like John W. Kingdon [19].

controlled process. The control structure models the in- and outputs of each control component and generates virtual container. These functional relationships of in- and outputs can be quantified by empiric data of the systems. By doing so, the static control structure can be translated into a dynamic system dynamics model which can be simulated. The basic structure of the system dynamics model is defined by the specific control structure. Thus STAMP is able to achieve the following main intentions to increase safety:

- Determining control limits for safe behavior
- Generating awareness of permissible behavior towards human and automated controllers
- Developing of strategies for coping with hazardous states
- Supporting of optimization and adaptation processes on contextual influences
- Admitting fault tolerances
- Ensuring visibility and reversibility of errors
- Liberating decision makers and system's operators of performance pressures [5]

Instead of seeing accidents traditionally as the result of event-chains, STAMP defines an accident as an inadequate implementation of safety constraints within the system's structure. The causal factors of accidents lie within the differing mental models about the system's structure and behavior of the system's controllers. The mental model of human beings can vary significantly to the models implemented within the automated controllers. The result of that are inadequate and conflicting controlling actions. STAMP is capable of analyzing technical and human errors within systems, but also the meshed control actions between humans, hardware, software, organizational factors, sociology, and management. [10]

4. CREATING A HYBRID METHODOLOGY USING STAMP AND MSP

In order to generate a hybrid model melding STAMP and MSP, one must get a basic understanding of how systems are designed and where the basic problems of operating process are located: Systems are generally designed by a system's designer with a certain mental model. Within this hybrid methodology the system's designer is the policy maker, e.g. the legislative authority. The designer develops the original design specifications of a socio-technical system, which is the basis for the manufacturers to translate the designer's mental model into an actual system. The mental model of the designer represents the ideality of a system. Controversy the actual system is implemented in the reality, which makes it a complex system so that a complete anticipation is impossible. Manufacturing and constructions variances generate discrepancies between the designer's mental model and the actual system. The policies are created on the designer's mental model, thus the developed policies do not fit the requirements of the actual systems, because they differ significantly from the mental model of the designer. Furthermore the operators create their own mental models of the system which is based firstly on the operating instructions (policies) and the experience with the actual system. The system's operators are according to MSP the policy entrepreneurs, which are trying to influence also the policy makers (designer) to achieve their own goals. The mental model of the operator is also differing from the designer's model and furthermore does not fit the requirement to understand the whole actual system but a few aspects in which the operator is involved. Thus the operator cannot foresee what consequences the individual control actions may generate at another place in the system. A single decision may be safety-compliant in one context of the system's operation. But at another context it may be hazardous (figure 1).

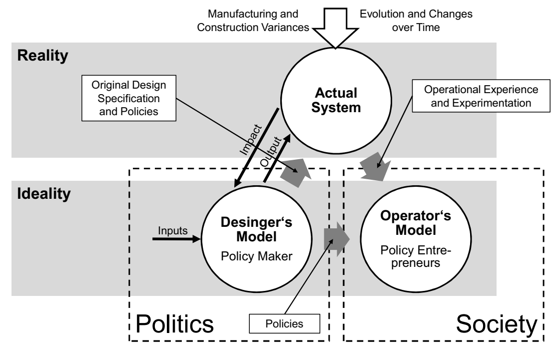


Figure 1: Principles of hybrid methodology, according to [4]

MSP highlights the difficulties of policy making under conditions of ambiguity. In policy processes information is neither "value-neutral" nor an unused instrument for manipulation [20] One principle of STAMP is trying to make all relevant information to each control component accessible. Thus the individual mental models of the system can be updated continuously. Policy entrepreneurs are a crucial analytical figure within MSP and drivers for policy change. The combination of STAMP and MSP helps to understand policy maker's decision patterns and illuminate the consequences of policy processes for systems safety on the basis of a control structure representing all relevant policy makers (system designing control components) and entrepreneurs (system operating control components). In addition to the challenge of an anticipation of consequences of a policy due to complexity and differences between ideality and reality policy entrepreneurs might push a single pet policy which might actually obstruct system´s safety. This is also due to the inadequate mental models by the policy entrepreneurs.

After creating a control structure including all relevant policy makers and entrepreneurs the next step is to identify relevant variables which can be used to develop a system dynamics model of the political system. Making use of system dynamics in analyzing policy processes offers a significant improvement due to the following three aspects: Firstly, system dynamics is based on the feedback approach, modeling the effects of variables on themselves. Secondly, using control structures aggregates the relevant variables to a minimum and focuses on the main ascendancies. Thirdly, the research field of system dynamics provides umpteen simulation tools. By analyzing policy processes the formerly qualitative analysis can be upgraded to a quantitative analysis and hence become more profound.

5. PERCEPTIONS FOR POLICY PROCESSES ANALYZING THE GERMAN TRAFFIC SYSTEM

After describing the hybrid methodology, the next section will show selected results by applying STAMP and MSP on German Traffic Safety. The interdisciplinary ascendancies of the legislative authorities by the various players within the political system can be illustrated by the STAMP-MSP-analysis. The results of STAMP are based on expert interviews.

Analyzing the German traffic system one can identify 11 operative control components, e.g. the individual driver, and in sum 18 system designing control components including the legislative authorities. Within the analysis the legislative authority is represented by the German Politics and the European Union.

The legislative authority is located on the highest level within the policy giving institutions. Located directly under the legislative authority are the public institutions, like the BMVBS (Bundesministerium für Verkehr-, Bau- und Stadtentwicklung), KBA (Kraftfahrbundesamt) and the BAST (Bundesanstalt für Straßenwesen). The institutions take responsibilities for servicing the infrastructure, creating engineering standards, performing research for increasing traffic safety, prosecuting traffic offenders, etc. On the one hand the legislative authority supplies the public institutions with personnel and monetary resources. On the other hand the legislative authority receives information about the status within the traffic system by the public institutions. The basis of any policy is defined by the German Basic Law (Grundgesetz), which is located within the control component engineering standards. Any kind of traffic related policy must be created in accordance to the Basic Law. Minding the policy process under ambiguity, one can see within the control structure (see figure 2) that the processes are underlying lobbying influences. Also the insurance companies, private and professional, execute an influence on the legislative authority. These ascendancies may hinder the policy processes in their effectiveness. For instance the automobile industry does have an interest in bringing new innovations to the market, which may be hampered by regulatory hurdles.

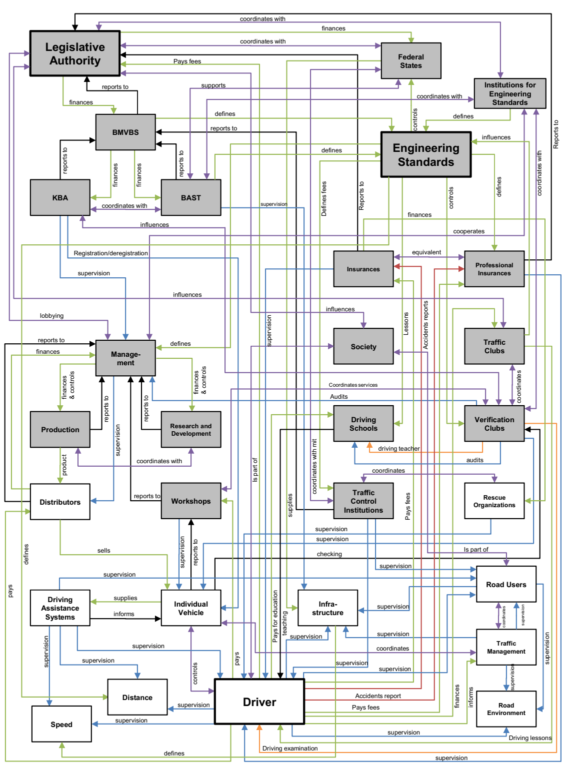


Figure 2: Control structure of the German traffic system, according to [29]

Also the legislative authorities are shaped by society. Minding the multiple streams perspective, different influences do shape the focus of politics to certain areas of interest. This is hard to model due to the numerous aspects having an impact on the attention of politics. Furthermore in Germany, the different federal states do have different policies defining the rules of traffic. Therefore the local influences do also shape policies and have an influence on the policy-processes. One more important aspect about the underlying control structure of the German traffic system is that the legislative authorities receive mostly their monetary resources by taxes which come from the people. But also the citizens do elect the politicians; hence they have an impact on the people defining the policy system.

A crucial part during the hazard analysis by STAMP is identifying missing safety constraints within the control structure. Focusing on the political aspects of the German traffic system one can see that it is mandatory to adapt the policies by the EU. Herein lays a hazard, because the policies of the EU may have a negative impact within the traffic system due to local and/or cultural aspects. Other aspects like the driver education are also hazardous. For instance it is possible to keep the driver license for decades but without any driving experience. There are only mandatory educational provisions at the beginning of a driver participating to the traffic, but no ongoing tests assuring the driver's capabilities.

Another aspect analyzing the control structure is that according to the number of control-relations within the system's structure, the legislative authority is inferior to automobile managements. Even though the legislative authority is strongly connected to other control components within the traffic system, it has just 11 relations to other components. But automobile managements have 13 relations within the traffic system, which are enabling them to perform more control actions influencing traffic safety (see figure 3).

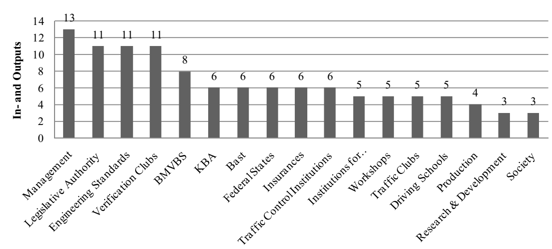


Figure 3: In- and outputs of system designing control components, according to [29]

The inferiority of the legislative authorities exemplifies that politics is hampered performing the optimal policies within the traffic system. But it shows also how interconnected the control loops are affecting safety and policy-processes within socio-technical systems. Furthermore the multiple streams can be visualized by analyzing the meshed relations of the different control components.

6. CONCLUSION

Very little research has been performed in order to analyze and improve policy processes by cybernetics, especially by theory [30]. Based on the MSP the policy process is described as ambiguous and often irrational phenomenon. This paper identifies two main aspects of the MSP that are crucial for a combination of this model of the policy process and cybernetics: the policy entrepreneur (system's operators) and the policy maker (system's designers) that by analogy can be modeled as components in a STAMP control loop structure. Furthermore the findings of this paper have shown that the incorporation of MSP into STAMP produces an enhanced analytical tool for the development of policies within socio-technical systems. We find strong support for our hypothesis that policy is a key variable for system´s safety.. STAMP-MSP illustrates that system designers' and operators' mental models of the system diverge from ideality thus creating a deterioration of system´s safety.

