San José State University
Computer Science Department
CS 123B Bioinformatics II, Sec 01, Fall 2019

Course Information

Instructor: Leonard P. Wesley
Department: Computer Science
College of Science, San Jose State University.
Fall Semester, 2019

Course and Contact Information

Instructor:	Leonard Wesley
Office Location:	MH 212
Telephone:	408.924.5287
Email:	Leonard.Wesley@sjsu.edu
Office Hours:	Tuesdays 2:00PM to 4:00PM
Class Days/Time:	Tuesdays and Thursdays 4:30PM – 5:45PM
Classroom:	DH 450
Prerequisites:	Prerequisite: CS 123A.

Official SJSU Catalogue Course Description

Computational methods used for searching, classifying, analyzing, and modeling protein sequences. Tools for analyzing DNA and RNA sequences. More advanced topics, such as genetic algorithms and simulated annealing, which can be used to address folding problems.

Expanded Course Description

The course investigates the main algorithms for solving computational problems in bioinformatics. Methods will include HMMs, sequence and structure alignment, and comparative structure analysis to identify and classify protein folds. Students will be given programming and/or web-portal projects that provide practice with using

bioinformatics related tools and algorithms. The advanced topic of gene editing using the CRISPR-Cas technology will be covered. In addition, the course will cover Next Generation Sequencing (NGS) technologies. Students will have an opportunity to provide samples of their own DNA, have it sequenced, and then learn how to assemble the sequenced DNA back into their original DNA sequence.

Learning Outcomes

Upon successful completion of this course, students will:

1. SLO-1 Have a basic understanding of molecular biology and the biology central dogma, the transcription and translation process, regulation of transcription and translation, genetic crossover and translocation, genes and alleles, genotype and phenotype.
2. SLO-2 Know how to design, build, and implement Genetic Algorithms and HMMs work, and how they along with various web-based portals can be used to help carry out gene finding and protein fold analysis.
3. SLO-3 Know how the CRISPR-Cas system works, find CRISPR arrays, and find genomic repeats.
4. SLO-4 Know how various Next Generation Sequencing (NGS) technologies work, and how to perform genome assembly and analysis.

Each SLO above corresponds to a learning module that is described in the course calendar below. That is, there are four (4) learning modules that corresponds to each of the SLOs described above.

Required Texts/Readings

All required text, publications, reference material, and so forth will be provided to the class.

Other Optional Reading Material

Developing Bioinformatics Computer Skills, Cynthia Gibas and Per Jambeck, O'Reilly & associates. (A good book for beginners)

Introduction to Computational Biology: Maps, Sequences and Genomes, Michael S. Waterman, CRC Press. (A statistical oriented view of bioinformatics)

Bioinfromatics: A Practical Guide to the Analysis of Genes and Proteins, Andreas D. Baxecvanis and B.F. Francis Ouellette, John Wiley & Sons 2nd Ed. (Includes contributions from several authors providing a wide perspective)

Course Requirements and Assignments

Course Logistics

Students should expect to spend approximately nine (9) hours per week (on average) outside of the classroom preparing for and completing the assigned course work. This includes reading papers, viewing videos as appropriate, completing homework and programming exercises, and so forth. The amount of time that a student actually spends studying and completing course work will depend on individual skills and the time that the student actually allocates to the course. The nine (9) hours per week estimate is based on previous experiences of the instructor and students. So please plan and schedule accordingly.

Previously, some students have asked for special exceptions to policies and procedures for this course. An example includes asking the instructor for extra assignments or work to help improve a grade. Even if such a request is reasonable in the opinion of the instructor, no exception will be given to a student unless the same opportunity can be made available to the entire class, and does not constitute significant extra work on the part of students, instructors, graders and so forth. Students should have no concern that other students will receive special exceptions that will not be made available to the entire class.

NOTE: University policy (F69-24) states that "Students should attend all meetings of their classes, not only because they are responsible for material discussed therein, but because active participation is frequently essential to insure maximum benefit for all members of the class. Attendance per se shall not be used as a criterion for grading." However, attendance will be required in order to complete and submit many in-class exercises, quizzes, and exams. Should students miss or leave early from one or more classes, students are responsible for knowing and understanding any and all course subject matter, assignments, exercises, instructions and so forth that are presented or discussed during official scheduled class time.

Success in this course is based on the expectation that students will spend, for each unit of credit, a minimum of 45 hours over the length of the course (normally three hours per unit per week) for instruction, preparation/studying, or course related activities, including but not limited to internships, labs, and clinical practica. Other course structures will have equivalent workload expectations as described in the Syllabus/Greensheet.

Quizzes and Exams

There will be three quizzes, one midterm and a final exam all of which will count toward a student's final grade as specified in the "Grades" section below. During quizzes and exams, communication with other individuals via any means is strictly prohibited without the express permission of the instructor. Violations will be met with the full impact of SJSU's academic integrity policy and procedures.

