CSE549: Computational Biology

Professor: Rob Patro
Office: 1422 Computer Science
Email: rob.patro@cs.stonybrook.edu

My Office Hours: Tu/Th 3:00 — 4:30. If you really cannot make the scheduled office hours and you need to meet with me, send me an e-mail to set up an appointment.

Lectures: 205 Frey Hall, Tu/Th 5:30 — 6:50 (sorry; I didn’t choose this time 😞)

TA: To be determined.

Website: The website for this course is www.cs.stonybrook.edu/~rp/class/549f14. Currently, it contains nothing of interest beyond what is in this syllabus. However, it will be updated throughout the semester with relevant announcements, references, and course material.

Textbook(s): The book Understanding Bioinformatics (Zvelebil and Baum 2008) is the officially listed text for this course. The lectures given in class and the material on the exams will not necessarily follow this book. Rather, this book acts as a good, broad, practical introduction Bioinformatics and Computational Biology and it provides a nice supplement to the in-class material (i.e. I’m not requiring that you buy this book, but you may find it to be an excellent reference). There are two other exceptional introductory texts that I can recommend if you want to do some external study / reading. They are An Introduction to Bioinformatics Algorithms (Jones and Pevzner 2004) and Bioinformatics Algorithms: An Active Learning Approach (Compeau and Pevzner 2014). The latter book (Compeau and Pevzner 2014) is very new. In addition to the book, the accompanying website (http://bioinformaticsalgorithms.com) has videos, slides and integrated set of small programming assignments (http://rosalind.info/problems/list-view/?location=bioinformatics-textbook-track) that some may find very helpful practice.

Finally, if you’re from a discipline where you’ve not had a formal algorithms course, I find “Algorithms”(Dasgupta, Papadimitriou, and Vazirani 2006) — PDF available at (http://beust.com/algorithms.pdf) — to be a nice introduction to the basics. I also recommend “Algorithm Design”(Kleinberg and Tardos 2006) and, of course, “The Algorithm Design Manual”(Skiena 2008) 😇.

As we cover different material in the course, the course website will be updated with references to other texts and articles relevant to the topics covered in the lectures.

Course Objectives: The main objective of this course will be to provide a broad overview of the major areas of Bioinformatics and Computational Biology (B/CB). Our perspective will be a computational and algorithmic one, though we will take the time to understand the necessary Biology and motivation for the problems we discuss. We will touch upon many areas of B/CB, including phylogenetics, genome structure and Biological network analysis. However, there will be a significant concentration on genomics and related problems such as high-throughput read alignment, gene finding, genome assembly and transcriptome assembly and analysis. At the end of this course, you should have a good understanding of the types of problems people work on in B/CB, and a fairly in-depth knowledge of the computational tools and techniques used to address some foundational problems in the field.
Course Work (& grading):

Homework(s) — ∼ 10% (I anticipate 1 or 2 of these)
Midterm — ∼ 20%
Final — ∼ 20%
Course Project — ∼ 50%
  • proposal — ∼ 5%
  • progress report — ∼ 10%
  • final report & deliverables — ∼ 35%

These percentages are tentative and may vary by some reasonably small amount. The final course project — since it constitutes a significant fraction of the grade — should represent a substantial investment of time. I expect that this project will be done in small groups. The deliverables will include a short (8 — 10 page) writeup, which should be a *publishable quality document*, even if we do not end up publishing it. The deliverables will also include any related software (if applicable) and / or analysis scripts & data (if applicable). Students are encouraged to come up with their own project ideas (we can discuss which ideas are feasible / appropriate). I will also distribute a list of potential project ideas in a couple of weeks. As in previous versions of this course, there will be a proposal and progress report due prior to the final project writeup. Both of these will be graded. The purpose of this is to help set project expectations, allow me to advise you on any stumbling blocks you may encounter, and to provide real motivation to not leave work on the project until the last couple weeks of the course.

Course-specific policy

**Excused Absences:** If you miss a class for a medical or health-related reason, please provide me with a record of this in writing or via e-mail (I do not need to know the specifics, just the date of your absence and that it was for a medical or health reason). If, for a health-related or medical reason, you will miss two or more consecutive classes, or will miss class on a recurring basis, or were unable to meet a particular academic obligation of this course, I will require a written note from the Student Health Service or a healthcare provider documenting the range of dates for which you were unable to meet your academic obligations. This note need not contain any diagnostic information.

If you will miss any classes or scheduled exams as a result of religious observances, you must submit this information to me, in writing, within the first two weeks of the semester to make necessary accommodations to complete the work that will be missed.

**Final Grades:** The grade you receive in this class will reflect, as much as possible, the degree to which you have mastered the necessary material. How much somebody “needs” an ‘A’ will have no bearing on whether or not (s)he receives an ‘A’, other than how this need or desire is reflected in the work that (s)he does. I want *everyone* to do well in this course, and will make every reasonable effort to help you understand the material as well as possible. However, barring errors in the grading of assignments, the grades you receive at the end of the semester are final, and I will not alter them for personal or non-academic reasons, so please do not ask me to!

**Enjoy:** The purpose of this class (and, ideally, any graduate-level class) is for you to learn. Computational Biology is an active, vibrant, and exciting field of research. It concerns a broad array of interesting problems that stretch from the primarily theoretical to the eminently practical. This course aims to introduce you to this exciting field and helps prepare you, should you choose to do so, to perform research in Computational Biology and tackle some of these interesting problems. Try to be inquisitive, engaged, and interested; try to enjoy learning about this exciting field.
University policy

**Academic integrity:** From the University’s Academic Integrity Syllabus Statement [http://www.stonybrook.edu/commcms/academic_integrity/syllstate.html](http://www.stonybrook.edu/commcms/academic_integrity/syllstate.html):

> Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person’s work as your own is always wrong. Any suspected instance of academic dishonesty will be reported to the Academic Judiciary. For more comprehensive information on academic integrity, including categories of academic dishonesty, please refer to the academic judiciary website at [www.stonybrook.edu/academicintegrity](http://www.stonybrook.edu/academicintegrity).

Academic integrity is a very serious issue. Any assignment, project or exam you complete in this course is expected to be your own work. If you are allowed to discuss the details of or work together on an assignment, this will be made explicit. Otherwise, you are expected to complete the work yourself. It is always *much* better to turn in an incorrect or incomplete assignment representing your own efforts than to attempt to pass off the work of another as your own. I have a lot of tolerance for those who are making a significant effort but may be having trouble understanding a particular concept or completing a certain assignment. However, there will be no tolerance of academic dishonesty.

**Learning disabilities:** If you have a physical, psychological, medical or learning disability that may impact your course work, please contact Disability Support Services, ECC (Educational Communications Center) Building, room 128, (631) 632-6748. They will determine with you what accommodations, if any, are necessary and appropriate. All information and documentation is confidential.

Bibliography


