CSC 448: Bioinformatics Algorithms
Fall 2015
Course Syllabus

September 21, 2015

Instructor: Alexander Dekhtyar
email: dekhtyar@calpoly.edu
office: 14-215

What	When		Where
Lecture	TR	3:10 – 4:30pm	192-224 (Engineering IV)
Lab	TR	4:40 – 6:00pm	14-301
Final Exam time	Tursday, December 10	4:10 - 7:00pm	192-224

Office Hours

	When	Where
Tuesday	9:10pm - 10:00am	14-215
Wednsday	9:10am - 10:00pm	14-215
Wednesday	11:10 - 12:00pm	14-215
Thursday	9:10pm - 11:00pm	14-215

Additional appointments can be scheduled by emailing the instructor at *dekhtyar@calpoly.edu*.

Description

This course studies the application of computing techniques to solving problems in the field of bioinformatics and computational molecular biology.

The key learning outcomes for the course are as follows. Upon completion of the course, you will:

- Know the main problems in the field of bioinformatics and computational molecular biology.

- Understand the key algorithms used to solve computational biology and bioinformatics problems
- Model computational biology problems
- Apply algorithmic techniques to solve problems in computational biology and bioinformatics
- Gain experience working on software projects on multidisciplinary teams.

Texbook

The material for this course comes primarily from these three books:

- N.C. Jones, P. Pevzner. *An Introduction to Bioinformatics Algorithms*, MIT Press, 2004, ISBN:978-0-262-10106-6.
- D. Gusfield, *Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology*, Cambridge University Press, 1st Ed., 1997, ISBN: 978-0521585194.
- Phillip Compeau, Pavel Pevzner, *Bioinformatics Algorithms: An Active Learning Approach*, Active Learning Publishers, 1Ed, 2014, ISBN: 978-0990374602

Topics

This is the tentative list of topics for the course. The order of topics is not quite fixed and some topics may be altered to accommodate the laboratory assignment themes.

No.	Topic	Number of Lectures
1.	Introduction: Biology, Bioinformatics resources	2
2.	Statistical Analysis of DNA	2
3.	Pattern and String Matching	4-6
4.	Genome Alignment	4
6.	Gene Prediction	3-4
7.	Clustering	3-4

Grading

Interdisciplinary Labs	50-60%
Homeworks	5-10%
Midterm	10-20%
Final Exam	20-30%

Course Policies

Course Prerequisites.

The prerequisite for the course is CSC 349: Algorithms. It is going to be **enforced**.

Exams

Content-wise, this is essentially an advanced algorithms course. The course will have a midterm exam and a final exam. The tentative date for the midterm exam is *October 27 (Tuesday) or October 29 (Thursday).*

The date for the final exam is *Thursday, December 10, 4:10 - 7:00pm.*

Both exams will be written tests designed to test your knowledge of the algorithms and data structures covered in the course.

Labs, CHEM 441

The course is taught *in concert* with CHEM 441, Bioinformatics Applications. The two courses discuss the field of bioinformatics from two different perspectives. CSC 448 is a course targeted at computing sciences majors (CSC, CPE, SE). It will concentrate on algorithms and software design and development for bioinformatics problems. CHEM 441 is targeted at life sciences and biotechnology majors (BIO, BCHM, CHEM, BMED, ANSC). It concentrates on the biological nature of the problems for which computational solutions are required, and on the use bioinformatics software in the field of biology.

Teaching the courses in concert means the following:

- Course content is coordinated between the CSC 448 (Alex Dekhtyar) and CHEM 441 (Ed Himelblau) instructors.
- All major hands-on laboratory assignments are shared between the two courses. While the requirements, expectations and deliverables for the assignments will be different, the work on these assignments will be performed by multidisciplinary teams formed jointly out of students from both courses.

Both classes are taught at the same time, with the labs periods being scheduled in nearby labs.

All major labs will be performed in joint teams. Teams will be formed during the first/second week of the classes, and will persist over the course of the quarter. The majority of the lab assignments is designed to contribute to a single overarching course project the students of CHEM 441 will be working on (hence, the team persistance).

