CMPSC 300
Bioinformatics
Syllabus

Fall 2022

Course Instructor

Dr. Oliver BONHAM-CARTER (said and written as "Bonham-Carter," not "Carter")
Classroom: Alden Hall 109
Office Location: Alden Hall 104
Email: obonhamcarter@allegheny.edu
Web Site: http://www.cs.allegheny.edu/sites/obonhamcarter/
Exam Code: G
Final deliverable due: 19th December 2022, 9:00am
Distribution Requirements: *QR* and *SP*
Syllabus updated on: July 31, 2022

Instructor's Office Hours

To schedule a meeting with me during my office hours, please visit my Web site and click the "Schedule" link in the top right-hand corner. Now, you can view my calendar or by clicking "schedule an appointment" link browse my office hours and schedule an appointment by clicking the correct link to reserve an open time slot.

Technical Leaders

- https://www.cs.allegheny.edu/teaching/technicalleaders/

Course Meeting Schedule

Lecture, Discussion, Presentations, and Group Work:
Duration: 30 August 2022 - 20 December 2022
Monday, Wednesdays and Fridays, 10:00 AM - 10:50 AM, Alden Hall, 109

Laboratory Session:
Duration: 30 August 2022 - 20 December 2022

Monday, 2:30 PM - 4:20 PM, Alden Hall, 109

Calendar

The calendar link is provided below to allow you to add the course and lab meeting times into your own Google calendar. Note, the whole link fits onto one line.

`https://calendar.google.com/calendar/u/0?cid=`

`Y182aWs5ZGpub3Y0NzNqcDZmZ2YxZjUxMG00NEBncm91cC5jYWxlbmRhci5nb29nbGUuY29t`

Slack Channel

The below link will expire on the 21st March.

`https://join.slack.com/t/cs300s2021/shared_invite/zt-mleieaad-i6G1To2GRVjMqoOJcak5qQ`

Academic Bulletin Description

An introduction to the development and application of methods, from the computational and information sciences, for the investigation of biological phenomena. In this interdisciplinary course, students integrate computational techniques with biological knowledge to develop and use analytical tools for extracting, organizing, and interpreting information from genetic sequence data. Often participating in team-based and hands-on activities, students implement and apply useful bioinformatics algorithms. During a weekly laboratory session students employ cutting-edge software tools and programming environments to complete projects, reporting on their results through both written documents and oral presentations. Students are invited to use their own departmentally approved laptop in this course; a limited number of laptops are available for use during class and lab sessions.

Prerequisite: BIO 221 and FSBIO 201, or CMPSC 100.

Distribution Requirements: QR, SP.

Course Objectives

Students successfully completing this class will have developed:

1. A "big-picture" view of bioinformatics.

2. An understanding of the objectives and limitations of bioinformatics.

3. An understanding of the biological foundations of bioinformatics (genes and genomes, gene expression, etc.).

4. An understanding of the computational foundations of bioinformatics (programming, databases, etc.).

5. An understanding of how genetic information is obtained and processed.

6. The ability to use basic bioinformatics software tools to study genetic information.

Throughout the semester students also will enhance their ability to write and present ideas about bioinformatics in a clear and compelling fashion. Students will gain practical experience in the design, implementation, and analysis of bioinformatics research during laboratory sessions and a final project. Finally, students will develop a richer understanding of the fascinating connections between biological systems, analysis and automation.

Required Textbooks

- *Exploring Bioinformatics: A Project-based Approach, second edition*, by Caroline St. Clair and Jonathan E. Visick.

- *Think Python, first edition*, by Allen B. Downey.

    - Textbook: `http://greenteapress.com/thinkpython/thinkpython.pdf`
    - Publisher: `http://greenteapress.com/wp/think-python/`

Suggested Reading

The below reading list is strongly recommended to improve students build technical writing skills and to gain a firm understanding in how to conduct responsible research in computer science.

- *BUGS in Writing: A Guide to Debugging Your Prose.* Lyn Dupré. Second Edition, ISBN-10: 020137921X, ISBN-13: 978-0201379211, 704 pages, 1998.

- *Writing for Computer Science.* Justin Zobel. Second Edition, ISBN-10: 1852338024, ISBN-13:978-1852338022, 270 pages, 2004.

