Interdisciplinary Teaching and Research at a Liberal Arts College on Computational Biology

S. Visa*, S. Strand+, R. D. Pasteur*, W. Morgan+

* Mathematics and Computer Science Department
+ Biology Department
{svisa, sstrand, rpasteur, wmorgan}@wooster.edu
The College of Wooster
1189 Beall Avenue, Wooster, Ohio 44691

ABSTRACT

This paper focuses in presenting several aspects of the interdisciplinary collaboration in teaching and research at a liberal arts college. In particular, it presents courses and research projects in the computational biology area, which naturally call for collaborative work between faculty members from mathematics & computer science and biology departments. It further discusses (1) challenges arising when doing interdisciplinary work, (2) the interaction between interdisciplinary teaching, research, and communication, (3) possible outcomes of such interdisciplinary training for both students and faculty members, and (4) how interdisciplinary training fits within the liberal arts philosophy.

Keywords: interdisciplinary communication, teaching, and research.

1. INTRODUCTION

In January 2006, Project Kaleidoscope (PKAL) published a report titled, "Transforming America's Scientific and Technological Infrastructure: Recommendations for Urgent Action" [6] which urges all stakeholders in research and education to make fundamental changes and to lend support to educators in their training of future scientists in solving complex real world problems at the interface of various disciplines

The complex nature inherent in real world problems requires investigators in various disciplines to work together to make discoveries and find solutions. From the perspective of faculty, the intellectual excitement of collaboration with colleagues in other disciplines can be a powerful motivator for interdisciplinary work. Yet there are many barriers to such collaborations, including the need for more time to plan and develop ideas, to learn another discipline's terminology and conceptualization, for access to appropriate instrumentation, and for flexibility in the structure of majors within disciplines. In terms of preparing students for scientific careers, learning how to work in an interdisciplinary team will be an essential skill for their future success. The challenge in terms of student learning is ensuring interdisciplinary work does not sacrifice preparation in an individual's primary scientific field. For instance, it is impossible to train an undergraduate student to be both an expert biologist and computer scientist in a short period of time. Exposing students to other fields, and developing methods to interact and communicate in productive ways is the major goal we wish to accomplish through the interdisciplinary courses and research projects described below.

The College of Wooster successfully fosters interdisciplinary collaborations. Currently, these types of collaborations are growing at the interfaces of biology/psychology, and biology/chemistry, largely due to funding that the college of Wooster has received from the Howard Hughes Medical Institute (HHMI) [5]. There are similar connections in their infancy with biology/computer science, and biology/physics. Here we will focus on presenting several case studies of interdisciplinary teaching and research between biology and math & computer science departments. From here on we will refer to this interdisciplinary collaboration as computational biology.

This paper is organized as follows. In sections 2 and 3 we present several examples of interdisciplinary teaching and research. Section 4 exemplifies how interdisciplinary communication supports interdisciplinary teaching and research, and section 5 points out the main outcomes of doing interdisciplinary activities. In section 6 we shortly illustrate how interdisciplinary learning aligns within the liberal arts philosophy. Finally, section 7 concludes our major observations about the interdisciplinary activities presented here.

2. INTERDISCIPLINARY TEACHING

Two examples of interdisciplinary courses are presented below: Intro to Bioinformatics (an upper-level course targeted toward computer science (CS) and biological science majors) and Intro to Math Biology (which has no prerequisites and was intended for a nonmajors audience). Both of these courses are co-taught by a two-member faculty team: a biologist and a computer scientist, and a biologist and a mathematician, respectively.

2.1 The Intro to Bioinformatics Course

Bioinformatics is one of the most exciting areas of science today – and its practitioners are in increasingly high demand. Researchers in bioinformatics must be well versed in computer science and molecular biology / biochemistry. Unfortunately, very few College of Wooster students have been able to explore both of these areas through their course work, and even rarer are students pursuing Independent Study projects at the interface of these two disciplines.

Therefore the goal of this course is to give students majoring in computer science or molecular biology an interdisciplinary experience that will allow them to:

• Explore the kinds of questions and opportunities available in this emerging field.

