Course Outline

BINF3020
Computational Bioinformatics
Semester 2 2017

Contents

- Course Staff
- Course Details
- Course Summary
- Course Aims
- Student Learning Outcomes
- Assumed Knowledge
- Teaching Rationale
- Teaching Strategies
- Assessment
- Student Conduct
- Course Schedule
- Resources for Students
- Course Evaluation and Development

Course Staff

Contact Details

Staff Name	Role	Email	Phone
Mike Bain	Course Convenor	mike at cse.unsw.edu.au	56935

Consultations

Staff Name	Day/Time	Location

Course Details

Course Code:	BINF3020
Course Title:	Computational Bioinformatics
Units of Credit:	6
Course WebSite:	http://www.cse.unsw.edu.au/~bi3020/
Handbook Entry:	http://www.handbook.unsw.edu.au/undergraduate/courses/2017/BINF3020.html

Course Summary

This course explores the computational basis of bioinformatics and complements and extends understanding of bioinformatics tools and resources acquired in other courses.

Course Aims

BINF3020/BINF9020 aims to build on the foundations laid in BINF2010 and BINF3010/BINF9010, and covers a range of core and advanced topics and current research areas in bioinformatics. The course demonstrates the application of computational, mathematical and statistical methods to problems in modern life sciences, through examples in the analysis of nucleotide and protein sequence, gene expression and function, networks of interaction and systems modelling.

As a result of successfully completing this course students will have a working knowledge of key topics in computational bioinformatics, and will be able to demonstrate their knowledge both by describing aspects of the topics and by solving problems related to the topics using well-founded engineering approaches. They will have practical experience with the topics covered in the assignments undertaken.

Student Learning Outcomes

As a result of successfully completing this course students will be able to work independently and in collaboration on advanced-level topics in bioinformatics. By the end of the subject, students should be able to:

1. Reproduce, explain and apply the major methods of pairwise and multiple sequence analysis in bioinformatics
2. Define key computational and statistical concepts in bioinformatics approaches to pairwise and multiple sequence analysis
3. Define Hidden Markov models, describe their associated algorithms, and discuss their application to biological sequence analysis
4. Define and apply methods of clustering and classification in analysis of genome-scale data
5. Express and apply key concepts from graph theory to model biological networks
6. Research, use and explain selected engineering approaches to the representation, execution and evaluation of models of biological processes and their application in systems biology
7. Research and present on a range of open source software tools and techniques available to bioinformatics software developers and use at least one of these in their own software development projects

This course contributes to the development of the following graduate capabilities:

Graduate Capability	Acquired in
scholarship: capable of independent and collaborative enquiry	1-7
scholarship: rigorous in their analysis, critique, and reflection	1-7
scholarship: able to apply their knowledge and skills to solving problems	1-7
scholarship: ethical practitioners	7
scholarship: capable of effective communication	2,6,7
scholarship: information literate	6
scholarship: digitally literate	1-7
leadership: enterprising, innovative and creative	6,7
leadership: capable of initiating as well as embracing change	5-7
leadership: collaborative team workers	6,7
professionalism: capable of independent, self-directed practice	2,6,7
professionalism: capable of lifelong learning	1-7
professionalism: capable of operating within an agreed Code of Practice	7
global citizens: capable of applying their discipline in local, national and international contexts	7
global citizens: culturally aware and capable of respecting diversity and acting in socially just/responsible ways	1-5,7
global citizens: capable of environmental responsibility	1-7

Assumed Knowledge

Before commencing this course, students should:

- have a basic knowledge of mathematical, computational and biological concepts, including basic probability, statistics, logic, computational complexity, basic biology (biochemistry, molecular biology, genetics) and bioinformatics
- be able to program in a programming language such as C, Java or Python and a scripting language such as Perl, R, Matlab, etc. in a Unix shell environment, and access and operate bioinformatics tools and resources
- have a working knowledge of fundamental principles and methods in engineering design, development and evaluation

These are assumed to have been acquired in the courses COMP1917, COMP1927, COMP2041, MATH1131, MATH1231, MATH1081, BINF2010 and appropriate BABS courses, or equivalent alternatives.

Teaching Rationale

This course is taught to establish links between the topics being covered and the current state of the art in bioinformatics research and applications, with emphasis on proper theoretical foundations coupled with practical experience based on sound computational approaches, designed to engage active learning in lectures and assignments.

