

What should ISCB's role be in accrediting bioinformatics programs?
A CRUDE FIRST DRAFT

Kevin Karplus

January 9, 2005

Abstract

This is a very preliminary draft of the report of the Accreditation Task Force. It is intended only for initiating discussion—almost everything in it is likely to change before the final report.

The following suggestions are made to the Board of Directors of the International Society for Computational Biology:

1. Do nothing.
2. Do something.
3. Regret it.

Note: the suggestions above are facetious—what do we really want to recommend?

Contents

1	Introduction and overview	1
2	What is bioinformatics?	2
3	Should bioinformatics programs be accredited?	2
4	Which bioinformatics programs should be accredited?	3
5	What role should ISCB play in accreditation?	5
6	Who are the other key players in accreditation?	5
6.1	United States of America	5
6.2	Germany	6
6.3	United Kingdom	6
6.4	Japan	6
6.5	Canada	6
6.6	Australia	6
6.7	other Asian countries	6
6.8	other American countries	7
6.9	other European countries	7
6.10	African countries	7
7	Conclusions	7
A	Existing example curricula for different styles of bioinformatics programs	9
A.1	University of California, Santa Cruz (a tool-builder approach)	9
A.1.1	B.S. program	9
A.1.2	The Bioinformatics Minor	12
A.1.3	Graduate Program	13
A.1.4	The Bioinformatics Combined BS/grad Degree Program	14
A.1.5	Courses for Non-majors	15
A.2	somewhere (a tool-user approach)	15
A.3	somewhere (a tool-maintainer approach)	15

1 Introduction and overview

The International Society for Computational Biology (ISCB), the leading professional society in bioinformatics, has been considering various roles for itself. One that has come up for consideration is whether the Society should be involved in accreditation of bioinformatics training programs, and, if so, what role it should play in accreditation.

This report will discuss the issues and propose specific actions for the Board of Directors to decide on. Here are some of the key questions:

- What is bioinformatics?
- Should bioinformatics programs be accredited?
- Which bioinformatics programs should be accredited?

- What role should ISCB play in accreditation?
- Who are the other key players in accreditation?

It is expected that no consensus will be reached on some of these questions, as ISCB spans several different academic cultures, both geographically and by discipline.

2 What is bioinformatics?

There are many different definitions of bioinformatics possible, and different programs may well choose different definitions, resulting in distinctly different programs.

One definition is that “bioinformatics is the use of computers and statistics to make sense out of the huge mounds of data that are accumulating from high-throughput biological and chemical experiments, such as sequencing of whole genomes, DNA microarray chips, two-hybrid experiments, and tandem mass spectrometry.” [3].

This definition, though broad, may not be inclusive enough, as it does not cover some fields in computational biology that are not inundated by data from high-throughput methods, and not all approaches are based on statistics. The emphasis on data, however, is useful for distinguishing bioinformatics from other computational fields, which may be more focussed on theoretical models.

There is an excellent set of pointers to different definitions on the web page *What is Bioinformatics?* at <http://www.binf.umn.edu/whatsbinf.html>, for example, the European Bioinformatics Institute describes the field as follows:

The ultimate goal of bioinformatics is to uncover the wealth of biological information hidden in the mass of sequence, structure, literature and other biological data and obtain a clearer insight into the fundamental biology of organisms and to use this information to enhance the standard of life for mankind. [2]

In this description, the emphasis is more on the goals, which overlap substantially with other sciences, than the methods, but the focus on the data is still central.

There are at least three different approaches to bioinformatics:

Tool building Creating new programs and methods for analyzing and organizing data.

Tool using Using existing programs and methods to answer biologically interesting questions.

Tool maintenance and installation Setting up databases, translating biologists’ questions into ones that programs can answer, keeping the tools working and the databases up to date.

None of these approaches is the uniquely “right” one, and we should be very careful not to create a prescriptive accreditation mechanism that unduly favors one approach over the others.

The educational model for bioinformatics might be similar to that for statistics—almost all scientists need to learn some statistics, but relatively few professional statisticians are needed to advance the field. In the same way, most biologists and biochemists in the 21st century will need to learn bioinformatics, but relatively few research bioinformaticians will be needed to advance the field. Because bioinformatics is not as mature a field as statistics, the ratio of research to teaching will be higher than in statistics for some time, but training biologists and biochemists to use the tools will probably be a major focus of bioinformatics education for the next 20 years.

3 Should bioinformatics programs be accredited?

The question of whether bioinformatics programs should be accredited probably does not have a universal answer, as different educational systems assign different importance to accreditation in general.