• Better understand the interface between molecular biology and computer science.

• Develop the conceptual framework, tools, and techniques needed for bioinformatics.

• Acquire the skills and confidence needed to work effectively on teams to confront problems and construct solutions relating to bioinformatics research. It is also hoped that some students will ultimately engage in additional interdisciplinary work, possibly pursuing Independent Study projects and/or graduate work in bioinformatics.

W. Morgan of the Biology Department and S. Visa of the Math & Computer Science Department have co-taught the Bioinformatics course twice (2011, 2012). The course organization and the challenges encountered while teaching this course are presented next.

The instructors acknowledge that the most challenging part is teaching students with a very diverse background. Typically, this class consists of two distinct groups: (1) students who have completed at least one year of introductory biology, but have no programming experience; and (2) students who have completed at least one year of introductory computer science, but have no biology background. Of course, a diverse body of students is the challenge of any interdisciplinary course.

To bring each of the two groups to a common ground, from which the class as a whole could start investigating bioinformatics techniques for biological data, S. Visa offered for the first three weeks a crash course in programming for the biology students, and at the same time W. Morgan taught the basics of Genetics to the computer science students. During these three weeks all students worked on homework from both groups such that they could successfully complete their first exam, which comprised both components. After developing a common knowledge basis, the class proceeded for the remainder of the course discussing central bioinformatics problems from [2] (W. Morgan) and working on case studies that required hands-on computer programming (S. Visa).

Another challenging part of this course, in particular for students, stems in its interdisciplinarity nature: this course is inherently difficult due to the nature of problems it tackles and because it combines biology, probability, statistics, and programming. In addition, it covers many topics and there is a lot of information to be absorbed from multiple textbooks. To overcome this difficulty, we constantly illustrate concepts and algorithms with concrete examples in ways that the students can connect the various disciplines. The case studies from the laboratories and the hands-on in class exercises are particularly useful here.

After the first offering (2011), based on students' feed-back and new ideas brought back by W. Morgan from an undergraduate conference on teaching Bioinformatics we made one major adjustments in the 2012 course offering. We moved from Matlab and Perl programming to Python, as the latter software is free, relatively easy to learn, and increasingly used in the bioinformatics area. In addition, Python can integrate a free library dedicated to Bioinformatics, namely the Biopython module [1].

In a one-semester course, we obviously can't develop bioinformatics experts; however, we can achieve our goal of developing scientists who are comfortable and confident working on interdisciplinary team projects. To foster this, students work frequently on small teams composed of computer science and biology majors to complete their hands-on lab assignments, as well as a more extensive bioinformatics project. This project requires students to work together to address an open-ended bioinformatics problem using the computational tools and programming skills developed during the course. For example, students recently annotated a novel bacterial genome sequence and then compared its structure to those of closely related species to identify shared as well as unique features. Within each team, the biology students bring their knowledge and insights to comprehend the problem and suggest possible solutions, while the computer science students complement this by helping to implement the necessary procedures. This collaborative work requires them to develop interdisciplinary communication skills.

2.2 The Intro to Math Biology Course

S. Strand, a biologist, and R. D. Pasteur, a mathematician, jointly developed and taught a non-major course in mathematical biology, which ran for the first time in 2011. Aimed at students pursuing majors in disciplines apart from natural or mathematical sciences, this course had no prerequisites. Students could take this course to satisfy institutional general education requirements in science/math and/or quantitative skills; however, nearly half of the students noted in an end-of-course evaluation that "personal interest" (in the content) was among their reasons for enrolling.

Given the opportunity to educate non-scientists about important scientific issues of our time, we chose a few such issues, based on current relevance, potential interest to students, and the associated mathematical/statistical content that could readily be included. These central topics included infectious disease (focusing on influenza, HIV, and malaria), pharmacokinetics, population and sustainability, and climate change.

The goal of this course was to give non-sciences and mathematics majors an opportunity to:

• Understand the basic biology of several contemporary scientific issues.

• Apply mathematical models related to these issues.

