

COURSE DESCRIPTION

Department and Course Number: CSCI 490

Course Title: Special Topics
(Topic described here is Bioinformatics.)

Current Catalog Description: Study of topics in computer science according to the interests of the instructor and students.

Total Credits: 3 hours

Coordinator: H. Conrad Cunningham, Chair & Associate Professor of Computer and Information Science. Developer of Bioinformatics topic is Dawn Wilkins, Associate Professor of Computer and Information Science.

Textbook: Cynthia Gibas and Per Jambeck. *Developing Bioinformatics Computer Skills*, O'Reilly, 2001.

References: <http://www.cs.olemiss.edu/~dwilkins/bioinf/fall01/bioinf.html>

Course Goals: To make students aware of computer science as a supporting discipline in the life sciences. Applications are presented that make use of databases, Web interfaces, visualization of data, and managing large data sets.

Prerequisites by Topic: Consent of instructor. For Bioinformatics: minimally the introductory sequence (through CSCI 211); preferably Algorithms (CSCI 433).

Major Topics Covered in the Course: Time estimations are approximations only.

1. Introduction (3 weeks)
 - a. What is Bioinformatics?
 - b. Central dogma of molecular biology, DNA, mRNA, etc
 - c. The Human Genome Project
 - d. Public databases and accessing them via the Web
2. Web programming (client/server architecture, HTTP, CGI, Perl, PHP) (2 weeks)
3. Databases (relational DB concepts, normalization, MySQL, Web-to-DB programming) (2 weeks)
4. Computational biology (3 weeks)
 - a. Sequence alignment algorithms, pairwise and multiple
 - b. Profiles and motifs
 - c. Predicting protein structure and function from sequences
5. Visualizing data (visualizing sequences and SNPs, tools for visualization) (1 week)
6. Working with large datasets (2 weeks)
 - a. Whole genome analysis
 - b. DNA microarrays
 - c. Biochemical pathway databases
 - d. Automating tasks with shell scripting, Perl
 - e. Data mining needs
7. Miscellaneous topics (1 week)
 - a. Medical literature searches
 - b. Bioethics
 - c. Future Challenges

Laboratory projects: There are three-four major programming assignments.

1. Web-based sequence manipulator
2. Creation and querying of a sequence database. Integrated with sequence manipulator.
3. Visualization project
4. Writing scripts to manage large data sets, with statistics.

Estimate of ABET/CAC Category Content:

	CORE	ADVANCED		CORE	ADVANCED
Data Structures	_____	_____1_____	Computer Organization and Architecture	_____	_____
Algorithms	_____	_____1_____	Concepts of Programming Languages	_____	_____
Software Design	_____	_____1_____		_____	_____

Oral and Written Communications:

Not a significant focus of this course.

Social and Ethical Issues:

A brief discussion on bioethics and genetic engineering, as well as privacy of personal medical data.

Theoretical Content (Foundations):

Approximately three weeks of the course is devoted to computational biology, which is algorithmic in nature. The typical solutions of the problems studied involve dynamic programming, and string matching. Many of the problems are NP-hard.

Problem Analysis:

The programming assignments, which were completed by the students individually, required significant analysis and design. Students had to understand the larger context in which their programs had to function.

Solution Design:

See problem analysis above.