In those systems where essentially all programs must be accredited, then indeed bioinformatics programs should be accredited, by the same mechanisms as other programs.

In those systems where accreditation is rare and is intended mainly for fields in which the practitioners need to be certified for reasons of public safety (medical professions, civil engineers, police, . . .), there is no pressing need for bioinformatics accreditation, as little that bioinformaticians do has direct life-threatening impact. Our impact on critical health decisions is usually filtered through medical practitioners.

In the middle ground, where some programs are accredited and some not, the need for accreditation in bioinformatics is unclear. In the United States, for example, almost all engineering programs are accredited, but programs in the natural sciences are rarely accredited, except through general institutional accreditation. If bioinformatics is viewed as an engineering discipline, like electrical engineering or computer science, then accreditation would be expected both by the institutions offering programs and their students. If bioinformatics is viewed as a science, like biochemistry or molecular biology, then there is unlikely to be much demand for accreditation, as these fields generally have no discipline-specific accreditation in the United States.

The culture clash between engineers and scientists is undoubtedly familiar to most bioinformaticians, who must communicate across the cultural divide in both directions. The difference of opinion on the value of accreditation is just one more instance of this cultural divide, with many biologists seeing it as a waste of time, both for the accreditors and the programs being accredited, and many engineers seeing accreditation as a standard quality-control measure for academic programs. Both sides of the argument have good points.

The effort of a meaningful accreditation process is significant—the self-study needed for ABET accreditation for a small computer engineering program takes approximately a person-year of faculty time every three–six years. A bioinformatics program may only have 5–6 faculty, and dedicating 3–7% of faculty resources to accreditation is a high price to pay.

On the other hand, the programs that are springing up around the world are so new that there has not been time for the reputation of programs to become established, and so it is very difficult for students and those who hire them to assess the value of the programs. An accreditation process could establish minimum standards that would provide some quality control for the new programs, ensuring that students are not misled by inadequate programs. It could also help underfunded programs make a case to their administrators that more faculty or facilities are needed to gain accreditation.

The self-study needed for accreditation can also be of significant value to a new program, as it forces them to define their objectives and outcomes clearly, and design a curriculum to meet those goals, rather than randomly adding courses based on the whim of instructors.

4 Which bioinformatics programs should be accredited?

Bioinformatics can be taught at several different levels, from short courses through minors added to biology or computer science bachelor's degrees to post-doctoral research training. Accreditation is not suitable for all such programs. Which ones would benefit the most and which ones the least?

Here are some possible programs and opinions about how much accreditation would be useful for such a program:

individual short courses Many institutions are offering short courses (often 1-week or 2-week intensive courses, but also evening classes) aimed at teaching specific skills in bioinformatics. These courses may be aimed either at tool users (generally researchers in the life sciences) or tool maintainers (generally computer-support staff).

Such short courses are generally aimed at a very specific group and teaching a very specific skill, and are usually not standardized enough to be suitable for accreditation.

There have been accreditation efforts in tool-maintenance classes for general-purpose computer tools, but these have been run mainly by the manufacturers of the tools, who needed to have an army of tool-maintainers around to ensure the market dominance of their tools. (Microsoft certification comes to mind as models for this.)

There are vendor-neutral industrial certification programs (such as the exam-based CompTIA network certification), but these are generally driven by industrial demand for certified practitioners, not by professional societies. There is not yet such a large class of similar jobs for bioinformaticians that standardization of training offers much benefit.

clusters of classes Some institutions offer a cluster of classes that does not lead to a degree but to a less formal certificate of completion. For example, the Extension program of the University of California, Santa Cruz has a certificate program in bioinformatics [1] requiring 9 or 10 classes. The classes are approximately at the undergraduate college level, but the courses are taken primarily by people who already have baccalaureate degrees, looking to change fields. The courses are generally less in-depth and graded less strictly than courses for matriculated students.

It is rare in the USA for such non-degree programs to be accredited. Quality varies enormously, even within a single institution, and the resulting certificates are generally only evidence of good-faith effort to learn the material.

Since non-degree programs can be offered (in the USA) by anyone and there is no tradition of accreditation, it would be difficult to accredit such programs.

B.S. minor in bioinformatics Given that the largest group of students who need to learn some bioinformatics are wet-lab biologists, and the second largest are computer scientists interested in new applications fields, it makes sense to have minors in bioinformatics for students who major in these fields.