Lab time will be split between lab periods designated for joint work between CSC 448 and CHEM 441 students and lab periods designated for independent software development work by the CSC 448 subteams and independent work by CHEM 441 teams.

The major joint labs are (note: this list is based on last year's course. Some of the later labs may change somewhat to accommodate changes in the content of CHEM 441):

Lab Number	Lab Topic	Lab Duration
Lab 1	Ice-breaking activities	1 class
Lab 2	Simple DNA analysis	1.5 weeks
Lab 3	Statstical analysis of DNA	1.5 weeks
Lab 4	Search for repeated sequences	1.5 weeks
Lab 5	Genome Annotation (alignment)	2.5 weeks
Lab 6	Gene Prediction	2.5 weeks

Homeworks

Paper-and-pencil homeworks can be given out to better prepare you for the written exams. Usually the homeworks are not graded and teamwork is encouraged.

Communication

The class will have an official mailing list. The email address for the mailing list is *csc-448-01-2158@calpoly.edu*. All students enrolled in the class are automatically subscribed to the mailing list (using the email addresss that the CS department has on file).

We will also have a joint mailing list for CSC 448/CHEM 441 students.

I encourage questions during classtime and questions via email. My answers to email questions may be broadcast to the entire class via the mailing list, if the answer may be relevant to everyone (e.g. a correction in a text of a handout, or a clarification of a homework problem), and may also appear on the web page. The questions can also be posted to the mailing list directly. The mailing list will also be used for all annoucements related to the course. It is your responsibility to read your class-related email. Failure to read email posted to the mailing list cannot be used as an excuse in the class.

Web Page

Class web page can be found at

http://www.csc.calpoly.edu/∼dekhtyar/448-Fall2015

Through this page you will be able to access all class handouts including homeworks, lab assignments, project information, lab/project data and lecture notes.

Links to additional information, and notes and announcements will also be posted.

Piazza

To facilitate information sharing within each team, we will be using Piazza. The address for the Piazza web page for CSC 448/CHEM 441 is:

https://piazza.com/class#fall2015/csc448bio441

Once the teams are organized, each team will have its own Piazza group, and will be able to share documentation/software with the teammates.

Academic Integrity

University Policies

Cal Poly's Academic Integrity policies are found at

http://www.academicprograms.calpoly.edu/academicpolicies/Cheating.htm

In particular, these policies define *cheating* as (684.1)

> "...*obtaining or attempting to obtain, or aiding another to obtain credit for work, or any improvement in evaluation of performance, by any dishonest or deceptive means. Cheating includes, but is not limited to: lying; copying from another's test or examination; discussion of answers or questions on an examination or test, unless such discussion is specifically authorized by the instructor; taking or receiving copies of an exam without the permission of the instructor; using or displaying notes, "cheat sheets," or other information devices inappropriate to the prescribed test conditions; allowing someone other than the officially enrolled student to represent same.*"

Plagiarism, per University policies is defined as (684.3)

> "... the act of using the ideas or work of another person or persons as if they were one's own without giving proper credit to the source. Such an act is not plagiarism if it is ascertained that the ideas were arrived through independent reasoning or logic or where the thought or idea is common knowledge. Acknowledgement of an original author or source must be made through appropriate references; i.e., quotation marks, footnotes, or commentary."

University policies state (684.2): "Cheating requires an "F" course grade and further attendance in the course is prohibited." (appeal process is also outlined, see the web site above for details.). Plagiarism, per university policies (684.4) can be treated as a form of cheating, although a level of discretion is given to the instructor, allowing the instuctor to determine the causes of plagiarism and effect other means of remedy. It is the obligation of the instructor to inform the student that a penalty is being assessed in such cases.

Course Policies

Homeworks can be completed by students working jointly. Submit the names of all people who worked with you when submitting homework. Lab assignments are to be completed by the appropriate units (individual, pair, group), and no code/solution-sharing between units is permitted. Students are encouraged to discuss class content among themselves but NOT in a manner that constitutes plagiarism and cheating as defined above (e.g., you can solve together a problem from the textbook that had not been assigned in the homework, but you should solve assigned problems individually).