Bioinformatics: An Example of a Cooperative Learning Course[*]

Namyong Lee[1], Dept. of Mathematics & Statistics, Minnesota State University, Mankato
Ernest Boyd, Dept. of Mathematics & Statistics, Minnesota State University, Mankato

Institution	Minnesota State University, Mankato
Size	about 15,000 students
Institution Type	medium size state university with Master of Science
Student Demographic	recent high school graduates with high potential and interests in math, biology or computer science
Department Structure	Mathematics, Biology, and Computer Science are separate departments in the College of Science, Engineering & Technology

Abstract

In this paper, we describe how our bioinformatics course was born as an interdisciplinary and cooperative learning course. Three departments---Mathematics and Statistics, Biology, and Computer and Information Science---participated in this project. We describe how our team designed the course, prepared group projects, and secured funds for it.

Course Structure

- Weeks per term: 15-week regular semester course
- Classes per week/type/length: Two75 minute lecture/computer lab (hybrid) per week
- Labs per week/length: Two 75 minute lecture/computer lab (hybrid) per week
- Average class size: 12-20 students
- Enrollment requirements: Each student must interview one of the three instructors to enroll in the course
- Faculty/dept per class, TAs: Team-taught by three instructors, one from each of the mathematics, biology, and computer science departments

---

[1] namyong.lee@mnsu.edu

[*] supported by MnSCU CETL Grant

- Next course: Students are encouraged to take Mathematical Models in Biology (Math 490/ Biol 490). Both Bioinformatics and Math Models in Biology courses focus on interdisciplinary undergraduate research
- Website: https://d2l.mnsu.edu/ (need class account)

Introduction

Colleges and universities face significant challenges when they plan to offer bioinformatics courses to undergraduate students. The literature strongly suggests the necessity of truly interdisciplinary courses where students from different majors work closely in small teams on bioinformatics research problems (National Research Council 2003; Bialek and Botstein 2004; Steen 2005; Fetrow and John 2006). Moreover, the courses ought to be taught by a team of instructors from different disciplines, an unusual situation for most traditional institutions of higher education.

The departments of Mathematics and Statistics, Biology, and Computer and Information Science at Minnesota State University, Mankato, have created an interdisciplinary and cooperative learning bioinformatics course. It is taught by three faculty members (one from each department) and draws advanced undergraduate and graduate students from other departments as well as those three. The course is taught as a problem-solving seminar focused around group projects. Students and instructors are expected and encouraged to learn from each other. We do not expect each student to have a complete background in all three disciplines. Students work in interdisciplinary teams so that each student can contribute prior expertise in one of the fields.

The course is listed for three credits in each department so that students can register within their own disciplines. The dean of the College of Science, Engineering, and Technology agreed to pay a one-credit overload to each faculty member. This kind of course credit split is one of the constraints that make a team teaching project more difficult to administer. Fortunately, our university provides some financial support from small internal grants as an incentive to faculty members to create interdisciplinary courses.

Designing a Course

The course is still evolving. Here we describe the current design.

Faculty: Our design team consisted of six faculty members, two each from Mathematics and Statistics, Computer and Information Science, and Biology. In a given year, primary responsibility for the course is given to a team of one faculty member from each department. Having more people on the design team than are needed to teach the course allows us to deal with retirements and sabbatical leaves, and initially helped build the knowledge base necessary to prepare the course materials and group projects.

Students and Prerequisites: Initially we wanted to require introductory courses in statistics, discrete mathematics, data mining, programming, molecular biology, and physiology. We realized, however, that we might not have enough students who could satisfy all the prerequisites. Thus, we posted our course announcement without any prerequisites, but required students to obtain an instructor's permission before registration. In this manner we could make sure that each student would bring enough advanced knowledge to be a productive member of a team. Unexpectedly, students came from more than the three departments. The enrollment ranged from 12 to 20 each semester.

Scheduling Logistics: The next challenge was developing a course listing that would be relevant to students from each department. Fortunately, we have a generic seminar course in each department and could use those course names. Students can register for their own department's seminar course with the course title Bioinformatics Seminar on their transcripts without a lengthy procedure to create a new course. Scheduling a classroom is another challenge in multi-disciplinary courses, because many departments do not communicate with each other when they schedule their courses and classrooms. We have been lucky to have a classroom at the Academic Computer Center, which solves the space problem and also makes it easier to do computer labs.

Lectures and Projects: The course has two major parts. The first half is an introduction to standard bioinformatics topics, such as molecular biology, sequence alignment, phylogenetic trees, microarrays, statistical analysis for biological data, and data mining. Students have reading assignments to broaden their knowledge and to discuss as a group. Each student is also asked to build a bioinformatics glossary journal to learn the terminology to communicate with a multi-disciplinary team.

Before the second half of the course, we introduce several research problems and divide the class into groups, each with at least one biology student, one mathematics student, and one student familiar with computer programming. In the second half, we focus on group projects, covering material necessary for the projects. In the last week of the course, each group presents the results of their research project orally and submits a written report for a course grade.