If the goal of bioinformatics education is to promote the use of bioinformatics tools, then minors in bioinformatics should be the main focus of our education efforts.

Minors are generally not suitable for accreditation, as they are incomplete programs by design.

B.S. major in bioinformatics A Bachelor of Science in bioinformatics is becoming a more popular degree offering. Such a degree is suitable either as preparation for graduate school or for students planning to become tool maintainers and installers.

The requirements of a program intended primarily as preparation for grad school and a program intended primarily for producing tool maintainers are somewhat different. It would be useful for students to know which sort of program they were signing up for!

The B.S. degree is the level at which accreditation is traditionally applied and is most likely to succeed.

M.S. in bioinformatics The M.S. programs in bioinformatics serve two overlapping functions: to train people to be tool-builders for industrial positions, and to retrain life sciences researchers and computer scientists to be productive bioinformaticians.

The M.S. in bioinformatics is analogous to the M.S. in engineering disciplines, where it is the primary working degree for designers. Accreditation is less common for master's degree programs than for bachelor's programs, as the master's degree is generally more specialized and more varied between programs, and so more difficult to devise accreditation standards for.

It is not clear yet whether the biotechnology industry will recognize the M.S. in bioinformatics as a working degree, though, as the oversupply of biology Ph.D.s for past 20–30 years has resulted in biotech companies only hiring post-doctoral researchers for positions above the level of technician.

Ph.D. in bioinformatics A Ph.D. degree in bioinformatics is intended for researchers in the field, to certify that they have done some original research and are prepared to do more. Accreditation is unusual for Ph.D. programs and probably not appropriate, as such programs by their very nature should be on the edges of knowledge and not standardized.

postdoctoral training Many life-science Ph.D.s and some in other fields (particularly mathematics and theoretical computer science) are seeking bioinformatics training to improve their job prospects, to escape the wet lab, or to compensate for holes in their previous education.

Post-doctoral training is a traditional mechanism for such learning in the life sciences, but is not really well-suited for the task of retraining in bioinformatics, as there are often large holes in previous training (programming and statistics, for example) that are not easily filled with the informal learning of traditional post-doc positions.

A better approach for many Ph.D.s is to take a master's degree in bioinformatics.

Postdoctoral training is basically an apprenticeship model, for which accreditation is wholly unsuitable.

5 What role should ISCB play in accreditation?

There are several different roles that ISCB could play in an accreditation process:

- ISCB could become an accrediting agency. (not recommended)
- ISCB could collaborate with existing accrediting agencies. (recommended, if ISCB wants to be deeply involved in accreditation)
- ISCB could publish model curricula that accrediting agencies and institutions could use to base their decisions on. (recommended if ISCB does not want to be involved in accreditation)
- ISCB could express support for accreditation efforts while not getting involved in them. (not recommended)
- ISCB could oppose accreditation for bioinformatics programs. (not recommended)

There are probably ISCB members who support each of these positions.

There is an immense amount of paperwork and bureaucracy involved in any accreditation process. I do not believe that ISCB has the resources to take on such a large task with the current size and structure of the society.

ISCB could collaborate with existing accrediting agencies (indeed, in those areas where accreditation is a government prerogative, there may be no other way for ISCB to be involved). Several other professional societies have adopted this approach, including the IEEE, which is the world's largest professional society. Generally, the professional society provides model curricula, minimal standards, and professional advice to the accrediting agency, and serves as a source for the site visitors needed by most accreditation processes. This level of commitment is feasible for ISCB, but would require a fairly large pool of volunteers who are interested in creating and maintaining the accreditation process.

It would be relatively easy for ISCB to publish model curricula, not as constraints on what can be accredited, but as inspiration for new programs and to assist existing programs in identifying where they have holes in their curriculum. This requires only modest amounts of volunteer effort (collecting detailed descriptions of curricula from existing, well-respected programs and collating them). This level of involvement seems to be well within the capabilities of the Society, and would produce a valuable product even if no other involvement with accreditation is desired.

For ISCB to express support for accreditation, without being willing to provide the staff and volunteer time needed to make it work, seems rather pointless to me.

I do not see any reason for ISCB to oppose accreditation efforts, as long as the accreditation agencies involved have enough input from respected bioinformaticians to avoid accrediting programs named bioinformatics but which actually have a different focus.

6 Who are the other key players in accreditation?

Not written yet: I'll need help in identifying the organizations (governmental and non-governmental) that are key players in accreditation world-wide.

