Establishing a master’s degree programme in Bioinformatics: challenges and opportunities

N.V. Sahinidis, M.T. Harandi, M.T. Heath, L. Murphy, M. Snir, R.P. Wheeler and C.F. Zukoski

Abstract: The development of the Bioinformatics MS degree program at the University of Illinois, the challenges and opportunities associated with such a process, and the current structure of the program is described. This program has departed from earlier University practice in significant ways. Despite the existence of several interdisciplinary programs at the University, a few of which grant degrees, this is the first interdisciplinary program that grants degrees and formally recognises departmental specialisation areas. The program, which is not owned by any particular department but by the Graduate College itself, is operated in a franchise-like fashion via several departmental concentrations. With four different colleges and many more departments involved in establishing and operating the program, the logistics of the operation are of considerable complexity but result in significant interactions across the entire campus.

1 Introduction

Across engineering and science, there has been a rapidly expanding interest in the role of biology as a foundational science. Many chemical engineering departments, for instance, have recently changed their names to chemical and biomolecular (or biological or biomedical) engineering, in recognition of the fact that biology has reached equal standing with chemistry as a chemical engineering foundational science. More broadly, research involving biological and biochemical systems is now commonplace in most engineering disciplines. Furthermore, the complexity of biological problems has attracted the interest of many computer scientists and mathematicians to problems ranging from protein structure prediction to biochemical reaction modelling and biomolecular imaging.

Concurrently with this biological revolution, the last few decades have witnessed ever-increasing computer power which has led to the accumulation of vast amounts of chemical, biological, and medical data. As a result, many scientific fields are becoming increasingly data-driven, as opposed to model-driven. For instance, in parallel with research towards understanding the etiology of cancer at the molecular and cellular levels, efforts are now undertaken to mine gene and patient databases with the goal of increasing our ability to correctly diagnose cancer patients and design efficient drugs based on data mining and other search algorithms.

These changes in the role of biology and computing in science, medicine, and engineering have led to a strong demand for a new generation of scientists and engineers with a strong grounding in both biology and computer science. In their 1999 report to the Alfred P. Sloan Foundation on bioinformatics and computational biology hiring patterns, economists Stephan and Black concluded that ‘the number of individuals currently enrolled in formal programs falls far short of the number of positions that have recently been advertised’ [1]. In the five years that followed, a dramatic increase occurred in the number of academic training programs in bioinformatics as well as the number of individuals enrolled in these programs, while the number of bioinformaticians produced was still fairly small relative to the number of position announcements [2]. As the amount of biological data grows at an ever-increasing pace, it can be expected that demand for bioinformaticians will continue to be very strong.

These research and technology trends have provided a strong impetus for changes in undergraduate and graduate education and led to the creation of many new programs in academic institutions. In contrast to traditional programs, the interdisciplinary nature of bioinformatics has required the formation of interdepartmental collaborations towards building new bioinformatics programs. The complexity of the subject is further compounded by the fact that the very term bioinformatics means different things to individuals with different academic backgrounds. Despite these major challenges, the development of bioinformatics programs also provides a unique opportunity for cross-fertilisation between different academic departments, leading to new educational and research initiatives.

The remainder of this paper is structured as follows. We next provide a brief assessment of the current status of bioinformatics education in the United States, followed by an examination of various definitions of the term bioinformatics. Our own definition motivates the structure of our degree program and its implementation. Finally, we outline the process that led to establishing the Bioinformatics MS degree program at the University of Illinois, describe the program’s current structure, and detail its first two departmental concentrations.

© IEE, 2005

IEE Proceedings online no. 20050033
doi:10.1049/ip-syb:20050033
Paper first received 15th June and in revised form 22nd September 2005
The authors are with the University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
E-mail: nikos@uiuc.edu


269
2 The status of bioinformatics programs in the United States

The major challenge in establishing bioinformatics research and educational programs is that trainees must be exposed to a very wide body of knowledge that is rapidly evolving at the interfaces among the life sciences, mathematical and computer sciences, and engineering. Many universities are responding to this challenge by establishing interdisciplinary centres, programs, and degrees. A recent survey of 78 bioinformatics programs in US universities is presented by Hemminger et al. [3]. The specifics of these bioinformatics programs vary considerably, including degree names and requirements. However, the following general trends can be observed:

- Bioinformatics degrees or concentrations are at this point offered in all parts of the country and by all types of academic institutions, including predominantly research universities, teaching colleges, private and state schools.
- There is considerable variation in terms of program contents, with certain programs having stronger algorithmic content and other programs having a more applied focus. No two programs are very similar.
- Because bioinformaticians must be trained in biology as well as computer science, most programs require a significant number of prerequisites. Many students wishing to enter these programs have to take remedial courses that do not count towards their degree requirements.
- The majority of the bioinformatics programs currently offered arose through the interaction of faculty from two or more traditional departments.

3 What is bioinformatics?

The wide diversity in terms of the contents and structure of existing bioinformatics programs is not a coincidence but rather a manifestation of the current lack of a firm consensus on what exactly bioinformatics comprises. As of the time of writing, the term bioinformatics does not appear in the on-line edition of Encyclopedia Britannica. Definitions found in other sources include the following:

- A field of biology concerned with the development of techniques for the collection and manipulation of biological data, and the use of such data to make biological discoveries or predictions. This field encompasses all computational methods and theories applicable to molecular biology and areas of computer-based techniques for solving biological problems including manipulation of models and datasets [4].
- Bioinformatics is the science of developing computer databases and algorithms for the purpose of speeding up and enhancing biological research [5].
- Bioinformatics is conceptualising biology in terms of molecules (in the sense of physical chemistry) and applying ‘informatics techniques’ (derived from disciplines such as applied mathematics, computer science and statistics) to understand and organise the information associated with these molecules, on a large scale [6].

The National Library of Medicine definition [4] uses the term ‘computational biology’ as equivalent to bioinformatics. Certain definitions restrict the term bioinformatics to apply only to molecules, whereas others include applications of computer science to living things above the molecular level, to population biology, systematics, epidemiology, and health informatics. For recent introductory surveys and further discussion of alternative definitions of bioinformatics, the reader is referred to works by Cohen [7] and Fenstermacher [8]. Despite the lack of a unified definition of bioinformatics, very few would disagree on what are some of the most important milestones in the development of the field. A timeline is presented in Table 1.

Where all existing bioinformatics definitions agree is that bioinformatics involves biology and informatics. Therefore, at the very least, a bioinformatics program needs to cover some elements of biology and informatics. The specific subjects that should be covered from biology and informatics, and perhaps other sciences, would be a consequence of the specific definition of bioinformatics adopted by the offering institution. For our purposes, we define bioinformatics as follows:

Bioinformatics is the discipline dealing with all aspects of collecting, analysing, and using data of biological origin to draw inferences.

This definition has several important implications:

- The ‘collection of data’ need not involve computers. In practice, a variety of experimental techniques are used to generate bioinformatics data. Hence, a bioinformatics program needs to offer coverage of these experimental techniques, preferably in terms of a lab course that familiarises students with the interplay between laboratory experiment and computation.
- The ‘use of data’ implies applications. Hence, the field extends to aspects of applied informatics in engineering and medicine.
- The discipline deals with ‘data’ and, hence, the definition at first may give the impression that bioinformatics is an entirely data-driven science. Yet, bioinformatics data can be gathered via quantum chemistry calculations or molecular dynamics simulations, and analysed via other model-driven scientific approaches.

Hence, our definition implies that bioinformatics could employ a wide range of tools from modern computational science and engineering. In this sense, chemistry, physics, biology, mathematics, and computer science are all foundational sciences of the field of bioinformatics, a field that can therefore be as much data-, model-, and hypothesis-driven as are its foundational sciences.