The results on implementing STAMP and MSP on the German traffic system show a strong dependency of policies by the automobile industry. One rationale for this phenomenon within the MSP is the successful use of resources by the automobile industry, which are influencing the policy process. Furthermore the distribution of all necessary information for policy makers is hampered due to a broad spectrum of different motifs, ideas, beliefs and other patterns of thinking. Thence the conversion of opportunities into policies is not determined and influenced by policy entrepreneurs. Modeling and simulating a system by STAMP and MSP can help constraining these effects and therefore increase the logical reasoning of policy processes. The results are based on semi-quantitative analysis (literature research, expert interviews) and must be fully quantified in order to verify the findings. The hybridization between STAMP and MSP shows significant improvement in the understanding of policy processes, especially applying cybernetics to policy-processes. Illustrating the policy process within control structures used by STAMP can help to identify all relevant actors and entrepreneurs engaged in the processes. This hybrid approach based on STAMP and MSP can be used for improving policies within the legislature, management and engineering standards and help increase compliance to safety-critical systems. Illustrating which information must be used for developing adequate policies helps to understand how a system is controlled by its system's resources.

7. REFERENCES

[1] K. Lim, **Enhancing Vehicle Safety Management in Training Deployments**: **An application of system dynamics. Dissertation**, Massachusetts, 2008.

[2] K.W. Deutsch, **The Nerves of Government: Models of Political Communication and Control, with a new introduction**, New York et al.: Free Press, 1966.

[3] C.A. Ericson, **Hazard analysis techniques for system safety**, Hoboken, New Jersey: Wiley, 2005.

[4] N.G. Leveson, **Engineering a safer world: Systems thinking applied to safety**, Massachusetts: MIT Press, 2011.

[5] N. Dulac, **A Framework for Dynamic Safety and Risk Management Modeling in Complex Engineering Systems. Dissertation**, Massachusetts, 2007.

[6] N.G. Leveson, **A New Accident Model for Engineering Safer Systems**, Massachusetts, 2004.

[7] J. Schũffele, **Automotive Software Engineering: Grundlagen, Prozesse, Methoden und Werkzeuge effizient einsetzen**, 4th ed. Wiesbaden: Vieweg + Teubner, 2010.

[8] M. Kuder, **Kundengruppen und Produktlebenszyklus: Dynamische Zielgruppenbildung am Beispiel der Automobilindustrie**, 1st ed. Wiesbaden: Dt. Univ.-Verl., 2005.

[9] N.G. Leveson, **Safeware: System safety and computers a guide to preventing accidents and losses caused by technology**, 5th ed. Boston, Mass.: Addison-Wesley, 2001.

[10] N.G. Leveson, **A New Accident Model for Engineering Safer Systems**, Massachusetts, 2004.

[11] N. Leveson, **Model-Based Analysis of Socio-Technical Risk**, 2002.

[12] D. Easton, "An Approach to the Analysis of Political Systems", **World Politics**, Vol. 9, No. 4, 1957, pp. 383–400.

[13] M.S. Reed, A.C. Evely, G. Cundill, I. Fazey, J. Glass, A. Laing, J. Newig, B. Parrish, C. Prell, C. Raymond, L.C. Stringer "What is social learning?", **Ecology and Society**, Vol. 15, No. 4, 2010.

[14] J. Walter , **Verkehrspsychologie für die Praxis: Wie sich das Verhalten im Straßenverkehr beeinflussen lässt: Ein Leitfaden für Praktiker**, Kröning: Asanger, 2009.

[15] E.L. Thorndike, **Animal intelligence: Experimental studies**, New Brunswick: Transaction Publishers, 2000.

[16] D. Easton, **A Systems Analysis of Political Life**, New York et al.: John Wiley & Sons, 1965.

[17] J.A. Thurber, "Foreword", In: J.W. Kingdon, **Agendas, Alternatives and Public Policies**, updated 2nd edition Boston et al.: Longman, 2011, pp. vii–xi.

[18] S. Kundolf, K. Lindloff, "Interest Intermediation in European transport policy: a case study of agenda-setting processes in public transport", **Paper tabled at the 6th ECPR General Conference**, University of Iceland, Reykjavik; 2011.

[19] S. Brunner, "Understanding policy change: Multiple streams and emissions trading in Germany", **Global Environmental Change**, Vol. 18, No. 3, 2008, pp. 501–507.

[20] N. Zahariadis, "The Multiple Streams Framework: Structure, Limitations, Prospects", In: P.A. Sabatier, editor, **Theories of the Policy Process**, 2nd edition, Boulder, Colorado: Westview Press, 2007, pp. 65–92.

[21] J.W. Kingdon, **Agendas, Alternatives, and Public Policies**, 2nd ed., New York et al.: Longman, 2003.

[22] S.E. Robinson, W.S. Eller, "Participation in Policy Streams: Testing the Separation of Problems and Solutions in Subnational Policy System", **Policy Studies Journal**, Vol. 38, No. 2, 2010, pp.199–216.

[23] N. Zahariadis, **Ambiguity and choice in public policy: Political decision making in modern democracies**, Washington, D.C: Georgetown University Press, 2003.

[24] N. Zahariadis, "Ambiguity, Time and Multiple Streams", In: P.A. Sabatier, editor, **Theories of the Policy Process**, Boulder, Colorado: Westview Press, 1999, pp. 73–93.

[25] M. Brugnach, H. Ingram, "Ambiguity: the challenge of knowing and deciding together", **Environmental Science & Policy**, Vol. 15, No. 1, 2012, pp. 60–71.

[26] F.W. Rüb, "Multiple-Streams-Ansatz: Grundlagen, Probleme und Kritik", In: K. Schubert, N.C. Bandelow, editors, **Lehrbuch der Politikfeldanalyse 2.0**, München: Oldenbourg, 2009, pp. 348–376.

[27] J.W. Kingdon, **Agendas, Alternatives, and Public Policies**, updated 2nd edition, Boston et al.: Longman, 2011.

[28] N. Zahariadis, "Ambiguity and choice in European public policy", **Journal of European Public Policy**, Vol. 15, No. 4, 2008, pp. 514–530.

[29] R.S. Hosse, **Modellierung von Regelkreisen der Verkehrssicherheit mit einem systemtheoretischen Ansatz. Diplomarbeit**, Braunschweig; 2011.

[30] N. Ghaffarzadegan, J. Lyneis, G. Richardson, "How small system dynamics models can help the public policy process", **System Dynamics Review**, Vol. 27, No. 1, 2011, pp. 22-44.

Morphological Studies of Bismuth Nanostructures Prepared by Hydrothermal Microwave Heating.

Oxana V. Kharissova
Universidad Autónoma de Nuevo León, Monterrey, México
E-mail bkhariss@hotmail.com

Mario Osorio
Universidad Autónoma de Nuevo León, Monterrey, México

Boris I. Kharisov
Universidad Autónoma de Nuevo León, Monterrey, México

Edgar de Casas Ortiz
Universidad Autónoma de Nuevo León, Monterrey, México

ABSTRACT

Elemental bismuth nanoparticles and nanotubes were obtained via microwave hydrothermal synthesis starting from bismuth oxide (Bi_2O_3) in the range of temperatures 200-220^{o}C for 10-45 min. The formed nanostructures were studied by scanning electronic microscopy (SEM) and transmission electronic microscopy (TEM). Relationship between reaction parameters and shape of the formed nanostructures is discussed.

Keywords: bismuth, nanotubes, truncated nanospheres, microwave hydrothermal synthesis.

1 INTRODUCTION

Metallic bismuth is an important element, having a lot of distinct industrial applications as a component of low-melting alloys, catalysts, for production of polonium in nuclear reactors and tetrafluorhydrazine, among others. High-purity metal is used, in particular, for measuring super-strong magnetic fields. Bismuth in nanostructurized forms has been mentioned in some recent monographs [1-4], reviews [5-7], patents [6-7], and a lot of experimental articles. Bismuth nanoparticles, nanopowders, nanowires, nanofilms and other nanostructurized forms have been produced by a host of methods, among which, microwave heating (MW) has been also used to obtain bismuth nanostructures [8]. Thus, microwave treatment of bulk bismuth in air in a domestic MW-oven (power 800 W and frequency 2.45 GHz) led to formation of bismuth nanoparticles (60-70 nm) with Bi_2O_3 impurities [9], in a difference of a similar treatment in vacuum [10, 11], when Bi nanotubes formed for 5-15 min. The optimal MW-heating process time was 60 min; the process was found to be highly reproducible and easy. The ***objective*** of this report is the study and comparison of various bismuth nanostructures, observed as a result of the synthesis in the conditions of microwave hydrothermal procedure.

2 EXPERIMENTAL

The reactions were carried out in a Teflon autoclave (equipment MARS-5), using bismuth(III) oxide as a precursor and ethyleneglycol (EG) as a reductant. Due to the necessity to exceed boiling point of EG (187^{o}C) and security limits of the equipment, the syntheses were made in the range 200-220°C, reaching pressure close to 300 psi. The reaction times were 10, 15, 30 and 45 min. The formed nanostructures were studied by scanning electronic microscopy (SEM, equipment Hitachi S-5500) and transmission electronic microscopy (TEM, equipment JEOL 2010-F).

3 RESULTS AND DISCUSSION

The samples, heated for 10-15 min at 200^{o}C, were analyzed by high-resolution TEM, where 5 nm nanoparticles were observed (Fig. 1a). Fig. 1b shows a nanostructure having 5 nm diameter and length of 58 nm, constituted of various aligned more thin structures. It can be affirmed that these nanoparticles grow accordingly to the bottom-up type: the nucleation process takes place, in which, meanwhile the reduction of bismuth oxide occurs, the bismuth atoms are being added to the particle constructing different nanoforms.

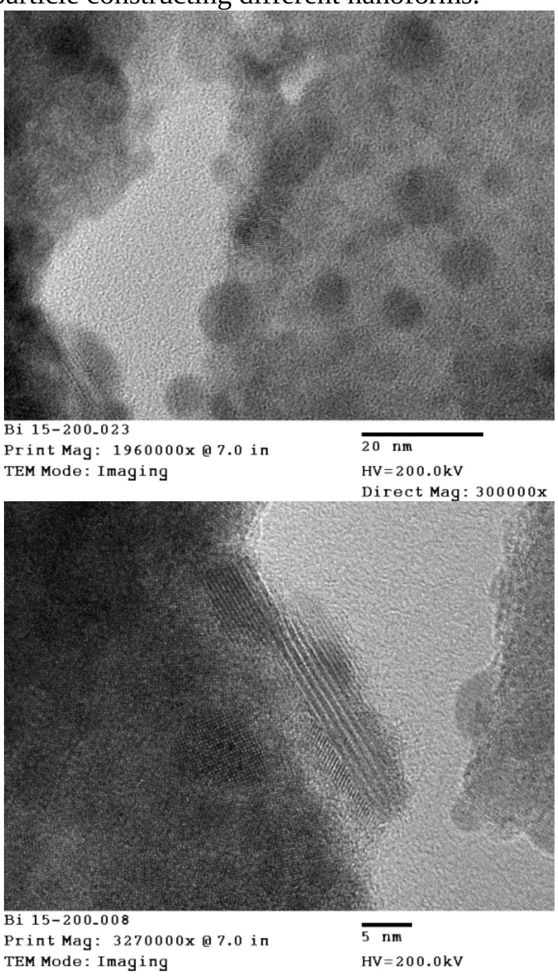


Fig. 1. TEM images of the formed nanostructures (15 min heating at 200°C). The image a) shows nanoparticles with <10 nm size; the image b) shows the growth of nanotube agglomerates.

