Computational Issues in Molecular Biology
CSE 397 / 497

Spring Semester 2004
Tuesdays and Thursdays, 1:10 pm – 2:25 pm   ~   Room: PL 208

Professor Daniel Lopresti
Department of Computer Science and Engineering
Lehigh University, Bethlehem, PA 18015
Email: dal9@lehigh.edu   ~   Homepage: http://www.cse.lehigh.edu/~lopresti
Office: PL 404B   ~   Ext: 85782
Office Hours: 2:00 – 4:00 Wednesdays (or by appointment)

Course Overview

1. Course Description

The field of bioinformatics is becoming one of the most exciting application areas for theoretical and applied techniques from computer science, as well as a major driver of new research in computational methods. The past decade has witnessed an explosion in genetic data, with entire genomes for many organisms (including humans) now available online. Computer science plays a central role in the creation, analysis, and management of this data: from the sequencing and assembly of DNA, to searching vast databases of known nucleic acid or protein sequences, to the calculation of evolutionary trees, to attempts to predict the structure, and ultimately the function, of these macromolecules.

CSE 397 / 497 will cover algorithms and computational models applied to molecular biology. The course will be run in a seminar format. After a set of introductory lectures by the instructor, students will deliver the remaining presentations, prepared with help from the instructor, and take responsibility for leading discussions. Assigned readings will be drawn from a textbook and the current literature.

2. Requirements

All students in the course will be required to present a lecture on a topic of their choice selected from the syllabus. As preparing a high quality lecture takes considerable time and effort, students will choose their topics well in advance and work closely with the instructor on content and delivery. Several one-on-one meetings will be scheduled prior to the student's lecture date. In addition to giving a lecture, each student will also serve as a scribe for one other lecture to record notes that arise as a result of the discussion that ensues.

Beyond this class participation component, there will be several modest homework assignments and a final exam. Students taking the course at the graduate level as CSE 497 will also be required to write a short (10 page) paper on a topic decided in consultation with the instructor. (In terms of the Graduate Core Requirements, CSE 497 will fall in the "Computer Applications" and the "Theory" categories.)

3. Prerequisites

CSE 340 (Math 340) Design & Analysis of Algorithms
No background in biology is assumed.

4. Textbook

Title: "Introduction to Computational Molecular Biology"
Authors: João Setubal and João Meidanis
Publisher: PWS Publishing Company
ISBN: 0-534-95262-3

5. Tentative Syllabus (subject to change)

* Introduction to molecular biology for non-biologists
  DNA, RNA, proteins, splicing, sequencing, gene structure

* Pairwise sequence alignment
  Homology, global and local alignment, dynamic programming
  Advanced alignment techniques: gap penalties, saving time and space

* Multiple sequence alignment
  Multidimensional dynamic programming
  Approximation techniques

* Physical mapping
  Restriction maps
  Double and partial digests

* Sequencing
  Shotgun sequencing
  Sequencing by hybridization, DNA microarray analysis, clustering
  Genome assembly

* Advanced topics
  Genome rearrangements
  Evolutionary relationships, phylogenetic trees
  RNA structure prediction, protein folding

6. University Policy on Disabilities

"If you have a disability for which you are or may be requesting accommodations, please contact your professor and the Office of Academic Services, Room 212, University Center or call (610-758-4152) as early as possible in the semester. University policy states that you must notify your professor seven (7) days prior to the exam."