

CSCI1950-Z Computational Methods for Biology: Course Missive

January 20, 2009

1 General

Meetings

Time MW 2-3:20
Location CIT Room 345

Course Staff

		Login	Office Hours
Instructor	Ben J. Raphael	braphael	CIT505
Head TA	Allan P. Stewart	apstewar	Sundays 7-9, CIT271
Grad TA	Crystal L. Kahn	clkahn	CIT423 (See webpage
UTA	Brendan M. Hickey	bh	for hours schedule)

Staff email

ALL STAFF (nb) cs195zheadtas@cs.brown.edu
Just apstewar cs195ztas@cs.brown.edu

Course email

Mailing list cs195z@list.cs.brown.edu
WWW <http://www.cs.brown.edu/courses/csci1950-z> (Note the spelling)
Course dir </course/cs195z/pub>

2 Description

This course will focus on algorithms and computational techniques for solving problems in phylogenetics, functional genomics, and biological networks / systems biology. These algorithms will rely on methods from graph theory (e.g. trees, cliques, isomorphism), linear algebra (matrix decompositions), and probability/statistics. The course will include a mixture of both the theory underpinning each algorithm and the application of the algorithms to real biological data. Assignments will include written homework and programming tasks.

3 Textbook

There is no required textbook for the course. Course notes will be supplemented with online resources. The following textbooks are useful for background reading and reference:

1. *Inferring Phylogenies*, J. Felsenstein. (On reserve in the Science library)
2. *An Introduction to Bioinformatics Algorithms*, N. Jones and P. Pevzner. (Available online from the library. Link on website)

4 Syllabus

4.1 Phylogeny (~ 50% of lectures)

- Inference of phylogenetic trees and ancestral states
 - Parsimony techniques
 - Large and Small Parsimony Problems
 - Sankoff & Fitch's algorithms (dynamic programming)
 - Perfect Phylogeny and Compatibility of Characters
 - Distance-based techniques
 - UPGMA & Clustering
 - Neighbor joining
 - Probabilistic techniques
 - Maximum likelihood and Felsenstein's algorithm
- Testing and comparing trees: Bootstrap, resampling, permutation test.
- Supertrees, consensus trees, and tree metrics.
- Whole-genome phylogeny; gene trees vs. species trees; genome rearrangements

4.2 Functional Genomics (~ 25% of lectures)

- Clustering: hierarchical clustering; graph-based methods; principle components analysis (PCA) and matrix-based methods.
- Data integration and Classification: HMMs, EM, Support Vector Machines, Bayesian networks.

4.3 Biological Networks and Systems Biology (~ 25% of lectures)

- Networks in biological systems: signaling, regulatory, and protein-protein interaction networks.
- Network motifs
- Network alignment
- Data integration with networks
- Random network models and network evolution

5 Prerequisites

Although there are no official listed prerequisites, you should have exposure to at least some of the following:

- **Mathematics:** Basic probability and/or discrete math (e.g. at the level of CSCI0220). Linear algebra: definition of a matrix.
- **Computer science:** Experience with computer programming. Knowledge of basic data structures/algorithms a plus.
- **Biology:** Basic definitions of molecular biology (e.g. DNA, RNA, proteins, etc.).

Deficiencies in a few of the above can be overcome with a little extra effort. Please talk to me if you have questions about your background.

Due to the interdisciplinary nature of this course and the varied background of the students, the difficulty level will vary by student throughout the semester. Please seek extra help on the sections that are most demanding for you. Alternatively, if you are bored, relax and wait for the next topic (or ask me for an additional challenge – there are plenty of options.)

6 Assignments and Grading

- 4 written assignments = 40%
- 3 programming assignments = 40%
- Final Exam (take home) = 20%

7 Administrivia

7.1 Collaboration Policy

All classwork is subject to the CS195Z Collaboration Policy, which is found on the “Course Docs” section of the website and must be filled out separately. You must sign a collaboration policy for us to grade your work. In addition, students are expected to abide by the Brown Honor Code.

7.2 Graduate Course Credit

By arrangement with the professor. You must meet Ben in person by March 2, 2009 in order to pursue graduate credit.

7.3 Programming Assignments

We will design the programming assignments for completion in Java, although it is possible to implement them in other reasonable languages. Please contact the course staff immediately if you opt not to use Java or are uncomfortable with programming.

The course staff will prepare a `jar` repository of relevant libraries as support code. We will provide you the requisite documentation when external libraries are used (to the best of our ability). We hope you enjoy them – they are meant to make the coding more visual and intuitive.

7.4 E-mail Contact

If you wish to contact the course staff, the best way to do so is `cs195zheadtas@cs.brown.edu`. Only this address will be read by the entire staff. **Note:** that it is intended for further inquiry into the course material and clarifications of the assignments.

Additionally, we are operating the `cs195z@list.cs.brown.edu` address for student subscriptions. This listserv is meant for announcements made by the course staff. The reasons for a student posting to the listserv will be rare, and allowed on a per e-mail basis. **PLEASE** do not attempt to post to the listserv any material that may assist other students on their assignments. Use the course staff address if your message should be confidential with us.