

New York City College of Technology
School of Arts and Sciences
Department of Biological Sciences

Programming for Biologists (BIO 2110) SYLLABUS

Course Information

Course Title: Programming for Biologists (Lecture and Laboratory)

Course Code: BIO 2110

Credits Hours: 2 hours lecture and 4 hours laboratory per week; 15 weeks total.

Prerequisite: BIO2000 (intro).

Required Lecture Text: *Practical computing for biologists*. SINAUER.
Haddock and Dunn

Supplemental optional reading:

Bioinformatics Programming Using Python. O'REILLY. Mitchell L. Model

Course Description: Foundational training in programming for Biomedical Informatics students. Hands-on exercises will introduce students to the usage of UNIX, scripting language and biomedical informatics tools needed for their research. Lectures cover theoretical concepts and the labs will provide direct application to biomedical informatics systems.

Course Mechanics: All the concepts and techniques taught in this course are computer-based. Assignments will be assigned periodically, and regular and active participation in discussions is required. Timely completion of assignments is critical to success in the course.

Attendance is absolutely required. Aside from serving as the venue to

introduce new topics, it will also provide an opportunity for students to discuss any difficulty they are having regarding the course.

Grading Procedure (see Grading Policies for details)

The grade is based on weekly quizzes and assignments, a group project, and a final exam.

Course Coordinators

Name: Joanne Weinreb

Office: P410

Email: jweinreb@citytech.cuny.edu

Name: Mai Zahran

Office: P313

Email: mzahran@citytech.cuny.edu

Course Objectives and Student Expectations

Students are expected to be able to work independently and regularly, as well as collaborate with fellow students on group projects if required. This course is fast paced, and covers a diverse set of topics, and therefore students must be able to keep up with the work assigned in order to be successful in the course.

Course Objectives

Having successfully completed this course, the student will be able to:

1. Understand the UNIX command line environment.
2. Use built-in UNIX commands to manipulate files and data.
3. Use bash shell scripts to drive pipelines of biomedical informatics programs.

4. Use scripting language to read, manipulate and write biomedical informatics data files.
5. Understand the utility of UNIX and scripting language in computational biology and biomedical informatics.
6. Apply visualization techniques to biomedical informatics datasets.
7. Access remote computers and clusters.
8. Learn about relational databases and how to access them.
9. Learn how to work as a team on an *application*.

Technology Prerequisites

1. Students should have access to and be able to use Internet Explorer, Firefox, or any appropriate web browser. Internet Explorer and Firefox work best with Blackboard.
2. Students will need a City Tech email account and should be comfortable using it. Students will also need access to CUNY's Blackboard service. Accounts and passwords to the CUNY Portal should be arranged prior to the beginning of the semester.
3. Students should check if their e-mail address on Blackboard is the e-mail address they check most. The instructor will send e-mail announcements only via Blackboard.

Schedule

At the beginning of the semester, the instructor will explain in detail the policies and procedures of this course. In addition, a short introduction on the features of Blackboard will be given on the first day of class. Instruction on other features of Blackboard (discussion board, wiki) will be provided too.

Some of the class material will be found in the "Assignments" section of the course in Blackboard. Because of the demanding amount of

material that will be covered in the class, deadlines will be followed strictly.

Online Resources

CUNY's Blackboard resource can be accessed via the CUNY Portal, at:

<http://portal.cuny.edu/portal/site/cuny/index.jsp>

The National Center for Biotechnology Information, which hosts all the databases that will be used in this course, as well as tutorials on how to navigate around the website and the databases, can be found here:

<http://www.ncbi.nlm.nih.gov/>

Easy access to all the online resources for Bioinformatics can be found here: <http://www.ncbi.nlm.nih.gov/guide/all/>

Lecture Schedule

Week	Lecture	Laboratory
1	Introduction to UNIX and biology review <i>Required reading Biology review found on Blackboard</i>	Basics of Unix, editing environment setup (Unix vs Linux)
2	Regular expression	Reformatting molecular data
3	Intro to shell scripting Advanced UNIX, SED, AWK	Manipulating/Parsing biological data files
4	Intro to Python environment and programming Variables, syntax, conditionals, comparators	Create DNA sequences, Find Open Reading Frame

