FACULTY of ENGINEERING

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING,

BINF3020/BINF9020

COMPUTATIONAL BIOINFORMATICS

TERM 3, 2023

Course staff

Course Convener/Lecturer in Charge:
Mike Bain (MB) – m.bain@unsw.edu.au
School of Computer Science and Engineering: K17-401H
Lecturer
Ray Louie (RL) – r.louie@unsw.edu.au

Course details

This course is worth 6 units of credit and includes 4 hours of lectures and 1 or 2 hours of tutorial/lab time per week, although tutorial/labs are not held every week, and some slots are set aside for group presentations. Note that weeks without scheduled tutorial/lab slots are to allow time for work on the practical components of the course.

Course summary

This course explores the computational basis of bioinformatics and complements and extends understanding of bioinformatics tools and resources acquired in other courses, with a focus on the analysis of complex biological datasets and the application of mathematical and computational methods to problems in modern life science. Example domains include sequence analysis, gene expression and function, networks of interaction, and systems modelling. See the course schedule for the syllabus.

Course aims and learning outcomes

To provide an overview of the field of computational methods of analysis and modelling in modern biology, together with its research and industrial context. By the end of this course you should be able to:

L1. Reproduce, explain and apply the major methods of pairwise and multiple sequence analysis in bioinformatics
L2. Define key computational and statistical concepts in bioinformatics approaches to pairwise and multiple sequence analysis
L3. Define Hidden Markov models, describe their associated algorithms, and discuss their application to biological sequence analysis
L4. Define and apply machine learning methods of clustering and classification in analysis of genome-scale data
L5. Express and apply key concepts from graph theory to model biological networks
L6. Use and explain selected engineering approaches to the representation, execution and evaluation of models of biological processes and their application in systems biology
L7. Research and present on a range of open-source software tools and techniques available to bioinformatics software developers

Assumed knowledge

Before commencing this course, students should:

- Be familiar with basic concepts in modern life science and tools in bioinformatics
- Be able to program and produce working code in at least one general-purpose programming language such as C/C++, Java, Perl, Python, etc.

This is assumed to have been acquired in the courses BINF2010 or COMP2041, or equivalent.

Teaching strategies and rationale

- Lectures are structured to emphasise the computational biology analysis and modelling process. Lecture are related to relevant biology problems and context, then focus on mathematical and computational concepts, algorithms and implementation issues.
- Lectures are supplemented by tutorial topics and computer lab work, and assignments that involve the implementation and application of software and interpretation of outputs.

UNSW graduate attributes especially developed in the course include:

- *The skills involved in scholarly enquiry* – students need to research, compare and evaluate different computational biology methods as part of coursework *[L1-7]*.
- *An in-depth engagement with the relevant disciplinary knowledge in its interdisciplinary context* – computational biology is presented in the context of its applications to biology, and of the computer science methods it draws on *[L1-7]*.

- *The capacity for analytical and critical thinking and for creative problem-solving* – tute/lab and project work requires students to solve a range of problems by choosing appropriate computational biology methods and applying them *[L1-7].*
- *The ability to engage in independent and reflective learning* – the final examination especially requires students to reflect and provide a critical synthesis of the course contents *[L1-7]*.
- *A respect for ethical practice and social responsibility* – the course relates to issues of ethical considerations in data collection and analysis in computational biology applications *[L4-6]*.
- *The skills of effective communication* – written and oral communication is assessed principally through assignment reports and presentations, and the final examination *[L1, 2 & 7]*.

Delivery Mode for Term 3, 2023
In 23T3 this course will be delivered in person. All lectures will be in person, and lectures will be recorded. The tutorial and lab activities, including class presentations, will be in person. All course materials will be available via the Moodle course website. Students are expected to be in Sydney and able to attend campus to complete this course.

Assessment
- Programming assignment in 3 parts: sequence alignment (20%)
  A programming assignment focusing on the dynamic programming alignment algorithm. *[L1–2]*.
- Group presentation: open-source bioinformatics software environments (16%)
  Research and present on a selected topic. Marks will include a peer-assessed component. *[L7]*.
- Homework: systems biology (14%)
  A two-part self-directed data analysis and programming task on methods from lectures. *[L4–6]*.
- Final exam 50%
  A formal exam in the exam period covering the lecture content of the course. *[L1–6]*.

Marking for the programming assignment, homework and group presentation is done with respect to a rubric and feedback will be provided with the online assessment.

Supplementary exams and assessment
A student will be offered a supplementary exam for the final exam only if they have missed the exam due to a well-documented medical reason. Students who sit a supplementary exam will have their marks calculated in the same way as other students. No supplementary exam will be awarded if the student has already sat the exam. Students must apply for special consideration in order to be considered for a supplementary exam.

Course schedule

Lectures in Term 3 will be **In Person** and will be held on **Tuesdays 11am-1pm in H13 Lawrence Theatre** (Lecture 1) and on **Thursdays 1-3pm in Ainsworth G02** (Lecture 2). Each week has three scheduled tutorial/lab times on **Tuesdays 3-4pm** and **5-6pm** and **Wednesdays 12-1pm in Quad G055**. **Please note:** we will use **only** the following tutorial/lab times:
Weeks 2-5: a 1 hour tutorial **Tuesday 3-4pm**
Weeks 7-8: a 2 hour lab **Tuesday 3-4 and 5-6pm**
Week 9: three 1 hour group presentation slots **Tuesday 3-4pm** and **5-6pm** and **Wednesday 12-1pm**

However, remember times may be subject to change !