- *On Being a Scientist: A Guide to Responsible Conduct in Research (Third Edition).* Committee on Science, Engineering, and Public Policy, National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. ISBN: 0309119715, 82 pages, 2009. References to the textbook are abbreviated as "OBAS".

- Along with reading the required books, you will be asked to study many additional articles from a wide variety of conference proceedings, journals, and the popular press.

The `ClassDocs/` Repository

All materials given out in class will be accessible using the `classDocs/` repository. Note: The HTTP link works in absence of SSH keys.

Main site on GitHub:

- `https://github.com/Allegheny-Computer-Science-300-S2021/classDocs`

HTTP:

- `git clone https://github.com/Allegheny-Computer-Science-300-S2021/classDocs.git`

SSH:

- `git clone git@github.com:Allegheny-Computer-Science-300-S2021/classDocs.git`

Class Policies

Grading

The grade that a student receives in this class will be based on the following categories. All percentages are approximate and, if the need to do so presents itself, it is possible for the assigned percentages to change during the academic semester.

Class Participation	10%
Exams	20%
Laboratory Assignments	40%
Final Project	30%

Definitions of Grading Categories

- *Class Participation*: All students are required to actively participate during all of the class sessions. Your participation will take forms such as answering questions about the required reading assignments, completing in-class exercises, asking constructive questions of the other members of the class, giving presentations, leading a discussion session in class.

- *Exams*: The exams will cover all of the material in their associated module(s). The finalized date for each of the exams will be announced at least one week in advance of the scheduled date. Unless prior arrangements are made with the course instructor, all students will be expected to take these exams on the scheduled date and complete the exams in the stated period of time.

- *Laboratory Assignments*: These assignments invite students to explore the concepts, tools, and techniques associated with the field of bioinformatics. All of the laboratory assignments require the use of the provided tools to study, design, implement, and evaluate informatics systems that solve biology problems. To ensure that students are ready to utilize and develop appropriate software in

both other classes at Allegheny College and after graduation, the instructor will assign individuals to teams for some of the laboratory assignments. Unless specified otherwise, each laboratory assignment will be due at the beginning of the next laboratory session. Some of the assignments in this course will expect students to give both a short presentation and a demonstration of the bioinformatics solution that they created.

- *Final Project*: This project will present you with an opportunity to design and implement a correct and carefully evaluated bioinformatics solution for a specific problem. Completion of the final project will require you to apply all of the knowledge and skills that you have acquired during the course of the semester to solve a bioinformatics problem. The details for the final project will be given approximately two months before the project due date (during finals week).

Assignment Submission

We will be using GitHub Classroom to collect all assignments. It is expected that you are able to effectively use `git` to submit your work. If you require help, please see your peers, the Technology Leaders, or your instructor.

All assignments will have a stated due date. The electronic version of the class assignments are to be turned in at the beginning of the lab session on that due date. Submissions after the beginning of class are counted as being late. Assignments will be accepted for up to one week past the assigned due date with a 15% penalty. All late assignments must be submitted at the beginning of the session that is scheduled one week after the due date.

Extensions

Unless special arrangements are made with the course instructor, no assignments will be accepted after the late deadline. If you are requesting extensions for a lab assignment, then you are to email me with your request and also provide a *valid reason* for your extension. This request must come before the due date of the lab and not on the due date. Requests will not be granted where the reason appears to be insignificant. Extensions are 24 hours of extra time (after the original due date) and are given out at my discretion. The decision to provide you with an extension (or not) will be weighed in light of fairness to your peers who are still able to complete their labs, regardless of their own busy schedules.

The submission of homework comprises the Honor Code pledge of the student(s) completing the work. For any assignment completed in a group, students must also turn in a one-page reflection that describes each group member's contribution to the submitted deliverables.

Attendance

Classes will be attended by in-person and online students. Each class will be recorded to produce videos for online students and to enhance learning for the class.

If you will not be able to attend your session, then please email the course instructor at least one week in advance to describe your situation. Students who miss more than five unexcused classes, laboratory sessions, or group project meetings will have their final grade in the course reduced by one letter grade. Students who miss more than ten of the aforementioned events will automatically fail the course.

Labs: The laboratory sessions are in-person. It is the student's responsibility be present for each meeting to ensure a comprehension of materials.