Week	Lecture	Laboratory
5	Python loops, files manipulation	Central Dogma in Biology using Python
6	Application of modules – Part 1	Applied Numpy modules in Biomedical Informatics data
7	Application of modules – Part 2	Applied Numpy modules in Biomedical Informatics data
8	Introduction to Relational Database	Learn the basics of SQL to query an EHR database
9	Interacting with SQL from Python	Embedding SQL within python to access EHR database
10	Creating functions and objects	Analyzing EHR data
11	Graphics Visualization – Part 1	Introduction to Data Visualization Matplotlib
12	Graphics Visualization – Part 2	Data Visualization Matplotlib
13	Create an open access for an application	Group work on the project
14	Access to remote computers SSH, SCP, FTP	Access NCBI computer server and computer clusters
15	Group project presentation	FINAL exam

College Policy on Absence/Lateness

A student may be absent without penalty for 10% of the number of scheduled class meetings during the semester as follows:

Class Meets Allowable Absence

1 time/week - 2 classes

2 times/week - 3 classes

Students are responsible for making up any missed work on days that they are absent. If a student's class absences exceed this limit the instructor will alert the student that a grade of WU may be assigned. Unless otherwise indicated by the instructor, two times late is treated as one absence.

Grading Policies

Please bear in mind that this course is a **4-credit** course. Student performance on this course will be evaluated as follows:

ASSIGNMENTS	DESCRIPTION	POINTS
Programming assignments	Expected timely completion of each lab assignments	35%
In-Class Quiz	Weekly Lab Quiz	20%
Group Project	Open Access	20%
Final	Cumulative Final exam	25%

Note:

Letter grades will be determined using a standard percentage point evaluation as outlined below:

Letter Grade	Point Range
A	93-100
A-	90-92.9
B+	87-89.9
B	83-86.9
B-	80-82.9
C+	77-79.9
C	70-76.9
D	60-69.9

Letter Grade	Point Range
F	Below 60

Policy on Academic Integrity

Academic dishonesty includes any act that is designed to obtain fraudulently, either for oneself or for someone else, academic credit, grades, or any other form of recognition that was not properly earned. Academic dishonesty, which will not be tolerated in this course and at City Tech, encompasses the following:

Cheating: Defined as intentionally giving, receiving, using or attempting to use unauthorized materials, information, notes, study aids, including any form of unauthorized communication, in any academic exercise. It is the student's responsibility to consult with instructors to determine whether or not a study aid or device may be used.

Plagiarism: Plagiarism is intentionally and knowingly presenting the ideas or works of another as one's own original idea or works in any academic exercise without proper acknowledgement of the source. The purchase and submission of a term paper, essay, or other written assignment to fulfill the requirements of a course, and violates section 213-b of the *State Education Law*. This also applies to the submission of all or substantial portions of the same academic work previously submitted by the student or any other individual for credit at another institution, or in more than one course.

Course Policy on Academic Integrity: Cheating and plagiarism will not be tolerated in this course. Penalties are the following. Cheating in in-class exams or quizzes will merit an automatic zero for the exercise. Copying from classmates' lab worksheets and other take-home or online assignments will also merit an automatic zero for the exercise. Repeated violations will be reported to the Chair and the Dean, and

may result in a final grade of “F” in the course, or even expulsion from the College. If you are unsure whether any of your actions constitute cheating or plagiarism, please consult the instructor for guidance.

College Policy on Academic Integrity: Students and all others who work with information, ideas, texts, images, music, inventions, and other intellectual property owe their audience and sources accuracy and honesty in using, crediting, and citing sources. As a community of intellectual and professional workers, the College recognizes its responsibility for providing instruction in information literacy and academic integrity, offering models of good practice, and responding vigilantly and appropriately to infractions of academic integrity.

Accordingly, academic dishonesty is prohibited in The City University of New York and at New York City College of Technology and is punishable by penalties, including failing grades, suspension, and expulsion.