Course Syllabus
EECS 600: Systems Biology & Bioinformatics
Fall 2007

Description

Bioinformatics is the science of making sense of biological information. In the genomic era, efforts on developing algorithmic and computational methods for organizing, integrating, analyzing, and querying biological data proved invaluable. Today, availability of high-throughput data relating to the interactions between biomolecules, coupled with past accomplishments in molecular biology, make it possible to study the cell at the systems level. The organization of the cell is abstracted using various models, including protein interaction networks, gene regulatory networks, signaling pathways, and metabolic pathways, each representing a different aspect of the same system. The data relating to these models is obtained via various high-throughput techniques, including DNA microarrays, yeast two-hybrid, affinity chromatography, mass spectrometry, and nuclear magnetic resonance. Integration of sample-specific molecular data with species-specific interaction data opens new doors for a range of medical applications, through analysis of genetic and phenotypical association from a systems perspective. In this course, we cover algorithmic, analytical, and statistical techniques used to effectively organize, integrate, and analyze these novel sources of biological data. This course targets graduate students with research interests in diverse areas, including bioinformatics, computational biology, algorithms, data mining, database systems, and machine learning.

Course Format

The course includes lectures by the instructor, guest lectures, discussions, and student projects/presentations.

Instructor

Mehmet Koyutürk
Office: Olin 512
Phone: (216) 368-2963
Email: koyuturk@eecs.case.edu
Office hours: MWF 1:00pm-2:00pm.

Class Meeting

MWF 2:00pm-2:50pm

Textbooks

No textbooks are required for this seminar. The course material is composed of research and survey papers selected by the instructor, as well as the students themselves.

Prerequisites

None.

Participation: The students are expected to review the material to be covered before each class, attend the class meetings, and actively participate in class discussions (10%).

Assignments: There are two assignments that will allow the students to get hands on experience on small samples from real datasets related to systems biology. (20%).

Presentation: Each student is expected to give a lecture about a topic of their choice related to algorithmic/analytical/statistical techniques for analyzing data that relates to systems biology. The student is expected to provide a thorough understanding of the current state-of-the-art on the selected topic. The breadth and depth of the covered material depends on the topic and student's preference (30%).

Project: Each student will select a problem (possibly related to the presented topic), perform innovative research on this problem, implement their ideas, and obtain experimental results. The student is expected to report the outcome of this project in a term paper and present this work in the class. (40%) .

Calendar (preliminary)

- Week 1 - Introductory cell biology, central dogma, gene expression.

- Week 2 - Regulation of gene expression, cell signaling, molecular interactions.

- Week 3 - Unsupervised analysis of gene expression data: Clustering, biclustering, differential gene expression.

- Week 4 - Supervised analysis of gene expression data: Classification, learning, etc.

- Week 5 - Inference of regulatory networks from gene expression data.

- Week 6 - Prediction, validation, and inference of protein-protein interactions.

- Week 7 - Integration and organization of molecular interaction data: graph databases.

- Week 8 - Topological properties of molecular interaction networks, network evolution.

- Week 9 - Analysis of PPI networks: identification of topological motifs, functional modules, signaling pathways, hierarchical modules.

- Week 10 - Comparative analysis of PPI networks: Graph mining, graph alignment, graph matching.

- Week 11 - Medical applications, analysis of sample-specific data from a systems perspective, identification of active pathways, discriminative and differential subgraphs.

- Week 12-15 - Presentations.