

Algorithms in Computational Genomics (COM SCI C122/COM SCI C222) Winter 2025

Instructors: Eleazar Eskin (eeskin@cs.ucla.edu) and Jason Ernst (jason.ernst@ucla.edu)

Monday and Wednesday 2:00-3:50

TAs:

Jingyuan Fu (fuyj2038951@ucla.edu)

Daniel Lee (seungmo6@g.ucla.edu)

Eric Liu (ericliu2023@g.ucla.edu)

Emily Maciejewski (eamaciejewski@g.ucla.edu)

Course Website: <https://bruinlearn.ucla.edu/courses/198840>

Office Hours:

Profs. Eleazar Ernst/Jason Ernst – Wed. 4-5pm in EVI (room #286)

Jingyuan Fu – Mon 4-6pm, Boelter 3256S

Daniel Lee – Thur 2-4pm, Boelter 3256S

Eric Liu – Tue 12-2pm, Boelter 3256S

Emily Maciejewski – Wed 10am-12pm, Boelter 3256S

Discussion Sections:

Discussion 1A F 2:00P-3:50P– Young Hall 4216 – Emily Maciejewski

Discussion 1B F 2:00P-3:50P– Haines Hall 118 – Jingyuan Fu

Discussion 1C F 4:00P-5:50P– Boelter Hall 5249 – Eric Liu

Discussion 1D F 4:00P-5:50P– Boelter Hall 5264 – Daniel Lee

Catalog Description: Designed for engineering students as well as students from biological sciences and medical school. Databases of genomic sequence data are among the largest datasets in all of science. Assembling, indexing, and querying such tremendous datasets is computationally challenging yet critical for many areas of biomedical research. Focus on development of scalable algorithms for analysis of genomic sequence data, with additional focus on formulating biologically relevant problems as computational problems and then solving these problems by developing new algorithms.

Course Learning Objectives: Upon completing the course, you will be able to formulate biological problems as algorithmic problems and have an advanced background in applied data structures and algorithms. You will also gain experience in implementing software that can scale to large datasets.

Required Text:

Phillip Compeau and Pavel A. Pevzner. Bioinformatics Algorithms: An Active learning Approach. 3rd Edition

For the course it is required that you have access to the interactive online version through our course instance to do some course assignments. We have worked out an arrangement with the textbook authors where if you buy the print textbook (new or used) and provide a receipt they will grant free access to the interactive online version.

You thus have two options:

Option 1: Purchase only the interactive online textbook directly through our course instance:

<https://cogniterra.org/course/556/>

Note the purchase needs to be done our course instance and not the generic version.

Option 2: Purchase the print textbook (the textbook site <https://www.bioinformaticsalgorithms.org/> has links to a direct purchase or Amazon option) and then upload your receipt as a pdf through this form <https://forms.gle/kTUQKgUoZnovLiZw9>

Uploading of the receipt should be done by the end of Wed. January 7th, and then access to interactive online version for our course instance will be granted by the textbook authors soon after that. Receipts uploaded after January 7th will still be accommodated but may cause some delays.

If you previously bought (or mistakenly buy) the online book not through our course instance, then fill out this form:

<https://forms.gle/UFwuHSNPL83syt6q8>

to request free access to our course instance from the textbook authors.

Note the textbook website <https://www.bioinformaticsalgorithms.org/> has additional resources including videos from the textbook authors and answers to chapters FAQ.

Use the same name for the online textbook as you use in Bruinlearn so credit can automatically be transferred.

Grading Basis:

Letter Grading. Students are graded primarily on the basis of successful completion of homework assignments, a midterm, a non-cumulative final exam, projects, and completion of paper reading question and responses.

Graduate students are expected to complete a more difficult programming project and answer more questions on the exams. The course will require the reading of a few papers and responses to the papers and/or guest speaker will be required.

Grade Breakdown: Projects 25%. Homeworks 20%. Midterm Exam 25%. Final Exam 25%. Paper/Guest Speaker Question and Responses 5%.

Requisites:

Students taking the course should be familiar with programming.

