

BUILDING A BIOINFORMATICS BRIDGE ONE COURSE AT A TIME

Lynn R. Heinrichs, *Elon University*, lheinrichs@elon.edu
Linda M. Niedziela, *Elon University*, lniedziela@elon.edu

ABSTRACT

The enrollment downturn faced by today's information systems programs is serving as a catalyst for change. Faculty members are seeking ways to re-chart their curricula and broaden their potential student audiences. One opportunity for reaching new audiences is building bridges with other programs. This paper describes the experience of one information systems program in forging a partnership with biology to develop a bioinformatics course for non-majors in an honors curriculum. The partnership is unique in its involvement of information systems instead of computer science and its target audience of non-majors.

Keywords: Bioinformatics, IS curriculum, team teaching, non-majors, interdisciplinary

INTRODUCTION

Bioinformatics is “an interdisciplinary area that applies computer and information science to solve biological problems [10].” Bioinformatics is particularly important in genomics research because of the large amount of complex data involved.

Bioinformatics offers a natural partnership between the computing disciplines and biology. Many partnerships have been in place for years including undergraduate and graduate degree programs, minors, and concentrations. However, many of these partnerships require numerous prerequisite courses in the sciences; rarely can a non-major enroll without taking either an introductory biology or computer science course, or both. Additionally, the partnerships on the computing side are most often with computer science; information systems is much less likely to be involved in these interdisciplinary efforts.

The computer information systems program at the authors' institution has suffered from declining or flat enrollments over the past five years. Faculty members in the program have been actively seeking ways to reach new audiences. The institution supports an honors program that requires students to enroll in an interdisciplinary seminar during the sophomore year. When honors course proposals

were solicited in Spring 2007, a member of the CIS faculty and the biology department proposed an interdisciplinary course in bioinformatics. The proposal was accepted, and the first offering of the course took place in Spring 2009.

This paper describes the creation and delivery of the interdisciplinary bioinformatics seminar for honors students. The course was team-taught by two faculty members (the authors), one in computer information systems and the other in biology. Students enrolled in the course came from a diverse set of backgrounds including human services, history, communications, and biology. The paper provides an overview of the course design process, course implementation, and future curricular opportunities.

A REVIEW OF BIOINFORMATICS CURRICULA

The authors began their course design process by reviewing existing bioinformatics curricula and resources. Even though bioinformatics is a relatively new field, there are many programs of study, published papers, and course materials available to guide the curriculum development process.

The authors' situation was unique in comparison to most bioinformatics curricula described in the literature. First, most interdisciplinary partnerships are forged between computer science and biology, not information systems and biology. Second, most of the examples in the literature are geared toward majors that have prerequisite coursework either in computer science, biology, or sometimes both. The authors identified the following challenges in their own course design for which they hoped to find best practices in action:

1. How can a course be structured for a diverse group of non-majors with limited or no prerequisites?
2. What is the best approach for integrating the information systems (IS) and biology content? How can the expertise of both faculty members be utilized?
3. What types of laboratory assignments and projects are appropriate?

The following discussion highlights sources that influenced the authors' final course design.

As a starting point, the authors found a summary of bioinformatics degree programs developed by Hemminger, Losi, and Bauers [7] at the University of North Carolina (UNC), Chapel Hill, Web site [9]. By far, the majority of programs listed were advanced degree programs. However, some B.S. degrees as well as undergraduate minors were included. The site provided a summary of degree requirements, such as credit hours, and a sampling of research interests. The information was useful as a starting point for identifying colleges or universities that have some type of bioinformatics curricula or coursework. However, the list was compiled from responses to a survey so many institutions with bioinformatics programs were not listed.

A "best practice" program from the UNC list was offered by the University of Arkansas, Little Rock. Its bioinformatics minor is also described in a paper by Bruhn and Jennings [1]. The minor is geared towards biology or computer/information science majors with a slightly different structure depending on a student's program of study. Three characteristics of the minor were relevant to the current problem: (1) the role of information science in delivery of the program, (2) the lack of technical/scientific prerequisites (just college algebra and computer literacy), and (3) the inclusion of both programming (Perl) and database coverage in the introductory bioinformatics course.

