CSE 397/497: Structural Bioinformatics

Location: Mohler Lab 355, Tuesdays and Thursdays, 9:20 am - 10:35 pm
Professor: Brian Y. Chen, Department of Computer Science and Engineering, Lehigh University

1. Course Description

Solving problems at the leading edge of medical and industrial technologies depends, in many cases, on improving our understanding of protein function. For example, the debilitating side effects of cancer treatment could be reduced by developing drug molecules that selectively fit the unique structures of cancer proteins. In this case, and in many others, protein shape can yield many deep insights into how proteins function. This course is an exploration, through collaborative and interdisciplinary projects, of the biological, computational, and statistical ideas developed for protein structure alignment, finding functional sites, structure-function inference, molecular surfaces, and evolution in protein structure.

2. Textbook

Structural Bioinformatics, 2nd Edition
Edited by Jenny Gu, and Philip E. Bourne.
Publisher: Wiley-Blackwell, 2009.

3. Lecture Topics

Week 1: Introduction
    Lecture 1: Introduction to Structural Bioinformatics
    Lecture 2: Introduction to Proteins

Week 2: Motif Design and Geometric Matching
    Lecture 3: Motif Design
    Lecture 4: Geometric Matching

Week 3: Modeling Matching Catalytic Sites in Protein Structure Data
    Lecture 5: Match Scoring

Week 4: Algorithmic Motif Design
    Lecture 6: Algorithmic Motif Design
    Lecture 7: Composite motifs

Week 5: Protein Structure Alignment
    Lecture 8: Whole Structure Alignment
    Lecture 9: Geometric Hashing and GRATH

Week 6: Multiple Structure Alignment
    Lecture 10: Multiple Structure Alignment
    Lecture 11: Applications of Multiple Structure Alignment

Week 7: Protein Electrostatics
    Lecture 12: Protein Electrostatics

Week 8: Analyzing Molecular Surfaces
Lecture 13: Alpha Shapes and Protein Cavities
Lecture 14: Molecular Surfaces and Surface Analysis

Week 9: Analyzing Molecular Volumes
Lecture 15: Volumetric Analysis of Protein Surfaces

Week 10: Protein-Protein Interactions
Lecture 17: Protein-Protein Interactions
Lecture 18: Predicting Protein Protein Interactions and Hotspots

Week 11: Protein-DNA Interactions
Lecture 19: Protein-DNA Interactions
Lecture 20: Predicting Protein-DNA Interactions

Week 12: Molecular Simulation and Docking
Lecture 23: Molecular Simulation

Week 13: Computational Drug Design
Lecture 21: Structure Based Drug Design

Week 14: Quantative Structure-Activity Relationships
Lecture 22: From Hits to leads
Lecture 16: Using Machine Learning to classify functional sites

Week 15: Protein Structure Prediction
Lecture 24: Knowledge-based Protein Structure Prediction
Lecture 25: Ab Initio Protein Structure Prediction

4. Course Structure and Assessment

	CSE397	CSE497
Semester Project	75%	60%
CSE497 Report	0%	20%
Participation	25%	20%

Semester Project (CSE397: 75%) (CSE497: 60%)

The semester project will be based on a prototype bioinformatics pipeline with biological, algorithmic and statistical modules. For their semester projects, students, working alone or in interdisciplinary groups, will retrofit the prototype with one or more modules. See the project description and wiki for details.

Class Participation (CSE397: 25%)(CSE497: 20%)

a) **Asking questions in lecture.** The participation score, evaluated at the end of the semester, is equal to the fraction of lectures in which the student has asked a question, except for the first lecture, which does not count. Example: Suppose the lecture is covering protein structure alignment.

Questions like "is protein structure alignment used in current biological research?", "It seems like there are an infinite number of possible structural alignments, how is an optimal alignment even possible?", "Is an optimal structural alignment even a biologically valid piece of information?" are discussion-starters that will be noted for class participation. Participating in these discussions will, likewise, be noted. There is no distinction in score between starting and participating in a discussion. Never be concerned about interrupting the lecture, until the lecturer says it is time to move on from the discussion – always ask the question immediately.

b) **Cell phones and laptops** are to be set on silent mode. Tactful use of electronic devices respects the interests of your peers.

Review Paper (CSE497: 20%)

Students taking CSE497 must prepare a 7-10 page review paper on the current research relating to one of the fields discussed in class. CSE450 Reports are individual work, and not to be discussed with other students. Formatting instructions: One page should be used for the title and your name. This title page does not count towards the page length, but bibliography pages do count towards page length. **Only full pages count towards page length.** Pages must be single spaced, and the first page must start at the top of the page. Use 12 point times fonts, 1 inch margins, and numbered bibliography references, on letter sized paper. Submit your document via email as a pdf.

5. Prerequisites

This course draws from three primary subjects: biochemistry, algorithms, and statistics. Knowledge of all subjects is unnecessary, but a familiarity with at least one of these subjects is expected. Having taken courses in one of the following groups is highly recommended, **though only instructor permission is necessary to enroll**. Students will be asked to state which set(s) of prerequisites they fulfill, and be thus associated with one of the primary subjects.

Group 1: BIOS 371+372 or CHM 371+372
Group 2: MATH 205, CSE 109, CSE 340
Group 3: MATH 312, MATH 334.

6. Outcomes

By completing this course, students will:

1. Understand the basic design and purpose of several major computational technologies in the field of structural bioinformatics
2. Be aware of how biological, algorithmic, and statistical concepts can be integrated to draw meaningful conclusions from multi-faceted biological data.
3. Have experience in the implementation challenges relating to these major technologies.
4. Have experience in technical communication with collaborators with technical expertise outside of their own field.

This course supports program missions to educate students that will:

1. Apply their education in computer science to the analysis and solution of scientific, business, and industrial problems.
3. Function effectively in a collaborative team and effectively communicate with members of the team.
4. Engage in continued education in their field of expertise

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.

Lehigh Student Senate Academic Integrity Statement:
We, the Lehigh University Student Senate, as the standing representative body of all undergraduates, reaffirm the duty and obligation of students to meet and uphold the highest principles and values of personal, moral and ethical conduct. As partners in our educational community, both students and faculty share the responsibility for promoting and helping ensure an environment of academic integrity. As such, each student is expected to complete all academic course work in accordance to the standards set forth by the faculty and in compliance with the university's Code of Conduct.