University of Colorado at Colorado Springs

CS 509: Syllabus

Bioinformatics

MW 4:30-5:45 p.m., SENG B216

Summary: Bioinformatics is the science and engineering of organizing and analyzing complex biological data. Bioinformatics primarily deals with computational processing of biological sequence data representing genes and the contents of *genomes*, the complete genetic materials of organisms, and with predicting the function and structure of organic macromolecules such as proteins. It also deals with simulation of life processes at various levels of detail. We will learn about the minimal biological basics, string comparison algorithms, computational learning algorithms in the context of bioinformatics, and also read and perform independent research.

Instructor: Jugal Kalita (255-3432)
Office: Engineering 178
Hours: MW 3:00-5:30 p.m., or by appointment

Text Books

I looked around a lot but didn't find *the single book* that discusses all the topics that I want to cover in the class at the level I want. I decided that I will recommend two text books.

The first book is *Introduction to Bioinformatics*, Third Edition, by Arthur M. Lesk (L). The objective of this book is to generate an understanding of the biological background of bioinformatics and to integrate it with an introduction to the use of computational skills. It does not describe the algorithmic or computer science details, but is excellent in its description of the computational tools used in bioinformatics. One of the most interesting topics discussed is the analysis and prediction of protein structures with an aim to drug discovery.

The second text book is *Biological Sequence Analysis* by Durbin et al (D). It is a mathematical book. We will refer to sections of this book, but it's quite likely that we will not work out all the math.

I will also direct you to various Web sites, and you should also surf the Web to find relevant sites. If you find a site that is particularly useful, please let me know so I as well as the rest of the class can learn from it. It is important that you go to the Web sites I mention in class or are listed in the books to get a feel for, and learn to use the many databases and programs that are on the Web, and the numerous tutorials and papers that are available.

Schedule of Topics

The following schedule is tentative. I will give out handouts when appropriate.

Weeks	Topics
3	Introduction to Genomes and Bioinformatics, Resources on the Web (Chapters: L-1, L-2, L-4)
3	Data Searches and Pairwise Alignment, String algorithms (Chapters: L-5, D-2, D-3, D-4)
2	Multiple Sequence Alignment (Chapter: D-6)
3	Phylogenetic Tree Construction (Chapters: L-5, D-7, D-8)
4	Protein Structure Prediction, Proteomics and Drug Discovery (Chapters: L-6, L-7)

Grading Scheme

The grading scheme is unconventional, at least for this campus. I seek your cooperation in making it work. That is, I need motivated students who are willing to be self-reliant. I have designed the grading scheme to allow for independent research and implementation, and development of presentation skills. The grading scheme reflects that it is a graduate class.

Grading for the class will be based on a class project and class participation. Each student will complete a substantial class project during the semester. Go to the UCCS library Web site and peruse recent issues of the journals *Bioinformatics*, *BMC Bioinformatics*, *IEEE/ACM Transactions on Bioinformatics* in addition to reading the recommended books to get project ideas. Of course, I am available for consultation. I know your projects are likely to be much simpler than what you see in these journals because this is an introductory class to Bioinformatics.

From time to time, you will be required to present your progress on the project to the class. Class participation will be evaluated in terms of physical presence in the class, reading the assigned material and participating in discussions in a manner that reflects the knowledge acquired from reading. You may be asked to read material and present it to the class as well. Please note that if you miss classes without permission, you will lose class participation credits.

In particular, the action items are given below.

- *Proposal:* A 2-3 page class project proposal is due at the end of the fourth week of classes. You will talk to the class for 5 minutes regarding what you want to accomplish in the class project. You must strictly comply with this time limit. The proposal write-up and talk are worth 15% of the class grade. You may be able to talk to professors in Biology or Chemistry, physicians in town, or anyone else about project ideas. The project can deal with any topic that relates life and information that goes with it. There must be a substantial computational aspect to the project.

- *First Mid-term Exam*: The first mid-term exam will be worth 15% of the class grade. The mid-term exam will be comprised of a presentation, a write-up and a demo of your class project accomplishments so far. You will do a 7-8 minute presentation in the class updating your proposal. The write-up will be 4-5 pages. The write-up must be well-written and well-integrated with any material from the proposal that you may use. You will do the demo to me during my office hours. You *must* make visible progress by the first mid-term date to get full credit.

- *Second Mid-term Exam*: The second mid-term exam will be worth 15% of the class grade. The second mid-term exam will also be comprised of a presentation, a write-up and a demo of your class project showing progress made so far. You will do a 7-8 minute presentation in the class bringing the class up to speed with your work. The write-up will be 5-6 pages. The write-up must be well-written and

well-integrated with any material from prior write-ups. You will do the demo to me during my office hours. You *must* make visible progress from the first mid-term to the second mid-term get full credit.

- *Final Exam*: The final exam will be held on the scheduled date. It is worth 45% of the class grade. The final exam will be comprised of a 10 minute presentation, a write-up and a demo of your class project. The write-up will be at most 7 pages long. The write-up must read like a paper ready to submit to *Bioinformatics*, the journal from Oxford University Press! You will demo the final accomplishments for the project to me before the day of the final exam.
- 10% of the grade will be based on your regular attendance and participation in class discussion. This will mean leading the discussion in the class.

Format

Use *Bioinforamtics* journal author style for all your reports.[1] Each report must have academic-type references from conferences and journals. Use Powerpoint or a similar tool for presentations. You will be penalized for not following the formats. The presentations should look professional. The final presentation will be like that in a computer science conference.

Important Dates

Semester Project Proposal:	9/17/2009 (8 presenters, alphabetically by last name), 9/22/2009 (rest of presenters)
Project Proposal Write-up due:	9/17/2009, 9/22/2009 (bring your write-up to your presentation)
First Midterm Exam:	10/15/2009 (8 presenters, reverse alphabetical order), 10/20/2009 (rest of presenters)
First Midterm Project Report:	10/15/2009, 10/20/2009
First Midterm Project Demo:	The week of 10/15/2009
Second Midterm Exam:	11/12/2009 (8 presenters, alphabetical order), 11/15/2009 (rest of presenters)
Second Midterm Project Report:	11/12/2009, 11/15/2009
Second Midterm Project Demo:	The week of 11/12/2009
Final Exam:	12/15/2009, 7:15-9:45 PM (all presenters, alphabetical order)
Final Report:	Must be emailed to me by the morning of 12/14/2009 Monday; I am leaving the country on 12/17 and must finish grading before then.
Final Demo:	The week of 12/7/2009

[1] Go to `http://www.oxfordjournals.org/our_journals/bioinformatics/for_authors/general.html` to find the style files.