Fig. 2 (in this case, the heating time was 10 min and temperature 220^{o}C) shows high-resolution TEM images of nanoparticles with 15-20 nm diameters, one of which is shown in a larger scale in Fig. 3. The nanoparticles are of a spherical or truncated-spherical form. No nanotubes were observed for 10 min heating.

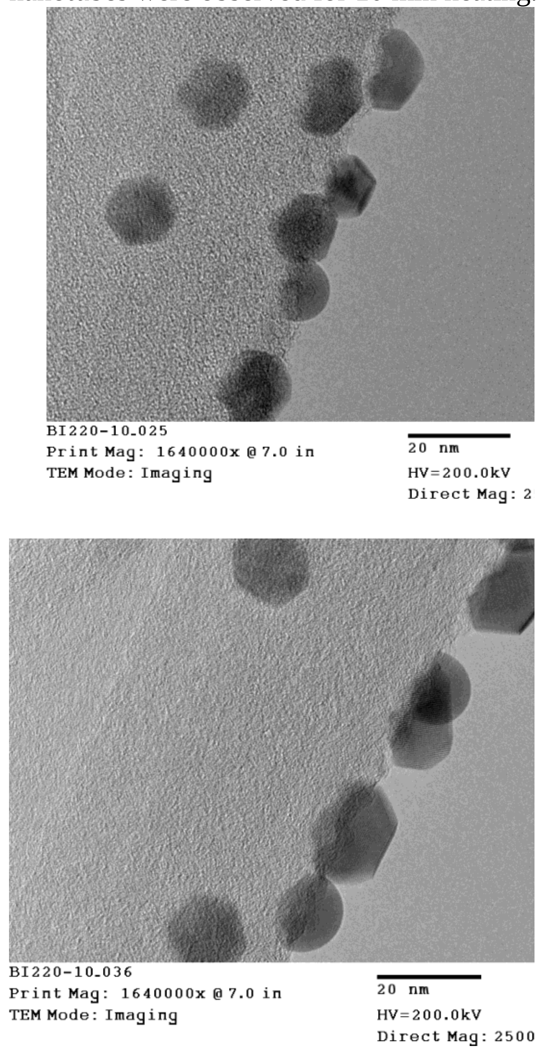
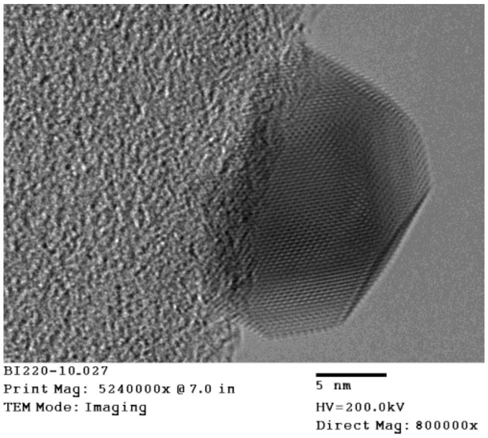
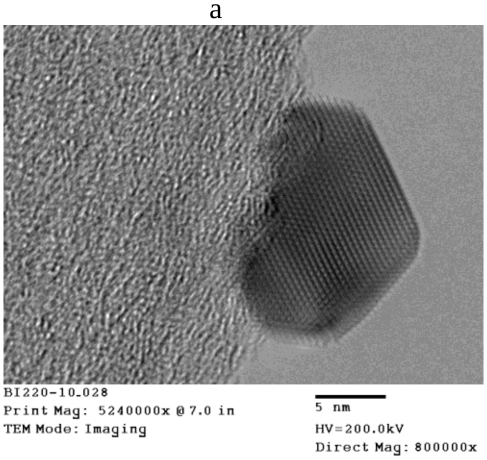


Fig. 2. TEM images of samples, heated for 10 min at 220^{o}C.



a



b

Fig. 3. High-resolution TEM images of the samples, heated for 10 min at 220°C: a) nanoparticle with a visible diameter of 18.8 nm; b) nanoparticle with sizes of 13.52 nm x 17.94 nm and interatomic distance of 3.5 Å

Further increase of heating time to 30-45 min leads to formation of two types of nanostructures, depending on temperature: spherical nanoparticles are observed in the samples heated at 200^{o}C and 220^{o}C (Fig. 4), as well as multi-wall nanotubes (Fig. 5), observed heating at 220^{o}C only. The maximum observed diameter of spherical nanoparticles reaches 500 nm.

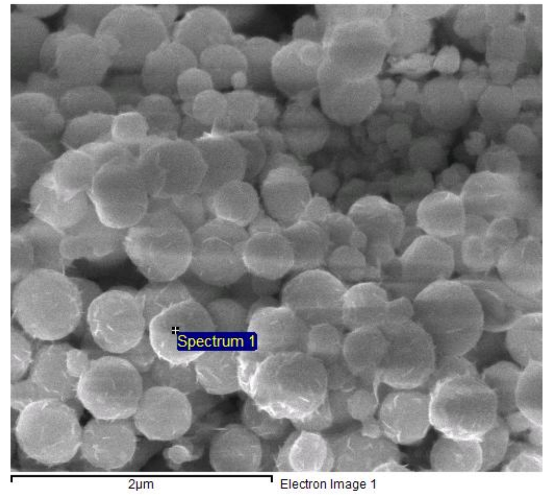


Fig. 4. SEM image of spherical nanoparticles (the sample, heated for 45 min at 220°C).

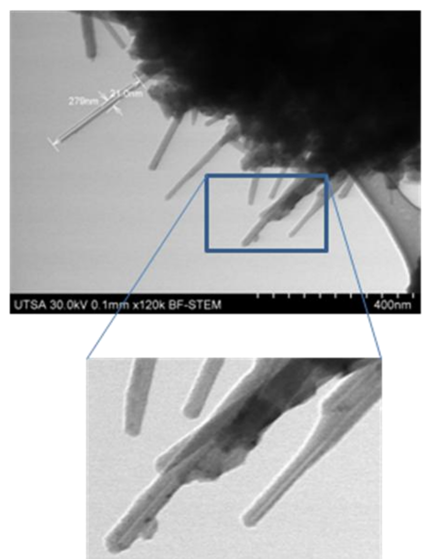


Fig. 5. S-TEM image of nanotubes (the sample, heated for 45 min at 220°C).

4 CONCLUSIONS

Elemental bismuth was obtained in the form of nanoparticles and nanotubes in the conditions of microwave hydrothermal heating. Complete reduction of bismuth oxide to metallic bismuth was observed starting from 10 min of treatment. Agglomerates and blocks of metallic nanotubes were observed at intermediate heating times (15-30 min), meanwhile short (10 min) and large (45 min) treatment durations led to spherical and truncated spherical nanoparticles. Additionally, multi-wall nanotubes were observed at large heating times and higher temperature.

5 ACKNOWLEDGMENT

The authors are very grateful to Professor *Miguel José-Yacamán* (University of Texas at San Antonio, TX, USA) for access to electronic microscopy equipments at the UTSA Department of Physics & Astronomy.

REFERENCES

[1] Sergeev, G. B. (2006). **Nanochemistry**. Elsevier Science; 1 edition, 262 pp.
[2] Koch, C.; Ovid'ko, I.; Seal, S.; Veprek, S. (2007). **Structural Nanocrystalline Materials: Fundamentals and Applications**. 1 edition, Cambridge University Press, 364 pp.
[3] Fryxell, G. E. and Cao, G. (2007). **Environmental Applications of Nanomaterials: Synthesis, Sorbents and Sensors**. Imperial College Press, 520 pp.
[4] Asthana, R.; Kumar, A.; Dahotre, N. B. (2005). **Materials Processing and Manufacturing Science**. 1 edition, Butterworth-Heinemann, 656 pp.
[5] Soderberg, B. C. G. (2003). **Transition metals in organic synthesis: highlights for the year 2000**. Coord. Chem. Rev. 241 (1), 147-247.
[6] Cao, G. and Liu, D. (2008). Template-based synthesis of nanorod, nanowire, and nanotube arrays. **Advances in Colloid and Interface Science**, 136 (1), 45-64.
[7] Dresselhaus, M. S.; Lin, Y. M.; Rabin, O.; Jorio, A.; Souza Filho, A. G.; Pimenta, M. A. et al. (2003). Nanowires and nanotubes. **Mat. Sci. Engin., C**, 23 (1), 129-140.
[6] Penner, R. M; Zach, M. P., Favier, F. (2007). Methods for fabricating metal nanowires. United States **Patent 7220346**, http://www.freepatentsonline.com/7220346.html
[7] Guo, T. Nanoparticle radiosensitizers. (2006**). Patent WO2006037081**.
[8] Kharissova, O. V. and Kharisov, B. I. (2008). Nanostructurized forms of bismuth. **Synth. React. Inorg. Met.-Org. Nano-Met. Chem**., 38 (6), 491-502.
[9] Kharissova, O. V. and Rangel Cardenas, J. (2007). The Microwave Heating Technique for Obtaining Bismuth Nanoparticles, in **Physics, Chemistry and Application of Nanostructures**, World Scientific, pp.443-446.
[10] Kharissova, O.V.; Osorio, M.; Garza, M. (2007). Synthesis of bismuth by microwave irradiation. **MRS Fall Meeting**. Boston, MA. (November 26-30, 2007). Abstract II5.42. p.773.
[11] Kharissova, O. V.; Osorio, M.; Kharisov, B. I.; José Yacamán, M.; Ortiz Méndez, U. (2010). A comparison of bismuth nanoforms obtained in vacuum and air by microwave heating of bismuth powder. **Mater. Chem. Phys**., 121, 489-496.

Traffic Control System: A Conceptual Framework of Developing an Urban Airport Design Intelligent Parking Guidance System for MM International Airport, Lagos, Nigeria

JO Oluwoye, PhD
Professor of Transportation and Environmental Health
Alabama A&M University, Normal, AL 35762
and
OO Oluwoye, BEET
System Engineer,
Bowen University, Iwo, Osun State, Nigeria
and
OA Oluwoye,
Co-op Engineer and Undergraduate Computer Science UG Student
ADTRAN,INC, Huntsville, AL and Alabama A&M University

Abstract

When people complain about traffic congestion in cities, the non-specialist imagines a lot of car drivers, who may be in each others way, but who at least all want to get to a certain destination. There is an increasing interest from motorists to have access to real-time information while en-route to a particular destination, advances in Intelligent Transportation Systems (ITS) has focused on the dissemination of real time information. As central business districts, airports, transit stations, and shopping centers continue to become more crowded during peak times, demand for real-time parking information is increasing. With decreasing parking supply and increasing MMI airport patrons, Federal government of Nigeria needs to beginning to realize the importance of properly allocating available parking. Intelligent Parking Systems (IPS) can provide the positive guidance necessary to help MMI airport patrons find available parking quickly and safely without missing the plane. The purpose of this paper is to develop a conceptual model of parking guidance information system (PGIS). The reality is quite different: Up to 40 percent of car drivers are just cruising around hoping to find a parking space. It should be noted that, looking for parking spaces is a major cause of inner-city traffic especially Lagos. Keeping urban traffic flowing therefore means helping drivers find parking spaces. As experience has shown, the complexity of traffic means that real improvements can only be achieved with carefully harmonized complete solutions taking all factors into account. An Intelligent Parking System could help the Federal government of Nigeria reallocate parking and reduce congestion and illegal parks within the airport and its environment.