In different parts of the world, different organizations are involved in accrediting university programs. In some places, these organizations may have a monopoly on accreditation, and any accreditation effort will have to be done through them. In other places, accreditation may be a competitive commercial enterprise, and we would have to either set up a collaboration with a reputable accrediter, or find out how to become accrediters within the system.

6.1 United States of America

In the United States, there are several different accreditation agencies with different (sometimes overlapping) missions. There are two major forms of accreditation:

institutional accreditation looks at the overall offerings of a school and accredits the school as a whole. For example, the University of California is accredited by the Western Association of Schools and Colleges (WASC).

It is unlikely that ISCB will want to be involved in institutional accreditation.

discipline-specific accreditation looks at one specific program and accredits it.

Engineering programs in the United States are generally accredited by the Accreditation Board for Engineering and Technology (ABET), psychology programs by the American Psychological Association (APA), education programs by the National Council for Accreditation of Teacher Education (NCATE), specialized health-science programs by the Commission on Accreditation of Allied Health Education Programs (CAAHEP), environmental health by the National Environmental Health Science and Protection Accreditation Council (EHAC), and so forth.

There does not appear to be an accrediting agency covering research biology programs (as opposed to nursing, medicine, or other health-related programs), and most college and university biology programs in the United States seem to be accredited only through the institutional accreditation.

Computer science programs are generally accredited by the Computing Accreditation Commission of ABET, but discipline-specific accreditation for computer science is not universal, even among good computer science programs.

If the ISCB were to get involved in discipline-specific accreditation in the United States, there would be two main methods: to become an independent accrediting agency (which doesn't seem to have any specific legal requirements), or to co-operate with ABET to create accreditation standards that ABET would then administer.

Since the Computer Society of the Institute for Electrical and Electronic Engineers (IEEE) has become interested in bioinformatics and the IEEE has historically worked with ABET in accrediting other engineering programmes, it is likely that an accreditation process for bioinformatics will be created in the US, with or without ISCB involvement.

6.2 Germany

Volunteer needed to write this section!

6.3 United Kingdom

Volunteer needed to write this section!

6.4 Japan

Volunteer needed to write this section!

6.5 Canada

Volunteer needed to write this section!

6.6 Australia

Volunteer needed to write this section!

6.7 other Asian countries

Volunteer needed to write this section!

6.8 other American countries

Volunteer needed to write this section!

6.9 other European countries

Volunteer needed to write this section!

6.10 African countries

Volunteer needed to write this section!

7 Conclusions

References

- [1] UC Santa Cruz Extension. Certificate program in bioinformatics. <http://www.ucsc-extension.edu/ucsc/public/category/courseCategoryCertificateProfile.do?method=load&certificateId=60&selectedCategory=> December 2004.
- [2] European Bioinformatics Institute. 2can bioinformatics educational resource: Bioinformatics—the genomic revolution. <http://www.ebi.ac.uk/2can/bioinformatics/>, December 2004.
- [3] Santa Cruz University of California. Frequently asked questions—bioinformatics program. <http://www.soe.ucsc.edu/programs/bioinformatics/undergraduate/faq.html>, December 2004.

A Existing example curricula for different styles of bioinformatics programs

There are many different approaches to bioinformatics education, which can be loosely categorized into three main approaches:

tool-builder This approach focuses on creating researchers who can develop new computational tools to address problems in biology. It requires a substantial programming, algorithms, and statistics component, and a moderate knowledge of biology.

tool-user This approach focuses on creating biological researchers who can use existing bioinformatics tools effectively to solve problems in biology. It requires some statistics, only modest amounts of programming (mainly scripting), and substantial knowledge of biology.

tool-maintainer This approach focuses on creating support personnel who can install bioinformatics tools, maintain them, set up databases and web sites, and provide some assistance in using the tools. It requires substantial training in database and web site creation, but little scientific expertise.

The subsections of this appendix are intended to provide existing examples of each of these approaches. They are not intended as prescriptive curricula, but only as examples of how different institutions can create quite different reasonable programs.

A.1 University of California, Santa Cruz (a tool-builder approach)

This section has been borrowed directly from the UCSC catalog copy, and has not yet been rewritten as a model curriculum. That rewrite will require explanations for the different portions of the curriculum, and not just lists of courses.

The minor should be split off into a separate subsection, and the Master's and PhD programs either omitted or moved to a separate subsection.