Table 1: Timeline of discovery and innovation in bioinformatics

<table>
<thead>
<tr>
<th>Decade</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950s</td>
<td>First protein (insulin) sequenced</td>
</tr>
<tr>
<td>1970s</td>
<td>Protein Data Bank created</td>
</tr>
<tr>
<td>1980s</td>
<td>First viral genome (bacteriophage φX174) sequenced</td>
</tr>
<tr>
<td></td>
<td>Dynamic programming algorithms introduced</td>
</tr>
<tr>
<td></td>
<td>Use of the term ‘genomics’ began</td>
</tr>
<tr>
<td></td>
<td>SWISS-PROT created</td>
</tr>
<tr>
<td>1990s</td>
<td>First bacterial genome (H. influenzae) sequenced</td>
</tr>
<tr>
<td></td>
<td>BLAST program made available</td>
</tr>
<tr>
<td></td>
<td>Use of term ‘bioinformatics’ started</td>
</tr>
<tr>
<td>Today and tomorrow</td>
<td>Human genome sequenced</td>
</tr>
<tr>
<td></td>
<td>Proteomics unfolds</td>
</tr>
<tr>
<td></td>
<td>Metabolomics becomes focus area, . . .</td>
</tr>
</tbody>
</table>
4 Challenges and opportunities in establishing a bioinformatics program

4.1 Bioinformatics and chemical engineering

In 2001, the faculty of the Department of Chemical Engineering at the University of Illinois decided to change their departmental name to the ‘Department of Chemical and Biomolecular Engineering’. The change reflected the increasing role of biology in departmental research activities and prompted a discussion of relevant educational changes at the undergraduate and graduate levels. As a result, the development of a master’s degree program in the area of bioinformatics began. Chemical engineers have long been involved in bioinformatics, especially in those aspects of systems biology that address metabolic network modelling and design (cf. [9]). It was quickly felt that a program covering traditional bioinformatics topics (database management, string matching, and so on) would satisfy most immediate needs of trainees. However, modelling and design of metabolic networks has been considered as an important bioinformatics area that is expected by many to become the next major focus area in bioinformatics. The initial plan was to propose a bioinformatics MS degree with a strong emphasis in systems biology.

4.2 A branch-and-bound algorithm

As discussions began in the Fall of 2001 to establish the degree, it was quickly recognised that potential contributors to a bioinformatics degree at an institution like the University of Illinois could come from a great variety of departments, including Animal Sciences, Biochemistry, Bioengineering, Chemical Engineering, Chemistry, Computer Science, Mathematics, Microbiology, Physics, Molecular and Integrative Physiology, Plant Sciences, Statistics, Veterinary Biosciences, and Veterinary Pathobiology. It was also quickly recognised that what bioinformatics meant to one of these disciplines was very different to the understanding of bioinformatics in another discipline. Hence, instead of calling for a meeting involving representatives from all these different departments, two of us (NVS and CFZ) began a long sequence of separate meetings with individuals and small groups from separate departments.

The approach we followed was in essence an implementation of the well-known branch-and-bound algorithm, an algorithm that is also known as ‘divide-and-conquer’ and ‘separation-and-progressive-elimination’. The algorithm partitions a large search space into subsets that are optimised separately until the entire search space is either explicitly or implicitly searched for an optimal solution. It is well known that such an algorithm may require the investment of significant levels of time resources, but guarantees that the search will not be trapped prematurely in local optima. Indeed, this algorithm is guaranteed to identify a globally optimal solution even in the presence of many local optima (cf. [10]).

During the academic year 2001–2002, we had over 30 meetings with over 50 individuals from two different educational institutions and an industrial group, seven different departments, and four different colleges. The original degree proposal was modified several times in the process and, as a result, additional rounds of meetings were necessary. In Fall 2002, discussions began with the Graduate College on how to implement the degree. These discussions involved the Departments of Animal Sciences, Bioengineering, Chemical and Biomolecular Engineering, Computer Science, Molecular and Integrative Physiology, and Veterinary Pathobiology. An agreement on the degree specifics was reached in February 2002. The proposed degree was finally approved by all pertinent departmental and college committees in May 2002, the University Senate in September 2003, the Board of Trustees in January 2004, and the Illinois Board of Higher Education in June 2004.

Approximately 1000 e-mails were exchanged in the process, most with more than four recipients, generating a total of over 3.6 MB of text, which is equivalent to about 600 pages of this manuscript text. Clearly, the amount of effort to establish a bioinformatics degree program through the branch-and-bound algorithm is comparable to that required to write a textbook.