Aside from the UNC list, a literature review of bioinformatics courses also revealed some relevant undergraduate model programs. San Jose State University offers a bioinformatics track in computer science [9]. The track includes two courses - Bioinformatics I and II - with both biology and programming prerequisites. Programming projects and research papers are utilized as learning tools. Providence College offers a bioinformatics course with no formal prerequisites for computer science and biology majors. One CS student is paired with one biology student. The biology component of the course focuses on an on-going research problem while the CS component emphasizes Perl programming labs [12].

COURSE DESIGN

A major theme that emerged from the review of bioinformatics curricula was the role of software tools – either creating them using a programming

language (e.g, Perl) or using them to search on-line biological databases. The authors believed that focusing on the creation and utilization of tools (1) could be adapted to a non-majors audience, (2) provided a reasonable approach to integrating the IS and biology content, and (3) leveraged the expertise of both faculty members.

Final Course Structure

The general course design for the honors bioinformatics course entitled "Bioinformatics: The Digital Code of Life," is shown in Table 1. The course was divided into four modules that provided for an introduction, separate units on the creation and utilization of tools, and a component relating to social, ethical, and legal issues.

The authors chose to introduce programming before on-line tool utilization for two reasons. First, the programming problems could be used to introduce/review basic biology and genetics terminology. Some students would not have a strong biology or science background. Introducing terminology through programming problems seemed like a reasonable way of giving students working knowledge of such terms as "amino acid" and "nucleotides." Additionally, programming in Perl would allow students to see some of the underlying techniques and data formats that provide the foundation for using online Web resources.

In addition to deciding on the best course design for delivering the bioinformatics content, the authors also had to address three conditions set forth by the institution's honors program. First, the course needed to satisfy two areas of study in the general studies program (expression, civilization, society, or science). Second, it had to be truly team-taught by two faculty members present in the classroom every day. Finally, the course was responsible for improving students' research skills in preparation for the senior honors thesis.

General Studies Contributions

Designing the course to meet two areas of general study was not straightforward. Biology met the science requirement. However, unlike computer science, information systems coursework did not count for any of the general studies areas. Therefore, the authors had to make a case that the content of the course fulfilled either an expression, civilization, or society requirement.

Since the authors planned to emphasize the application of problem-solving to real-world scenarios that considered the societal, legal, and ethical issues, they successfully made the case that the course also addressed the society area of general studies. The final project for the course became one of the major vehicles for ensuring that the society area was adequately addressed.

Team Teaching

Creating a course structure that leveraged the expertise of both faculty members was challenging. Both biology and information systems require technical expertise. The IS faculty member had not taken a formal biology course in forty years. The biology faculty member had no programming experience.

In the world of bioinformatics, it is generally considered more difficult for the computing person to learn the necessary biology than vice versa [11]. Therefore, in preparation for teaching the course, the IS faculty member audited the genetics course of the biology faculty member. This also provided an opportunity to understand and discuss some of the potential issues related to what to teach and how to teach bioinformatics.

As the authors began plotting out the details of the course design, there were natural points at which one faculty member served as lead in organizing and delivering instruction. However, the instructors were always present, working together in the classroom. At no point did delivery become a tag-team effort.

Thesis Preparation

Students in the honors program are required to complete a senior thesis, and the second-year honors seminar should help students prepare for that challenge. Therefore, the seminar is expected to provide some type of research project. The authors struggled with how to approach this requirement, and when the course started, all that was specified on the syllabus was a final project assignment. After the course got underway and the instructors had an opportunity to work with the students, a final research project was developed where students created a public information product examining the impact of the Genetic Information Nondiscrimination Act (GINA) on a topic related to the course.

COURSE IMPLEMENTATION

The first offering of the course took place in Spring 2009. Even though the instructors spent numerous hours planning and preparing for delivery of the course, many uncertainties loomed as the start of the semester approached.

Honors students were well-known for their concerns about grades. They would prefer a well-defined, very predictable system of assessment. The instructors, on the other hand, felt that they needed some latitude in conducting the course and assigning work. A system of grading was chosen that relied on percentages instead of absolute points, allowing for flexibility in the number of assignments and point values for each module. No tests were planned; laboratory assignments and the final project were selected as the primary means of assessment. With the exception of the final project, students were expected to complete most assignments individually rather than in groups.

Honors students also are often uncomfortable in open-ended, less traditional learning environments. Since this course was designed to provide students with just enough information as needed to understand the tools they were creating and using without becoming “programmers” or “bioinformaticians” the course structure needed to provide adequate support so students could be challenged without becoming disenfranchised. The instructors accomplished this through intentional design of assignment and projects, scaffolding of projects, in-class work sessions and instructor feedback.