Teaching Strategies

- Lectures introduce concepts, show examples and are supplemented by scheduled tutorial or lab activities and computer assignments that require students to implement algorithms discussed in lectures, design solutions to practical problems, and analyse software environments and published research
- Group assignments and class presentations assess research, teamwork, leadership and communication skills

Assessment

Details of assessment items

Assessment component	Description	Marks available	Due date
Mid-semester exam	Covers material up to week 5	10	Week 6
Final exam	Covers potentially all material, but with emphasis on material not covered in mid-semester exam	40	Exam period
Assignment 1: protein sequence evolution	Design, implement, test and evaluate a protein sequence evolution simulator and optimal alignment system, with a written report	20	Three parts, due end of weeks 2, 4 and 5
Assignment 2: software environments for bioinformatics	Group assignment to select software environment and tools (e.g., BioPerl, BioJava, BioConductor, etc.), use and critically evaluate them, report and present results to the class to assess how these resources can assist in bioinformatics software development.	10	Week 9
Assignment 3: bioinformatics systems engineering	Group assignment to research into, design and implement a web-based platform to enable the engineering evaluation of a biological system model; results to be demonstrated and summarised by the group in a written report.	20	Week 13

Note: all dates subject to change !

The course mark will be the sum total of the component marks.

Programming assignments will be marked on the ability of the software to generate correct output with both provided and new test input, plus subjective evaluation of code quality including structure, organisation, use of comments, and demonstrated adherence to a standard and consistent style. Reports and class presentations will be marked on content, organization and quality of writing and oral presentation skills.

Assignments involving software construction or other written material will usually require submission via `give`. Presentations or demonstrations will be during scheduled class times.

Late submission penalties for assignments will be given as part of each specification.

If your work in this course is affected by unforseen adverse circumstances, you should apply for Special Consideration through MyUNSW, including documentation on how you work has been affected. If your request is reasonable and your work has clearly been impacted, then you may be allowed an extension to an assignment deadline or a supplementary exam.

Student Conduct

The **Student Code of Conduct** (Information, Policy) 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, Grievance Officer, 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
- UNSW Plagiarism Procedure

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

The pages below describe the policies and procedures in more detail:

- Student Code Policy
- Student Misconduct Procedure
- Plagiarism Policy Statement
- Plagiarism Procedure

You should also read the following page which describes your rights and responsibilities in the CSE context:

- Essential Advice for CSE Students

Course Schedule

Schedule of main topics, organised by week:

Week	Topic
01	Optimal methods of sequence alignment
02	Heuristic methods of sequence alignment
03	Statistics of sequence alignment
04	Multiple sequence alignment and analysis
05	Phylogeny and tree-building algorithms
06	Hidden Markov Models
07	Gene expression and functional genomics 1
08	Gene expression and functional genomics 2
09	Biological Network Analysis
10	Systems biology modelling 1
11	Systems biology modelling 2
12	Biomedical semantics, ontologies and data integration

Note: all topics and dates subject to change !

Resources for Students

Texts and recommended readings:

The course has two recommended textbooks (you don't absolutely need to buy these, but they have have a lot of excellent content and if you continue in bioinformatics they will serve as useful reference books). Copies should be available through the UNSW bookshop and in the Library.

- "Bioinformatics Algorithms: An Active Learning Approach (Volumes I and II)", Phillip Compeau and Pavel A. Pevzner. Active Learning Press, 2015. (also available to order from here). If you can't source a copy then the earlier version will suffice: "An Introduction to Bioinformatics Algorithms", Neil C. Jones and Pavel A. Pevzner, MIT Press, 2004.
- "Computational Modelling Of Gene Regulatory Networks - A Primer", Hamid Bolouri, Imperial College Press, 2008.

A number of other bioinformatics textbooks are available through the UNSW library for reference reading. One starting point for assistance is the University library catalogue. Pointers to other resources will be given during the course.

Course Evaluation and Development

This course is evaluated each session using the MyExperience system.

In the previous offering of this courses, students noted that more time on coding algorithms could be useful.

Based on these comments, we will aim to extend the use of the online resource Rosalind, a platform for learning bioinformatics and programming.