Keyword: Urban design, IPS, ITS, Advanced traveler systems, Urban parking guidance

1. Introduction

There is an increasing interest from motorists to have access to real-time information while en-route to a particular destination, advances in Intelligent Transportation Systems (ITS) have focused on the dissemination of real time information. As central business districts, airports, transit stations, and shopping centers continue to become more crowded during peak times, demand for real-time parking information is increasing. With decreasing parking supply and increasing MMI airport patrons, Federal government of Nigeria need to beginning to realize the importance of properly allocating available parking. Intelligent Parking Systems (IPS) can provide the positive guidance necessary to help MMI airport patrons find available parking quickly and safely without missing the plane.

Stout et.al. (1997) reported that for over thirty years, traffic information has been provided to help motorists make en-route decisions. In more recent years, the development of Intelligent Transportation Systems (ITS) and Advanced Traffic Management Systems (ATMS) have begun to improve transportation through the use of technology.

Along the same lines, Intelligent Vehicle Highway Systems (IVHS), systems that "acquire, analyze, communicate, and present information to assist surface transportation travelers in moving from a starting location (origin) to their desired destination," can now be utilized for en-route assistance as well as traffic data collection (Stout et.al. (1997).

However, Polak et.al. (1991) also reported that technology is beginning to recognize the importance of post-trip information dissemination by providing information on the location and availability of parking. Real-time information can be accurately provided to motorists through Intelligent Parking Systems (IPS) to reduce congestion in or near parking areas, insufficient utilization of the available parking space stock, road congestion caused by space-searching traffic, access problems and safety hazards caused by illegal parking, and environmental strains.

The best MMI airport IPS application will provide real-time parking information, reduce congestion, and reallocate parking for all patrons creating more efficient use of airport parking supply.

2. Purpose of the Paper

The purpose of this paper is to develop a conceptual model of parking guidance information system (PGIS) that will assist drivers to locate a car parking space within airport and its environs

3. Conceptual Model of Parking Guide Information System

The figure 1 below illustrates the operation of the parking guide information system (PGIS). The system includes the following components: Outstation unit, Car park site, Vehicle detectors, variable message sign and control center.

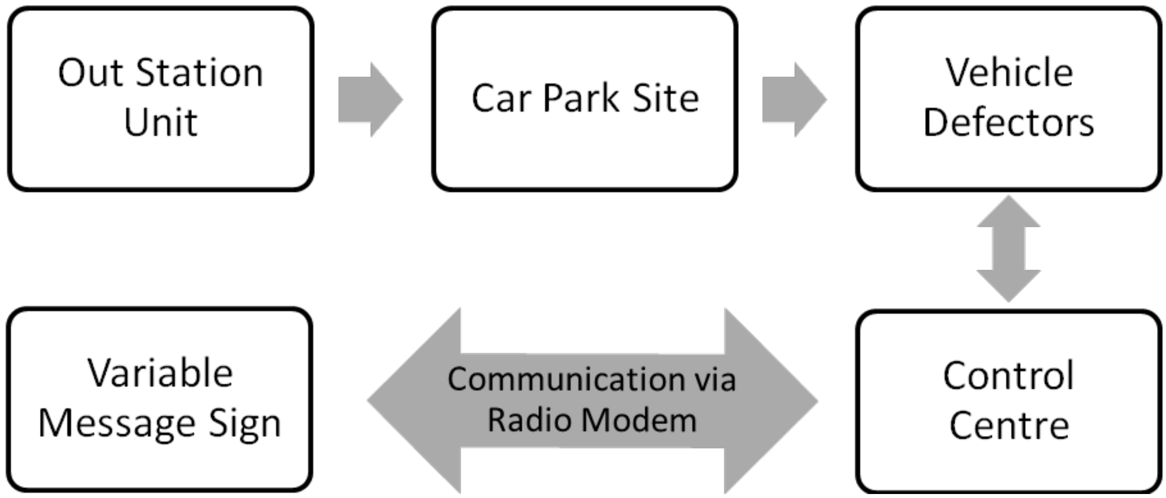


Figure 1: Conceptual Model of PGI System Operation

4. Analysis of Conceptual Model of Parking Guidance Intelligent System Operations

Basic System that Coordinates a Multiple Facility Intelligent Parking System

The basic system that coordinates a multiple facility intelligent parking system contains the three main elements of (1) parking facility equipment, (2) central computer and connections, and (3) signage. The parking facility equipment includes the vehicle counters, space monitors and processing units found on site to monitor ingress and egress traffic, which is sent usually by modem to a central computer. The central computer then controls the variable message signage producing the desired LED displays to direct traffic to open garages or other parking areas. The central computer can also be programmed to send simultaneous messages by radio frequency, dedicated phone line, or Ethernet connection to the local radio, television station, or Internet.

Design of the System

The system includes the following components:

- vehicle detectors will be located at the entrances and exits to car parks, which are capable of detecting the passage of a vehicle;
- variable Message Signs (VMS) will be located at strategic points on the road network, which provide

information to drivers on the occupancy status of car parks or car park areas and direct drivers to car parks or car park areas with available spaces;

- a central computer which calculates the number of available spaces in each car park and car park area and commands the VMS to display the appropriate legend;
- a communications subsystem over which the central computer communicates with the vehicle detectors and VMS.

Possible Applications of Intelligent Transportation Systems

Intelligent Transportation Systems has a number of possible applications in relation to parking alert to drivers. Parking guidance systems have traditionally linked counters (microwave, inductive loop or infrared) at the entrances and exits of off-street car parks (which monitor occupancy and queueing) to variable message signs on key links into and around the town or city center, in order to advise drivers where they are most likely to find a space, close to their final destination.

In an effort to balance competing transportation objectives as well as meet urban design criteria, the MM International Airport Parking Guidance and Information will be strategically designed to group car parks into distinct parking areas to restrict cross-town movement around the airport and its environ.

5. Conclusions

The paper concludes that an Intelligent Parking System could help the Federal government of Nigeria reallocate parking and reduce congestion and illegal parks within the airport and its environment. However, even the best parking guidance system will not help if a tailback in front of a multi-storey car park blocks the road. Car park guidance must work just as smoothly as the equipment installed on the periphery and must be completely integrated.

In order to meet the efficiency demands, the Intelligent Parking System needs to be utilized; if not utilized, IPS will have little or no affect on the parking problems. The real challenge for the Federal government is to begin to develop a parking system that meets the demands of the workers and visitors that utilize parking. Current parking policy may hinder IPS effectiveness; however, if progressive changes are made to MMI airport parking policy, IPS could have positive effects on the supply and efficiency of parking at MMI airport.

References

1. Stout, Tom, Christopher Poe, and Michael Dixon. FHWA Report TX/981751-1. Overview of the State-of-the-Practice in Disseminating Traffic Information by Advanced Traffic Management Systems. Texas Transportation Institute: 1997: 5.
2. Stout 5.
3. Polak, John W., Ian C. Hilton, Kay W. Axhausen, and William Young. "Parking Guidance and Information Systems: A European Review." The Parking Professional. 1991: 16.
4. Southwest Region University Transportation Center Project Proposal. UT Austin. 1999: 1.

Improving Adherence to Complex Medical Regimens through the Integration of Knowledge, Technology and Informatics.

Roberto Badaro; Camille Cruz; Rosa Garcia.
Federal University of Bahia Salvador, Bahia, Brazil
badaro@ufba.br
EliahAronoff-Spencer;Hyeoneui Kim; Sara Browne; Robert T. Schooley
University of California San Diego, La Jolla, CA, USA
rschooley@ucsd.edu

ABSTRACT

About 33.3 million people were living with HIV at the end of 2009. Over the past several years a significant investment in the scale up of antiretroviral therapy (ART) has contributed to both a reduction in AIDS-related morbidity and mortality and in a decline in new HIV infection. A total of 5.2 million people in low- and middle-income countries now receive antiretroviral therapy. Within several months of initiating ART most patients become asymptomatic and then must closely adhere to unforgiving drug regimens for life. Incomplete adherence to ART regimens results in treatment failure, the emergence of drug resistance and the transmission of drug resistant virus to uninfected individuals. This requirement for lifelong regimen adherence has proven to be the Achilles' heel of antiretroviral therapy. Advances in the capability and affordability of mobile technologies present a new perspective on this paradigm. Combined with human and computer networks, mobile devices such as smart phones and tablets bring patients and providers together in previously impossible ways and may present real and lastingremedies for patient outreach and communication. In order to integrate the three major components of adherence: Information, Motivation and Behavioral Modification, we are developing open source pictographic software enabled for web and mobile platforms. We will use these tools to study ART adherence and motivation in low-literacy patients with the goal of enhancing patient comprehension and health.Cybernetic technologies utilizing mobile computing and pictogram interfaces can serve to communicate highly tailored information, evaluate regimen-specific knowledge and assess motivation and behavioral skills of patients requiring therapy. We expect this workshop will bring together all of the requisite expertise in informatics and cybernetics required to develop powerful computerized tools to ensure ART treatment success at the level of both individual patients and the population at large.

Key words: ART adherence, pictographic software, HIV adherence, tablets.

INTRODUCTION

The drug armamentarium developed to control HIV replication is unique in the history of medicine. Incomplete adherence to the treatment of HIV with the recommended drug regimens results in treatment failure, emergence of drug resistance and the transmission of drug resistant virus to uninfected individuals [1, 2]. Once treatment failure has occurred, the emergence of drug resistant viral variants usually requires the initiation of a new drug regimen that is selected on the basis of laboratory testing [3-5]. Unfortunately, after several rounds of treatment failure, more extensive drug resistance drives the selection of less well-tolerated and less convenient regimens. If patients are not able to adhere sufficiently to treatment, over time with successive treatment failures, multidrug resistant virus emerges and, ultimately, it is no longer possible to craft a fully suppressive treatment regimen [3, 4].

Several attempts have been made in the 30 years of the HIV epidemic to improve HIV adherence. Many studies have been performed and others are still ongoing to enhance adherence to antiretroviral treatment regimens (ART) [5].The challenges to improving adherence to ART is based on several factors that have been categorized by Ickovics [6] into four groups: a) patients characteristics; b) patient-provider relationship; c) treatment regimen; and d) additional variables related to the treatment regimen(including associated illness, contextual or environmental factors). Other reports have emphasized additional challenges in antiretroviral therapy. Fear of disclosure of being HIV infected has been deterrence to starting therapy for many patients. Other factors that must be considered are the extent to which medications are perceived to interfere with daily life and the fact that both depression and forgetfulness are critical factors that affect adherence in medicine [7]. Today, over 30 different products in five different drug class targets are available for use in combinations designed to achieve optimal HIV viral suppression [8]. Although much progress has been made in simplifying drug regimens – especially first line therapy – antiretroviral therapy is unforgiving and it remains difficult for patients to adhere to drug treatment regimens. Many tools have been developed and tested that are designed to improve adherence. We are proposing a self-reporting electronic device that will be used by the patient and the providers to learn, instruct and evaluate patient readiness to begin and to adhere sufficiently to treatment to control viremia and improve immunological function over the longer term.

