CPSC 445 **Algorithms in Bioinformatics** **Spring 2016**

Course Description

Bioinformatics involves the application of computational methods to answer or provide insight on questions in molecular biology. This course provides an introduction to the design and analysis of algorithms for bioinformatics applications.

Topics covered will include sequence alignment, phylogenetic tree reconstruction (parsimony and distance-based methods), prediction of RNA and protein structure, gene finding, and motif finding. Algorithmic techniques that will be covered include dynamic programming and heuristic search methods, as well as combinatorial algorithms for exploration of graphs and trees. Statistical models of molecular sequence and structure, such as hidden Markov models and stochastic context free grammars, and associated algorithms, will also be covered.

Text: Biological sequence analysis: Probabilistic models of proteins and nucleic acids, R. Durbin, S. Eddy, A. Krogh, and G. Mitchison. Cambridge University Press, 1998.

Web Page: http://www.cs.ubc.ca/~hoos/cpsc445/

The web page will point to lecture notes, homework assignments, additional literature and other useful information.

Prerequisites

CPSC 320 and six credits of BIOL beyond BIOL 111, or equivalent. Experience with design and analysis of algorithms is essential for this course.

The additional reading listed below, such as “Molecular Biology of the Cell”, will be useful for those who need to catch up on biological background. The algorithms texts by Cormen et al. and Kleinberg and Tardos provide useful background on combinatorial algorithms, including graph algorithms and dynamic programming algorithms. The book by Hoos and Stützle is a more advanced text which provides very useful background on local search techniques.

Further Reading:

- Alberts, Bruce et al. Molecular Biology of the Cell (any edition) Garland Science
- Baldi, Pierre and Brunak, Søren. Bioinformatics - The Machine Learning Approach. MIT Press, 1998/2001.
- Cormen, Thomas H. et al. Introduction to Algorithms. The MIT Press 2009.
- Gusfield, Dan. Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology. Cambridge University Press, 1997.
- Hoos, Holger H. and Stützle, Thomas. Stochastic Local Search: Foundations and Applications. Morgan Kaufmann, 2005.
- Kleinberg, Jon and Tardos, Eva. Algorithm Design. Addison Wesley, 2005.

Student Evaluation and Grading:

A student's grade will be based on ca. 5 homework assignments (accounting for approx. 45% of the final grade), in-class quizzes (accounting for approx. 20% of the grade) and a final exam (accounting for approximately 35% of the total grade). Discussing homework problems with your classmates is fine, but all write-ups must be done on your own. Depending on final enrolment, homework assignments may or may not include a mini-project, and the final exam may be oral or written (this will be decided at least 4 weeks before the exam and announced in class and via the course web page).

Missed Course Work and Academic Misconduct:

University policy and departmental guidelines on incompletes and academic misconduct will be followed strictly. All work on the quizzes and exams must be entirely your own, with no discussion or aid from anyone else. Late hand-ins of assignments will not be accepted. Missed course work (assignments, quizzes, etc.) can only be excused in the case of officially documented medical reasons (doctor's note required).

Office Hours: Weekly office hours will be posted on the course web page. Additional consultation time with Holger can be arranged by appointment.