Bio331 Fall 2024 Syllabus

Bio331: Computational Systems Biology

In the era of high-throughput genomics, proteomics, and transcriptomics, we can now computationally explore biological processes at a systems level. Networks, or graphs, have become a dominant mathematical representation in this area of research. We will draw on the vast amount of established graph theory to learn about network models currently applied to biological systems.

About this syllabus

This syllabus is a public, searchable document that contains all of the details for this iteration of the course. It serves as a contract for the students taking Bio331 - if you take this course, you agree to abide by the policies described here. As the instructor, I commit to following this syllabus and administering it fairly and equitably. The main parts of the syllabus are:

- Goals: the course goals and learning objectives.
- Support: the support systems offered in the course.
- Policies: policies surrounding attendance, deadlines, collaboration, and generative AI (e.g., ChatGPT).
- Schedule: an overview of the course topics and assignemnts.
- Assessment: an explanation of how components in this course are evaluated and contribute to the final course grade.
- Learning Environment: my commitment to making the classroom an inclusive learning environment.
- Keys for Success: tips for succeeding in the course.

The final version of this syllabus will be set by Monday, Sept 2. Any modifications during the semester will be noted in the change log and will be communicated to the class via Moodle.

Course details

Instructor: Anna Ritz (Biology 200B, see the support page for student hours)

Lecture: MWF 10-10:50am in ETC 205

Labs: M 1:10-4:00 in ETC 208 and Tu 1:40-4:30 in ETC 205.

Materials: All reading, videos, and other materials will be freely available via Moodle.

Communication & Technology

Official communications about all course materials and assignments will be done through the Moodle page.

- Check your email every day to stay updated.
- We will use GitHub for the programming assignments.
- We will use Moodle for written assignment submissions.
- We will make a Slack workspace for informal communications and questions.
- If you have a question, others likely have the same question. Reach out at any time to me at aritz@reed.edu.

Goals & Objectives

At the end of Bio331, you will be able to:

1. Represent biological systems as networks and understand how they are built from data.
2. Implement graph algorithms and apply them to biological networks (including simulated and real-world datasets).
3. Interpret the results of an algorithm run on a biological network.
4. Read computational biology literature to understand the main points and claims.
5. Use technologies that enable collaboration and communication.
6. Present work through oral updates and written reports.
7. Design and execute a research project motivated by a question in systems biology.

Bio331 satisfies the Group III and Group III-Data Collection and Analysis distribution requirements, and you can find the relevant group distribution learning outcomes in the course catalog description.

Course support

Bio331 has a number of support systems to help you in this class.

Student Hours: Wednesday 12-1:30 and Friday 11-12pm, B200B

Student hours are times where I am in my office (Biology 200B) waiting for students to come talk with me. Drop in during these times to talk about anything (class related or otherwise), no appointment needed. Email me to set up an appointment if you cannot make student hour times.

Tutoring (DoJo)

Tutoring is offered by Academic Support Services, and gives students additional times outside of class to get help with assignments and the course material. Each student can get up to one hour per week *per class* of individual (one-on-one) tutoring. If someone requests individual tutoring, I will work to find a student who can tutor Bio331 material (this may take a bit of time because Bio331 has not been offered recently). Please reach out to the Office of Academic Support and Cc me if you are interested in tutoring.

Academic Accommodations

It is my goal to create a learning experience that is as accessible as possible. If you anticipate any issues related to the format, materials, or requirements of this course, please meet with me outside of class so we can explore potential options.

Please discuss any documented accommodations with me. Disability Accommodation Notification Letters can be obtained from Disability & Accessibility Resources. All discussions will remain confidential.

Illness and Exceptional Circumstances

There may be cases where you cannot attend class for stretches at a time (due to illness, for example). In this case, email me to strategize a successful plan for learning the material. If you miss multiple classes and do not email me, I may reach out to make sure you are OK.

In rare cases, there may be circumstances beyond your control where you are unable to meet the expectations outlined in this syllabus. If there are extenuating circumstances that might impact your academic performance in this class, please inform me as soon as you can to strategize a plan for completing all parts of the course. Additionally, you may want to reach out to the Student Life Office at student-life@reed.edu, phone: (503) 517-7396, office: Eliot 218. The Student Life Office is dedicated to assisting students in accessing campus and support resources, especially during challenging situations.

Obligated Reporting

I am happy to talk with you about anything that concerns you, but note that I (and many faculty) are obligated reporters who must report possible violations of the Title IX and DHSM policies which govern discrimination and harassment on the basis of sex, gender, and gender identity. Confidential resources are available at Health & Counseling Services and through the SHARE program; more information is available on this page.