ATTEMPTS TO IMPROVE ART ADHERENCE

Nonadherence is a serious and challenging problem that complicates individual viral suppression in individual patients and compromises the global effort to prevent HIV transmission [9]. Several strategies it has been developed to enhance HIV medication adherence. These strategies can be grouped into five categories: 1-Cognitive strategies; 2- Behavioral strategies; 3- Affectivestrategies; 4-Drug treatment simplification.

1-*Cognitive strategies:* These are intervention strategies designed to instruct, clarify and teach the patient about HIV disease, how drug treatment works and why treatment interruption leads to treatment failure. This approach also, in addition, attempts to tailor specific education to ARV therapy (ART) intervention strategies. For example, nursing case management and individualized counseling by pharmacists have been reported to improve adherence above 90% in over 70 % of patients in several clinical trials [10, 11]. Group education sessions appear to be difficult to implement because of high refusal rates [13-15].A great challenge of cognitive aspect of a patient's treatment adherence is the nature of the provider-patient dialogue in HIV cares centers. A recently reported multisite study in a developed country revealed that only 10% of utterances were devoted to ARV treatment and only 23% of patients had any ARV problem solving with dialogue [12]. Apparently, single strategies are not enough to significantly improve adherence to ART. In Uganda, for example, a recent report of improving clinic attendance by comparing a standard adherence intervention package (SAP) and monthly self-monitoring of medication taking using adherence diaries versus a group that received additional treatment supporter initiatives that included regular individual and group education revealed no difference in the mean adherence between the two groups [13].

2-*Behavioral strategies:* These are designed to shape and reinforce the patient's ability to take the treatment as recommended by their physician and to adapt this long term treatment commitment to their life style. It has been documented that HIV infected patients with substance abuse challenges, mental incompetence, dietary restrictions, homelessness and lack of education are often non-adherent to ARV treatment [14].

3-*Affective strategies:* These are designed to improve emotional and social support. Social economic factors such as social support from family and friends, psychological distress, depression and inadequate confidence in the treatment providers, which are the most common factors challenging adherence. Nonadherence to HAART was independently associated with worse depression rating scale scores, but not associated with neurocognitive impairment in a study of HIV patients in Europe [15]. Depressive symptoms were assessed among 324 people living with HIV/AIDS. Greater adherence to psychotropic medication regardless of medication class of medication class was positively related to higher antiretroviral adherence [16].

4-*Drug simplification:* ARV therapy for HIV disease is used in a strategic order with a well-defined combination to ensure efficacy. The first line ARV currently available consists in one to three pills taken once daily [17, 18].Regimens for individuals with drug resistance are usually require twice-daily regimens with a larger number of pillsper day [8]. The ADONE (Adherence to ONE pill) study was performed to verify the effect of a reduced number of pills on adherence and quality of life (QoL) in HIV-infected patients on highly active antiretroviral therapy (HAART) [19]. Patients chronically treated with emtricitabine (FTC) + tenofovir (TDF) + efavirenz (EFV) or lamivudine (3TC) +TDF +EFV and with a HIV-RNA < 50 copies/mL were switched to the single-pill fixed-dose regimen (FDR) of FTC +TDF +EFV. The study revealed one month post switch to FDR the adherence rate increased significantly to 96.1% from a baseline value of 93.8% (P<0.01). ARV regimens design for multi-experienced patients are more complicated and third and fourth line regimens, in particular, often cause more side effects and toxicities [8] resulting in lower rates of adherence. In a cross-sectional study we performed in Brazil among 182 subjects on ARV therapy, 40% were highly adherent were in their first (and simplest) drug scheme compared to only 15% of those on the second or additional regimens [20].

As previously mentioned stigma about HIV can also be a barrier to adherence as can, food insecurity, health literacy and organization of the health care. In several African countries individuals who self-report missed medication over time also report higher levels of perceived HIV stigma [21]. In Malawi, hunger during HAART initiation emerged as a leading obstacle to ARV adherence [22]. Also, food insecurity is emerging as an important barrier to ARV adherence in Sub-Saharan Africa [23]. Similar data are reported in US peri-urban areas (defined as being at least 5miles distant from a city) where 45% of participants with less that 85% adherence reported experiencing hunger during their ARV treatment [24]. The association between health literacy and ART adherence appears to be robust. Low literacy skills are closely related to understanding medication instruction and the need to plan for pharmacy refills in order to keep the long term commitment to antiretroviral therapy after AIDS symptoms disappear [25]. Evaluation of the effectiveness of ARV of countries such as Brazil that have implemented free access to ARV revealed that 24% of patients on ARV in a public clinic in Rio de Janeiro lacked supply of medication for more than a month at least once during the year [26].

METHODS FOR MEASURING ADHERENCE

An array of tools has been used to assess the level of adherence to ARV therapy [27]. Adherence has been measured in various ways: Two major categories are commonly employed: a) Patient -derived information (self-reported) and b) Methods for independently monitoring drug intake are electronic devices, pharmacy refill logs, pill counting or a combination of these methods.

a) Patient-derived information: Questionnaires are the most commonly used method in many intervention studies on adherence. The Patients Medication Adherence Questionnaire (PMAQ) is one of the most widely used questionnaires for assessing adherence in HIV infected patients, especially in clinical trials [28].Patients may also self-report via interview with a member of the health care team. Walsh and colleagues [29] described a patient self-reported adherence instrument called the Medication

Adherence Self-Report Inventory (MASRI) that collects patient reports of missed doses electronically and a visual analogue scale that correlates with virologicaltreatment outcome. SERAD (Self-Reported Adherence) questionnaire, a qualitative and quantitative self-reported instrument designed to provide an easier adherence measurement has shown good correlation with pill count, electronic monitoring, and plasma drug monitoring [30]. Several structured sets of questions on self-reported adherence are useful to measure adherence. However, the recognition of factors which are determinants of non-adherence must systematically be included in self-reported questionnaires; the most common of these factors are related to patient or medication-centered issues.

b) *Patient independent drug monitoring techniques*: There is no gold standard test for adherence monitoring. Directly Observed Therapy (DOT) is the most reliable one. It has been used with success in treating tuberculosis (TB) [31, 32]. It is less feasible in HIV since HIV requires lifelong daily drug therapy It is thus extremely costly to fully apply DOT in clinical practice [33]. In addition, patient and family concerns about confidentiality also make it difficult to employ DOT for prolonged periods in HIV infection [34]. Community-based modified directly observed therapy (MDOT) is a promising approach that led to lower plasma HIV-1 RNA levels and higher CD4 cell counts in a cohort of HAART experienced substance users [35]. In one preliminary report six women with substance abuse and histories of poor adherence received daily antiretroviral directly observed therapy (DOT) and achieved undetectable plasma HIV-1-RNA levels by 6 months [36]. The Medication Event Monitoring System (MEMS) has been available for decades. A microchip housed in a plastic cap that fits on standard medication bottles records the date and time of each opening and closing of the medication bottle [29, 37]. MEMS adherence studies results were strongly correlated with concurrent viral load [38]. In one study, the mean self-reported 1-day adherence was 79% compared to electronicmonitoring (MEMS) of 52%. However, MEMS correlated better with HIV viral load [39]. A major reason that MEMS adherence is lower than self-reported adherence is that the MEMS system is programmed to register opening the bottle. Unannounced pill counts conducted by telephone were demonstrated to be a viable alternative for objectively assessing medication adherence [40].A high degree of concordance was observed between phone and home-based number of pills counted.

READINESS FOR OPTIMAL ADHERENCE ART

Several factors are correlated with the inability and ability of patients to take ARV medication. Many of these have been outlined in the manuscriptsreviewed .Enriquez and Mckinsey [9] summarized in a table the majority of those factors divided in two columns:.

Barriers: Substance abuse, Fear of disclosure of HIV status, Denial of the HIV diagnosis, Speaking a different language than the health care provider, Stigma, Depression, Forgetfulness, Suspicions about ARV treatment, Perceived unpleasant side-effects from ARV medications, High number of pills in an ARV regimen, Sleeping through medication dosing time, Decreased quality of life, Work and family responsibilities, Limited access to ARV medications, in addition to change daily routine, interference of ARV in their life, difficulties to understand prescriptions, inability to read written instruction, apathy, high stress, problems swollen tablets, inability to distinguish colors or identify makings medication are barriers to adherence ARV treatment.

Facilitators: Sense of self-worth, Seeing/feeling positive effects of ARV therapy, Strong will to live, Acceptance of the HIV diagnosis, Understanding the need for adherence, Making use of reminder tools such as pill organizers, Having an ARV regimen that "fits" into one's daily schedule, Once daily dosing of ARV medications, ARV regimens that are considered to be too complicated, Presence of motivational readiness, Perception of a positive health care provider–patient relationship and Having social support.

Another important factor related to treatment success is the selection of the optimal time to initiate the ARV treatment. Most treatment guidelines strongly emphasize the need to fully prepare each patient for the initiation of therapy. Assessing the readiness for initiation of antiretroviral therapy has a special significance to physicians who provide daily care to people living with HIV/AIDS. Most HIV care providers already have a large file of patients with viruses that are resistant to most classes of ARV drugs. In the case of many of these patients treatment failure could have perhaps been avoided if a careful motivational readiness evaluation had been done. One challenge to the evaluation of *Readiness* is the lack of precise readiness measurement tools that can reliably predict which patients will fail therapy and, therefore, should delay treatment. In their comprehensive review of this topic Grimes et al., concluded that there is currently no viable predictor of adherence [41].

We have summarized the complexity of providing ARV therapy to people living with HIV/AIDS in the format of a Trilogy [27] that includes Information, Motivation and Behavioral skills.

Information: Information is the mainstay of adherence tools. Patients must have comprehensive knowledge of any disease including its cause and the undesirable consequences that will occur if no intervention is provided. In the case of HIV-1 infection the knowledge base includes the meanings of viral load and CD4 cell count as well as how ARV therapy will change the natural course of disease. Although the digestion of this information by a patient might appear to be very simple to the health care specialist , it is often extremely complex for individuals without a background in science. As treatments become more complex even moreinformation must be explained to patient.

Motivation: Motivation is the second fundamental in the adherence trilogy. Even though a patient might know about the disease, what HIV does to the immune system and how drugs will help control his or her disease, this knowledge will not necessarily lead to optimal adherence to therapy, if the patient is not also motivated. Motivation acts in the threshold limit of psychological disturbance. Depression is the biggest impediment to motivation. Often when a physician feels a patient is not motivated, the real problem is that the patient is depressed. A depressed patient is rarely enthusiastic about receiving therapy. Depression may be caused by external factors but it may also be an intrinsic aspect of the patient's individual personality. Depression is common – nearly everyone has experienced a period or moment of low motivation or even a more profound depression. Depressed patients may fail to participate actively in care, may act negatively and may mirror the bad experience that they had with a friend or relative. Individual values are usually structured by motivation. In ancient times philosophers held that values are external to people. In the 1890's a new philosophy concluded that values are intrinsic to each individual. Things have no value; people's wishes are what give value to everything. Thus, the wishes built the values, and motivation transforms these values to action. In other words, patients define their values, and if they are treated with such a level of consideration in the hierarchy of his life, then they will be motivated to adhere to treatment.

