CNG 465 Introduction to Bioinformatics

2012-2013 Spring Semester

Syllabus

Course Info

Catalog Description

This course covers computational techniques for mining the large amount of information produced by recent advances in biology, such as genome sequencing and microarray technologies. Main topics of the course include: DNA and protein sequence alignment, sequence motifs/patterns, phylogenetic trees, protein structures: prediction, alignment, classification microarray data analysis: normalization, clustering and biological networks.

Credits 3

Category Content Departmental Technical Elective

Prerequisites None.
(However, an understanding of algorithm analysis and design and advanced programming is required to get the most from the course.)

Lecture Time and Place

- Friday 08:40-11:30 (Three Hours) [TZ-20]

Textbook


Reference Books


**Instructor**

- Tolga Can, TZ-30, tcan@ceng.metu.edu.tr

**Teaching Assistant**

- None

**Course Conduct**

The course is structured on the following materials:

1. Text book
2. Lectures
3. Homeworks

There will be 4 written/programming assignments throughout the semester. You will be free to use a programming language of your choice in the programming assignments. By following the textbook, you are expected to learn the biological terminology as well.

Check the previous year’s online course material for information about homeworks, lectures, etc.

http://www.ceng.metu.edu.tr/~tcan/ceng465_s1112/overview.shtml

**Course Outline**

1. Introduction to biology, biological databases, and high-throughput data sources.
   
   Overview of bioinformatics problems.

2. Pairwise sequence alignment algorithms:
   
   - Dynamic programming

3. Statistical significance of alignments

4. Suffix Trees, Suffix Arrays

5. Patterns, Profiles, and Multiple Alignments
   
   - Hidden Markov Models
6. Phylogenetic trees  
   o Neighbor joining, UPGMA methods
7. Introduction to protein structures  
   o Structure Prediction  
   o Structural Alignment of Proteins
8. Microarray data analysis  
   o Clustering techniques
9. Introduction to Computational Systems Biology  
   o Gene regulatory networks  
   o Construction and Analysis of protein networks

**Weekly Coverage**

- **Week-1**  
  o Introduction to biology, biological databases, and high-throughput data sources.  
  o Overview of bioinformatics problems.
- **Week-2**  
  o Pairwise sequence alignment algorithms:  
    • Dynamic programming
- **Week-3**  
  o Statistical significance of alignments
- **Week-4**  
  o Suffix Trees, Suffix Arrays

Distribution of Homework #1

- **Week-5**  
  o Patterns, Profiles, and Multiple Alignments
- **Week-6**  
  o Hidden Markov Models

Distribution of Homework #2

- **Week-7**  
  o Multiple Sequence Alignment Algorithms
- **Week-8**  
  o Phylogenetic trees

Midterm Exam (April 19 Friday, in class) (Week 9)

- **Week-10**  
  o Introduction to protein structures  
  o Structure Prediction
Distribution of Homework #3

- **Week-11**
  - Protein Structure Prediction (continued)
  - Structural Alignment of Proteins
- **Week-12**
  - Microarray data analysis
  - Clustering techniques

Distribution of Homework #4

- **Week-13**
  - Introduction to Systems Biology
  - Gene regulatory networks
- **Week-14**
  - Construction and Analysis of protein networks

**Grading Policy**

Homeworks 20%
Midterm Exam 40%
Final Exam 40%

If you are caught cheating in an exam or homework, you will get a grade of 0 (zero) for that exam or homework.