CSci 5461: Functional Genomics, Systems Biology and Bioinformatics

Spring 2014 Syllabus
Prof. Rui Kuang (kuang@cs.umn.edu)
TA: Huanan Zhang (huanan@cs.umn.edu)

Course Description

The new biotechnologies developed in the past decade are transforming molecular biology research into a quantitative science based on informatics. New platforms are now capable of collecting large amount of DNA/RNA sequences and measuring their activities in cell at a population scale. For example, a single second-generation sequencing machine can generate 6 billion basepairs of DNAs (approximately double-size of human genome) in about 10 days, and microarray platforms can measure the intensities of hundreds of thousands of gene expressions in several days. Without doubt, the contents of the collected data offer unprecedented opportunities for a deep understanding of the molecular mechanisms in cell, which in turn will increase the hope towards a complete mapping of the linkage between individual’s phenotypes and genetic/genomic patterns. Facing the challenges of making sense out of the sheer volume of data, computer algorithms and data analysis models are playing the central role.

This course covers an introduction to various types of functional genomic data available and current computational/statistical methods used for analyzing the data to answer questions in functional genomics, systems biology and their applications in disease research. We will cover the analysis of gene expression data, proteomic data and protein-protein interaction data, with a special focus on how they can be used to understand and infer networks. The topics are organized in the following categories: 1) Introduction to Molecular Biology and Genomic Techniques, 2) Statistical Analysis of Microarray Gene Expressions, 3) Data Mining and Machine Learning Methods for Microarray Gene Expressions, 4) Analysis of Nextgen Sequencing Data, and 5) Biological Network Analysis.

The objective of this class is to prepare students for undertaking bioinformatics research that is biologically driven and develop their skills for critical evaluation of computational biology literature. The specific goals include

- Acquire a solid background in fundamental concepts of functional genomics and systems biology.
- Learn the state-of-the-art computational methods for biological data analysis.
- Develop a general understanding of the current state of the functional genomics field, and learn how to formulate and solve current biological questions with advanced computational methods.

Homework and project: There will be 4~5 homework assignments, each requiring the implementation of a computational method and its application to a real functional genomic dataset. Students will also choose a topic for the final project, and present the results in-class. Group work is encouraged. All the homework assignments are required to be programmed in matlab/java/C/C++ and submitted through moodle.

Grading: Homework Assignments: 60%, Course Project: 30%, and Participation: 10%.

Prerequisites: Some programming skills are required for this course. Biology or other non-CS students are required to take Csci 3003 or an equivalent programming course as a prerequisite or get instructor approval. Prior knowledge of biology is highly recommended.

Intended Audience: This course is primarily for graduate and senior undergraduate students in computer science, math, statistics and biological sciences with interest in computational biology.

Time and Location: 01:00 P.M. - 02:15 P.M., M,W (01/21/2014-05/09/2014); KHKH 3-111.