CSc 87100 Topics in Computational Biology

Ph.D. Program in Computer Science, Graduate Center, CUNY

Last Updated: August 20, 2003
Fall 2003

Course Information

- Meet in room TBA, Tues. 2:00 - 4:00 pm
- Instructor: Dr. Mingzhou Song, msong@cs.qc.edu, 718-997-3584
- Offices:
  - Room 4420, CUNY Graduate Center, 365 Fifth Ave., New York, NY
  - New Science Building A326, Queens College, 65-30 Kissena Blvd., Flushing, NY
- Office hours: 1-2pm Tuesdays at Graduate Center office or by appointment.
- Course web page: http://www.cs.qc.edu/~msong/87100
- Prerequisites: graduate standing in biological, computer, mathematical or statistical science, or permission of instructor.

Text

Bioinformatics and Computational Biology

- [recommended] Gregory R. Grant, Warren J. Ewens, Statistical Methods in Bioinformatics, Springer Verlag, 2001.
- [recommended] Dan E. Krane and Michael L. Raymer, Fundamental Concepts of Bioinformatics, Benjamin Cummings, 2003.
- [recommended] Richard Durbin, Sean R. Eddy, Anders Krogh, Graeme Mitchison, Biological Sequence Analysis : Probabilistic Models of Proteins and Nucleic Acids, Cambridge University Press, 1998.

- [recommended] Dan Gusfield, Algorithms on Strings, Trees and Sequences, Cambridge University Press, 1997.
- [recommended] Pierre Baldi and Soren Brunak, Bioinformatics - The Machine Learning Approach, Second Edition, MIT Press, 2001.

Statistical Learning

- [recommended] Trevor Hastie, Robert Tibshirani and Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference and Prediction, Springer-Verlag, 2001.

Molecular Biology and Genomics

Current journal articles and conference papers

- Journals: Bioinformatics, J. of Computational Biology, J. of Bioinformatics and Computational Biology,
- Conferences: ISMB, RECOMB, CSB, ECCB, PSB

Grading Policies

- About 5 sets of homework (50%).
- One team project (50%).
- Grading policies:
  - Homework is due two weeks after posted. Late homework will not be accepted.
  - No incomplete grades.

Topics

Lecture 1. (9/2/03) Course overview and introduction. Self-introduction and description of research topics by each student.

Lecture 2. (9/9/03) Basics of molecular biology and human genome composition

Lecture 3. (9/16/03) Sequence alignment – dynamic programming and BLAST

Lecture 4. (9/23/03) Multiple sequence alignment

Lecture 5. (9/30/03) Expectation maximization algorithm

Lecture 6. (10/7/03?) Transcription regulation with motif discovery

Lecture 7. (10/14/03) Bayesian network

Lecture 8. (10/21/03) Hidden Markov model and gene finding

Lecture 9. (10/28/03) Gene finding by comparison of human and mouse genomes.

Lecture 10. (11/4/03) Clustering techniques

Lecture 11. (11/11/03) Classification techniques (support vector machine)

Lecture 12. (11/18/03) Gene expression analysis

Lecture 13. (11/25/03) RNA secondary structure prediction

Lecture 14. (12/2/03) Mass spectrometry analysis

Lecture 15. (12/9/03) Protein folding

Final week: project presentation.