Course Policies

There are four policies for this course: one for attendance, one for deadlines, one for collaboration (which is encouraged for the programming assignments) and one for the use of generative AI (which is discouraged).

Attendance Policy

Why Attend? Attending the lectures and completing work during labs will put you in the best position to succeed in the course. Lectures will provide valuable information about the assignments, and labs will include some group work. Labs also offer time for me to provide guidance - they could be framed as small-group office hours working through problems together.

What's the Policy? Email me if you cannot attend lecture or lab. There is no penalty for missing class or lab, but you may miss out on some discussion/worksheet components. If you are doing group work in lab, you have a responsibility to communicate with your group and catch up with work asynchronously. You are expected to review the missed material and meet with me if anything is unclear.

There are some religious holidays and other events in the fall that might affect student attendance - please note these in your calendars and let me know if your attendance will be affected.

- October 2-4: Rosh Hashanah
- October 11-12: Yom Kippur
- October 29 - November 3: Diwalai
- November 5: Presidential Election

If you miss multiple classes, I may reach out to make sure you are OK. Communication is key. Please see "Illness and Exceptional Circumstances" on the support page for more information about extended absences.

The Main Takeaway: There is no penalty for missing class, but attending and participating is expected and will set you up for success.

Deadline Policy

The the programming, written, and research project assignments all have deadlines. Refer to the schedule overview and the Moodle page for a detailed schedule with deadlines. The purpose of these deadlines is for you to complete the assignments at a steady pace throughout the semester.

Why Submit Work on Time? The programming and written assignments are designed so that you are working on them after you have learned the relevant material in lecture and lab. If you are working on older assignments when another assignment is out, you might lose ground with the current coursework.

What's the Policy? You must submit whatever work you have by the deadline, but there is no penalty for re-submitting programming and written assignments. If you plan to resubmit your work, you must include a comment or statement to that effect and schedule a time to meet with me within one week of the submission in order to make a plan. These details will be provided on every written and programming assignment.

You have until Finals Week to re-submit all assignments (but note that the assignments are substantial enough that you cannot wait until Finals Week to complete them).

The Main Takeaway: Assignment deadlines are intentional, and you should submit your work by the deadline. You will be able to resubmit any of the work, but the class is not designed so you can submit *all* the work at the very end of the semester.

Collaboration Policy

This collaboration policity is specific to Bio331 only.

1. Unless otherwise specified, collaboration on lab activities, class worksheets, and programming assignments is highly encouraged. When you collaborate, properly cite it in your work (see below).
2. You can always look up how to use specific functions in Python, as long as you know what it is doing and you properly cite it (see below).
3. You can always attend posted office hours and email me questions.
4. Do not use any graph library packages (e.g., `networkx` or `igraph`) or other math/stats packages (e.g., `scipy` or `numpy`) unless directed. You *may* use `pandas` to manage datasets, but check with Anna before using any of the built-in functions.
5. **You must write all of your own code for the programming assignments**. The research project can be done in groups, and you can also use graph library packages and other tools (see below).
6. Unless otherwise specified, **the written assignments must be completed on your own.**

Breaking the rules outlined in numbers 5-6 violates the Honor Principle and may result in severe consequences.

Citing Code

Code plagiarism is a real thing. Identical code is just as bad as copying and pasting entire paragraphs of an essay from another source - **always write your code in your own style.** You have a lot of flexibility in naming variables, including print statements and comments in your code.

Citing collaborations. Just like when writing papers, there is a way to cite others' work in programs. If you write functions with other students in the class, add their names as comments at the beginning of the function, e.g.,

```
def file2graph(file):
   """
   Function to convert a file to a JSON graph.
   Written with Jane and Rob for Lab 5
   """
```

When in doubt, cite discussions with other students. Here are some examples.

```
# I talked about the structure of this code with Marco
# Julie and I worked on the function post_network() together.
# I modified the convert() function from Lab 3
```

Remember, this type of collaboration is expected and encouraged! People who worked together should list each other as collaborators, since this will help us determine why some code might look similar.

Citing Previous Code from Class. If you copy your own code from previous assignments or labs, cite in the comments where you copied the code from (e.g., `# read_graph() function from Lab1.`). I will post solutions to the labs on Fridays; you are free to copy pieces of these solutions for future assignments, clearly state that it is from a posted solution (e.g., `# read_graph() function from posted Lab1 solution.`).