• Comprehend data given in various forms (charts, tables, graphs, descriptions).

• Make qualitative interpretations of quantitative data.

• Produce charts, tables, and graphs that aid in the interpretation of biological data.

• Read non-technical scientific articles with greater comprehension.

A typical week in the course would include short lectures on the biological background of one of the central topics, viewing of documentary excerpts, small-group discussion of non-technical articles (from Scientific American, The New York Times, Newsweek, etc.), and computer-based explorations of mathematical models of the phenomena being studied. The faculty emphasized that mathematics can be used to make future predictions about biological systems (rather than just as a means for analyzing collected data), and that these predictions can then be used for other purposes such as developing methods of treatment, or informing governmental policy decision-making.

Nearly forty students signed up for the course, and they represented many intended major fields and ranged from first-year students through seniors. The students' heterogeneous backgrounds made it challenging to pitch the material at an appropriate level for all students, especially in terms of using scientific and mathematical terminology that was representative of the disciplines, but not overwhelmingly technical. Another obstacle was that many standard models associated with biological phenomena involve upper-level mathematics (such as systems of differential equations), so helping students comprehend the key ideas, without losing them in technical details was a balancing act. Despite these challenges, we found anecdotally that the level of participation was higher than expected for a non-majors science course, and students self-reported substantial gains in their "understanding of the link between math and biology" and their interest in biological science topics.

3. INTERDISCIPLINARY RESEARCH

Three interdisciplinary research examples in the Computational biology area are presented in this section. Besides a short description of each project and its outcomes, challenges encountered are also discussed.

3.1 HHMI Sponsored Summer Research

For three consecutive summers S. Strand and S. Visa have co-advised an interdisciplinary project funded by the HHMI grant [5]. They have been working with three undergraduate students in comparing two related Histoplasma capsulatum genome sequences with a goal toward comparing similar and disparate regions within the genome The students wrote Matlab code and used several free Bioinformatics software programs to make initial gene predictions in these two organisms, to find common and unique genes, and to identify potential genes involved in the pathogenesis of the organism. All these students have taken the Bioinformatics course, which meant the students had the computations skills to begin work on the very first day of the program. The students successfully identified common and distinct genes, and they expanded their initial analysis to predict and compare groups of gene families within the two organisms. The initial findings have been published with two undergraduate students as co-authors at an international conference [8]. Currently we are working on a journal paper co-authored with two undergraduates that presents additional findings (to be submitted in 2012).

Several challenges arise in conducting interdisciplinary research. First, synthesizing the expertise of the faculty members involved was more challenging than we anticipated. Each professor is proficient in their own field of research, yet exploring a question interdisciplinarily requires quickly acquiring a basic understanding of a new research field. While learning a new discipline is intellectually stimulating and rewarding, competing academic demands mean there is often little time for this exploration, and therefore the research can seem more compartmentalized than synthesized. Weekly project meetings were useful in synthesizing our research goals and findings by bringing both faculty and students together to discuss ideas, analyze results, and plan future experiments. A second challenge was the difficulty in to gauging a project of an appropriate scope. A genomic comparison of two eukaryotic organisms is a project that is quite large in magnitude, and we were naïve in our understanding of what we could reasonably accomplish during a single summer of research. An unintended benefit of this initial miscalculation in the scope of our genomic comparison has led to three consecutive summers of collaboration between the same students and faculty thereby deepening both the professor's and the student's understanding of the interdisciplinary questions being addressed.

3.2 Interdisciplinary and Inter-institution Research Collaboration

S. Visa collaborated for the past three years with E. van der Knaap from The Ohio State University – OARDC (Ohio Agricultural Research and Development Center) on an interdisciplinary research project that addresses several problems: (1) improving the Tomato Analyzer (TA), a feature extraction software which computes around 40 morphometric attributes from tomato fruits images [4]; (2) modeling tomato fruits in various categories; and (3) finding genes that regulate size and shape in tomatoes. Of great importance is that this research involves undergraduate students as well.