Projects

Several life-science related project topics will be described near the start of the course. Projects will involve applying the skills and knowledge learned in the course to the project. Projects in this course will be individual (not team) projects. Project scores will count toward the final grade as specified in the "Grades" section below.

In-Class Exercises

There will be four in-class exercises where groups of two to four will be formed to work on an assigned exercise. In-class participation is mandatory, and an attendance sign-up sheet will be passed around to verify participation. The assigned exercises are intended to reinforce learning and understanding of previous lecture, homework, and programming assignment subject matter by providing hands-on experience with completing the provided assignment. A supplement document named "In-Class Exercise Procedure.pdf" is available on Canvas in the same location as the course Syllabus/Greensheet. The "In-Class Exercise Procedure.pdf" document describes the general organization of all in-class exercise assignments as well as the procedure for completing and submitting all in-class exercises. The "In-Class Exercise Procedure.pdf" document should be treated as part of the Greensheet for this course.

Reading, Homework, Programming, Participation Assignments

Graded reading, homework, programming, class participation and brief course feedback assignments will be given almost weekly. All graded assignments will count toward a student's final course grade.

Computational Resources

Students are required to make sure that they have access to sufficient UNIX, Windows, or Mac based computational resources (e.g., computers and software) to carryout assignments in the course. An attempt to offer the course in a classroom with sufficient computation resources will be made by the department to support classroom instruction and demonstrations. However, students should be prepared to bring their portable laptops to class.

Questions and Regrade Requests

All questions about grading and re-grade requests must be presented to the instructor within two weeks from the date that graded assignments, exercises, and exams are returned to the class or by the last day of instruction for the semester (whichever is sooner). Assignments, quizzes, and exams will typically be returned (i.e., posted) to Canvas, or manually handed back in class. General questions about the topics covered in assignments, exams, exercises, programming assignments, and the course are permissible at any time.

Tentative course calendar of assignment due dates & exam dates:

(Please note that course calendar below, and its content is "subject to change with fair notice")

Week and Class Mtg #	Thur	Tue	Module # & Name	TOPIC	Assignment See Canvas For Module & Weekly Assignment Details and Due Dates
Week 1 Class Mtgs 1 & 2	8/22	8/27	#1 Biology Basics	**8/22:** Intro To Course: -Topics, learning objectives, course logistics, Instructor background - Greensheet **8/27:** - Intro to molecular biology, DNA, RNA, and the central dogma. - DNA Replication	Module #1 Week #1
Week 2 Class Mtgs 3 & 4	8/29	9/3	#1 Biology Basics	**8/29:** - Transcription, Translation - Regulation of transcription and translation **9/3:** - Genes, alleles, genotype, phenotype, genetic crossover, translocation, types of mutations	Module #1 Week #2 September 3rd Last Day To Drop Classes
Week 3 Class Mtgs 5 & 6	9/5	9/10	#1 Biology Basics #2 Comp Algorithms	**9/5:** Module #1 - **In-Class Exercise 1** Topics Covered 8/22 – 9/3 **9/10:** Module #2 - Genetic algorithms	Module #1 & Module #2 Week 3

4	9/12	9/17	#2 Comp Algorithms	**9/12:** - Project Descriptions - Finish Genetic Algorithms **9/17:** - HMMs	Module #2 Week 4
5	9/19	9/24	#2 Comp Algorithms	**9/19:** - **Quiz 1 (~35 mins): Covers Topics Week 1 thru Week 4** - HMMs **9/24:** - HMMs	Module #2 Week 5 Project Selection Due 9/24
6	9/26	10/1	#2 Comp Algorithms	**9/26:** - Clustering: Hierarchical, k-means, Knn **10/1:** - Clustering: Hierarchical, k-means, Knn	Module #2 Week 6
7	10/3	10/8	#2 Comp Algorithms	**10/3:** - **In-Class Exercise 2** Topics Covered 9/10 – 10/1 **10/8:** - **Midterm (Full period): Covers Topics Week 1 thru Week 6**	Module #2 Week 7
8	10/10	10/15	#3 CRISPR	**10/10:** - CRISPR-Cas **10/15:** - CRISPR-Cas	Module #3 Week 8

9	10/17	10/22	#3 CRISPR	**10/17:** - CRISPR-Cas **10/22:** - **Quiz 2 (~35 mins): Covers Topics Week 7 thru Week 8** - CRISPR-Cas	Module #3 Week 9
10	10/24	10/29	#4 NGS	**10/24:** - NGS: Various Technologies **10/29:** - NGS: Various Technologies	Module #4 Week 10
11	10/31	11/5	#4 NGS	**10/31:** - **In-Class Exercise 3** Topics Covered 10/3 – 10/29 **11/5:** - NGS: Various Technologies	Module #4 Week 11
12	11/7	11/12	#4 NGS	**11/7:** - NGS: Assembly **11/12:** - **Quiz 3 (~35 mins): Covers Topics Week 9 thru Week 11** - NGS: Assembly	Module #4 Week 12
13	11/14	11/19	#4 NGS	**11/14:** - NGS: Assembly **11/19:** - **In-Class Exercise 4** (Work on Projects, Q&A)	Module #4 Week 13
14	11/21	11/26	#4 NGS	**11/21:** - NGS: Annotation **11/26:** - **Thanksgiving Holiday**	Module #4 Week 14