The program in bioinformatics is a multidisciplinary program sponsored by the Biomolecular Engineering Department. The program currently offers a B.S. in bioinformatics (Section A.1.1) a minor in bioinformatics (Section A.1.2), and M.S. and PhD. degrees in bioinformatics (Section A.1.3). Because of overlap in requirements between the B.S. and grad degrees, there is a special B.S.-grad program for UCSC bioinformatics undergrads who continue in grad school (Section A.1.4). There are several courses offered by the department that are suitable for non-majors (Section A.1.5).

Bioinformatics combines mathematics, science, and engineering to explore and understand biological data from high-throughput experiments, such as genome sequencing, gene expression chips, and proteomics experiments. The program builds upon the research and academic strengths of the faculty in the Center for Biomolecular Science and Engineering.

The Human Genome Project, the international collaboration to determine the sequence of human DNA and understand its function, had its origin in a conference that took place at UC Santa Cruz in 1985. One notable output from our research is that UCSC is the primary release site for the public version of the human genome and its annotation: <http://genome.ucsc.edu>. We are also a major player in protein-structure prediction, and have a strong research group in DNA microarray analysis.

A.1.1 B.S. program

The undergraduate bioinformatics degree program prepares students for graduate school or a career in the fast-paced pharmaceutical or biotechnology industries. The program is something of an honors program, although the campus does not allow us to use those terms. Entrance requirements are stringent, and the course requirements more demanding than MCD biology or computer engineering. The curriculum is designed more for students planning to go to graduate school than for students working towards a terminal bachelor's degree.

The immense growth of biological information stored in computerized databases has led to a critical need for people who can understand the languages, tools, and techniques of statistics, science, and engineering. A classically trained scientist may be unfamiliar with the statistical and algorithmic knowledge required in this field. A classically trained engineer may be unfamiliar with the chemistry and biology required in the field. Thus, this program strives for a balance of the two, an engineer focused on the problems of the underlying science or, conversely, a scientist focused on the use of engineering tools for analysis and discovery.

The undergraduate degree program in bioinformatics builds a solid foundation in the constituent areas of the field. Students complete core sequences in mathematics (including calculus, statistics, and discrete mathematics), science (including biology, chemistry, and biochemistry), and engineering (including programming, algorithms, and databases). The core topics are brought together in two bioinformatics courses: BME 110, Computational Biology Tools, and BME 100+L, Introduction to Bioinformatics+Laboratory. Students have two electives for specialization within the fields of bioinformatics and are required to take a bioethics course (either BME 80G, Bioethics in the Twenty-First Century: Science, Business, and Society or PHIL 145, Brave New World: Ethical Issues in Genetics) to study the ethical, legal, and social implications of this new technology.

As a capstone requirement, all students complete a graduate project course: BME 220+L, Protein Bioinformatics or BME 230+L, Computational Genomics. Note: students who work on independent research projects with faculty may substitute a senior thesis, BME 195, for the graduate project course.

Preparation for the Major

Students applying for admission to the bioinformatics major should have completed four years of high school mathematics (through advanced algebra and trigonometry) and three years of science, including one year of chemistry and one year of biology. Comparable college mathematics and science courses completed at other institutions may be accepted in place of high school preparation. Students without this preparation may be required to take additional courses to prepare themselves for the program.

Major Requirements

Every bioinformatics major must have a faculty adviser, assigned by the Baskin School of Engineering Undergraduate Advising Office, and with that adviser must formulate a program of proposed course work that meets the major requirements. Because of the enormous breadth of requirements, bioinformatics majors are urged to take honors courses or sections whenever possible, to get as much as possible out of the courses they take in each field.

Lower-Division Requirements (courses usually taught in the first 2 years)

UCSC has an unusual system, whereby students take 13-week semester classes squeezed into a 10-week quarter, typically taking only 3 courses at a time, and normally taking a total of 36 courses in four years (plus a few partial courses for labs and seminars). UCSC has general-education requirements that add about nine more courses to the requirements specific to the major.

Majors must complete the following lower-division courses:

Biology The standard freshman biology courses are required, as they are prerequisites to later courses in biochemistry, bioinformatics, and popular electives in biology.

- 21A, Accelerated Cell and Molecular Biology; or 20A, Cell and Molecular Biology
- 21B, Accelerated Development and Physiology; or 20B, Development and Physiology

Biomolecular Engineering 80G, Bioethics in the Twenty-First Century: Science, Business, and Society; or Philosophy 145, Brave New World: Ethical Issues in Genetics

Chemistry Standard freshman general chemistry courses are required, as they are prerequisite to required organic chemistry and biochemistry classes.

