Course Syllabus Spring 2009

EECS 359: Bioinformatics in Practice

Instructor: Mehmet Koyutürk

1 Course Objectives

This course provides an introduction to the analysis of biological data using computational methods, as well as investigating problems in molecular and biology from a computational perspective. It is expected that, upon completion of this course, the students will achieve the following objectives:

- Develop an understanding of the basic principles of molecular and cell biology.
- Become familiar with existing tools and resources for computational analysis of biological data, including sequences, phylogenies, microarrays, ontologies, and biomolecular interactions.
- Develop an awareness of the computational problems that arise in the modeling and analysis of living systems.
- Understand basic abstractions and computational approaches used to formulate and address these problems.
- Be able to use and extend existing computational infrastructure for analyzing biological data.

2 Class Meeting

Tu-Th 10:00 AM – 11:15 AM, Bingham 304.

3 Instructor

Mehmet Koyutürk

- **Office:** Olin 512
- **Phone:** 368-2963
- **e-mail:** `koyuturk@eecs.case.edu`
- **Office hours:** Tu-Th 1:30 PM – 2:30 PM

4 Teaching Assistant

Yoon Soo Pyon

- **Office:** Olin 514
- **e-mail:** `yoon.pyon@case.edu`
- **Office hours:** W-F 2:00 PM – 3:00 PM

5 Textbook

S. Gopal, A. Haatke, R. P. Jones, and P. Tymann. *Bioinformatics: A Computing Perspective*, McGraw-Hill Higher Education, New York: 2008.

6 Prerequisites

None.

7 Course Work & Grading

Participation: (10%) Attendance is obligatory. This has three reasons: (i) The class meets in the morning. (ii) The class size is relatively small, so we have to make sure that we have enough people to make discussions interesting. (iii) The main objective in this course is for the studens to acquire a vision in biology and develop research skills that involve critical thinking. For this reason, in-class discussions are an essential part of the course work. The students are expected to review the material to be covered before each class (the text book is fairly easy and fun to read), attend the class meetings, and actively participate in class discussions. Students are not allowed to use laptops in class (the reason for this should be clear).

Assignments: (30%) There are five assignments related to practical use of available tools, as well as development of new methods for biological data analysis. In these assignments, the students will be expected to retrieve data, analyze the data to answer specific problems, and return detailed, but concise reports summarizing their work and results.

Midterm: (20%) There is one in-class midterm scheduled for approximately the middle of semester. The midterm will consist of several short-answer questions attempting to assess the students' understanding of the material covered in class.

Presentation: (15%) In the final weeks of the semester, each student will present and discuss a recent research paper related to the material covered in class.

Final: (25%) The final exam will include several short-answer questions and more detailed problems, attempting to assess the students' understanding of the material covered in class comprehensively.

8 Calendar

1. **Road Map.**
   (a) Jan 13: A vision on biology and computing, and their "marriage"
2. **Biological Basics.**
   (a) Jan 15: Evolution, domains of life, chemistry of life, structure of the cell.
   (b) Jan 20: Genome, central dogma, DNA, RNA. [*Assignment 1 Out*]
   (c) Jan 22: Proteins.
3. **Wet and Dry Lab Techniques.**
   (a) Jan 27: Sequencing technology, DNA microarrays, proteomics.
4. **Fragment Assembly.**
   (a) Jan 29: Longest common subsequence problem. [*Assignment 2 Out*]
   (b) Feb 3: Greedy algorithms. [*Assignment 1 Due*]
   (c) Feb 5: Improved heuristics, sequencing by hybridization.[*List of Research Papers Out*]
5. **Sequence Alignment.**
   (a) Feb 10: Pattern matching. [*Assignment 3 Out*]
   (b) Feb 12: Alignment algorithms: Needleman-wusch, Smith-Waterman. [*Assignment 2 Due*]
   (c) Feb 17: Heuristic approaches, BLAST, scoring matrices.
   (d) Feb 19: Multiple sequence alignment.
6. **Simulating and Modeling Evolution.**
   (a) Feb 24: Genetic Algorithms.[*Assignment 4 Out*]
   (b) Feb 26: Modeling evolutionary relationships. [*Assignment 3 Due*]
   (c) Mar 3: Phylogenetic tree reconstruction.
   (d) *Mar 5: Midterm.*
7. **Gene Finding.**
   (a) Mar 17: Finding patterns in sequence data. [*Assignment 4 Due*]

(b) Mar 19: ORF Finding.

(c) Mar 24: Markov Models.[*Assignment 5 Out*]

8. **Gene Expression.**

   (a) Mar 26: Transcriptional profiling and applications. [*Selection of Research Papers Due*]

   (b) Mar 31: Preprocessing gene expression data: Normalization, transformation.

   (c) Apr 2: Mining gene expression data: Differential expression, clustering, classification.

9. **Systems Biology.**

   (a) Apr 7: Molecular networks: protein-protein interactions, metabolic pathways, transcriptional regulation.

   (b) Apr 9: Topology of molecular networks: network motifs, degree distribution. [*Assignment 5 Due*]

   (c) Apr 14: Module identification, network alignment, functional inference.

10. **Class Presentations.**

    (a) Apr 16, Apr 21, Apr 23: Research papers presented by students.

    (b) *May 6: Final Exam.*

9 Plagiarism Policy

Zero-tolerance policy on plagiarism is enforced. Following the departmental plagiarism policy, cheating on homeworks or tests will result in an F grade for the whole course and appropriate disciplinary action, independently of the extent of plagiarism. In case of doubt, the students are responsible for checking with the TA or the instructor on what is allowed and what is not.