Week	Starting	Lec 1 (2hrs)	Lec 2 (2hrs)	Tute/Lab	Assessment
1	Sep 11	Optimal and heuristic sequence alignment	Optimal and heuristic sequence alignment	–	–
2	Sep 18	Optimal and heuristic sequence alignment	Optimal and heuristic sequence alignment	Tutorial (substitution scoring)	Assignment part 1 due 7%
3	Sep 25	Multiple sequence alignment	Phylogeny and tree-building algorithms	Tutorial (sequence alignment)	–
4	Oct 2	Hidden Markov Models	Hidden Markov Models	Tutorial (sequence search)	Assignment part 2 due 9%
5	Oct 9	Sequence assembly and genome informatics	Sequence assembly and genome informatics	Tutorial (sequence assembly)	Assignment part 3 due 4%
6	*Oct 16*	*Flexibility week*			
7	Oct 23	Gene expression and functional genomics	Gene expression and functional genomics	Lab – HW part 1	–
8	Oct 30	Biological network analysis	Biological network analysis	Lab – HW part 2	–
9	Nov 6	Systems biology modelling	Systems biology modelling	Open source software (presentations)	HW due (14%)
10	Nov 13	Bio-ontologies and data integration	Bio-ontologies and data integration	–	Presentation slides due

Resources for students
There is no required textbook for this course. However, the following books can be recommended as sources of introductory or deeper background reading and for further study on some of the course topics:

- "Bioinformatics Algorithms: An Active Learning Approach (Volumes I and II)" by Compeau and Pevzner, Active Learning Press (2015)
- "Biological Sequence Analysis – Probabilistic models of proteins and nucleic acids" by Durbin, Eddy, Krogh and Mitchison, Cambridge University Press (1998)
- "Introduction to Systems Biology – Design Principles of Biological Circuits (2nd Edition)" by Uri Alon, CRC Press, 2019.
- "Systems Biology – A Textbook (2nd Edition)" by Edda Klipp *et al.*, Wiley-VCH, 2016.
- "Networks (2nd Edition)" by Mark Newman, Oxford University Press, 2018.
- "Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry and Engineering (2nd Edition)" by Steven Strogatz, CRC Press, 2015.
- "Computational Modeling of Gene Regulatory Networks – A Primer" by Hamid Bolouri, Imperial College Press, 2008.

Lecture slides, discussion forums, announcements and specifications for assessed material will be made available on the course Moodle website (accessible through myUNSW).

Course evaluation and development
This course will be evaluated through the online MyExperience process at the end of session. Individual lecturers may also distribute surveys on their own teaching. Feedback from these surveys is taken seriously and you are encouraged to respond. This is the second year in which the course is offered; feedback indicated that having additional course staff would be helpful, and this is planned for 2023.

Student Conduct

The Student Code of Conduct (Information at https://student.unsw.edu.au/conduct, Policy at https://www.gs.unsw.edu.au/policy/documents/studentcodepolicy.pdf) sets out what the University expects from students as members of the UNSW community. As well as the learning, teaching and research environment, the University aims to provide an environment that enables students to achieve their full potential and to provide an experience consistent with the University's values and guiding principles. A condition of enrolment is that students *inform themselves* of the University's rules and policies affecting them, and conduct themselves accordingly.
In particular, students have the responsibility to observe standards of equity and respect in dealing with every member of the University community. This applies to all activities on UNSW premises and all external activities related to study and research. This includes behaviour in person as well as behaviour on social media, for example Facebook groups set up for the purpose of discussing UNSW courses or course work. Behaviour that is considered in breach of the Student Code Policy as discriminatory, sexually inappropriate, bullying, harassing, invading another's privacy or causing any person to fear for their personal safety is serious misconduct and can lead to severe penalties, including suspension or exclusion from UNSW.
If you have any concerns, you may raise them with your lecturer, or approach the School Ethics Officer (ethics-officer@cse.unsw.edu.au), Grievance Officer (grievance-officer@cse.unsw.edu.au), or one of the student representatives.
Plagiarism is defined as using the words or ideas of others and presenting them as your own. UNSW and CSE treat plagiarism as academic misconduct, which means that it carries penalties as severe as being excluded from further study at UNSW. There are several on-line sources to help you understand what plagiarism is and how it is dealt with at UNSW:

Plagiarism and Academic Integrity: https://student.unsw.edu.au/plagiarism

UNSW Plagiarism Procedure: https://www.gs.unsw.edu.au/policy/documents/plagiarismprocedure.pdf

Make sure that you read and understand these. Ignorance is not accepted as an excuse for plagiarism. In particular, you are also responsible that your assignment files are not accessible by anyone but you by setting the correct permissions in your CSE directory and code repository, if using. Note also that plagiarism includes paying or asking another person to do a piece of work for you and then submitting it as your own work.

UNSW has an ongoing commitment to fostering a culture of learning informed by academic integrity. All UNSW staff and students have a responsibility to adhere to this principle of academic integrity. Plagiarism undermines academic integrity and is not tolerated at UNSW. Plagiarism at UNSW is defined as using the words or ideas of others and passing them off as your own.

If you haven't done so yet, please take the time to read the full text of UNSW's policy regarding academic honesty and plagiarism (https://student.unsw.edu.au/plagiarism)
The pages below describe the policies and procedures in more detail:

Student Code Policy:
https://www.gs.unsw.edu.au/policy/documents/studentcodepolicy.pdf

Student Misconduct Procedure:
https://www.gs.unsw.edu.au/policy/documents/studentmisconductprocedures.pdf

Plagiarism Policy Statement:
https://www.gs.unsw.edu.au/policy/documents/plagiarismpolicy.pdf

Plagiarism Procedure:
https://www.gs.unsw.edu.au/policy/documents/plagiarismprocedure.pdf

You should also read the following page, which describes your rights and responsibilities in the CSE context:
https://www.engineering.unsw.edu.au/computer-science-engineering/about-us/organisational-structure/student-services/policies/essential-advice-for-cse-students