- 1B+M and 1C+N, General Chemistry+Laboratory

Computer Engineering The freshman discrete math class is a prerequisite for later programming classes (Abstract Data Types) and provides a first introduction to probability, that is useful preparation for later probability and statistics classes.

- 16H, Honors Applied Discrete Mathematics; or 16, Applied Discrete Mathematics

Computer Science The standard freshman programming series is required. This is taught in an object-oriented language (currently Java), so that students are exposed to object-oriented programming and data abstraction from the beginning.

- 13H, Introduction to Programming and Data Structures (Honors); or both 12A+L, Introduction to Programming+Laboratory and 12B+M, Introduction to Data Structures+Laboratory

Mathematics Students are required to take single-variable differential and integral calculus, and multi-variable differential calculus. The more advanced series (intended for mathematicians, engineers, and physicists) is required. Multivariable integral calculus is not required, primarily because of lack of room in the schedule.

- 20A-B, Honors Calculus or 19A-B, Calculus for Science, Engineering, and Mathematics (Credit for one or both can be granted with adequate performance on the CEEB calculus AB or BC Advanced Placement examination.)
- 23A, Multivariable Calculus

Upper-Division Requirements (courses usually taught in the third and fourth years)

Majors must complete the following upper-division courses:

Applied Math and Statistics Students are required to take a junior-level course on probability theory and a graduate-level course on Bayesian statistics. For much of the bioinformatics work done at UCSC, Bayesian statistics forms an essential underpinning.

- Computer Engineering 107, Mathematical Methods of Systems Analysis: Stochastic; or Applied Math and Statistics 131, Introduction to Probability Theory
- Applied Math and Statistics 206, Bayesian Statistics

Chemistry Organic chemistry is required primarily as a prerequisite for biochemistry. We obtained special permission from the instructor of the biochemistry class for our students to take one course less than the usual organic chemistry pre-requisite, as long as they got an A or B grade. Students who passed with only a C grade are required to take the full organic chemistry series.

- 108A+L, Organic Chemistry+Laboratory; or 112A+L and 112B+M, Organic Chemistry+Laboratory

Biochemistry and Molecular Biology The biochemistry course is the first course in a three-part sequence for biochemistry majors (not the reduced version intended for biologists), and covers primarily nucleic acids and proteins, concentrating on the mechanisms of replication, transcription, and translation.

- 100A, Biochemistry (first in three-part sequence)

Bioinformatics Bioinformatics majors need to have both a user's view and an algorithmic view of bioinformatics tools, and so are required to take two courses, one focussing on each aspect. In addition, students are required to do a project, either as part of a graduate bioinformatics class, or (as an honors option) as a senior thesis.

- Biomolecular Engineering 110, Computational Biology Tools. This course was originally developed for biologists and biochemists, but turned out useful enough that we required it of bioinformatics majors as well. It provides a tool-user's view of bioinformatics.
- Biomolecular Engineering 100+L, Introduction to Bioinformatics+Laboratory. This course covers dynamic programming (particularly the various sequence alignment algorithms), Markov chains, hidden Markov Models, Dirichlet mixtures, neural nets, and other standard algorithmic tools used at UCSC. Several Perl programs are required (including alignment of sequences using affine gap costs), but students are expected to learn Perl on their own, without explicit classroom instruction. The course was originally developed for graduate students, and is still required of first-year grad students, although it has been given an undergraduate number.

- One of the following:
Biomolecular Engineering 220+L, Protein Bioinformatics+Laboratory; or 230+L, Computational Genomics+Laboratory; or 195, Senior Thesis Research
One of the graduate classes with projects are the standard way for a student to complete the campus's capstone requirement, but students who have been actively engaged in research with a faculty member may substitute a senior thesis.

Computer Engineering More students fail in grad school from inability to put their research into thesis form than from inability to do the research. Standard freshman composition courses are not sufficient to teach students how to write technical papers and grant proposals, so we require an additional course specifically on technical writing.

- 185, Technical Writing for computer engineers and scientists.

Computer Science • 101, Abstract Data Types. This is the core course on abstract data types and analysis of algorithms with some programming, which is a prerequisite for almost all upper-division computer science classes.

- 180, Database Systems This course covers theory and practice of relational databases—an important skill both for tool maintainers and tool builders.