Behavioral skills: Health care workers usually forget this third component of the trilogy. A prescription always is based on the most appropriate pharmacokinetics of the drugs, but often conflict with the patient's habits. Usually a prescription will require that a patient change his life style. If the patient is able to adjust to this change in life style, the regimen may work temporally. Dietary patterns, job activities and biorhythms are peculiar to each individual. For many individuals, these issues are considered to be private matters. Treatment must be adapted to the life style of the patient.

Thus, structuring treatment regimens to optimize adherence is quite complex and is far beyond the external attachment. Optimal adherence requires an integration of several critical factors that must thenceforth be part of life.

PATIENT COMPUTERIZED SELF-ADHERENCE TOOL (PCSA)

We are developing preliminary computerized software that that includes three major parts Figure 1.

A-Patient evaluation of HIV knowledge. A series of questions related to knowledge of the disease and its therapy will be asked of each patient in a friendly manner using a computerized touch screen or a voice driven system. These questions will explore patient knowledge level about HIV infection and its therapy. What are the symptoms? How should it be treated? Which types of laboratory tests should be used and in what frequency? A fully validated questionnaire would also include measurements of motivation to fight against the disease as well as treatment preferences, beliefs and fears about the disease and its therapy. Each factor that might complicate adherence to a proposed ARV treatment will be specifically assessed.

B-Educational aspects for those living with HIV infection. The second aspect of the electronic interaction is constructed with a maximum of self-explanatory animated pictograms that patients can access as many times as desired. By the end of the electronic interaction the patient will understand the requirement of a minimum of 95% adherence for treatment success. When this is demonstrated, the patient is deemed to be ready to initiate therapy.

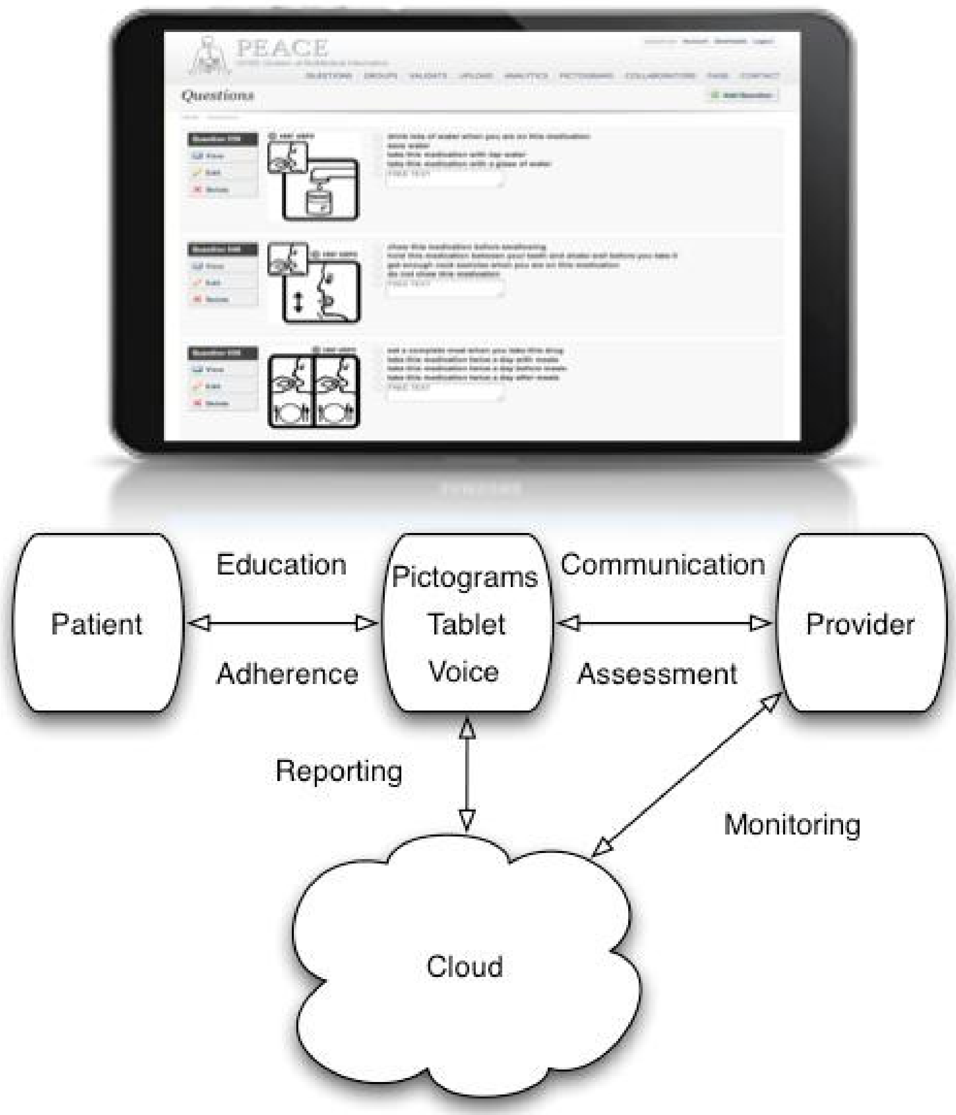
C-Proposed ARV treatment analysis. Based on the Bruin and Hospers [42] Theory of Planned Behavior (TPB), we will explore TPB by using specific pictograms with information of three determinants factors: ***Subjective Norm***. We will explain subjective norms using animated pictograms that illustrate what the care team expects of the patient if treatment success is to be achieved.

Attitude: Using the same strategy we will access the patient's beliefs in terms of how difficult it will be to start and maintain ARV treatment in the context of a personal cost and benefits outcome analysis. What are the required behaviors for successful adherence to therapy? The instrument, in short, provides the patient with objective feedback about self-perceived behavioral control. We will construct a structured algorithm using pictograms that will ultimately help each conclude either: "Yes, I can!" or "No, I am not yet prepared to successfully undertake optimal adherence behaviors for my own treatment regimen.

Feasibility. This system will be built on tablets that the patient can carry to his home or use at the hospital, clinic or public sector. This will allow each patient to take time necessary to assimilate the required adherence knowledge, motivation and behavior to initiate ARV treatment. Each patient will have access to a training assistant if needed to instruct in the correct operation of the tablet. In most settings, at 80% of patients are already proficient in the use of touch screen "smart" phones or computer interfaces and, thus, we expect that the ability to master use of the adherence assessment and instruction platform will be rather intuitive.

CONCLUSION

With this self-reporting tool for assessment of and instruction about the adherence trilogy we expect to have together all of the requisite expertise in informatics and cybernetics required to develop this powerful computerized tool to ensure ART treatment success at the level of both individual patients and the population at large



Footnote figure 1.

Promoting patient adherence and care through pictogram enanced- tablet based communication. Above: Preliminary cloud tools will be based on the *Pictogram Evaluation Authoring and Collaboration Enviromnment (peace.ucsd.edu)* currently under development by our US colleagues. 3.5, 5, 7 or 10.1 inch tablet devices (Android or iOS) with cellular connection may be used to access the system for patient assessment, pictogram validation or patient training. Future iterations will consist of native mobile clients specificitly designed for HIV adherence. Bottom. Such a system will promote patient education and adherence through enanced knowlede and ongoing feedback and communication. A cloud based database and reporting structure will enable efficient moniting by providers as well as aggregate system reporting.

REFERENCES

[1] Knobel H, Urbina O, Gonzalez A, et al. Impact of different patterns of nonadherence on the outcome of highly active antiretroviral therapy in patients with long-term follow-up. HIV Med **2009 Jul**; 10(6):364-9.

[2] Maggiolo F, Airoldi M, Kleinloog HD, et al. Effect of adherence to HAART on virologic outcome and on the selection of resistance-conferring mutations in N. HIV Clin Trials **2007 Sep**; 8(5):282-92.

[3] Garcia F, Alvarez M, Fox Z, et al. Predicting antiretroviral drug resistance from the latest or the cumulative genotype. Antivir Ther **2011**; 16(3):373-82.

[4] Wang D, Hicks CB, Goswami ND, et al. Evolution of drug-resistant viral populations during interruption of antiretroviral therapy. J Virol **2011 Jul**; 85(13):6403-15.

[5] Simoni JM, Frick PA, Pantalone DW, Turner BJ. Antiretroviral adherence interventions: a review of current literature and ongoing studies. Top HIV Med **2003 Nov**; 11(6):185-98.

[6] Ickovics JR, Meisler AW. Adherence in AIDS clinical trials: a framework for clinical research and clinical care. J Clin Epidemiol **1997 Apr**; 50(4):385-91.

[7] Mills EJ, Nachega JB, Bangsberg DR, et al. Adherence to HAART: a systematic review of developed and developing nation patient-reported barriers and facilitators. PLoS Med **2006 Nov**; 3(11):e438.

[8] Department of Health andHuman Services N. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. **2011**:1-167.

[9] Enriquez M, McKinsey DS. Strategies to improve HIV treatment adherence in developed countries: clinical management at the individual level. HIV AIDS (Auckl) **2011**; 3:45-51.

[10] Knobel H, Carmona A, Lopez JL, et al. [Adherence to very active antiretroviral treatment: impact of individualized assessment]. Enferm Infecc Microbiol Clin **1999 Feb**; 17(2):78-81.

[11] Martin J, Sabugal GM, Rubio R, et al. Outcomes of a health education intervention in a sample of patients infected by HIV, most of them injection drug users: possibilities and limitations. AIDS Care **2001 Aug**; 13(4):467-73.

[12] Laws MB, Beach MC, Lee Y, et al. Provider-patient Adherence Dialogue in HIV Care: Results of a Multisite Study. AIDS Behav **2012 Jan 31**.

[13] Kunutsor S, Walley J, Katabira E, et al. Improving clinic attendance and adherence to antiretroviral therapy through a treatment supporter intervention in Uganda: a randomized controlled trial. AIDS Behav **2011 Nov**; 15(8):1795-802.

[14] Uzochukwu BS, Onwujekwe OE, Onoka AC, Okoli C, Uguru NP, Chukwuogo OI. Determinants of non-adherence to subsidized anti-retroviral treatment in southeast Nigeria. Health Policy Plan **2009 May**; 24(3):189-96.

[15] Ammassari A, Antinori A, Aloisi MS, et al. Depressive symptoms, neurocognitive impairment, and adherence to highly active antiretroviral therapy among HIV-infected persons. Psychosomatics **2004 Sep**; 45(5):394-402.

[16] Cruess DG, Kalichman SC, Amaral C, Swetzes C, Cherry C, Kalichman MO. Benefits of Adherence to Psychotropic Medications on Depressive Symptoms and Antiretroviral Medication Adherence Among Men and Women Living with HIV/AIDS. Ann Behav Med **2011 Nov 11**.

[17] DeJesus E, Young B, Morales-Ramirez JO, et al. Simplification of antiretroviral therapy to a single-tablet regimen consisting of efavirenz, emtricitabine, and tenofovir disoproxil fumarate versus unmodified antiretroviral therapy in virologically suppressed HIV-1-infected patients. J Acquir Immune Defic Syndr **2009 Jun 1**; 51(2):163-74.

[18] Gallant JE, DeJesus E, Arribas JR, et al. Tenofovir DF, emtricitabine, and efavirenz vs. zidovudine, lamivudine, and efavirenz for HIV. N Engl J Med **2006 Jan 19**; 354(3):251-60.

