Comp 555 - Bioalgorithms

Bulletin Description

Bioinformatics algorithms. Topics include DNA restriction mapping, finding regulatory motifs, genome rearrangements, sequence alignments, gene prediction, graph algorithms, DNA sequencing, protein sequencing, combinatorial pattern matching, approximate pattern matching, clustering and evolution, tree construction, Hidden Markov Models, randomized algorithms.

General Course Info

Term:	SPRING 2015
Department:	COMP
Course Number:	555
Section Number:	001
Time:	M W, 1:25 – 2:40
Location:	SN 014
Website:	http://www.csbio.unc.edu/mcmillan/?run=Comp555S15

Instructor Info

Name:	Prof. Leonard McMillan
Office:	SN 311
Email:	mcmillan@cs.unc.edu
Phone:	919-590-6078
Web:	http://www.cs.unc.edu/~mcmillan
Office Hours:	M T, 4:00 – 5:00

Textbooks and Resources

An Introduction to Bioinformatics Algorithms
by Neil C. Jones and Pavel A. Pevzner
MIT Press © 2004, ISBN: 0262101068.

Bioinformatics Algorithms: An Active Learning Approach
by Phillip Compeau and Pavel Pevzner
Active Learning Publishers © 2014, ISBN: 978-0-9903746-0-2.

Course Description

Computational methods are fueling a revolution in the biological sciences. Computers are already nearly as indispensable as microscopes for analyzing and interpreting biological data. As a result, two new multidisciplinary fields, bioinformatics and computational biology, have emerged. This course will explore the computational methods and algorithmic principles driving this

revolution. It will cover basic topics in molecular biology, genetics, and proteomics. The course also addresses basic computational theory and algorithms including asymptotic notation, recursion, divide-and-conquer approaches, graph algorithms, dynamic programming, and greedy algorithms. These fundamental concepts from computer science will be taught within the context of motivating problems drawn from contemporary biology. Example biological topics include sequence alignment, motif finding, gene rearrangement, DNA sequencing, protein peptide sequencing, phylogeny, and gene expression analysis.

This course is suitable for both computer science and biology students at both undergraduate and graduate levels. Students taking this course should have some programming experience in a modern language.

Target Audience

This course is intended for advanced undergraduate computer science majors and graduate students who are interested in exploring the challenges of bioinformatics and computational biology. It is also well suited for graduate students from Biostatistics and the Curriculum in Computational Biology who would like to better understand the algorithms underlying the common bioinformatics analysis tools.

Prerequisites

All students are expected to have taken introductory courses in data structures and discrete math equivalent to COMP 410 and MATH 381.

Goals and Key Learning Objectives

Comp 555 emphasizes the design, performance, and application of algorithms. All algorithms are motivated by problems from modern biology. A major objective of Comp 555 is to develop a student's intuition for what is the appropriate algorithm to use in a given setting, what is a correct algorithm, and how to analyze the performance and scalability of an algorithm. Comp 555 also exposes the methods underlying many of the common tools used in modern quantitative biology.

Course Requirements

Students must complete assigned reading from the textbook specified in each lecture. Student will be assigned problem sets, with associated programming assignments to be completed on their own computers using a programming language designated by the instructor (Python in this offering). There will be five homework assignments, a midterm, and a final exam.

Key Dates

Midterm: March 4, 2015
Final Exam: April 30, 2015 (Thursday. Noon-3:00pm)

Grading Criteria

The final grade will be based on the follow weighting factors:

5 – Problem Sets/Programming Exercises	50% (10% each)
1 – Midterm Exam	25%
1 – Final Exam	25%

Course Policies

This section should address the following:

- Attendance is expected, but no roll will be taken
- Late problem sets will be penalized by a factor of 0.7071 for each class meeting after the assigned due date. Problem sets will only be accepted at the beginning of class meetings.

The course final is given in compliance with UNC final exam regulations and according to the UNC Final Exam calendar.

Honor Code

Collaboration on assignments is encouraged. However, what you hand in must be your own work. Good scholarship requires that all collaboration must be acknowledged. Thus, if you collaborate on the solution of a problem set, I expect that you list your collaborators at the top of the page.

Collaboration on tests (midterms, final) is, of course, a violation of the Honor Code. This includes discussion of questions on a midterm, or final with students that have not yet taken the test.

Using any unauthorized information sources on an exam is a violation of the honor code.

Course Schedule

A course schedule and handouts from each lecture will be posted on the course website.

Disclaimer

"The professor reserves to right to make changes to the syllabus, including project due dates and test dates. These changes will be announced as early as possible."