The following discussion highlights some of the assignments employed during each module of the course.

Introduction

The instructors were very concerned about getting student buy-in for the course early on. Students enrolled in the course had a choice between the bioinformatics seminar and a course on exploring consciousness. The other course quickly filled to capacity, and some students enrolled in bioinformatics because there was no other option. The authors were very concerned that the technical/scientific nature of their course would intimidate some students.

In addition to other general introductory material in the first week, students were able to get hands-on experience with how sequencing data is created before being stored as a record in a database. In a

series of wet labs students isolated their DNA and amplified a piece of their mitochondrial DNA using the PCR (polymerase chain reaction) technique. This DNA was sent out for sequencing so that it could be analyzed using online databases and bioinformatics tools during the third module of the course. The students were very excited about the laboratory exercise.

Creating Tools through Perl Programming

A questionnaire at the start of the course revealed that only one student had any previous programming experience, but many were concerned about learning a programming language. The lead instructor for this part of the course was concerned with trying to (1) minimize the number of learning curves students would need to traverse and (2) select problems that would be meaningful for introducing/reviewing biology concepts.

Two freely available products for the Windows environment were used for program development and testing: Crimson Editor [3] and Cygwin [4]. Crimson Editor (CE) is a simple to use text editor that provides Perl support. Cygwin (CW) is a Linux-like environment for the Windows platform. In order for students to use Cygwin, one class period was needed to teach the basic use of Linux commands.

Because of the limited timeframe for teaching Perl in a bioinformatics context, the instructor used a “programming by example” approach. Students completed four programming problems that introduced such topics as arrays, loops, conditionals, string functions, regular expressions, and file I/O. The last problem required students to search gene data in a file, providing a good transition to the third module of the course.

An open-content textbook entitled, *Beginning Perl*, by Simon Cozens [2] was selected for the course. The book provides a general introduction to Perl, not necessarily for bioinformatics problems. Chapters often contained more content than needed for the limited course coverage, and many examples were difficult to relate to the problems in class. For the next offering of the course, a Perl textbook written specifically for bioinformatics might make more sense.

In many ways, the problems honors students experienced with programming were no different than non-honors students. They were frequently frustrated by the strictness of programming syntax, and often experienced difficulty with the transfer of

learning from class to homework assignments. Perhaps the most noticeable difference between the honors students and a non-honors group was their commitment to finishing the problems. They loved the feeling of exhilaration that came with a working program, but they were impatient with the debugging process and required a substantial safety net of support.

Using Online Databases for Human Disease Analysis

The online database component of the course was built upon an inquiry-based laboratory curriculum developed at Washington University in St. Louis, through funding from the Howard Hughes Medical Institute [5] for a Principles of Biology III, Bioinformatics Lab. The following course description explains the approach:

This inquiry-based lab is designed around genetic diseases with a focus on protein structure and function. To allow students to work on their own investigatory projects, ten projects on ten different proteins were developed. . . Students use bioinformatics tools to investigate their proteins and form a hypothesis for the effect of the mutation on protein function.

The curriculum Web site [5] provided self-contained, instructional materials including prepared laboratory manuals for ten different human diseases. After completing an initial tutorial, students were able to follow the laboratory work associated with their assigned human diseases. Labs involved background readings, database search activities, and written questions.

The lead instructor for this module of the course assigned each student a disease to investigate. Since there were more students than diseases, some students investigated the same disease, but on an individual basis.

While the lab manuals provided definitive direction on what steps to complete and what information to obtain from the databases, the interpretation of the information required some analysis and synthesis. Because the instructional materials were developed for a majors course with prerequisites, additional class time was spent during this module introducing molecular genetics concepts as they were needed to progress through the stages of analysis. Also, since the completion of the human disease analysis labs occurred at the same time during the semester as project due dates from other courses, this impacted

students' schedules and the final project for the bioinformatics course. Like students in any other course, the honors students were challenged to complete all their end-of-semester work. The final, culminating assignment for the human disease analysis varied in quality from outstanding to good.

The Final Project

The final project became the primary vehicle for ensuring the course met the society requirement for general studies. On May 21 2008, then President George W. Bush signed into law the Genetic Information Nondiscrimination Act (GINA) to protect Americans and their genetic information. The bill paves the way for people to take full advantage of the promise of personalized medicine without fear of discrimination in such areas as health insurance and employment [6]. GINA provided a relevant connection between the scientific themes of the course and the social, legal, and ethical implications of bioinformatics.