[19] Airoldi M, Zaccarelli M, Bisi L, et al. One-pill once-a-day HAART: a simplification strategy that improves adherence and quality of life of HIV-infected subjects. Patient Prefer Adherence **2010**; 4:115-25.

[20] Garcia R, Badaro R, Netto EM, et al. Cross-sectional study to evaluate factors associated with adherence to antiretroviral therapy by Brazilian HIV-infected patients. AIDS Res Hum Retroviruses **2006 Dec**; 22(12):1248-52.

[21] Dlamini PS, Wantland D, Makoae LN, et al. HIV stigma and missed medications in HIV-positive people in five African countries. AIDS Patient Care STDS **2009 May**; 23(5):377-87.

[22] Hardon AP, Akurut D, Comoro C, et al. Hunger, waiting time and transport costs: time to confront challenges to ART adherence in Africa. AIDS Care **2007 May**; 19(5):658-65.

[23] Weiser SD, Tuller DM, Frongillo EA, Senkungu J, Mukiibi N, Bangsberg DR. Food insecurity as a barrier to sustained antiretroviral therapy adherence in Uganda. PLoS One **2010**; 5(4):e10340.

[24] Kalichman SC, Pellowski J, Kalichman MO, et al. Food insufficiency and medication adherence among people living with HIV/AIDS in urban and peri-urban settings. Prev Sci **2011 Sep**; 12(3):324-32.

[25] Kalichman SC, Pope H, White D, et al. Association between health literacy and HIV treatment adherence: further evidence from objectively measured medication adherence. J Int Assoc Physicians AIDS Care (Chic) **2008 Nov**; 7(6):317-23.

[26] Carmody ER, Diaz T, Starling P, dos Santos AP, Sacks HS. An evaluation of antiretroviral HIV/AIDS treatment in a Rio de Janeiro public clinic. Trop Med Int Health **2003 May**; 8(5):378-85.

[27] Garcia R, Schooley RT, Badaro R. An adherence trilogy is essential for long-term HAART success. Braz J Infect Dis **2003 Oct**; 7(5):307-14.

[28] Duong M, Piroth L, Grappin M, et al. Evaluation of the Patient Medication Adherence Questionnaire as a tool for self-reported adherence assessment in HIV-infected patients on antiretroviral regimens. HIV Clin Trials **2001 Mar**; 2(2):128-35.

[29] Walsh JC, Mandalia S, Gazzard BG. Responses to a 1 month self-report on adherence to antiretroviral therapy are consistent with electronic data and virological treatment outcome. AIDS **2002 Jan 25**; 16(2):269-77.

[30] Munoz-Moreno JA, Fumaz CR, Ferrer MJ, et al. Assessing self-reported adherence to HIV therapy by questionnaire: the SERAD (Self-Reported Adherence) Study. AIDS Res Hum Retroviruses **2007 Oct**; 23(10):1166-75.

[31] Yen YF, Rodwell TC, Yen MY, et al. DOT associated with reduced all-cause mortality among tuberculosis patients in Taipei, Taiwan, 2006-2008. Int J Tuberc Lung Dis **2012 Feb**; 16(2):178-84.

[32] Udwadia ZF, Pinto LM. Review series: the politics of TB: the politics, economics and impact of directly observed treatment (DOT) in India. Chron Respir Dis **2007**; 4(2):101-6.

[33] Mitty JA, Macalino G, Taylor L, Harwell JI, Flanigan TP. Directly observed therapy (DOT) for individuals with HIV: successes and challenges. MedGenMed **2003 Feb 4**; 5(1):30.

[34] Berg KM, Litwin AH, Li X, Heo M, Arnsten JH. Lack of sustained improvement in adherence or viral load following a directly observed antiretroviral therapy intervention. Clin Infect Dis **2011 Nov**; 53(9):936-43.

[35] Macalino GE, Hogan JW, Mitty JA, et al. A randomized clinical trial of community-based directly observed therapy as an adherence intervention for HAART among substance users. AIDS **2007 Jul 11**; 21(11):1473-7.

[36] Harwell JI, Flanigan TP, Mitty JA, et al. Directly observed antiretroviral therapy to reduce genital tract and plasma HIV-1 RNA in women with poor adherence. AIDS **2003 Sep 5**; 17(13):1990-3.

[37] Wagner GJ, Kanouse DE, Koegel P, Sullivan G. Adherence to HIV antiretrovirals among persons with serious mental illness. AIDS Patient Care STDS **2003 Apr**; 17(4):179-86.

[38] Liu H, Golin CE, Miller LG, et al. A comparison study of multiple measures of adherence to HIV protease inhibitors. Ann Intern Med **2001 May 15**; 134(10):968-77.

[39] Arnsten JH, Demas PA, Farzadegan H, et al. Antiretroviral therapy adherence and viral suppression in HIV-infected drug users: comparison of self-report and electronic monitoring. Clin Infect Dis **2001 Oct 15**; 33(8):1417-23.

[40] Kalichman SC, Amaral C, Swetsze C, et al. Monthly unannounced pill counts for monitoring HIV treatment adherence: tests for self-monitoring and reactivity effects. HIV Clin Trials **2010 Nov**; 11(6):325-31.

[41] Grimes RM, Grimes DE. Readiness: the state of the science (or the lack thereof). Curr HIV /AIDS Rep **2010 Nov**; 7(4):245-52.

[42] de BM, Hospers HJ, van den Borne HW, Kok G, Prins JM. Theory- and evidence-based intervention to improve adherence to antiretroviral therapy among HIV-infected patients in the Netherlands: a pilot study. AIDS Patient Care STDS **2005 Jun**; 19(6):384-94.

Clustering Hyperlink-Induced Topics for effective paper search

Hiroyuki Yamada
Faculty of Science and Technology,
Tokyo University of Science,
2641 Yamazaki, Noda,Chiba, 278-8510, Japan
k-yoshida@ohwada-lab.net

and

Hayato Ohwada
Faculty of Science and Technology,
Tokyo University of Science,
2641 Yamazaki, Noda,Chiba, 278-8510, Japan
ohwada@ia.noda.tus.ac.jp

Abstract

This paper presents a new search method whose targets are hyperlinked papers on the Web. In this sense, this method is similar to that of Web-search engines, and is based on reference-dereference relationships among papers that are clustered by user-supplied topics. This means that more flexible search can be achieved to find user-desired papers effectively. The experiments demonstrate that hyperlinked papers on the same topic can be obtained.

1. Introduction

Currently, many academic paper databases are accessible over the Web, but do not realize the full benefit of an online paper search. For academic research, it is important to understand the relationship among the various papers when examining the references and previous research. Existing methods make it difficult to enumerate the search results. This research, thus focuses on reference-dereference relationships among papers. We propose an effective method of paper search that utilizes the relationships among papers.

Structure mining a set of papers deals with the papers obtained from the given query and papers standing in a reference-dereference relationship to them.

A user issues queries, such as a keyword and an author, to a paper database, and narrows down the number of papers. By selecting important papers from the paper set based on criteria such as Web page ranking, the user can then efficiently retrieve the information that comes with important papers.

The Hyperlink-Induced Topic Search (HITS) algorithm is used to calculate a degree of association based on reference-dereference relationships. The HITS algorithm begins with a foundation, such as the page rank provided by Yahoo and is a method of extracting the Web community relevant to a specific topic.

If a topic with a broad meaning appears in a query, it can be assumed that there will be several communities related to the topic. However, since the HITS algorithm corresponds only to a single community, when several communities exist, structure mining cannot be conducted correctly. Cluster processing using Markov Cluster Algorithm (MCL) a graph clustering method based on a random walk on a graph , then employed.

After clustering several communities using MCL, the HITS algorithm is applied to each cluster. Rather than being a HITS algorithm independent technique, this technique can generate the user-desired community, and can perform exact structure mining.

The technique used in this research is described in Chapter 2. The results and a discussion of the experiment conducted based on the HITS algorithm are found in Chapter 3. Chapter 4 examines possible future research

2. Proposed Algorithm

In this research involving mining the structure of reference-dereference relationships among papers, searches for related papers are conducted, based on importance. With this approach, however, if there are several communities related to the topic, structure mining cannot be performed accurately. We propose a more precise structure mining method that is combination of topic-specific clustering method.

This section, about the HITS algorithm presents the underlying principles. A problem discussed after that. A combination with MCL, which is at the heart of the proposed technique, is discussed.

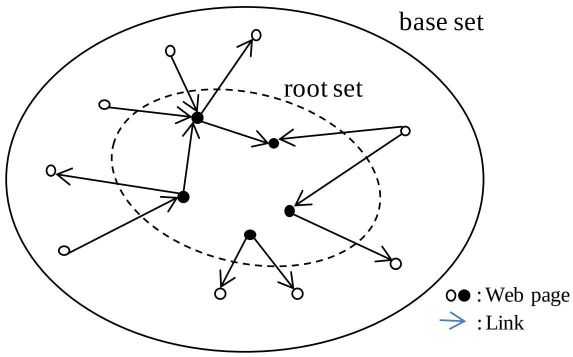


Fig.1 Expansion to base set from root set

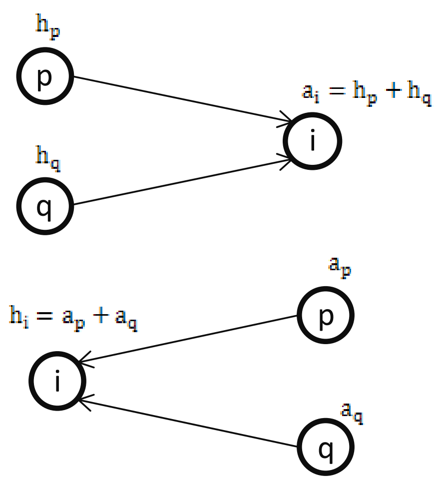


Fig.2 Renewal of score of authority/hub

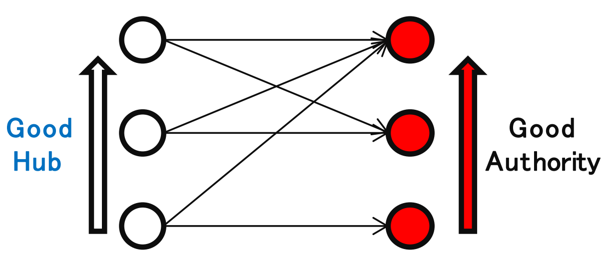


Fig.3 Relation between authority and hub

2.1 Structural mining of paper data

In this research, the HITS algorithm is used as a technique for determining the importance of a paper and extracting the paper set relevant to a specific topic. This extracts a paper (authority) set that includes exact information about a search topic, a paper (hub) set that has a reference/dereference relation with more than one authority set. A paper set consisting of an extracted authority and a hub is defined as a paper community. As mentioned above, the technique for discovery of a paper community relevant to a specific topic leads to an increase in the efficiency of a user's information retrieval from a paper database, and is thus useful.

The extended processing procedure of the HITS algorithm used by this research is outlined below.

Procedure (1) A set (root set) of a paper containing a search topic specified by a keyword is obtained.