In 2009 a team of four undergraduates worked under a CS faculty advisor (J. Strecker) on the TA project. They

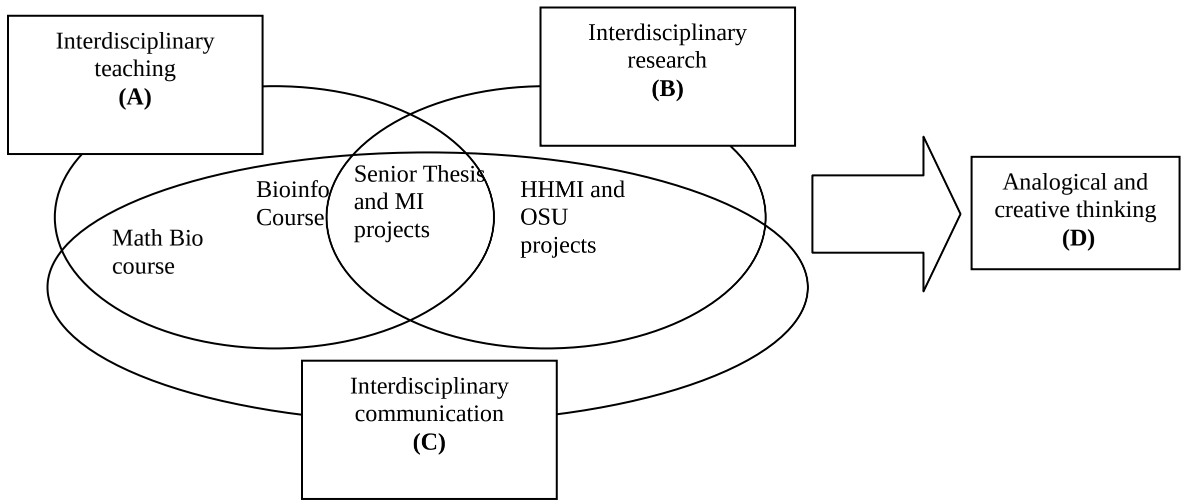

Fig.1 Interdisciplinary teaching, research, and communication develop analogical and creative thinking skills.

worked on fixing existing bugs in the TA software and on implementing new features. The CS students and the faculty advisor discussed in weekly meetings with the OSU biologists the requirements for the TA. The new and improved version was released and the students presented their interdisciplinary research experience at a conference [4].

Further, the data obtained through the TA software was used as class project for the Machine Intelligence class. This is not only an interdisciplinary teaching example, but also a case where teaching informs research and the other way around. This class is a higher-level class that teaches students to analyze data through Machine Learning techniques. It seemed natural to assign the tomato data analysis as class project, in particular because the students were able to immediately see how the theoretical algorithms they learned apply to solving problems in other fields. In a nutshell, they have used classification and feature extraction techniques to classify 800 examples of tomato fruits into nine shape classes such as round, heart, ellipse, etc. The results obtained by the students have been published and presented at a conference [7].

3.3 Senior Thesis Research

One distinctive of the institution's academic program is that every student engages in a year-long senior "Independent Study" research project, culminating in a written thesis and an oral defense. For double majors, these projects blend the two disciplines and are jointly mentored by two faculty members, one from each field. For example, a student completing majors in the biochemistry and molecular biology program and in mathematics used artificial neural network models in a portion of his senior research. He sought to identify patterns of amino acid patterns in arginine kinases (AKs) that differentiated them from other phosphagen kinases (PKs), to allow for prediction of whether a novel PK was also an AK. Given the increasingly computational nature of the natural sciences, increasing numbers of students at Wooster are choosing to pair a major in one of the natural sciences with a computer science or mathematics major. For those aspiring to careers in research, the combination of early interdisciplinary research experience and perspectives from two fields frequently prepare them well for graduate study.

One common difficulty encountered in all these interdisciplinary research projects is the lack of a common vocabulary and shared conceptual foundation between the collaborators. To promote interdisciplinary communication we used analogical examples and analogical thinking. For example, the CS students used their knowledge of programming with strings in framing a computational solution for gene analysis, by recognizing the analogy of strings from CS to the gene sequences in biology.

