



Cmpt857

Readings in Bioinformatics

Term 2 (Winter), 2016/17

Course Syllabus

Index

- [Overview](#)
 - [Instructors](#)
 - [Scheduling Details](#)
 - [Textbooks](#)
 - [Topics](#)
 - [Class Format](#)
 - [Grade Allocation](#)
 - [Policies](#)
-

Overview

Cmpt857 is designed to be a multi-disciplinary course allowing students from a wide range of disciplines to explore bioinformatics, where "bioinformatics" encompasses topics ranging from biological applications of information technology and computer science to models of computation inspired by biological systems. The topics covered will depend on the interests of the professor and the students in the class. The course is useful to students with prior schooling in bioinformatics, as well as students who are interested in bioinformatics but are uncertain that they have the necessary background to take a more traditional bioinformatics course (e.g. BINF 200 or CMPT 830).

Note: this is not a class on how to get / install / use particular software in biological applications. Rather, through readings of research papers, it will explore ideas behind such software and new ways of making use of computer science techniques and concepts in the procurement, analysis, and exploitation of biological data. It will also explore how ideas from biology can be used to perform computations in interesting new ways.

Instructors

Professor, Computer Science: Tony Kusalik

Office: Spinks 424
Phone: 306-966-4904
e-mail: kusalik@cs.usask.ca

Professor, Animal and Poultry Science: Matt Links
Office: Ag 6D30
Phone: 306-966-8737
e-mail: Matthew.Links@usask.ca

Scheduling Details

Time: The time for the class is Tuesdays and Thursdays from 10:00 until 11:20.
Room: Thorv S386 (in the Spinks Addition)

First Meeting: Thursday, January 5, 2017

Textbooks

There is no "required" text for the class, though there are two "recommended" ones. In the "quick reference" category is *Instant Notes in Bioinformatics* 2nd. edition, by Hodgman, French, and Westhead, published by Taylor & Francis, 2010, ISBN 978-0-415-39494-9. For students interested in a thorough grad-level textbook on Bioinformatics, the bookstore should have copies of *Understanding Bioinformatics* by Zvelebil & Baum, published by Garland Science, 2008, ISBN 978-0-8153-40024-9. This is currently the premier reference text in the field. It is also [available for purchase or rent in electronic form](#). These titles should be in the UofS bookstore, and are definitely available from the Library.

The University has many other excellent resources in the area of [bioinformatics or computational biology](#).

For students requiring remedial life sciences background, the following resources are readily available on the internet:

- Go to <http://aitopics.org/publication/artificial-intelligence-and-molecular-biology>. This is the web page for an early textbook in bioinformatics. The first chapter is an excellent informal introduction to molecular biology written for computer scientists, engineers, and mathematicians. Just click on the title of the first chapter to get the PDF.
 - The first chapter of the text by Zvelebil & Baum mentioned above is also a life sciences primer. It is available as a free sample from the publisher. Find it at http://www.garlandscience.com/res/pdf/9780815340249_ch01.pdf
-

Topics

The following is a list of possible topics that could be covered by papers taken up in the course.

- From the computer science perspective, application of techniques from the following areas to biological

information:

- computational linguistics, languages, and grammars
- simulation
- vision and image analysis
- graphics and visualization
- symbolic logic
- databases, information management
- machine learning and artificial intelligence
- probabilistic methods
- operations research
- computer networks
- graph theory
- parallel computation
- novel models of computation
- algorithms
- From a biological perspective, computer science techniques applied to:
 - DNA, RNA, and protein sequence analysis
 - DNA sequence assembly
 - gene or protein expression analysis
 - biomolecular structure prediction or determination
 - ligand-protein binding prediction
 - genome evolution and taxonomy
 - interaction networks
 - systems biology
 - cell differentiation
 - (molecular) physiology
 - probe and primer design
 - localization
 - diagnosis
 - literature search

Given the above, a particular topic or paper typically has two categorizations: one from a biological perspective and another from a computer science perspective.

Important note: this class is not intended as a course in any of the computer science disciplines or techniques listed above. Neither is it meant to be a course addressing specific biological ideas or problems listed. Students will be assumed to either know the particular technique, idea, problem, etc., or be willing and able to learn about it from relevant sources.

Class Format

Because of the nature of this class (multidisciplinary, and having no stringent prerequisites in either Computer Science or the life sciences), students in this class usually come with a wide range of backgrounds.

For the most part, the course will be readings-based. That is, most of the instruction will come through reading, presentation, discussion, and review of (research) papers. This allows students to build on their particular expertise, while learning new concepts in less familiar areas. To provide supplementary material, students may be asked to attend lectures of [BINF200](#) or research seminars given on campus. Occasionally, but especially near

the beginning of class, there may be partial, supplementary lectures given to provide background material (e.g. on local and global sequence alignment). Finally, there will be a take-home final exam. These components of the class are discussed further below.

