Course Syllabus

Jump to Today

Overview

Please note that most of this content also appears in **this PDF document**.

Contact info

- Instructor: Ian Holmes **ihh@berkeley.edu (mailto:ihh@berkeley.edu)**
- GSIs:
  - Chenling Xu **chenlingantelope@berkeley.edu (mailto:chenlingantelope@berkeley.edu)**
  - Zoë Steier **zsteier@berkeley.edu (mailto:zsteier@berkeley.edu)**

Grading

- 50% lab-based homework assignments and project
- 20% mid-term exam
- 30% final exam

Content

- Review of fundamental molecular biology, emphasizing **signals in biological sequences**
- **Biophysical principles** of RNA and protein folding
- Overview of biological databases. Common formats, data compression. SQL, NoSQL. GenBank, SwissProt, PDB.
- Introduction to Unix. Review of Python **programming**: strings, methods, lists; file manipulation
- Biophysics of **synthetic biology**: RNA folding kinetics & viral genome design. Bioinformatics issues arising: Ka/Ks, Nussinov algorithm, grammars, Metropolis-Hastings simulation. Specific examples may include riboswitches, allosteric DNA/RNA logic circuits, DNA origami
- **Sequence alignment** algorithms: Needleman-Wunsch, Smith-Waterman, Gotoh, BLAST. Extreme value distributions, automata, kmer-based indices
- Genome annotation with HMMs. **Biological ontologies**, and logical graph operations. The Gene Ontology, Fisher's exact test for term enrichment, hypergeometric distribution. Pathway databases
- **Probabilistic inference**; Bayes' theorem; experimental error; expectation and variance; basic properties of IID sequences. Binomial, multinomial, geometric, exponential, Poisson, Gaussian, Extreme-value, hypergeometric, mixture distributions...
- Quantitative measures of **information**; illustration via data compression. Ideal codes. Log-likelihood ratios and substitution matrices as information-theoretic scores. Data compression in bioinformatics (eg CRAM), cryptanalytics
- Probabilistic models for sequence motifs. **Sequence logos** as visualizations, as probabilistic models, or as outputs of other functions (e.g. neural nets)
- **Algorithmic complexity** & "big-O" notation. Review of basic data structures & their complexity. Application to time & memory complexity of algorithms in this class

- Finite state machines; multiple alignment; phylogenetic reconstruction. Commonality of dynamic programming algorithms
- Structural biology, **RNA & protein structure prediction**, RNA & protein design
- Measurements of evolutionary rate at the resolution of individual amino acids; ancestral sequence reconstruction; applications to **protein design**
- **Clustering** algorithms: K-means, K-medians; application to transcriptomic data analysis
- **Sequence assembly** (overlap-layout-consensus, de Bruijn). Pan-genome graphs. Metagenomics

Modules

Lecture slides for each module will be made available on bCourses shortly before the lectures are delivered. The PDF files available online will be for last year's slides, until the module finishes; then I'll update bCourses to reflect edits made this year. I will try to update the slides for a given topic in advance of any quizzes that are based on that material - but, in general, as a word to the wise, *please* do not assume that conveniently downloadable lecture slides are a substitute for taking your own notes and doing your own research!

Modules List

Course Summary:

Date	Details	
Wed Sep 11, 2019	**Lab 1** (https://bcourses.berkeley.edu/courses/1486407/assignments/8022594)	due by 11:59am
Fri Sep 13, 2019	**Roll call** (https://bcourses.berkeley.edu/courses/1486407/assignments/8021654)	due by 12pm
Wed Sep 18, 2019	**Lab 2** (https://bcourses.berkeley.edu/courses/1486407/assignments/8022519)	due by 11:59am
Wed Sep 25, 2019	**Lab 3** (https://bcourses.berkeley.edu/courses/1486407/assignments/8024841)	due by 11:59am
Wed Oct 2, 2019	**Lab 4** (https://bcourses.berkeley.edu/courses/1486407/assignments/8027148)	due by 11:59am
Wed Oct 9, 2019	**Lab5** (https://bcourses.berkeley.edu/courses/1486407/assignments/8028485)	due by 11:59am
Fri Oct 18, 2019	**Lab 6** (https://bcourses.berkeley.edu/courses/1486407/assignments/8030026)	due by 11:59am

Date	Details	
Wed Oct 23, 2019	**Midterm Exam (2019)** (https://bcourses.berkeley.edu/courses/1486407/assignments/8031040)	due by 3:05pm
Thu Oct 24, 2019	**Lab 7** (https://bcourses.berkeley.edu/courses/1486407/assignments/8031322)	due by 11:59pm
Wed Nov 6, 2019	**Lab 9** (https://bcourses.berkeley.edu/courses/1486407/assignments/8034746)	due by 11:59am
Wed Nov 13, 2019	**Lab 10** (https://bcourses.berkeley.edu/courses/1486407/assignments/8037555)	due by 11:59am
Wed Nov 20, 2019	**Project Checkpoint 1** (https://bcourses.berkeley.edu/courses/1486407/assignments/8037810)	due by 11:59am
Wed Nov 27, 2019	**Project Checkpoint 2** (https://bcourses.berkeley.edu/courses/1486407/assignments/8037813)	due by 11:59am
	Lab 10 resubmission (https://bcourses.berkeley.edu/courses/1486407/assignments/8039785)	due by 12pm
Thu Dec 5, 2019	**Final Presentation** (https://bcourses.berkeley.edu/courses/1486407/assignments/8037814)	due by 11:59am
	Final Project Writeup (https://bcourses.berkeley.edu/courses/1486407/assignments/8037817)	due by 11:59am
Thu Dec 19, 2019	**Final exam 2019** (https://bcourses.berkeley.edu/courses/1486407/assignments/8045001)	due by 5:05pm