Citing Online Code. You can always look up basic Python syntax online - if it's more than "how do I structure this FOR loop" or "how do I sort a list" then be sure to cite the URL in the comments. Online resources linked from Moodle don't have to be cited. Do not look up entire functions that are part of an assignment or lab (e.g., calculating the degree distribution of a graph in Python or shortest paths algorithm in Python). While it may take some time for you to write these functions yourself, they are important for the learning goals in the class.

For the Research Project assignment, you may use Python libraries we have not previously used in Bio331 and modules and borrow code from the web. If you do this, add the following information as comments to your program. This pertains to both Python modules and StackOverflow-like queries. 1. Title and author information (if it exists) 2. URL 3. Brief description of what the code/module does.

Warning: There are many incorrect programs floating around the internet. Make sure that the code you use actually does what you want it to do.

The Main Takeaway: Collaboration in class and on programming assignments is expected and encouraged! You should "cite" anyone you worked with on the programming assignments. Copying previous code from class is OK; you may look up Python syntax online but avoid copying entire functions.

Generative AI Policy

Generative AI is technology that is trained to generate text, images, or code from natural language prompts. Two examples of generative AI that you might have seen before are ChatGPT (which returns text based on prompts) and GitHub Copilot (which returns code based on prompts).

The fact is, generative AI could do a pretty darn good job with some of the **programming assignments**. But the assignments are designed for *you* to learn about graph algorithms, which can only happen if *you* write the code. The course goals of this class include implementing graph algorithms and applying them to biological networks - if you use generative AI, you are not demonstrating what you have learned.

The same goes for using generative AI to answer the **written assignments**. Generative AI might get some questions correct, but *you* would not be demonstrating that you have learned the content. The **research project** will also have some written components; you should also write these in your own words without the aid of generative AI.

Don't worry - you will have opportunities to revise all asssignments in order to demonstrate your learning. If you are feeling pressure to use generative AI to just *get the answer no matter what*, then come talk with me to strategize a way to help you learn the content without that pressure.

The Main Takeaway: Generative AI should not be used in this class. But you should never feel the need to use it anyway.

Course Assessment

The course assessment is designed for students to be able to demonstrate what they have learned through programming assignments, written assignments, and a collaborative research project. Each assignment will have a clear set of criteria for success (e.g. a rubric or some other assessment description).

- **Programming Assignments (30%)**: graded Python programming assignments which build upon content learned in labs and lectures. These can be done in collaboration with others as outlined in the collaboration policy.
- **Written Assignments (30%)**: a combination of multi-part questions (which are graded) and paper summaries (which are marked as complete/incomplete). These should be completed individually, and you can

use and cite any resources (notes, internet, course materials, etc.).

- **Research Project (30%)**: Individually or in groups, you will design and execute your own small research project - there are mutliple milestones throughout the semester that will demonstrate all learning objectives. You will have in-class lab time to work on this project.
- **Labs (10%)**: each lab will have a minimum requirement for completion, which is designed to be accomplished within the lab period. These are evaluated as submitted/missing. Labs must be submitted by Friday of that week - a solution will be posted by Friday evening.

The Honor Principle is in effect at all times. Written assignments are completed individually; for all other assignments you can work with each other. Refer to the policies pagefor more information.

Note: Bio131 S24 used contract grading, where all work was marked as Complete, Partially Complete, or Missing, and your grade was based on the number of Complete vs. Partially Complete vs. Missing assignments you had. Bio331 is not yet contract grading, but I plan to add labor based grading elements to Bio331.

Schedule Overview

This schedule may shift and the assignment timelines are approximate; refer to the Moodle page for up-to-date information and a full schedule.

	Lecture	Lab	Assignment
Week 1	Intro to CompSysBio	No Lab	P1: Programming Refresher
Week 2	Animal Social Nets & Net Properties	L1: Degree Distribution	W1: Paper Summary & Bio Refresher
Week 3	Random Graph Models & Centralities	L2: Network properties	P2: Centralities on Animal Social Nets
Week 4	Shortest Paths & Ecology Nets	L3: Shortest Paths	
Week 5	Community Detection & Signaling Pathways	L4: Spanning Trees	P3: Community Detection Food Webs
Week 6	Steiner Trees	L5: Steiner Trees	W2: Algs & Nets I
Week 7	Random Walks & Human Interactomes Fall Break_	Research: Data Collection	

	Lecture	Lab	Assignment
Week 8	Motifs & Neuroscience Nets	L6: Random Walks	P4: Rich Club Network in Brain Nets
Week 9	Co-Expression Nets	Research: Algorithms	
Week 10	Disease Networks	Group Work Time	W3: Algs & Nets II
Week 11	Protein Structure Networks	Research: Interpretation & Analysis	
Week 12	Special Topics	Group Work Time	Research Project
Week 13	Wrap Up	Research: Writing & Presentation	Research Project
Week 14	Project Presentations	No Lab	Research Project

There is no final exam. All work due by the end of Finals week - no exceptions.