Class meets twice a week, with each 75-minute class period typically consisting of 40-45 minutes of lecture and discussion and 30-35 minutes of computer lab. Thanks to our grant, we also have had outside speakers who have given inspiring overviews of topics in bioinformatics and ideas for new projects.

Computer Lab: An essential component of the course is the computer lab. Each instructor demonstrates software with tutorials. We have used R and Bioconductor for statistical analysis and visualization of microarray data, Orange and Weka for data mining, CLC Free Workbench, BioEdit, ClustalX, TreeV32, and Cn3D for sequence alignment and phylogenetic trees. Most of these programs are in the public domain. We make a course resource CD-Rom and distribute it so that students can get familiar with the tools on their own computers. We have also used web databases and their tools, such as NCBI, BioCyc, and ExPASy.

Suggestions

Through years of experience with the course, we have seen many encouraging possibilities as well as obstacles to interdisciplinary courses. One great discovery has been the joy of working together with faculty in other disciplines. Unlike experimental sciences, such as biology, or team-project-oriented computer science, mathematicians do not frequently work in teams of people with complementary expertise. Faculty members can demonstrate to students how to collaborate with each other and contribute to a team project.

It is important to present the three disciplines as interconnected and interwoven. It would be a mistake to divide such a course into three separate presentations. When the members of the team present mathematical concepts, we often ask the group “Is this mathematical assumption reasonable for biology?” When they present a computer algorithm, we ask, “Is this method robust and efficient for large data sets? How sensitive are the results to parameter values?” It is easier to understand how to use a BLAST search at NCBI if it is preceded by a discussion of the mathematical concepts behind the Smith-Waterman algorithm and the Needleman-Wunsch algorithm. To appreciate how to use ClustalX and TreeV32 for phylogenetic trees, students need a discussion of neighborhood joining trees, UPGMA, and maximum likelihood algorithms. Clustering algorithms to analyze microarrays need to be accompanied by mathematical foundations for distance measures.

Support is needed from the departments and the college, including compensation for extra or reassigned teaching, flexibility to adjust departmental schedules, and even moral support from colleagues. It works better to confer with our Dean as a team rather than as individuals. Many institutions now realize the need for such interdisciplinary work, so we have a better chance to gain support.

Computer lab and hands-on experiments are essential. Selection of software is also important. For instance, there are many different software packages for sequence alignments and phylogenetic trees, each with its own pros and cons. To conserve time, only one should be presented.

Collective efforts to apply and receive funding are important to building the faculty team and to giving life to the course. We started with a small grant from our university system for new course design and then applied for larger funds such as an NSF CCLI grant. The process of applying for an outside grant is worthwhile, as it provides an opportunity to discuss what we can do and what we wish to do with the teaching and learning project.

Building a quality term project is another time and energy consuming task. Experts in the field should be consulted. Good ideas are available and do not need to be reinvented. We invited a number of external speakers that range from mathematicians to ecologists. They delivered insightful lectures and suggested team projects.

Short courses or workshops are useful to extend knowledge and to exchange ideas. MAA’s PREP workshop (http://www.maa.org/prep), MBI (http://mbi.osu.edu), NIMBioS (http://www.nimbios.org), or DIMACS (http://dimacs.rutgers.edu) all have excellent programs. Joining a national society such as SMB (http://www.smb.org) and MAA’s BioSIGMAA (http://sigmaa.maa.org/bio) is another way to extend knowledge and network.

Conclusion

Creating an interdisciplinary seminar using cooperative learning strategies is rewarding but difficult. The team of faculty members must have the support of the administration. They must be able to interact and complement each other's work. Students need to be self-motivated. The curriculum must be planned so that the component disciplines are not separate strands, but interwoven. Bioinformatics is a rich and challenging field of study, and we need to bring those characteristics to the course.

References

Bialek, W. and D. Botstein, 2004: Introductory Science and Mathematics Education for 21st-Centry Biologists. *Science*, **303**, 788-790.

Campbell, A. M., and L. J. Heyer, 2007: *Discovering Genomics, Proteomics and Bioinformatics*. 2nd ed. Benjamin Cummings, 352 pp.

Fetrow, J. and D. John, 2006: Bioinformatics and Computing Curriculum: A New Model for Interdisciplinary Courses. *SIGCSE'06*, Houston, TX, Association for Computing Machinery, 185-189.

Isaev, A., 2004: *Introduction to Mathematical Methods in Bioinformatics*. Springer-Verlag, 312 pp.

National Research Council, 2003: *Bio 2010: Transforming Undergraduate Education for Future Research Biologists*. National Academy Press, 208 pp.

Steen, L. A. Ed., 2005: *Math & BIO2010, Linking Undergraduate Disciplines*. Mathematical Association of America, 150 pp.