4.3 What’s in a name?

Given the large number of stakeholders in the bioinformatics area, one possible approach to the development of bioinformatics programs is for each interested department to establish its own program as an option or concentration of one of its existing degree programs. For instance, one could offer a bioinformatics concentration under the existing MS in Computer Science, and a separate bioinformatics concentration under the existing MS in Chemical Engineering. No branch-and-bound is needed to design and establish the degree programs in this case – each department need only deal with its own definition of bioinformatics and a local search algorithm will suffice to produce an implementation plan. However, there are two downsides to such an approach. First, an opportunity is missed for cross-fertilisation between different departments. The resultant programs are very likely to be inferior to a globally optimal university-wide program that involves faculty with expertise from complementary areas. Moreover, students would be deprived of the opportunity to have an actual degree in bioinformatics. Instead, they would be forced to accept degrees in traditional academic areas with possibly an option or concentration in bioinformatics. Two such programs were already in existence at the University when we began development of the MS in Bioinformatics degree program: undergraduate and graduate concentration tracks in bioinformatics/computational biology in the Department of Bioengineering and an undergraduate emphasis in bioinformatics/computational biology in the School of Molecular and Cellular Biology. Yet, each of the departments that participated in our discussions expressed a strong interest in using the name ‘MS in Bioinformatics’ for students in their own department.

4.4 Overcoming the interdisciplinary barrier

In addition to organisational challenges, the deployment of a bioinformatics program must address the educational needs of students from a variety of backgrounds. Many students from life sciences wish to enter a bioinformatics program with computing skills that do not go much beyond standard word processing software and spreadsheet application programs. On the other extreme of the scale, there are many computer scientists that wish to enter a bioinformatics program but have very little understanding of what enzymes and proteins are.

4.5 Opportunities

The above discussion has outlined the organisational and educational challenges that result from the diversity of the
field of bioinformatics. The same diversity, however, presents a major opportunity. The field requires interdisciplinary interaction, provides opportunities for group learning where students from different fields learn from one another, and facilitates collaborations across campus, leading to research and teaching experiences that can happen only after the interdisciplinary barrier has been crossed.

5 A university-wide degree with departmental concentrations

Several departments in an academic setting would like to offer a bioinformatics degree, each with its own flavour to better suit its own students. At the same time, students in all these departments would like the term ‘bioinformatics’ in their degree name. For these reasons, the University has decided to have only one university-wide bioinformatics degree, with potentially many departmental concentrations. The Graduate College administers the overall degree structure and contents, whereas departments administer the concentrations that must adhere to strict rules to guarantee a common core and uniform program quality.

5.1 Degree goals and requirements

The goals of the MS degree in bioinformatics are to train students in:
- analysing and modelling DNA microarray data, proteomic data, and other bioinformatics data;
- understanding biological databases and algorithms for extracting useful information from these databases;
- modelling metabolic and other biochemical networks with complex feedback and control mechanisms;
- applying informatics tools, especially to processes involving biomolecules.

The degree is offered in both thesis and non-thesis options. For the thesis option, a minimum of 32 hours are required, including a minimum of 28 hours of coursework (the standard one-semester graduate course is a four-hour course). For the non-thesis option, a minimum of 36 hours are required. For either option, at least 12 hours must be taken at the advanced graduate level. In addition, a minimum of 12 hours must be taken from courses approved for degree credit in each of three core bioinformatics areas.

5.2 Common core

We envision students entering this program from three potential entry points and exiting from three potential exit points, as shown in Fig. 1. Students with different academic backgrounds could choose to enter the program through any of the participating departments, irrespective of their background. We expect that most computer scientists will choose to enter the program through the Department of Computer Science and most engineers will choose to enter the program through the Department of Chemical and Biomolecular Engineering. However, the applications we have thus far received for this program have confirmed what we anticipated when establishing the degree, namely that students may choose to go for graduate studies in a department other than the one where they pursued their undergraduate studies. It is not unlikely for undergraduate students from engineering departments to be interested in a graduate degree with an algorithmic orientation and for students with an undergraduate training in computer science to be interested in a graduate degree with an applied flavour.

Irrespective of student backgrounds and offering department, each departmental concentration must require that a minimum of 12 hours be taken from courses approved for degree credit in each of the three core bioinformatics areas:
- Fundamental Bioinformatics (courses in genomics, proteomics, bioinformatics, and computational biology);
- Biology (courses in general biochemistry and cell biology);
- Computer Science (courses in database systems and combinatorial algorithms).