Learning Environment

Bio331 brings together students from different majors to tackle inherently interdisciplinary topics, which is exciting but can also feel overwhelming at times. My goal is have a learning environment where students are motivated to learn the material without the fear of being incorrect or penalized for mistakes.

Feedback for the Instructor

Please give me feedback on anything in the syllabus or course, especially with respect to making Bio331 an inclusive learning environment. Let me know if anything makes you uncomfortable in class, if you would like more instruction on a topic, or if you are experiencing a hardship outside of class. Anonymous feedback is available on Moodle, though with a small class your comments may be identifiable.

Diversity & Inclusion

Bio331 is a combination of biology and computer science, and each field has been claimed to be free of racism and prejudice. This is simply not true. Historically, biological discoveries and advances in computer science have been dominated by privileged voices, namely those of white men. Computational biology, while more recent of a field, also lacks diversity along many important axes (including

race, gender, nationality, class, sexuality, religion, ability, etc.). To foster an inclusive learning environment:

1. I acknowledge the bias in course materials that stem from systemic privilege, and I aim to make Bio331 content more inclusive each time I teach the course.
2. I expect that students come to Bio331 with varying levels of knowledge in both biology and computer science, and I aim to provide course materials to help fill in gaps in both areas.
3. Many students will have personal circumstances that may affect their performance in the course. I will work with you to make adjustments to the course schedule as needed, and students should be encouraged to seek guidance on any anticipated or realized issues.
4. Mental and physical health is more important than attending all parts of class. If you miss multiple days of class, I will reach out to make sure you are okay.

Bio331 is a better course when students have a diversity of lived experience and previous knowledge.

Land Acknowledgement

As we come together to learn at Reed College, we acknowledge that the territory on which Reed stands is that of Indigenous and Native peoples. The Portland Metro area rests on traditional village sites of the Multnomah, Wasco, Cowlitz, Kathlamet, Clackamas, Bands of Chinook, Tualatin, Kalapuya, Molalla, and many other tribes who made their homes along the Columbia River (Portland Indian Leaders Roundtable, 2018). Native lands have been taken through hundreds of breaches of treaty laws, and in 1953 more then sixty tribes in Oregon were terminated by the federal government which removed millions of acres of land from Indigenous stewardship (The Native American Community in Multnomah County: An Unsettling Profile). This statement is one small step in acknowledging the history that brought us to reside on this land. Portland also has a vibrant and diverse Indigenous community and has the ninth largest urban Native American population in the US. We honor the present Indigenous communities in the area. Visit native-land.ca to explore an interactive map of Indigenous lands.

- **Start the assignments early.** The assignments are designed to be done over 1-2 weeks (and you may spend many hours debugging your programs). Read the assignments when they are assigned and follow the suggested guidelines to make sure you are on track to complete all the work.
- **This course is cumulative** – topics build upon each other throughout the semester. If you are having trouble with a concept or assignment, talk to me early before you find that you need that concept in a later assignment.

- **When in doubt, ask.** Asking for help is sometimes unintuitive, but a little clarification goes a long way. If you have the question, others will probably have the same question. Ask early so I can communicate the answers to the rest of the class. Also, sometimes I make mistakes! Your question might help me fix a typo.
- **When you don't even know where to start, ask.** This can be the toughest thing when working on problems that do not have a nice, clean solution - where to start? Coming to office hours when you don't even have a properly-formed question is still valuable.
- **Working collaboratively is fun!** Try to talk through complicated concepts with others, and make use of the group time in lab to get clarification. You can work with others outside of class on the programming assignments as outline in the collaboration policy.
- **Start the assignments early.** Really.

Change Log

The purpose of a change log is to document all major changes to some code/website/etc. Starting on September 2, all major changes will be documented in this change log. If you want to see *all* the changes made to this website since it was first created, take a look at the GitHub commit log.

9/4/2024

Syllabus is finalized.

9/11/2024

Two changes based on feedback: - In the description of the Labs, Labs must be submitted by Friday of that week - a solution will be posted by Friday evening. -In the Collaboration Policies, I added a "Citing Previous Code from Class" section and a "Citing Online Code" section.