Cpt S 580 (05): Computational Genomics

Spring 2009, 3 cr.

Where: Sloan 9

When: Tu,Th 9:10-10:25

Course Objectives

1. To introduce the set of algorithms and data structures that have applications to computational genomics
2. To be able to formulate and/or model a biological problem/system as a computer science problem
3. To be able to design algorithms using appropriate data structures to solve the underlying biological problem
4. To be able to appreciate the role of computer science in modern day biological sciences (interdisciplinary training)
5. To see applicability of algorithms & techniques in other domains such as text mining, pattern matching, etc.

Course Organization

- Topics:
  1. Approximate string matching
  2. Exact string matching
  3. Probabilistic modeling for biological sequence analysis
  4. Genome-scale applications
  5. Phylogenetics

Course Material

- Lecture Notes & Handouts
  - Scribes distributed in class
- Textbook References:
  1. Edited by S. Aluru. Handbook of Computational Molecular Biology, 2005. ISBN: 1584884061 (available through WSU digital library)
  2. Durbin, et al. Biological Sequence Analysis: Probabilistic Models of Protein and Nucleic Acids, 1999. ISBN: 0521629713
  3. D. Gusfield. Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology, 1997. ISBN: 0521585198
- Other Useful References:
  1. C. Setubal and J. Meidanis. Introduction to Computational Molecular Biology, 1997. ISBN: 0534952623
  2. M.S. Waterman. Introduction to Computational Biology, 1995. ISBN: 0412993910

Prerequisites

- Familiarity with Algorithmic Design & Analysis (ie., Cpt S 450 equivalent)
- Familiarity with basic Probability fundamentals
- Biological background NOT required!!
- C/C++/Java programming experience

Grading

- 3 Homeworks (15% each)
- 2 Programming projects (15%)
- 1 Mid-term Exam (20%)
- Survey project (20%)
  - Propose papers and source material
  - Present an 18-minute overview talk to the class (during the last 2 weeks of class)
  - Prepare a 5-page review manuscript and submit during finals week (will get 6 weeks time)
- Grading policy: *curved*

Course Webpage

- *http://www.eecs.wsu.edu/~ananth/CptS580*

- Contents to watch out for:
  - Homeworks, projects
  - Survey project details
  - Lecture notes
  - Tentative course schedule
  - Links to several reference papers, handouts, and other useful web resources

- Class Mailing List

Course Announcements

- All course announcements will be made through the *eLearning* web portal
- *http://elearning.wsu.edu*
- To contact the instructor by email, eLearning is preferred

Homeworks

- Due in-class on the due date
- Hardcopies
- Along with the cover sheet

Programming projects

- Submit through eLearning as an email attachment
- Due 11:59 pm on the due date

Late Submission Policy

- No late submission allowed
- Extensions *may* be allowed under extraordinary circumstances
- Contact instructor at least 1 week prior to permission

Collaboration Policy

- All assignments should be done *individually*. For homeworks, "collaboration" may be allowed on a *subset* of problems as explicitly specified in the problem set.
- "Collaboration" is defined as a discussion with other students in the same class (no outsiders allowed) aimed at obtaining a better understanding of the problem question and/or exploring potential approaches at a very high level that can lead to a solution.
- All collaborative efforts should be explicitly acknowledged/cited in the answer sheet by all the participants using the cover sheet.
- Regardless of whether you collaborate or not, the final writing in the answer sheet should be solely yours.
- No points will be deducted for collaboration as defined above.
- Any deviation from the above guidelines will be considered "cheating" and will be subject to academic dishonesty code. This includes sharing (or even showing of) your solutions, looking up solutions on the web and using them, etc.