Required unless approval of instructor:

- Computer Science 32 or Program in Computing 10C with grade of C- or better, and

Recommended:

- Civil and Environmental Engineering 110 or Electrical and Computer Engineering 131A or Mathematics 170A or Mathematics 170E or Statistics 100A.

Note: C121 is NOT a pre-requisite for C122.

Homework and Projects:

There will be online short homework problems sets due throughout the quarter integrated with the readings through the interactive textbook website. The deadlines are 12:00pm (noon) on their due date but we recommend doing them earlier as you do the reading. There will be four programming projects due throughout the quarter. Programming project outputs must be submitted to the Codalab project webserver by 12:00pm (noon) of the due date. Project credit will be based on the score displayed on the leaderboard. You must register your leaderboard handle on this form: <https://forms.gle/AeDVBsj5ve7Lwk5r7>. You can pick a handle that keeps you anonymous.

In addition, code must be submitted to the BruinLearn course website. Programs should be able to be run from standard command line interface as described in the programming assignment. There should be a shell command called runproject.sh which will run with the project input files located in the same directory. Do not upload the input files.

Students may discuss homework problems and projects with others but must write any code to solve the problems themselves. Standard built in software libraries can be used with the exception of libraries implementing the task of the homework or project unless instructed otherwise.

Papers and Guest Speaker:

Three research papers will be read over the course of the class. Each student is required to participate on the discussion of these papers on the discussion board. Each student is required to post one question about the paper by 12:00pm on Thursday when the reading is due and respond to two questions by 12:00pm of Tuesday of the next week. There will also be one guest speaker question which requires a question to be posted by 12:00pm on Thursday following the guest speaker and responses to two questions by 12:00pm of Tuesday of the next week.

Instructor in Charge:

If you have questions for a course instructor about textbook chapters, homework, paper readings, projects, or exams please direct them to the instructor in charge as follows:

Prof. Eskin: Chapters 1,3,5,9; HW1-4; Project 1-2; Paper 1-2; Guest speaker; Midterm;

Prof. Ernst: Chapters 2,8,10; HW5-7; Project 3-4; Paper 3; Final;

Note both TAs can be asked questions about any course content

Extension Policy:

If you have extenuating personal circumstances justifying an extension on a homework, project, paper/guest speaker question/response then email the instructor in charge of that assignment. Barring any extreme circumstances, extensions will be considered on those grounds only if requested in advance of the deadline. If you receive such an extension over email then record the extension you received in this form:

<https://forms.gle/TKYeX3bzd168GmjG6>

For projects, even without extenuating personal circumstances, using this form from a UCLA account <https://forms.gle/seUCUvjDKvXxy7JS9> you can request a 48hr project extension without extenuating personal circumstances. Note the form requires uploading the current code you have written to date. After the first 48hr period, if needed you can submit for another 48hr extension. If you are requesting an extension, you are expected to have been making an earnest attempt to complete the project on time. If you don't hear anything you can assume the extension is automatically granted. If an instructor/TA feels you may be abusing this extension system we will contact you. If you need an additional extension after 4 days, you must email the instructor in charge of the project explaining your progress and how long of an additional extension is needed. The instructor may consider your previous form entries in determining whether to grant an additional extension and for how long. If you receive an extension from the instructor past the 4 day period, then enter that extension on this form: <https://forms.gle/kVWKXxXPt8egFppV6>.

For homeworks, without extenuating personal circumstances, you can submit up to 1-week after the deadline and still receive some credit. The amount of credit awarded decreases linearly from the initial deadline.

Policy on Electronics during Lectures: Electronic devices including laptops and cell phones should not be used during lectures and discussion without permission from the instructor.

Policy on Lecture Attendance and Recordings: Lecture attendance is expected. Recordings of individual lectures are available upon request to the TA in charge of lecture recordings (Jingyuan Fu - fuyj2038951@ucla.edu) and copying the instructor that gave the lecture or in the case of the guest lecture Prof. Eskin.

Special Circumstances:

If you wish to request an accommodation due to a suspected or documented disability, please inform your instructor and contact the Office for Students with Disabilities as soon as possible at website: www.cae.ucla.edu.