Bring your own computer to class

The classrooms in the Department of Computer Science no longer provide machines for student use. You are to bring your own wifi-ready device to class to be able to follow along with course material. If the class is meeting online using Zoom, then please be sure that you machine is configured correctly to use these services to connect you to the class. As it is your responsibility to maintain your machine, please perform online research to determine how to configure your machine accordingly, or to install any necessary software to enable online meetings.

During the semester, you will be told which software to install on your machine to be prepared for class. Some of the prominent software that we may be using include;

- Git and GiHub (a software development software system): `https://github.com/`

- VsCode (an editor): `https://code.visualstudio.com/download`

- Docker (a software container system): `https://www.docker.com/`

  - Installing Docker: `https://docs.docker.com/get-docker/`
  - Basic tutorial from Docker: `https://www.docker.com/101-tutorial`
  - Play with Docker: `https://labs.play-with-docker.com/`

- SQLite (a database system): `https://www.sqlite.org/index.html`

  - Some machines have differing methods of installing this database software. Please complete online research to read documentation to determine how to get your machine ready to run this software.

Extensions

Unless special arrangements are made with the course instructor, no assignments will be accepted after the late deadline. If you are requesting extensions for an assignment, then you are to email me with your request and also provide a *valid reason* for your extension. This request must come before the due date of the lab and not on the

due date. Requests will not be granted where the reason appears to be insignificant. Extensions are 24 hours of extra time (after the original due date) and are given out at my discretion. The decision to provide you with an extension (or not) will be weighed in light of fairness to your peers who are still able to complete their labs, regardless of their own busy schedules.

A Note on extenuating circumstances

If you should find yourself in difficult circumstances that significantly interfere with your ability to prepare for this class and to complete assignments, please inform me immediately so that we can work something out together! Do not wait until the last day of class to ask for exceptions to what is stated in this syllabus. In such a situation, you may also find it helpful to contact one of the available resources on campus:

- The Maytum Learning Commons, Library/Academic Commons,
  http://sites.allegheny.edu/learningcommons/tutoring/,
  814-332-2898

- Counseling & Personal Development Center,
  https://sites.allegheny.edu/counseling/,
  814-332-2105

- Winslow Health Center,
  https://sites.allegheny.edu/healthcenter/,
  814-332-4355

Communication

Various digital channels will be used in this course for communication, including email, Discord, and the GitHub issue tracker. It is strongly advised for the student to install the Discord app on their computer and smart-phone to be sure to receive all communications from the instructor, as well as, the other members of the class.

Additionally, the course website will be used to store the syllabus, course schedule and information about the `classDocs/` repository using the GitHub. Your grades will be communicated to you by a Gradebook GitHub repository.

Special Needs and Disability Services

The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. Students with disabilities who believe they may need accommodations in this class are encouraged to contact Disability Services at 332-2898. Disability Services is part of the Learning Commons and is located in Pelletier Library. Please do this as soon as possible to ensure that approved accommodations are implemented in a timely fashion.

Honor Code

The Academic Honor Program that governs the entire academic program at Allegheny College is described in the Allegheny Course Catalogue. The Honor Program applies to all work that is submitted for academic credit or to meet non-credit requirements for graduation at Allegheny College. This includes all work assigned for this class (e.g., examinations, laboratory assignments, and the final project). All students who have enrolled in the College will work under the Honor Program. Each student who has matriculated at the College has acknowledged the following pledge:

> *I hereby recognize and pledge to fulfill my responsibilities, as defined in the Honor Code, and to maintain the integrity of both myself and the College community as a whole.*

It is recognized that an important part of the learning process in any course, and particularly one in computer science, derives from thoughtful discussions with teachers and fellow students. Such dialogue is encouraged. However, it is necessary to distinguish carefully between the student who discusses the principles underlying a problem with others and the student who produces assignments that are identical to, or merely variations on, someone else's work. While it is acceptable for students in this class to discuss their programs, technical diagrams, proposals, paper reviews, presentations, and other items with their classmates or other individuals, deliverables that are nearly identical to the work of others will be taken as evidence of violating the Honor Code.

Welcome to Computer Science Research!

Computer hardware and software are everywhere! Conducting research in computer science is a challenging and rewarding activity that leads to the production of hardware, software, and scientific insights that have the potential to positively influence the lives of many people. As you learn more about research methods in computer science you will also enhance your ability to effectively write and speak about a wide range of topics in computer science. I ask that you bring your best effort and highest enthusiasm as you pursue research in computer science this semester.