A significant part of this course will involve study and presentation (leading the discussion) of research papers. In particular, students in the class will be expected to read all selected papers, take their turn leading the discussion papers, participate in the discussion of all papers, and compose a review of each paper. The papers covered (discussed) will be selected by the students and made available well in advance of their presentation. All members of the class will be responsible for reading each paper (prior to its presentation!) and — except for the presenter — formulating discussion-oriented questions. The presenter will chair the discussion of the paper. If necessary, he or she may provide some background information for members of the class not familiar with an area. Students are welcome to ask their prepared questions during the discussion. Finally, students will be required to submit a review of the entire paper, or on a major theme from the subsequent discussion.

As for deadlines and due dates,

- each paper to be taken up must be made available to the class at least one week prior;
- prepared questions on a specific paper are due at the **start** of the class where it is to be discussed; and
- reviews are due one week after the the presentation being reviewed.

Each paper covered (discussed) will be selected by a student, but subject to approval of the instructors. Students should thoroughly read their prospective choice prior to approaching the instructors for consent since the quality and appropriateness of the papers selected contributes to the final grade.

The optimal situation is for each student to twice lead discussion of a paper. However, due to the fact that the class is over-capacity, this may not be possible. Therefore, the following scheme will be followed. Papers will be categorized as either "long" or "short" depending on the quantity of technical content in the paper, and the estimated length of time it will take to adequately discuss that content. For the first part of the class, each student will present one "long" paper. Once everyone has done so, the class will switch to two "short" papers per class period. Students will still be responsible for pre-class questions for both papers. However, they will need to submit a review for only one of the two short papers discussed that day. The choice of which one is up to the student.

The distinction between "long" and "short" papers will be somewhat, but not entirely, determined by the length of the paper. Typically a "short" paper will be 2 to 3 pages in length. However, the primary criterion is "how much time will it take to discuss the salient points in the paper?" If that can be done in about 40 minutes or less, then it will be categorized as a "short" paper. If it will take 80 minutes (or more), then it will be classified as a "long" paper. A good example of "short" papers would be contributions in the "Application Notes" section of the journal *Bioinformatics*, or some of the alternate paper types ("Database", "Methodology", or "Software") in *BMC Bioinformatics*.

Despite the inclusion of a "two short papers per class" component to the class, it still may not be possible for every student to lead the discussion of two papers. If this is the case, an extra class period will be scheduled.

At times, the lecture material in [BINF200](#) will be relevant to some or all the students in Cmpt857. In those instances, students in Cmpt857 will be notified, and they will be welcome to attend the BINF200 class (lecture). Credit will not be given specifically for attending BINF200 classes. However, material provided via BINF200 may be useful in performing well in Cmpt 857; e.g. writing a good review or answering questions on the final exam.

The final exam will be of the "take home" variety with discussion-style questions that will be answerable by

anyone who actively participated in the class and regardless of background discipline.

Questions, reviews, and answers to the final exam may be submitted in hard copy or electronic form, unless otherwise stated. If they are to be submitted in electronic form, they should be in a universal format such as raw text, PDF, PostScript, RTF, or HTML, and emailed to the instructors.

Grade Allocation

The final grade calculation will be:

- paper "presentations": 68%, apportioned as follows
 - selection of good-quality papers: 8%
 - leading discussion of papers: 25%
 - preparatory questions: 8%
 - reviews: 27%
- participation in discussion of papers: 7%
- final exam: 25%

Graded reviews will be marked out of 25 as follows:

- Summary (5 marks) - The synopsis of the paper must be given. What scientific work was being presented?
- Evaluation (5 marks) - The paper must be evaluated. Is the science sound? Were the conclusions justified? Was proper methodology followed? Were the findings presented in a coherent way? Why or why not? This is an important chance for students to show their ability to critique the science and how it is presented.
- Recommendation (5 marks) - Recommend whether the paper should be used in future offerings of this course. Most importantly, justify the recommendation. What evidence is there to support the recommendation?
- Clarity and organization (10 marks) - This criterion reflects the quality of the technical writing. Were the arguments in the review clear? Were points made in a manner that was easy to follow? Does the document adhere to the rules for proper English grammar and punctuation?

Due to the fact that the class is over-capacity, it will not be possible to grade all reviews. For approximately the first month, all reviews will be graded. After that, and at the discretion of the instructors, reviews for only randomly chosen classes will be marked. The others will be graded on a completed/not-completed basis. Students will still be responsible for submitting all reviews. Which reviews will be graded will not be announced in advance.

Recall that a presenter does not have to submit questions on the paper he or she is presenting.

Policies

- All students must be properly registered in order to attend class and receive credit for this course.

- Students activities involving University computational facilities are governed by the Department of Computer Science Laboratory Policies. These are available at <http://www.cs.usask.ca/support/index.php#all>, under "Computer Lab Policies".
 - The standard University of Saskatchewan Department of Computer Science [Academic Honesty Policy](#) and the [University's regulations on academic integrity](#) apply to this course. Therefore, all reviews and answers to the "take home" exam are to be completed on an individual basis.
-

Luxuriantly hand-crafted from only the finest HTML tagsby ...
kusalik@cs.usask.ca