5.3 Oversight committee

A Bioinformatics Steering Committee is appointed by the Dean of the Graduate College in consultation with the deans of the colleges sponsoring the degree (Liberal Arts and Sciences; Engineering; Agricultural, Consumer, and Environmental Sciences; and Veterinary Medicine). The Steering Committee’s role is to:
- establish a common academic core and program guidelines (including procedures for faculty affiliation);
- synchronise and harmonise concentrations offered by different departments in order to avoid duplication and provide uniform program quality;
- maintain a relevant Web site with pointers to the different departmental concentrations;
- advise the Graduate College with respect to proposed new and revised core bioinformatics courses, including the impact of projected enrollment; and
- advise the Graduate College with regard to approval of newly proposed concentrations.

Thanks to the role of the Steering Committee, even though students enroll into the program through different departments, all students end up pursuing a common bioinformatics core that justifies the common degree. At the time of this writing, the members of the Bioinformatics Steering Committee come from the following departments/schools:
- Department of Animal Sciences;
- Department of Bioengineering;
- Department of Chemical and Biomolecular Engineering;
- Department of Computer Science;
- Department of Crop Sciences;
- Graduate School of Library and Information Science; and
- Department of Veterinary Pathobiology.

Fig. 1 Nine educational paths
6 The Chemical and Biomolecular Engineering concentration

6.1 Concentration goals

By offering a Bioinformatics MS concentration, the Department of Chemical and Biomolecular Engineering ensures that students have access to bioinformatics training in the systems biology area. Modelling of metabolic circuits is widely considered as the next major focus area in mainstream bioinformatics. In addition, discussions with industrial contacts at Eli Lilly and Company indicated that there is a current need in industry for bioinformaticians trained in engineering systems. A second goal of the concentration is to offer trainees an opportunity to have laboratory experience that will help provide a strong appreciation for the interaction between algorithmic and experimental tools in bioinformatics.

6.2 Required courses and electives

In addition to 12 hours of core courses required by the campus-wide degree program in biology, computer science, and bioinformatics, the Chemical and Biomolecular Engineering concentration of the MS in Bioinformatics requires students to take additional courses in:

- metabolic systems engineering (four hours); and
- laboratory techniques in bioinformatics (two hours).

Table 2 provides a comparison between the Bioinformatics MS concentration in Chemical and Biomolecular Engineering and the traditional MS degree in Chemical Engineering. There are two main differences. First, the bioinformatics degree has been designed to be a professional degree that requires no thesis research. Secondly, only one-third of the Chemical Engineering degree courses are required to be in chemical engineering core courses. However, almost two-thirds of the bioinformatics degree courses are predetermined in order to satisfy the campus-wide core requirements as well as the departmental concentration’s goal to cover systems biology and laboratory techniques.

6.3 Bioinformatics core course

Along with the proposal to establish its departmental concentration, the Department of Chemical and Biomolecular Engineering also proposed to introduce a bioinformatics course. This course has been approved by the Bioinformatics Steering Committee as one of the acceptable core courses in the area of "bioinformatics". The course begins with a brief overview of molecular and cell biology, followed by a variety of algorithmic and application topics. Algorithms and applications will be intermixed for pedagogical reasons and will cover the following topics:

- **Algorithms**: linear, nonlinear, dynamic, and integer programming; branch-and-bound; computational complexity, approximability, approximation schemes; sampling, hypothesis testing; design of experiments; Markov chains and hidden Markov models.

- **Applications**: pairwise sequence alignment, multiple sequence alignment, phylogenetic tree construction; conformational energy calculations and potential energy functions; protein structure calculations from X-ray diffraction and nuclear magnetic resonance measurements; data mining in DNA microarray and other bioinformatics databases; design of experiments for protein design; computational methods for drug design; analysis, control, and design of metabolic pathways.