15	11/28	12/3	#4 NGS	**11/28:** - NGS: Variant Calling **12/3:** - NGS: Applications	Module #4 Week 15
16	12/5	No Class	#4 NGS	**12/5:** - Finish NGS - Review For Final Exam	Module #5 Week 16
			Final Exam Tuesday December 17, 2019, DH 450,2:45PM to 5:00PM		

SCHEDULE FOOTNOTES:
NONE AS OF AUGUST 2019

Grades *

WRITTEN HOMEWORK (6 at 10 points each)	60 pts
QUIZZES (3 at 50pts each)	150 pts
MIDTERM	100 pts
IN-CLASS EXERCISES (4 at 50pts each)	200 pts
WEEKLY COURSE FEEDBACK (14 at 5pts each)	70 pts
PROGRAMMING ASSIGNMENTS (3 @ 40pts each)	120 pts
FINAL EXAM	200 pts
FINAL PROJECT REPORT & CODE	200 pts
Total Course Points	= 1,100 pts Total

* The total points for each category might change depending on the number of project teams and assignments. The instructor reserves the right to adjust, with sufficient advanced notice, the above point distributionby± 5 pts. Such adjustments might be based on the difficulty or simplicity of assignments or quizzes or exams.

Grading Percentage Breakdown		
Percent of Total Points	**Points**	**Letter Grade**
96.66%	≥ 1063	A+
93.33%	≥ 1027	A
90.00%	≥ 990	A-
86.66%	≥ 953	B+
83.33%	≥ 917	B
80.00%	≥ 880	B-
76.66%	≥ 843	C+
73.33%	≥ 807	C
70.00%	≥ 770	C-
66.66%	≥ 733	D+
63.33%	≥ 697	D
60.00%	≥ 660	D-
59.99%	< 660	F

(NOTE: Ranges might change if point totals change)

How To Calculate/Estimate Your Grade

If students would like to calculate their numeric grade percentage, the formula is as follows:

Numeric Grade Percentage =

$$\frac{Total\ points\ from\ assignments}{Total\ course\ points} x\ 100\%$$

There is no guarantee that grades will be curved. If so, it will typically be done at the end of the semester. The instructor is already aware that graduate students need to maintain an overall GPA of B or better. Just because a student NEEDS a particular grade doesn't mean that the instructor will automatically GIVE the student that grade. Students must EARN a passing grade based on submitted and evaluated course work.

Extra Credit Options

There are no pre-planned extra credit assignments in this course. However, homework assignments and exams might, on occasion, contain extra credit options/questions. At times, the instructor might announce the availability of extra exercises or assignments. There is no guarantee that such extra credit exercises or assignments will be offered to

the class. If, in the opinion of the instructor, offering such extra credit options will be significantly advantageous to the learning process, they might be offered.

Late Assignment Submission

Late assignments will receive a 25% point deduction of a graded assignment for each 24hr period the submission is late. For example, if an assignment is worth 10 points, and the grade for the assignment is 8/10, and the assignment is submitted one day late, then the point deduction equals 2.5, and the final grade for the assignment is MAX(0, 8 – 2.5) = MAX(0, 5.5) = 5.5 .

Making Up Missed Assignments

An opportunity to makeup missed exams, homework, in-class exercises, programming assignments, and so forth will be provided if and only if verifiable documentation of a compelling reason (e.g., illness, accident, death in the immediate family) for missing the assignment is provided within one week from the student's ability to return to class. It is the student's responsibility to (1) contact the instructor if an assignment has or will be missed; (2) obtain verification from the instructor that the student will be allowed to make up the assignment, subject to acceptable and verified documentation; and (3) make arrangements with the instructor to submit all missing assignments by the end of the semester.

Receiving An Incomplete (I) Grade

Receiving a grade of Incomplete (I) is not automatic. Students must complete at least 80% of course assignments by the end of the semester to be eligible to receive a grade of incomplete. Students must also provide documentation to support the reason for the request to receive an Incomplete grade. The instructor has the final decision to give an Incomplete grade. If the instructor agrees to give a student an Incomplete grade, the instructor will enter the remaining work to be completed as part of the PeopleSoft grade submission process.

Grade Change Policy

It is a university policy (S09-7) that "A change of grade request must be submitted by the department office directly to the Office of the Registrar in a timely fashion. Normally, such requests must be received by the drop deadline of the following Spring or Fall semester … Requests for exceptions to this policy must be accompanied with a documented and compelling reason. …"

University Policies

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Programs' Syllabus Information web page at http://www.sjsu.edu/gup/syllabusinfo/. Make sure to review these policies and resources.