

Syllabus for *Computational Systems Biology* CS 5984, Spring 2007

Meeting times	230 McBryde Hall, TR 9:30AM–10:45AM
Instructor	T. M. Murali
Office	2160B Torgerson Hall
Office hours	1-3PM Tuesday and by appointment
Email	murali@cs.vt.edu
Mailing list	CS5984_17109@listserv.vt.edu
WWW	http://courses.cs.vt.edu/~cs5984/2007-spring-csb

Description

This graduate course will introduce students to Computational Systems Biology (CSB), a rapidly-growing field that tries to understand the behaviour and function of a living cell as a system of interconnected components and modules. This field studies fundamental high-level biological questions such as:

- What are the properties of the biological networks in a cell?
- How do these networks interact with each other?
- How does a biological system behave over time under various conditions?
- How does a cell maintain its robustness and stability?
- How can we modify or construct biological systems with desired properties?

This course will emphasise a data-driven approach to systems biology. This approach is in contrast to physics-driven methods that carefully construct models of (small) biological networks and simulate them. The data-driven methodology, or at the least what we focus on in this course, tries to uncover and understand the large scale structure of biological networks by analysing and integrating massive amounts of different types of data. We will use a selection of papers from the literature to discuss these topics. Lectures by me, student presentations, and one or two invited lectures at the end of the semester will drive the course.

Intended Topics

We will discuss research papers in the following topics:

- the analysis of DNA microarray data: basic clustering algorithms; computing condition-specific patterns of co-expression; identifying sets of co-regulated genes.
- high-throughput biological data: experimental techniques for detecting protein-protein interactions, protein-DNA binding, protein phosphorylation, metabolite concentrations; assessing the quality of this data.
- regulatory networks: determining their basic building blocks and structural properties.
- metabolic pathways: high-level structural properties; gleaning evidence of robustness in these pathways; reconstructing metabolic pathways from data.
- obtaining biological insights by combining diverse sources of data: functional annotation of proteins; improving the reliability of the data; refining our current knowledge about biochemical pathways.

The course web-site will contain a detailed list of papers on each of these topics that we will select from for discussing in the course. Students can suggest other topics and papers to me. If you have any such ideas, talk to me as soon as possible so that we can include them in the schedule before it is settled.

Pre-requisites

You must have graduate student standing in the Department of Computer Science, a Life Science department, the Department of Mathematics, or the Department of Statistics. The course is also open to undergraduate students; please speak to me first.

Computer Science students

- Required: Data and Algorithm Analysis (CS 4104).
- Preferred: Algorithms in Bioinformatics (CS 5124)
- Useful: Biological Paradigms for Bioinformatics (PPWS 5984), Applied Combinatorics (MATH 3134).

Life Science students

- Required: Undergraduate or graduate training in biology, genetics, biochemistry, or a similar subject.
- Preferred: Computation for the Life Sciences (CS 5045/5046).

Other students should meet me to discuss your qualifications. I will accept coursework at the undergraduate or graduate level that is equivalent to these pre-requisites. Computational students should also have prior programming experience and should be able to create a software package from a given project specification.

Course Format and Grading Policy

Lectures, presentations of papers from the literature by groups of students, and some invited lectures will drive the course. Class participation is extremely important. Your group presentation will make up 20%, your class participation 30%, and your final project 50% of your grade. This ratio is commensurate with the amount of effort I expect you to put in each of these items.

I encourage students to form groups of 2–3 people before the second class of the semester, go over the list of papers on the course website, and carefully consider what topics and the papers you want to present. Choose at least two or three topics in case your choice conflicts with somebody else's. We will assign papers and topics to groups in the second or third class (January 18 or January 23). Student presentations will begin about halfway through the semester.

Presentation The group of students assigned to present the papers should prepare a formal presentation and lead the discussion. Prepare your presentation like how you would for a conference or symposium. Use a set of prepared slides to drive your presentation but use the blackboard if necessary. I would like a copy of each group's slides. Please provide them to me in PDF format.

Participation I also expect every student to carefully read each paper assigned to each class and participate in the discussion. Given the interdisciplinary nature of CSB, it is critical that both life and computational scientists contribute to our understanding of the research we will all learn about.

Final Software Project I will describe the projects in class on January 18 or 23. Each group and I will decide on the specifications of the project. A significant component of your project must involve performing biological analysis of your results; if your group contains a life-student student, this analysis will primarily be his or her responsibility. A high-quality software project can result in a publication in a premier conference like ISMB, RECOMB, PSB, and/or an appropriate journal.

Announcement

If any student needs special accommodations because of a disability, please contact me during the first week of classes.