4. INTERDISCIPLINARY COMMUNICATION SUPPORTS INTERDISCIPLINARY TEACHING AND RESEARCH

From the case studies discussed above we observed that interdisciplinary communication is an essential component of interdisciplinary teaching and research (see Fig. 1). However, as noted above, this type of communication is the most challenging among the three activities.

We will now examine how interdisciplinary communication supports interdisciplinary teaching and research in the case studies from section 2 and 3.

• The Bioinformatics and the Math Biology courses involve (A) and (C). For example, in the Bioinformatics class we first established a common vocabulary and conceptual base to promote interdisciplinary communication.

• The Senior Thesis and the Machine Intelligence class-project illustrate the collaboration between (A), (B), and (C). The students working on these projects applied the theoretical knowledge acquired in courses to the research projects, and further communicated their research and results in a manner that satisfied both disciplines. This is particularly true for the Senior Thesis, as both major requirements must be fulfilled by the same project.

• The HHMI and OSU research projects combine in a higher proportion (B) and (C). The students and faculty working on these projects had to constantly communicate with the collaborators in the other field to advance the research and to obtain the desired results.

5. OUTCOMES OF INTERDISCIPLINARY TEACHING AND RESEARCH

We believe that interdisciplinary learning and research enriches students (and faculty members) in many ways, a few of which are discussed below.

Primarily, through interdisciplinary activities the students develop critical and analogical thinking skills (see Fig.1). They also learn to adapt and transfer their knowledge from one domain to another. In addition, they develop the ability to understand and explain concepts from less familiar domains. For example, at one HHMI research group meeting the faculty advisors (S. Strand, S. Visa) could not find a biological explanation to a computational result obtained when comparing GC content in the two Histoplasma strands. At the following group meeting one of the students (Matt Lambert'13) who is a CS major came with a research paper in biology that, to some extent, explained the controversial result.

By working in mixed teams (i.e. members having different theoretical backgrounds) the students develop the ability to use their complementarity in designing a solution. They also learn to appreciate ¬¬¬that a non-expert, through having a fresh and very different look at a problem, can bring innovating and interesting ideas into discussions.

Finally, faculty interaction on an interdisciplinary level will support institutional enrichment and the generation of new knowledge through innovative research. A deeper understanding of computational biology developed through the interdisciplinary research between S. Visa and S. Strand has lead one faculty member in biology (S. Strand) to integrate discussions of computational research as well as a short computational assignment into an upper-level microbiology course.

One open question is how we measure the learning outcomes from interdisciplinary research and teaching with respect to faculty and to students? At the present time, we do not have direct measures for the learning outcomes discussed here. However the outlined outcomes stem from direct observations from working with the students. In doing interdisciplinary research, successful outcomes are through publication and presentation of scientific findings in highly competitive conferences. In addition, the feedback received from students speaks about the success of these interdisciplinary projects: "Presenting one's hard work and engaging in discussions about it is such an invigorating feeling", stated J.Thomas'11, one of the student researchers who presented at MAICS2011 [7].

6. INTERDISCIPLINARY TEACHING AND RESEARCH WITHIN THE LIBERAL ARTS CURRICULUM

The case studies presented here are examples of interdisciplinary activities carried out at The College of Wooster, which is a liberal arts college. We believe that the liberal arts philosophy in particular fosters interdisciplinary learning as pointed out below.

Firstly, the liberal arts curriculum encourages students to explore various disciplines before focusing on a particular one. It also requires the students to take a generous number of credits outside their discipline of interest. Thus, this cross-discipline model gives the students a broader view, and better prepares them to take on interdisciplinary problems. It also spurs students' interest in more than one discipline. This is reflected in the high number of double majors at The College of Wooster and in the many interdisciplinary research projects existing in our campus.

Secondly, critical thinking, problem solving, and communication skills are central to the liberal arts philosophy as well as to the interdisciplinary learning. Students who participate in interdisciplinary courses or research activities learn that although technical language and content is discipline specific, the approach to problem solving and communicating in any discipline, requires similar skills.