The prerequisites for this course are introductory matrix theory, calculus of several variables, and introduction to probability. This course is cross-listed with Animal Sciences, Molecular and Cellular Biology, and Statistics. An algorithmically oriented variation of this course was co-taught by instructors from the Departments of Computer Science and Chemical and Biomolecular Engineering in the Spring Semester of 2004 [11] and is offered for the second time in the Fall Semester of 2005 [12]. The latter course utilises the Jones and Pevzner book [13], which is an algorithmically oriented textbook with biological motivations in all topics it discusses. These courses are complemented at the University by a more applied bioinformatics course that is offered by the Department of Animal Sciences and provides hands-on experience with computational biology databases and bioinformatics tools [14]. Students in the program are strongly encouraged to take both algorithmically oriented and applied bioinformatics courses.

6.4 Metabolic Engineering Course

The metabolic systems engineering course that is required for completion of this concentration addresses the modeling and redesign of metabolic circuits. Topics covered by this course include:

- a brief overview of molecular biology and cellular metabolism;
- graph representations of metabolic and signalling pathways;
- flux balance analysis via linear optimisation techniques;
- Michaelis-Menten kinetic models, S-systems theory, and other non-linear optimisation models of metabolism;
- modelling and design of metabolic networks via integer programming techniques;
- databases and software for metabolic pathways; and
- computational projects using genome-scale metabolic models.

The prerequisites for this course are introductory matrix theory, calculus of several variables, and introductory differential equations.

6.5 Bioinformatics lab course

In addition to computational bioinformatics courses, this concentration requires students to be exposed to the experimental techniques that provide bioinformatics data. During a half-semester course (two hours), the students are introduced to biological data collection and the interplay
between computation and experimentation in bioinformatics. Experimental techniques covered include PCR, DNA cloning, DNA sequencing, 2-D gel electrophoresis, and DNA microarrays.

The prerequisite for this course is introductory molecular and cellular biology.

6.6 Elective courses

Depending on students’ backgrounds and the specific degree concentrations they select, students are allowed to take electives from among a large number, currently over 80 courses. In this way, students can specialise in areas such as chemical engineering systems bioinformatics, information systems, numerical computing, and statistics. Students are strongly encouraged to take at least one of their elective courses from Computer Science, in particular from the numerical analysis area.

7 The Computer Science concentration

Informatics is synonymous with computer science in many parts of the world, with the United States being a notable exception. Computer and information scientists have made essential contributions to bioinformatics in a wide variety of ways, from combinatorial search techniques for sequence matching to optimisation techniques for structure determination. The analysis of biological data is often an extremely demanding computational task, for which advanced knowledge of techniques for organising and efficiently accessing massive quantities of stored information is essential. This motivates the development of a Bioinformatics MS degree concentration that offers specialisation in advanced computer and information science techniques.

The Computer Science concentration of the MS degree in bioinformatics provides advanced, specialised training in modern combinatorial, numerical, statistical, and database techniques that will prepare students with backgrounds in computer, physical, biological, medical, or engineering sciences for employment in biotechnology and pharmaceutical industries and research laboratories or for further graduate study in these and related fields.

This concentration requires that the 36 credit hours for the degree be distributed as follows:

- four hours in each of the following two areas:
  - database and information systems
  - combinatorial algorithms
- twelve additional hours of Computer Science, which can be chosen from any graduate level CS courses, but the following are especially recommended:
  - introduction to parallel programming
  - software engineering
  - pattern recognition and machine learning
  - numerical analysis
  - program verification
  - design of database management systems
  - data mining
  - scientific visualisation
  - parallel numerical algorithms
- four hours in biological science (the core biology requirement of the campus-wide program)
- four hours in statistics (introduction to mathematical statistics and probability)
- four hours in bioinformatics (the core bioinformatics requirement of the campus-wide program)
- four hours selected from any graduate level course, including the courses listed above.

Compared to the Chemical and Biomolecular Engineering concentration, the Computer Science concentration does not require a metabolic engineering course but requires additional computer science courses and a statistics course. The Bioinformatics Lab course discussed in the previous section can be used towards satisfying the core bioinformatics requirement of the Computer Science concentration.

8 Other departmental concentrations

Additional departmental concentrations are currently under development. In particular, the Bioinformatics Steering Committee has recently received proposals to establish concentrations of the Bioinformatics MS degree in

- the Department of Animal Sciences;
- the Department of Crop Sciences;
- the Graduate School of Library and Information Science.