Advanced Programming One of the following five courses is required to develop more sophisticated programming skills:

- Biomolecular Engineering 109, Resource-efficient Programming; or
- Computer Engineering 177, Applied Graph Theory and Algorithms; or
- Computer Science 104A, Fundamentals of Compiler Design I; or
- Computer Science 109, Advanced Programming; or
- Computer Science 115, Software Methodology

Required Electives Students must select two additional courses as electives, justify their choices in writing, and get the choices approved by their faculty adviser. A list of typical courses is provided in the catalog, from the departments of Applied Math and Statistics, Biochemistry, Biology, Biomolecular Engineering, Chemistry, Computer Engineering, and Computer Science, but there is no pre-approved list. Even students making fairly common choices (such as a genetics course and eukaryotic molecular biology) have to justify their choices.

A.1.2 The Bioinformatics Minor

The bioinformatics major is intended for people who wish to become bioinformaticians—to create the tools needed to solve new problems in computational biology. The bioinformatics minor is intended primarily for bioinformatics tool users who are majoring in a biological or chemical specialty. It is also appropriate for computer science or computer engineering majors who are considering graduate work in bioinformatics.

A bioinformatics minor consists of the following 15 courses:

Lower-division (10 courses):

- Biology (2): BIO20A+B or BIO21A+B
- General chemistry (2): CHEM 1B+M and CHEM 1C+N
- Calculus (3): (MATH 19A and MATH 19B and MATH 23A)
or (MATH 11A and MATH 11B and MATH 22)
or (MATH 20A and MATH 20B and MATH 23A)
- Programming (2): (CMPS 12A+L and CMPS 12B+M) or CMPS 13H

- Bioethics (1): BME 80G or PHIL 145

Upper-division (5 courses):

- Organic chemistry (1): CHEM 108A or CHEM 112A+B
- Biochemistry (1): BIOC 100A or BIO 100
- Statistics (1): CMPE 107 or AMS 131
- Bioinformatics (2): Two of the following three courses: BME 109, BME 100+L, or BME 110

A bioinformatics minor may count any of the courses of the minor toward the fulfillment of the requirements of their major. Majors with substantial overlap with bioinformatics include biochemistry, chemistry, computer science, computer engineering, and molecular, cellular, and developmental biology.

For example, a biochemistry and molecular biology major, chemistry major with biochemistry emphasis, or MCD biology major could double-count the biology, general chemistry, calculus, organic chemistry, and biochemistry courses. A chemistry major could double-count the general chemistry, calculus, organic chemistry, programming and biochemistry courses. A computer science major could double-count the programming, calculus, and statistics classes. A computer engineering major could double-count the chemistry, programming, calculus, and statistics classes.

A.1.3 Graduate Program

The graduate program in bioinformatics offers both M.S. and Ph.D. degrees.

Course Requirements

Both masters and doctoral students must complete nine, 5-credit courses (seven core courses and two electives; see below) and a 3-credit research and teaching course. In addition, M.S. students must complete four seminar credits, while Ph.D. students must complete eight seminar credits. M.S. students must complete two (1-credit or 2-credit) research project courses (such as BME 220L, BME 230L, BME 297F, or BME297), and Ph.D. students must complete three research lab rotations (BME 296) with different supervisors.

Core courses (5 credit)-seven are required

Biomolecular Engineering

- 100+L, Introduction to Bioinformatics+Laboratory
- 220, Protein Bioinformatics
- 230, Computational Genomics
- 80G, Bioethics in the Twenty-First Century: Science, Business, and Society; or PHIL 245 Brave New World: Ethical Issues in Genetics

One graduate course, approved by the faculty, in each of the following three areas:

- Statistics (Engineering 206 recommended)
- Biology (Biology 200B recommended)
- Chemistry (Chemistry 200B recommended)

Electives (5 credit)-two are required

The electives should be graduate-level courses selected with approval of the faculty to ensure a coherent, balanced program. For M.S. students, 5 credits of independent research (297) or thesis research (299) may count as electives toward the degree requirements upon approval of the faculty. For Ph.D. students, independent or thesis research cannot be counted as electives.

Students must choose their electives with faculty guidance and approval to balance their preparation and make up for deficiencies in background areas. In addition to fulfilling background needs, students may choose to emphasize one of the breadth areas: molecular biology, biochemistry, statistics, computational

biology, genetics, computer science, computer engineering, applied mathematics, cell biology, and computer graphics/visualization or may take a cross-sampling of the electives to achieve a broad knowledge base.