The search topic is input as a query into CiNii which is a paper search engine that can be used on the Web, and all papers containing the search topic in the paper database are collected. The paper set obtained by this procedure is called a root set.

Procedure (2) A paper set (base set) containing a paper that expands a root set and is related to the search topic is obtained.

The paper in the root set and the paper which has the reference/dereference relation are added to the root set, and a base set is obtained. This base set is the target of paper structural mining.

Procedure (3) A contiguity procession is created within the base set.

All of the reference and dereference relations between the pages contained in the base set are investigated, and contiguity procession $L = [l_{ij}]$ is created. A contiguity procession is sets to $[l_{ij}] = 1$, when reference and dereference relations to page j exist from page i. The procession indicates a connection relation based on the reference and dereference relations between the papers set to $[l_{ij}] = 0$ here.

Procedure (4) The reference and dereference structure of a base set is analyzed based on the contiguity procession, and an authority score and a hub score are determined.

In order to extract this authority and hub, with the HITS algorithm, the dignity a_i of the authority and the dignity h_i of the hub are assigned to paper i. It performs updates by repetition processing of these scores, and calculates a convergence score. That is, the dignity of the authority of paper i is updated by the total dignity of the hub of a paper that refers to paper i.

The dignity of the hub of paper i is updated by the total of dignity of the authority of the paper currently referred to from paper i. This repeats until a_i and h_i converge as a result of this processing. The ranks of the authority paper and the hub paper can be determined as seen in Fig.2.

In actual calculation, the convergence score of a is determined by the characteristic vector corresponding to the maximum characteristic score of procession $L^T L$, and the convergence score of h is determined by the characteristic vector corresponding to the maximum characteristic score of procession $L^T L$.

2.2 Problem with the HITS algorithm

A topic with a general meaning may be closely related to a topic differs in more than one meaning.

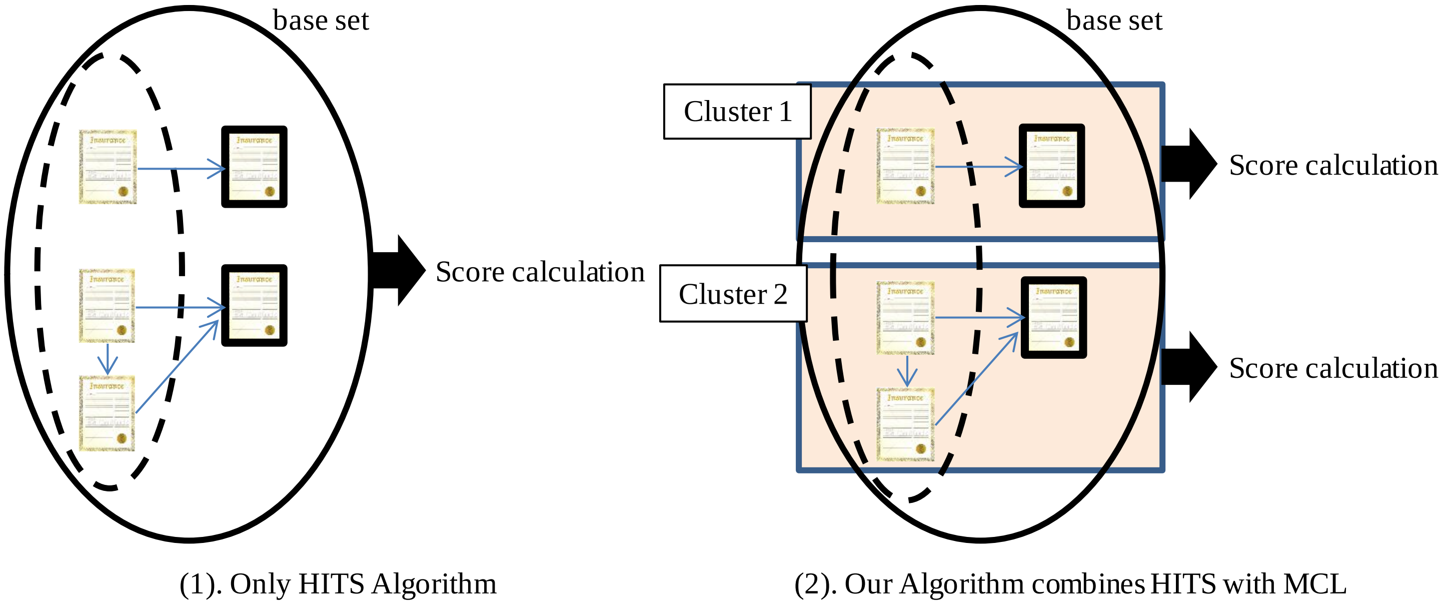


Fig.3 The difference between HITS and our algorithm

For example, the topic “apple” is related to the topic from which one more than meanings, such as Steve Jobs' Apple Company and apple the fruit. Therefore, it is possible that a paper group relevant to more than one topic has dense reference/dereference relations in the base set created based on a general keyword which generally has a large meaning.

The HITS algorithm extracts the paper set with the densest reference and dereference structure as a paper community consisting of an authority and a hub.

Therefore, when the base set consists of paper sets with more than one dense reference and dereference relations, only one can be used to extract the paper community relevant to a topic.

One or more meanings, exist for a general topic, so providing a result that satisfies a user with HITS algorithm that extracts a single community is of little value. Thus, a technique for discovering one or than paper communities in relation to a topic is proposed.

2.3 Combination with MCL

The HITS algorithm extracts the paper set with the densest reference and dereference structure as a community in the base set created based on the topic, so only one community relevant to a topic can be extracted. In the base set, though, a paper set treating the same topic has many references and dereference relations, but there are only a few reference and dereference relations between paper sets treating a different topic. Therefore, there are many references and dereferences relations inside a cluster, and few reference and dereference relation between classes. As a result, a cluster can be generated by clustering a base set for every topic from which a meaning differs. In this research, a base set is clustered using MCL and one or more clusters are generated. Authority and hub scores are computed by applying the HITS algorithm to each generated cluster.

MCL is a graph class ring technique based on a random walk on the graph. MCL is performed on a base set using the following procedures. First, a Markov procession M_G to the base set is created. When one thinks of a random walk model, it can be considered that the Markov procession M_G is a transition probability matrix of a Markov chain, and $(M_G)_{pq}$ expresses the probability of change from node q to node p. Moreover, the transition probability matrix after the kth step of a Markov chain can be denoted by the k-th power of processionM_G. Here, the transition probability$\left(M_G^2\right)_{pq}$ becomes so large that the density of the edge in the local domain where nodes p and q exist is high. The fundamental idea of MCL involves alternating two operations, expansion and inflation. The expansion operation for the procession M is given by $M = M^2$. Inflation operation Γ_r for procession $M \in \mathbf{R}^{k\times l}$ is defined by the following formula.

$$(\Gamma_r M_G)_{pq} = \frac{{(M_G)_{pq}}^r}{\sum_{i=1}^{k}(M_{ip})^r}$$

r is called an inflation parameter and $r > 1$. Usually, r is set to 2.

3. Experiment and Consideration

Calculating the authority and hub scores using the HITS algorithm, the foundation of this research, is explained. The experiment input "collaborative

Table1. Calculation result of the hub score base on HITS algorithm

Rank	Result	Paper Title
1	0.525	Proposal of Rating Contribution in Recommender Systems Based on Collaborative Filtering
2	0.477	Collaborative Filtering Based on Sequential Extraction of User-Item Clusters
3	0.423	Development of Recommendation System for Independent Label Artists using MySpace Information
⋮		

filtering" to the paper database CiNii as a search query, targeting papers whose publication year, which can provide reference and dereference information is before 2010 for inclusion in the root set for the 50 newest papers. The 354 papers that the 50 papers refer to and the 11 papers currently referred to are extracted, and the root set and the papers, or 415 papers in all are considered as the base set. As described in Chapter 2, the algorithm determines the characteristic vectors a and h by repetition processing after analyzing reference/dereference structure of the base set of 415 papers and doing normal-matrix L conversion. The hub score and the title of the papers with the top three hub scores acquired from the experiment are listed in Table 1.

The papers with the top two hub scores in Table 1 refer to many of the same papers, so the degree of association is considered to be high. Many papers refer to the third paper, so the hub score is considered high. This paper is the result of the "collaborative filtering" phrase used as a search query not being entered directly, but causing the extraction of circumferential information relevant to the target paper which was one of the purposes of this research.

4. Future Research

In this experiment, the keyword "collaborative filtering", which restricted the search to papers about an information system, was used as a search query. For this reason, the possibility that the obtained search-result root set and the base set expanded based on reference/dereference relation have more than one community was low. When a general keyword is made into a search query, use of the Markov Cluster Algorithm discussed in Chapter 2 is required.

Authority and hub score are computed for each paper contained in the base set that has a relation to the search topic that generated the search query. This will be evaluated using the relative measures from the paper communities obtained from each search query. Since it is relative evaluation, a hub score may be high compared with other communities in the paper whose reliability is not high. To process all of the paper data, it is necessary to change each relative evaluation into an absolute evaluation. By converting this evaluation to an absolute scale, a paper structure network can be formed for all of the paper data. The necessary processing can be handled by combining principal component analysis with the HITS algorithm.

Moreover, new evaluation approaches, such as evaluating the degree to which keywords are similar, can be added. When a user searches a paper, support for extracting the papers that a user usually cannot search can be provided in addition to support for reducing the number of papers returned.

REFERENCE

[1] J. M. Kleinberg: "Authoritative sources in a hyperlinked environment", Journal of the ACM, 46, 5, pp. 604-632 (1999).

[2] D. Gibson, J. Kleinberg and P. Raghavan: "Inferring web communities from link topology", Proc. 9th ACM Conference on Hypertext and Hypermedia (HyperText 98), pp. 225-234 (1998).

[3] A. Borodin, G. O. Roberts, J. S. Rosenthal and P. Tsaparas: "Finding authorities and hubs from link structures on the world wide web", World Wide Web, pp. 415-429 (2001).

[4] K. Bharat and M. R. Henzinger: "Improved algorithms for topic distillation in a hyperlinked environment", Proceedings of SIGIR-98, 21st ACM International Conference on Research and Development in Information Retrieval, Melbourne, AU, pp. 104-111 (1998).

[5] R. Kosala and H. Blockeel: "Web mining research: A survey", SIGKDD: SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery and Data Mining, ACM, 2, (2000).

[6] Kato Kazutani, Matsu Hiroshi" The Discovery Method of Multiple Web Communities with

Markov Cluster Algorithm " IPSJ SIG Technical Reports vol.2005;Nno22pp.87-93(2005) 2005(22), pp87-93,

[7] Nanba Hidetsugu, Okui manabu" Automatic Det ection of Survey Articles from a Multilingual Da tabase-Towards Automatic Generation of Survey Articles." IEIC Technical Report **vol**102 **no.**199(NLC2002 10-23)pp35-41(2002) 2002(66), pp35-41, 2002-07-15

[8] ICISE Ryutaro, FUJITA Setsu, MURAKI Taichi, TAKEDA Hideaki ""Experimental Evaluation of Topic Prediction by the Author-topic Model" SIG-KBS. KBSE,**vol**.76th pp75-80(2007)

IMCIC'12

AUTHORS INDEX
(Post-Conference Edition)