Whereas the campus-wide degree required approval from the Illinois Board of Higher Education, degree concentrations require approval only at the University, thus making possible the speedy implementation of new departmental concentrations.

9 Prerequisites

Students entering this program must ensure that they meet the prerequisites of all courses they take. Depending on student backgrounds and the specific degree concentration they select, some remedial courses may need to be taken.

The minimum background needed in order to enter the Chemical and Biomolecular Engineering concentration is

- five hours of molecular and cell biology;
- nine hours of general and organic undergraduate chemistry;
- nineteen hours of mathematics and statistics; and
- three hours of introduction to computing.

An additional four hours in data structures and software principles are highly recommended. Students with a BS in Chemical Engineering normally have the minimum background needed to enter the core courses. Students from other engineering disciplines typically come with only eight hours of undergraduate chemistry, in which case they will be required to take an organic chemistry course before they enter the core courses.

The minimum undergraduate background for entering the Computer Science concentration is

- five hours of molecular and cellular biology;
- six hours of general chemistry;
- nineteen hours of mathematics and statistics;
- four hours of data structures and software principles; and
- three hours of introduction to the theory of computation.

In addition, three hours of numerical methods and three hours of organic chemistry are recommended.

Overall, prerequisites vary somewhat for the different concentrations, but there is considerable overlap in terms of the required introductory chemistry, biology, and mathematics. Irrespective of the concentration followed, the program is designed so that all students will have developed strong grounding in mathematics, biology, chemistry, and computer science once they have completed the 24 hours in core courses that are required by the campus-wide degree program.

Clearly, students entering this program will benefit greatly from the development of new introductory
graduate-level courses that cover general chemistry, biology, and computer science topics in a shorter period of time than the currently available standard undergraduate courses on these topics.

10 Current program status

A few students have already been admitted to the program, even though the program has not been advertised. Formal advertising of the Bioinformatics MS will take place in the Fall 2005 semester with the goal of having the first group of students in Fall 2006.

11 Conclusions

Bioinformatics cuts deeply through the boundaries of traditional science and engineering disciplines. The challenges in developing a university-wide bioinformatics degree program are primarily two. First, there is the need to train students from diverse backgrounds on many diverse topics. Secondly, there are many stakeholders in a university setting, thus leading to a very time-consuming process for establishing the program. At the same time, developing such a degree program offers a major opportunity for fostering interdisciplinary interactions. Furthermore, a university-wide approach leads to a program whose value truly exceeds the sum of its parts.

The development of the Bioinformatics MS degree program at the University of Illinois has presented unusual organisational challenges and has provided us unusual rewards. For bioinformatics on our campus, this new degree program has laid the groundwork for the integration of research and teaching across many departments that did not have substantial prior interactions. For students, this program offers unique opportunities for interdisciplinary learning at an institution that has had a long tradition of excellence in research and education. For the institution, the establishment of its first campus-wide degree program with departmental concentrations now provides a model that can be used for developing similar campus-wide programs in other subject areas. The program could also inspire similar campus-wide efforts in bioinformatics and other subject areas in institutions across the country and elsewhere.

12 Acknowledgments

During the development of this degree program, we had discussions with and received input from a large number of individuals, whose contributions are far too many for them all to be listed here. We had discussions regarding degree requirements with Professors Jiawei Han, Robert Skeel, and ChengXiang Zhai in the Department of Computer Science. Professor Huimin Zhao from the Department of Chemical and Biomolecular Engineering developed the contents of the bioinformatics lab course. We also had numerous discussions on implementation issues with all the members of the Bioinformatics Steering Committee: Philip M. Best, Department of Molecular and Integrative Physiology; Gustavo Caetano-Anolles, Department of Crop Sciences; Karen M. Carney, Graduate College; Michael T. Heath, Department of Computer Science; Mark J. Kushn, Department of Electrical and Computer Engineering; Lamar Murphy, Graduate College; Sandra Luisa Rodriguez-Zas, Department of Animal Sciences; Nikolaos V. Sahinidis, Department of Chemical and Biomolecular Engineering; Bruce R. Schatz, Graduate School of Library and Information Science; Ronald M. Weigel, Department of Veterinary Pathobiology; and Bruce Wheeler, Department of Bioengineering.

13 References