Finally, Cronon [3] emphasizes that a liberal arts educated person is not someone that simply fulfills a list of required courses, but one that constantly works towards acquiring several important personal qualities. He lists and discusses ten such highly desired qualities, but in the end acknowledges that the liberal education is about "the freedom to connect". We acknowledge that interdisciplinary learning and research are connecting experiences as both require the students to connect information from different disciplines when solving problems. In addition the students must connect and effectively communicate with individuals in other fields.

The College of Wooster recognizes the advantages of creating interdisciplinarily trained young scientists and supports interdisciplinary activities. Below we highlight some of the support that has spurred our success in this area:

•The College’s generous Program in Interdisciplinary Studies (PIDS) fully credited each of the participating faculty members who team-taught Introduction to Bioinformatics (rather than splitting the teaching credit). This allowed both members to attend all classes, ensuring the full integration of course activities.

•Both departments (math & computer science and biology) counted the Intro to Bioinformatics as an upper-level elective towards the major, which has encouraged student enrollment.

•Interdisciplinary research projects received support from an HHMI Undergraduate Science Education grant to the College of Wooster [5]. This provided summer stipends for student and faculty researchers, as well as supply money and meeting travel funds.

7. CONCLUSIONS

We present several instances of interdisciplinary teaching and research at The College of Wooster, a liberal arts college. From the case studies presented here we conclude that the interdisciplinary communication is the main vehicle in performing the former two. However, effective communication between individuals having different backgrounds is difficult to achieve. For instance, it is challenging to communicate interdisciplinary lectures to students having different backgrounds. On the other hand, our case studies point out that different backgrounds are particularly useful in research teams, as the team members use their complementarity to find more innovative solutions to problems. In addition, working with individuals having different expertise promotes analogical and critical thinking.

REFERENCES

[1] P.J. Cock, T. Antao, J.T. Chang, B.A. Chapman, C.J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, M.J. de Hoon. Biopython: freely available Python tools for computational molecular biology and bioinformatics. **Bioinformatics** 2009 Jun 1; 25(11) 1422-3. doi:10.1093/bioinformatics/btp163 pmid19304878. PubMed HubMed.

[2] N.Cristianini, M.W. Hahn. **Introduction to Computational Genomics – A case studies approach**, Cambridge University Press, 2007.

[3] W. Cronon. **“Only Connect…”The Goals of a Liberal Education**. The American Scholar, Volume 67, No. 4, 1998.

[4] G. Rodriguez, D. Francis, E. van der Knaap, J. Strecker, I. Njanji, J. Thomas, A. Jack. New Features and Many Improvements to Analyze Morphology and Color of Digitalized Plant Organs Are Available in Tomato Analyzer 3.0. **In Proceedings of the 22nd Midwest Artificial Intelligence and Cognitive Science Conference**, Cincinnati, OH, pages 160-163, 2011, http://ceur-ws.org/#Vol-710.

[5] Howard Hughes Medical Institute (HHMI**) grant to The College of Wooster ($1,000,000) from the Undergraduate Science Education Program**, 2008. http://www.wooster.edu/Academics/Undergraduate-Research/HHMI

[6] **Transforming America’s Scientific and Technological Infrastructure: Recommendations for Urgent Action**, 2006, www.pkal.org/documents/2006 ReportOnReports-National.pdf

[7] J.Thomas, M.Lambert, B. Snyder, M. Janning, J. Haning, Y. Hu, M. Ahmad, S. Visa. Learning Morphological Data of Tomato Fruits. **In Proceedings of the 22nd Midwest Artificial Intelligence and Cognitive Science Conference**, Cincinnati, OH, pages 133-137, 2011, http://ceur-ws.org/#Vol-710.

[8] S.Visa, S.Strand, T.Vrolijk. Genomic Comparison of North American Strains of Histoplasma Capsulatum. **Proceedings of the 3rd International Conference on Future Computer and Communication Conference**, Iasi, Romania, June 2011.