Course Evaluations:

Completion of course evaluations is encouraged.

Tentative Course Syllabus (subject to change; All textbook reading assignments in Compeau and Pevzner (C&P); Some lecture material will not be in textbook)

	Date	Instructor	Topic	Textbook Readings	Assignment Due
Week 1	Mon Jan 6 th	Ernst/ Eskin	Intro to Course/Read Mapping	C&P Chapter 1	
	Tue Jan 7 th	-			
	Wed Jan 8 th	Eskin	Read Mapping	C&P Chapter 9	
	Thur Jan 9 th	-			
	Fri Jan 10 th	TAs	Discussion		
Week 2	Mon Jan 13 th	Eskin	Meta genomics/k-mer methods	C&P Chapter 9	
	Tue Jan 14 th	-			HW 1 – Chapter 1 Project 0
	Wed Jan 15 th	Eskin	Alignment	C&P Chapter 5	
	Thur Jan 16 th	-			Paper 1 – Question Project 1a
	Fri Jan 17 th	TAs	Discussion		
Week 3	Mon Jan 20 th	-	NO CLASS – MLK DAY		
	Tue Jan 21 st	-			HW 2 - Chapter 9 Paper 1 – Responses
	Wed Jan 22 nd	TBA	Guest speaker		
	Thur Jan 23 rd	-			Project 1b Guest Speaker – Question
	Fri Jan 24 th	TAs	Discussion		
Week 4	Mon Jan 27 th	Eskin	Alignment	C&P Chapter 5	
	Tue Jan 28 th	-			HW 3 – Chapter 5 Guest Speaker – Responses
	Wed Jan 29 th	TAs	Genome Assembly	C&P Chapter 3	
	Thur Jan 30 th	-			Project 1c
	Fri Jan 31 st	TAs	Discussion		
Week 5	Mon Feb 3 rd	Eskin	Genome Assembly	C&P Chapter 3	
	Tue Feb 4 th	-			HW4 – Chapter 3
	Wed Feb 5 th	Eskin	MIDTERM		
	Thur Feb 6 th	-			Paper 2 – Question Project 1d
	Fri Feb 7 th	TAs	Discussion		

Week 6	Mon Feb 10 th	Ernst	Clustering	C&P Chapter 8	
	Tue Feb 11 th	-			Paper 2 – Response
	Wed Feb 12 th	Ernst	Clustering	C&P Chapter 8	
	Thur Feb 13 th	-			Project 2a
	Fri Feb 14 th	Eskin	Discussion		
Week 7	Mon Feb 17 th	-	NO CLASS - President's Day		
	Tue Feb 18 th	-			HW 5 – Chapter 8
	Wed Feb 19 th	Ernst	Motifs	C&P Chapter 2	
	Thur Feb 20 th	-			Project 2b
	Fri Feb 21 st	TAs	Discussion		
Week 8	Mon Feb 24 th	Ernst	Motifs/Sequence Prediction	C&P Chapter 2	
	Tue Feb 25 th	-			HW 6 – Chapter 2
	Wed Feb 26 th	Ernst	Sequence Prediction		
	Thur Feb 27 th	-			Project 3a Paper 3 – Question
	Fri Feb 28 th	TAs	Discussion		
Week 9	Mon Mar 3 rd	Ernst	HMMs	C&P Chapter 10	
	Tue Mar 4 th	-			Paper 3 - Responses
	Wed Mar 5 th	Ernst	HMMs	C&P Chapter 10	
	Thur Mar 6 th	-			Project 3b
	Fri Mar 7 th	TAs	Discussion		
Week 10	Mon Mar 10 th	Ernst	HMMs	C&P Chapter 10	
	Tue Mar 11 th	-			HW 7 – Chapter 10
	Wed Mar 12 th	Ernst	Final Review		
	Thur Mar 13 th	-			Project 4
	Fri Mar 14 th	TAs	Discussion		
	Mon Mar 17 th	Ernst	FINAL (3:00pm-6:00pm)		