Other Curriculum Requirements

Biomolecular Engineering 200, Research and Teaching in Bioinformatics, 3 credits

Seminars

M.S. students: a minimum of two seminar courses, including at least one quarter of the 2-credit Biomolecular Engineering seminar, 280B (formerly Computer Engineering 280B)

Ph.D. students: a minimum of four seminar courses, including at least two quarters of the 2-credit Biomolecular Engineering Seminar, 280B

Research experience:

M.S. students: a minimum of two research project courses. This requirement can be met by taking BME 220L, BME 230L, and/or independent study (BME 297F or BME 297).

Ph.D. Students: three quarters of lab rotations (BME 296), generally within the first 12 months.

Thesis and Dissertation Requirements

In addition to completing the course requirements, students must fulfill the following thesis or dissertation requirements.

For M.S. students, a written thesis proposal must be submitted to a faculty member before the end of the fourth academic quarter. If the faculty member accepts the proposal, he or she will become the student's adviser and will be in charge of supervising the writing of the master's thesis. When the thesis is completed, it will be submitted to a faculty review committee consisting of the thesis adviser and at least two additional readers. The committee must include a School of Engineering faculty member, may include participants from the Division of Physical and Biological Sciences and from industry as appropriate, and must be approved by the bioinformatics program director. Students are required to present their thesis project in a public seminar.

Ph.D. students must select a faculty research adviser by the end of the second year. A written dissertation proposal is required before the end of the third year. A qualifying committee is then formed, which consists of the adviser and three additional members, and approved by the bioinformatics program director and the campus graduate dean. The student must submit his or her written dissertation proposal to all members of the committee and the graduate assistant one month in advance of the examination. The dissertation proposal is publicly and formally presented in an oral qualifying examination given by the qualifying committee.

Ph.D. candidates will submit the completed dissertation to a reading committee at least one month prior to the dissertation defense. The reading committee, formed upon advancement to candidacy, consists of the dissertation supervisor and two readers appointed by the program director upon the recommendation of the dissertation supervisor. The candidate will present his or her research in a public seminar. The seminar will be followed by a defense of the dissertation to the reading committee and attending faculty, who will then decide whether the dissertation is acceptable or requires revision.

A.1.4 The Bioinformatics Combined BS/grad Degree Program

Because our bioinformatics BS program provides excellent preparation for a graduate program in bioinformatics, we offer a combined BS/grad program that allows our BS students to complete the MS (or PhD) somewhat sooner than students with a less tailored preparation.

The current BS and graduate requirements have 4 courses in common:

- BME 80G, Bioethics in the 21st Century or PHIL 145/245 Brave New World: Ethical Issues in Genetics
- BME 100+L, Introduction to Bioinformatics
- BME 220, Protein Bioinformatics or BME 230, Computational Genomics
- AMS 206, Bayesian Statistics

Masters students take 9 courses, two seminars (4 credits), BME 200, and two independent project courses (such as BME 220L and BME 230L). The course work for PhD students is essentially the same, except that

8 credits of seminars are required and three research lab rotations are required in place of the two project courses.

The combined BS/grad program does not make any changes to the undergraduate program, except that students must pass the four overlapping courses listed above for a grade of B- or better.

The requirements at the graduate level are changed to remove the four courses that overlap with the BS and to add two graduate electives to be chosen by the students with the approval of their advisers. Thus, the total number of full courses required is reduced from 9 to 7.

To apply for the combined program, students apply to the MS or PhD program through the normal graduate admission process in the fall of their senior years. If admitted into the graduate program, they would automatically be included in the combined BS/MS or BS/PhD program.

A.1.5 Courses for Non-majors

- Biomolecular Engineering 60, Programming for Biologists and Biochemists, provides an introductory programming class using Perl and BioPerl to analyze, transform, and publish biological data.
- Biomolecular Engineering 80G, Bioethics in the Twenty-First Century: Science, Business, and Society, is particularly appropriate to all students interested in the societal issues surrounding the revolutions in bioinformatics and biotechnology.
- Biomolecular Engineering 110, Computational Biology Tools, provides an introduction to the tools and techniques of bioinformatics from a user's view. It is intended for biologists and biochemists who need to use bioinformatics tools, but are not primarily interested in building new bioinformatics tools.
- Biomolecular Engineering 100+L, Introduction to Bioinformatics+Laboratory, provides a detailed look at some of the important algorithms and theory that is used in bioinformatics tools. It may be of interest to majors in chemistry, biology, computer science, and mathematics.
- Biomolecular Engineering 109, Resource-efficient Programming, provides advice and practice for people working at the limits of their computer hardware. It is of use for bioinformaticians, game programmers, and embedded-system designers.

A.2 somewhere (a tool-user approach)

A.3 somewhere (a tool-maintainer approach)