

BIOL 5312.001 Programming in the biological sciences for graduate students

Fall 2020

Classroom MS Teams, TR 8:30am-9:45am

Instructor Dr. Faruck Morcos Office BSB 12.601
email: faruckm@utdallas.edu Office hours: by appointment
Department of Biological Sciences TA: Hong Wei
University of Texas at Dallas **TA's email:** Hong.Wei@utdallas.edu

Pre-requisites

BIOL 2281: Introductory Biology Laboratory, BIOL 2311: Introduction to Modern Biology I and BIOL 2312: Introduction to Modern Biology II.

Course Description

This course is an introduction to programming practices using Python designed specifically for students in the biological sciences. Special emphasis will be put in particular features of Python like object oriented programming, some data structures as well as applications to process, model and analyze biological data. The course will cover fundamentals of Biopython, a popular set of tools for biological computation. One goal of this course is to provide a strong background on programming skills on a basic level while leaving more advanced techniques of software development and algorithms for other advanced courses.

Outcomes

Students will be able to:

1. Be proficient with programming platforms and environments like Unix/Linux which will allow an easier transition to other languages like AWK or Matlab.
2. Identify and use the most important structures needed to develop basic programs. These structures include data types, iterations and control structures.
3. Write programs and algorithms that analyze biological data like nucleotide and amino acid sequence data, expression data and molecular interaction datasets.
4. Build programs to model biological processes using the principles of object oriented programming.
5. Write efficient scripts that can be used to analyze and visualize the large amounts of genomic and molecular data deposited in local and public databases.
6. Students will be able to apply programming techniques to solve problems in their area of research. Particularly focused on biological data analysis and modeling.

Textbook and references

- Sebastian Bassi. *Python for Bioinformatics. 2nd Edition, 2017*
- Martin Jones. Python for Biologists. 2015 (**reference**)
- Neil Jones & Pavel Pevzner. An introduction to bioinformatics algorithms. MIT Press. 2004. (**reference**)

Class Schedule

<i>Aug. 18: Tue</i>	Computation and Life Sciences
<i>Aug. 20: Th</i>	Introduction to Python
<i>Aug. 24: Tue</i>	Linux and data types
<i>Aug. 27: Th</i>	Flow control
<i>Sept. 1: Tue</i>	Loops & Iterations
<i>Sept. 3: Th</i>	Loops & Iterations (cont.)
<i>Sept. 8: Tue</i>	Biological examples: interactive lecture
<i>Sept. 10: Th</i>	Functions
<i>Sept. 15: Tue</i>	Function Parameters
<i>Sept. 17: Th</i>	Modules and Packages
<i>Sept. 22: Tue</i>	Modules exercise
<i>Sept. 24: Th</i>	Recursion
<i>Sep. 29: Tue</i>	Midterm
<i>Oct. 1: Th</i>	Strings and multidimensional arrays used in Biosciences and File I/O
<i>Oct. 6: Tue</i>	Sequence Conservation Interactive Lecture
<i>Oct. 8: Th</i>	Sequence Logos
<i>Oct. 13: Tue</i>	Input/Output
<i>Oct. 15: Th</i>	Introduction to Biopython
<i>Oct. 20: Tue</i>	Common Data structures in Biopython
<i>Oct. 22: Th</i>	Tuples, Vectors and Dictionaries
<i>Oct. 27: Tue</i>	Biopython functions
<i>Oct. 29: Th</i>	Common datasets in biology (in-class exercise)
<i>Nov. 3: Tue</i>	Introduction to Classes
<i>Nov. 5: Th</i>	Practical examples of Classes
<i>Nov. 10: Tue</i>	In-class exercise on Classes
<i>Nov. 12: Th</i>	Member functions and class concepts
<i>Nov. 17: Tue</i>	Class function definitions with examples
<i>Nov. 19: Th</i>	The Biological Molecule Class
<i>Nov. 24: Tue</i>	<i>Thanks giving break</i>
<i>Nov. 26: Th</i>	<i>Thanks giving break</i>
<i>Dec. 1: Tue</i>	The Biological Molecule Class (cont.)
<i>Dec. 3: Tue</i>	Logistics and Review of Applications to Biology
<i>Final Exam</i>	Date: TBD, Time: TBD , Place: TBD

Course Policies

Grading

The grade is composed by a weighted average of the grades in midterm exam (15%), in-class participation (5%), a final exam (30%), homework assignments (40%) and a research project (10%).

Homeworks

Programming homeworks will be assigned during the lecture. Homework guidelines will be posted on eLearning in a separate hand out.

Collaboration Policy

Collaboration during homeworks and study sessions is highly recommended. Exchange of ideas is part of the learning process and programming is not an exception. However, **there should be a clear line between a healthy collaboration and unethical practices**. Homework problems can be discussed, debugging strategies can be shared as well as general ideas that could improve understanding of a topic or homework. **Writing code must be an individual task** and every submission should be easily reproducible by its author. **Copying code**, from another student or the internet, is against the collaboration policy. *Changing the name of variables* is **not** writing your own code :-). Groups of 2 students can be formed to officially collaborate in assignments, you must put the name of the person in your team in every submission.

Exams

There will be **two exams**, a **midterm** exam and a **final** exam. Exams will have a combination of theory questions and programming problems. Given the nature of the material discussed in class, the final exam will be accumulative and will cover all the topics taught during the semester with an emphasis on the second part of the course.

This course will use Honorlock - an online exam proctoring tool. To successfully take an exam, you must have a web camera with microphone, a laptop or desktop computer (no tablets/phones), Chrome browser, a reliable internet connection and your photo ID. You will be prompted to install the

Final Project

A research project involves applying the concepts learned in class to a particular problem related to the graduate student interest or research area. Students will write a report with methods and results of the proposed work. More details about the final project are specified in a separate document.

Attendance and Participation

Lecture **attendance** is an important part of your learning experience. Attending a lecture should be seen as a resource rather than an obligation. Attendance will not be strictly enforced but rather encouraged to improve the exchange of ideas, solve questions and clarify concepts. **Participation**, on the other hand, enriches the class experience and interactivity greatly improves learning connections. Participation also sets the pace of the lecture, an important feature to avoid falling asleep! Attendance, participation as well as in class exercises or quizzes contribute to the final grade. If you *cannot attend to the lecture synchronously*, then you need to work on exercises asynchronously in order to receive credit for your work.

Academic Honesty

Students are expected to follow the rules and guidelines of academic honesty established by the University of Texas. Basically, be honest about your contributions to homeworks, work on your own during exams and spend enough time reading and trying to understand class materials. This creed was voted on by the UT Dallas student body in 2014. It is a standard that Comets choose to live by and encourage others to do the same:

“As a Comet, I pledge honesty, integrity, and service in all that I do”

Note: The descriptions and timelines contained in this syllabus are subject to change at the discretion of the Professor.