For the final project, students worked in small groups of two or three to research a topic within bioinformatics and/or human genomics and its relationship to GINA. Information products could involve any delivery system such as a hardcopy brochure, Web site, or video presentation. Because of the importance of balancing the technical and ethical, legal and social aspects of the topics, students had multiple opportunities to receive instructor feedback as they developed their final product. The most popular information product was the brochure/catalog, selected by all but two groups. The two exceptions were implemented using a Web site and video. Examples of project topics included designer children, genetic testing, and prenatal testing. On the day of the final exam groups delivered a presentation about their topic and demonstrated the informational product for the class.

LOOKING TO THE FUTURE

The authors demonstrated that a bioinformatics course could be successfully delivered to a non-majors audience. The next offering of the bioinformatics course will be in Spring 2011. The authors believe the general structure of the course worked well, although alterations certainly need to be made. The programming content in the first half of the course probably needed more time; content in the second half of the course may have needed less. Additional integration of programming skills into the bioinformatics modules would also be beneficial.

Several opportunities exist to move beyond the current experimental course. The university is committed to expanding and enhancing interdisciplinary science initiatives. This could take place through the introduction of general studies courses for all students or even an interdisciplinary bioinformatics minor. Additionally, the biology department recently hired a faculty member with bioinformatics expertise who will bring fresh ideas on how the curriculum can be developed.

Delivering the bioinformatics course required a tremendous amount of time and effort. The challenges were many: integrating content from two technical areas for a non-majors audience, meeting the specific requirements of the honors program, and developing a successful team-teaching approach. The authors believe the bioinformatics course was a success and that they have paved the way for future interdisciplinary efforts.

REFERENCES

1. Bruhn, R., & Jennings, S. F. (2007). A Multidisciplinary Bioinformatics Minor. *SIGSCE '07*, (pp. 348-421). Covington, KY.
2. Cozens, S. (2009). *Beginning Perl*. The Perl Foundation. Retrieved 3 30, 2009 from: <http://www.perl.org/books/beginning-perl/>.
3. Crimson Editor (n.d.). Retrieved 3 31, 2009, from Crimson Editor: www.crimsoneditor.com.
4. Cygwin Information and Installation. (n.d.). Retrieved 3 31, 2009, from Cygwin: www.cygwin.com.
5. Elgin, S. (2005). *Principles of Biology III Bioinformatics Lab*. Retrieved 3 30, 2009, from Genomics in Education, Washington University: <http://www.nslc.wustl.edu/elgin/genomics/bio3055.html>.
6. Genetic Information Nondiscrimination Act: 2007-2008. (2008, May 21). Retrieved 3 31, 2009, from National Human Genome Research Institute: <http://www.genome.gov/24519851>.
7. Hemminger, B., Losi, T., & Bauers, A. (2005). Survey of Bioinformatics in the US. *Journal of the American Society for Information Science and Technology*, 56 (5), 529-537.
8. Informatics Degree Programs Summary. (n.d.). Retrieved 3 31, 2009, from School of

Information and Library Science, University of North Carolina at Chapel Hill:
http://ils.unc.edu/informatics_programs/.

9. Khuri. (March 12-15, 2008). A Bioinformatics Track in Computer Science. *SIGSCE '08*, (pp. 508-512). Portland, Oregon.

10. Minnesota, U. o. (2009). Graduate Program in Bioinformatics. Retrieved 3 31, 2009 from University of Minnesota:
<http://www.binf.umn.edu/>.

11. Tisdall, J. (2001, October 15). *Why Biologists Want to Program Computers*. Retrieved 3 31, 2009, from O'Reilly:
http://oreilly.com/news/perlbio_1001.html.

12. Toth, C., & Connelly, C. (2006). A Bioinformatics Experience Course. *Journal of Computer Science in Colleges*, 21 (6), 100-107.

Table 1. Bioinformatics: The Digital Code of Life Course Outline

Course Module	No. of Weeks	General Topic	Biology Contribution	IS Contribution
1	1	Introduction to the Course	Introductory Biology and Programming Concepts	
2	5	Creating Tools with Perl	Biology Problems/Applications	Perl Programming
3	5	Using Tools for Human Disease Analysis	Human Disease	Using On-line Databases
4	3	Understanding Social, Legal, and Ethical Issues	Final Project – GINA Information Product	