

BIF 400- Bioinformatics Capstone Spring 2008

Instructor: Debra T. Burhans, Ph.D.	Office Hours held in WTC 207E
Office: Wehle 207E	Tue: 3:30-5:00pm
Phone: 888-2433	Wed: 9:00 – 11am
E-mail: burhansd@canisius.edu	Fri: 8:00 – 10:00am
Course Web Page: http://cs.canisius.edu/~burhans/400.html	Contact me if you need to set up an appointment at another time. If I'm not in my office I suggest you send me an email.
Course: BIF 400 A	T 8-9:45am OM 224 R 8-9:45am OM 115

Course Overview

This is an advanced course in bioinformatics problems and algorithms. The course will provide a brief review of bioinformatics tools and a more in depth examination of different types of algorithms employed in bioinformatics. The major course project will be carried out in student teams. Each team will work on solving a problem in bioinformatics that is selected in consultation with the instructor. Teams will design and implement a solution and evaluate the results. Documentation and oral presentation are required.

Text

- An Introduction to Bioinformatics Algorithms, Neil C. Jones and Pavel A. Pevzner, MIT Press, 2004.

Web Site

The URL for the BIF 400 web site is given above, course information will be posted there.

Course Objectives

- To review available on-line bioinformatics tools and materials.
- To explore new tools and software packages such as Matlab.
- To understand different algorithmic approaches to solving bioinformatics problems.
- To deepen student understanding of the nature of the information involved in bioinformatics and how it is represented, including databases and data structures.
- To work effectively in a team to produce at least a partial solution to a currently unsolved problem in bioinformatics.
- To create and deliver an effective presentation on the final course project.

Upon successful completion of this course, students will know:

- How to select an appropriate algorithm to solve a problem in bioinformatics
- How to compare different algorithms in terms of computational complexity
- How to write programs using Matlab
- How to locate and utilize bioinformatics databases
- How to use bioinformatics resources to work on solving a research problem in bioinformatics

Attendance

It is important that you attend class regularly. You will be allowed 4 absences before your grade is affected. Missed work must be made up. Absences for any reason whatsoever count towards your 4 absences. Your final course grade will be lowered by five percentage points for each absence beyond the four allowed.

Academic Misconduct

The definition of academic misconduct is given in the Canisius College Student Handbook. An on-line version is located at the following web site: <http://www.canisius.edu/studentlife/standards/appenda.asp>
If you have any questions or doubts regarding this policy as it pertains to this course please ask the instructor for clarification.

Special Accommodations

If you have any condition, such as a physical or mental disability, which will make it difficult for you to carry out the work of the course or which will require extra time for examinations, a note-taker, etc. please notify me in the first two weeks of the course so that we may make appropriate arrangements.

Grading

Your grade will be determined by a weighted average of the following items:

Midterm	20%
Final	20%
Homework, Quizzes, In-Class Exercises	20%
Group Project	15%
Lab	25%
Total	100%

From this a letter grade will be determined based on the following table:

At least 95	A
Less than 95 but at least 90	A-
Less than 90 but at least 85	B+
Less than 85 but at least 81	B
Less than 81 but at least 77	B-
Less than 77 but at least 73	C+
Less than 73 but at least 69	C
Less than 69 but at least 65	C-
Less than 65 but at least 60	D
Less than 60	F

The instructor may revise the cutoffs downward, which would work to a student's advantage.

Course Components

- Midterm -- There will be one 75-minute midterm exam that will cover the first half of the course. No make-up test will be given except for extenuating circumstances that the student can document. If you know ahead of the test date that you will need to be gone, please make arrangements with the instructor. You must do this at least 1 week in advance.
- Final Exam -- The final will focus on material covered in the second half of the course plus some topics from earlier in the course. Missing the final results in an automatic failure of the course with a grade of FX. Note that final exams will not be scheduled until part way through the semester, please do not make plans to leave campus until you find out when all of your final exams are offered. You must take the final exam during the scheduled time.
- Homework, Quizzes, and In-Class Exercises – You will be assigned reading and homework problems. We will have in-class discussions as well as in-class group exercises and quizzes.
- Group Project -- The project will involve investigation of a real problem in bioinformatics. Students will work in groups of 2-3 on the project, which will result in a written report and an oral presentation. The project will be introduced early in the semester and work on the project will be ongoing.
- Lab -- The lab will include exercises designed to reinforce concepts from lectures as well as an introduction to MatLab programming to Internet resources for bioinformatics and an introduction to the Perl programming language with a number of short programming assignments. The unix/linux operating system will also be covered.

Course Schedule

Tuesdays will involve lectures, homework review, and in-class exercises and quizzes
Thursdays will be held in the lab and will involve hands-on exercises

Date	Topic	Chapter
January		
Tues 15	Introduction to course, algorithms, review of biology/CS	1-3
Thur 17	Lab 1 – bioinformatics resources, Perl review	
Tues 22	Exhaustive Search	4
Thurs 24	Lab 2 – Motifs	
Tues 29	Exhaustive Search, Greedy algorithms	4-5
Thur 31	Lab 3 – Search, Greedy algorithms	
February		
Tues 5	Greedy algorithms	5
Thur 7	Lab 4 – Matlab	
Tues 12	Dynamic programming	6
Thur 14	Lab 4 continued – Matlab	
Tues 19	*** no class, President's Day ***	
Thur 21	Project	
Tues 26	Dynamic programming, review for midterm	6
Thurs 28	*** Midterm examination ***	1-6, lab exercises
March		
Tues 4	Divide and conquer algorithms	7
Thur 6	Lab 5 – multiple alignment and gene finding	
Tues 11	Graph algorithms	8
Thur 13	Lab 6 – graph problems, fragment alignment	
Tues 18	Graph algorithms	8
Thur 20	*** no class, Spring/Easter break ***	
Tues 25	*** no class, Spring/Easter break ***	
Thur 27	*** no class, Spring/Easter break ***	
April		
Tues 1	Combinatorial pattern matching	9
Thur 3	lab 7 – Hash tables, suffix trees	
Tues 8	Combinatorial pattern matching, Clustering and Trees	9-10
Thur 10	Project	
Tues 15	Clustering and Trees	10
Thur 17	Lab 8 – data analysis, clustering	
Tues 22	Finite state automata, Hidden Markov Models	11, additional notes
Thur 24	Lab 9 – HMMs	
Tues 29	Hidden Markov Models	11
May		
Thur 1	Group project presentations, review for final exam	

WEEK OF MAY 5: Final examination (exact date TBA)

* this schedule may be revised as necessary during the semester