Computational Biology: Genomes, Networks, Evolution  
Fall 2012 Course Information  
MIT: 6.047/6.878/HST.507  
Lectures: Tue/Thu, 9:30-11:00, Room 4-159  
Recitation: TBA  
Units: 3-0-9. Prerequisites: 6.006, 7.01, 6.041

Course staff

Lecturer:
- Manolis Kellis: 32D-524 (Stata Center), manoli@mit.edu, 617-253-2419. Office hours: any time upon request.

Teaching Assistant:
- Rachel Sealfon: Broad Institute 5 Cambridge Center, 13th floor, 1350F, rsealfon@mit.edu, (617) 714-8196

Course administrator:
- Derek Aylward: Stata Center, Dreyfus Tower, 4th floor, 32D-433B, derekaylward@csail.mit.edu, 617-715-4882 (x5-4882)

Web site and mailing lists

- Course wiki: [https://wikis.mit.edu/confluence/display/6DOT047fa12/](https://wikis.mit.edu/confluence/display/6DOT047fa12/)(for scribing, projects, presentations signup)
- Emails: You can reach the course staff through the stellar website or directly by email.

Grading

Your grade in this course will be based on the following:

- Problem sets (30%)
- In-class quiz (20%)
- Final project (40%)
- Scribing (10%)

Problem Sets

There will be four problem sets during the first part of the semester. Each problem set will include 3-5 problems for all students and a hands-on lab component that is optional for 6.047 students. The problem sets will include both theoretical and programming problems. For programming problems, we typically provide skeleton code in Python, but you may use a different programming language if you so choose.
In-class quiz

There will be an in-class quiz after the 4th module and before Thanksgiving (see course schedule), which will cover all material in the four modules. There will be no final exam. The quiz will include True/False and Short Answer comprehension questions, practical problems that use algorithms seen in class and usually one or two design problems asking you to develop a new methodology.

Final Project

You will complete a final project that is planned out across the whole term, with planning beginning at the start of the term and feedback and advice throughout the term. The project will become the primary focus in the second part of the term. You may either work alone or with one partner. Teams and 6.878 students will be expected to undertake more ambitious projects.

The most rewarding project topics are usually the most challenging (and possibly riskiest!). This might involve defining a biological problem, identifying relevant datasets, designing and implementing new algorithms, applying the methods, and interpreting the results. Alternatively, you can select a less risky project such as comparing several computational biology algorithms for solving the same problem and implementing, applying, and rigorously evaluating the results. You might also analyze a relevant conference or journal article, including criticism, corrections, and/or improvements. An element of the final project grading will depend on the challenge and originality of your project, so spend enough time to choose carefully.

Final Project Proposals

Prior to beginning work on the final project, there will be a proposal submission and review process modeled after the NIH grant/fellowship application process. You will submit a proposal following guidelines that mimic those of an NIH grant/fellowship. You will then be assigned three submitted proposals to critique and your proposal will be submitted to three reviewers. A more detailed handout describing the goals, process, and grading criteria for the final project is available on the course website.

Scribing

Each student will be required to scribe for one lecture. Several students may be assigned to work together on each lecture, depending on course enrollment.

This year’s scribes will get the exciting opportunity to take part in the creation of a set of course notes that will eventually be incorporated into a course textbook. The course has a rich set of material consisting of lecture slides, audio, previous years’ scribe notes, and additional resources that are being compiled into a comprehensive reference with each lecture forming a chapter of the book.

As a scribe, you will be responsible for familiarizing yourself with this material before the lecture. The specific scribe task will vary for each lecture depending on the state of existing notes. After the lecture you will need to go back and add things that were discussed in lecture but were not captured in the slides and notes on their own. For example: what were some particularly insightful questions and answers that we discussed? Were there any interesting digressions from the lecture topic that deserve elaborating on? Were there any common misunderstandings or points of confusion? How about alternative ways of explaining a concept or algorithm? Did we stumble upon any good ideas for a final project?
Recitation

A weekly recitation will be held on Fridays, during which we will discuss additional aspects of the lecture material and hold Q&A (time & location TBA). Since there will be only one recitation section, we will likely not be able to accommodate all scheduling conflicts. Therefore, attendance is not mandatory and recitation notes will be made available on the course web site. The material in the recitation notes may be covered in the exam. The TA will also hold office hours, time and location TBA.

Textbooks

The required text for the class are the course notes that are being compiled throughout the semester. We will also use the following three optional textbooks:

- Richard Duda, Peter Hart, David Stork, *Pattern Classification*.

Collaboration Policy

You are welcome to collaborate on problem sets and the final project. However:

- You must work independently on each problem before you discuss it with others.
- You must write the solutions on your own.
- You must acknowledge outside sources and collaborators.