CSE 308-408
Bioinformatics: Issues and Algorithms

Spring 2010 • TuTh 1:10 pm – 2:25 pm • Packard Lab 466

Instructor **Professor Daniel Lopresti**
Office PL 350 • Ext 85782 • Email dal9@lehigh.edu
Office Hours 2:30 pm – 4:00 pm Tu and Th (or by appointment)

Biology Advisor **Professor Jutta Marzillier**
Ext 84902 • Email jym2@lehigh.edu

Grading **Wen Cheng**
Assistant Email wec206@lehigh.edu

Textbooks *An Introduction to Bioinformatics Algorithms*, Neil C. Jones and Pavel A. Pevzner
MIT Press, ISBN 0-262-10106-8. (IBA)

Exploring Bioinformatics: A Project-Based Approach, Caroline St. Clair and Jonathan E. Visick,
Jones & Bartlett, ISBN 978-0763758295. (EB)

Blackboard Lecture slides, assignments, etc. will be available @ http://ci.lehigh.edu

Grading

	CSE 308	CSE 408
• Homework assignments =	20% of grade	15% of grade
• Programming assignments =	20% of grade	15% of grade
• Lab notebook =	10% of grade	10% of grade
• Final project or paper =	50% of grade	50% of grade
• Scribe duty (CSE 408 only) =	n/a	10% of grade

Notes
- Homework / programming assignments will generally be posted to Blackboard by 9:00 am and due by 5:00 pm on the specified day. Carefully follow all instructions when naming your programs and submitting your work.
- Your *Exploring Bioinformatics* textbook contains a number of Web Exploration projects to give you experience with bioinformatics tools and datasets. These sections of the book are listed under the "Other Activities" column below. Record your results and observations in a "lab notebook" (really, a word processor file) which is due the next-to-last week of classes.
- Assignments turned in up to one week late will received one-half credit. After that point, no credit will be given. Extensions must be approved by Professor Lopresti.
- If you already consider yourself proficient in Unix and Perl programming, you may choose to skip the lectures marked with an asterisk (*). Please skim the indicated reading, however, and review the lecture notes online in Blackboard.

Week	Topics	Readings	Other Activities
Jan. 18	No class – Prof. Lopresti away	EB 1.1-1.2	EB 1.3
Jan. 25	Course introduction		
	Intro to molecular biology	IBA 3, EB 2.1-2.4	EB 2.5.1
	genetics and genomes, analyzing DNA		
Feb. 1	Intro to algorithms	IBA 2, EB 1.4	HW #1 out Feb. 1
	correctness, recursion, iteration, time complexity		
	Intro to Unix; Perl programming 1 *	EB B.1-B.4, 2.5.3	
	getting started, control flow, variables, arrays, hashes		
Feb. 8	Perl programming 2 *	EB 3.5.3, 4.5.3	HW #1 due Feb. 8
	subroutines, scoping rules, file I/O		
	Perl programming 3	EB 5.5.3, 6.5.3	HW #2 out Feb. 12
	pattern matching, regular expressions, sorting		

Week	Topics	Readings	Other Activities
Feb. 15	Restriction mapping *biology, full and partial digests, brute-force and practical algorithms*	IBA 4.1-4.3	
	Motifs, search trees *regulatory motifs, profiles, search trees, motif-finding*	IBA 4.4-4.9, EB 6.1-6.4	HW #2 due Feb. 19
Feb. 22	Genome rearrangements *biology, sorting by reversals, greedy and approximation algorithms*	IBA 5	HW #3 out Feb. 22
	Sequence comparison & alignment 1 *biology, comparison models, dynamic programming, global alignment*	IBA 6.1-6.7, EB 3.1-3.4	EB 3.5.1
Mar. 1	Sequence comparison & alignment 2 *local alignment, gap penalties, multiple alignment*	IBA 6.8-6.10, EB 4.1-4.4	EB 4.5.1
	Sequence comparison & alignment 3 *saving time and space, divide-and-conquer*	IBA 7	HW #3 due Mar. 5
Mar. 8	No class – spring break		
Mar. 15	Sequencing & assembly 1 *biology, graph theory, shortest superstrings, sequencing by hybridization*	IBA 8.1-8.9, EB 5.1-5.4	HW #4 out Mar. 15 EB 5.5.1
	Genomics (Prof. Marzillier)	reading TBA	
Mar. 22	Genetic pattern matching 1 *repeat-finding, suffix trees*	IBA 9.1-9.5	
	Genetic pattern matching 2 *heuristics: FASTA and BLAST*	IBA 9.6-9.8	HW #4 due Mar. 26
Mar. 29	DNA microarrays (Prof. Marzillier)	EB 9.1-9.2, other reading TBA	HW #5 out Mar. 29
	Clustering for expression analysis *biology, hierarchical clustering, k-means*	IBA 10.1-10.4	EB 9.3.1
Apr. 5	Evolutionary trees 1 *biology, distance-based tree reconstruction, additive trees*	IBA 10.5-10.7, EB 8.1-8.4	Project proposals due Apr. 5
	Evolutionary trees 2 *character-based tree reconstruction, parsimony (small and large)*	IBA 10.8-10.11	EB 8.5.1 HW #5 due Apr. 9
Apr. 12	RNA and protein structure prediction *RNA secondary structure prediction, protein threading*	EB 7.1-7.4, other reading TBA	EB 7.5.1
Apr. 19	Bioethics	Reading TBA	
	History of the Genetic Code	Reading TBA	Notebooks due Apr. 23
Apr. 26	Student final project / paper presentations I		
	Student final project / paper presentations II		Final projects due Apr. 30

Accommodations for Students with Disabilities

If you have a disability for which you are or may be requesting accommodations, please contact both your instructor and the Office of Academic Support Services, University Center C212 (610-758-4152) as early as possible in the semester. You must have documentation from the Academic Support Services office before accommodations can be granted.

Academic Integrity

The work you submit in CSE 308-408 must be entirely your own. While we encourage you to discuss basic concepts and strategies with friends and classmates, the copying or sharing of solutions to homework or programming assignments, or of final papers or projects, is never acceptable. While you may use the Internet as a resource, the unattributed copying of online material is also a violation of academic integrity. Such cases will be referred to the University Committee on Discipline and, if found guilty, you may be given the failing grade WF in the course.

You should keep in mind that computer programs exhibit an individual's "style" just as much as other forms of authorship. Changing variable names, editing comments, or making other trivial updates in an attempt to hide plagiarism is rarely effective.

If you have questions about this policy at any point throughout the semester, ask. It is far better to be safe than sorry when your academic career may be on the line.