

BCB 410 – Applied Bioinformatics – Fall 2009
Course Project Information

Proposal due by email on Mon November 16 (by end of day).
Project due by email by end of day on Dec 18th (last day of exams).

Traditional Course Project. Topics for your project can come from any area of computational biology (even if we did not cover it in class). As is the case for most upper-level undergraduate course projects, my motivation is to give you an opportunity to develop a novel piece of research and write up your results. That means you should come up with a **simple** but new idea (i.e., I should not be able to do a web or literature search and find that someone else has already done exactly what you propose), implement your idea, and report your results. You can choose to do a more applied project (i.e., applying an existing algorithm to a new dataset or designing and applying a new algorithm) or a more theoretical project (i.e., proving complexity bounds of a problem). Successful student projects often incorporate ideas (and code) from the student's area of research interest.

The project write up should be 6-10 pages (not including references). Your write up should be in the format of a research paper – think of it as a conference submission (in fact, some students will be encouraged to submit their project for publication). Your write up should therefore introduce the problem, state why it's important, contain a brief previous work section, describe in detail your new method (including an explanation as to why your approach is reasonable), present your results, and have a conclusion. It is often difficult to develop an algorithm which improves upon previous approaches in such a short period of time. It is certainly ok if your project does not generate 'better' results than previous approaches – but it is important that to the extent possible, you compare your results to previous work. Your write up should discuss why your algorithm performed as well (or as poorly) as it did. You should test your approach on real data (when possible) and therefore you should make use of publicly available datasets. For example, if you present a new method for secondary structure prediction, you should test your results on proteins with solved structures. Your write up should include at least 4 references to previously published work. A few other tips: Everything takes longer than you anticipate – if you think the project you've selected may be too hard or may take too long it almost certainly will. Pick something easier and do a more thorough job. Do appropriate testing of your algorithm – i.e., don't put training data in your testing set. Incorporate figures – A single figure can explain a concept that might take a page to describe otherwise.

The course project grade will be based on clarity of the problem statement, presentation of your contribution in the context of previous work, design and execution of experiments, and your analysis of the results. *Note that your grade is not based on whether the project actually improved on previous methods.* This is thus your opportunity to play.

IMPORTANT DETAILS

Use of public or open source data or software: Use of public or open source data or software is certainly allowed. You must give credit and acknowledgement to these sources.

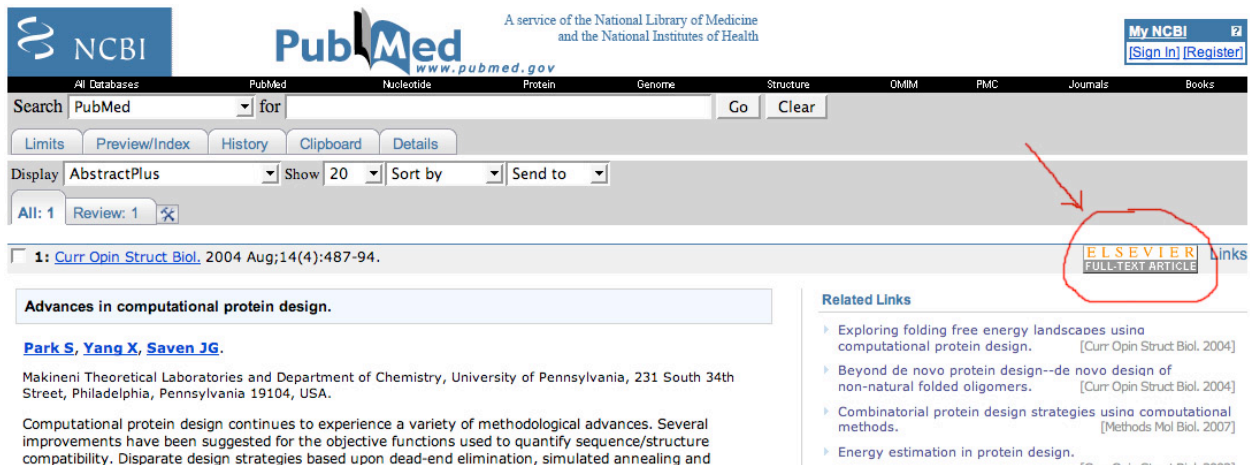
General discussion: You should feel free to discuss ideas with other students. Act as if you are all working in the same lab on different but related projects. Just as in life, you are likely to benefit greatly from interacting with another student.

Page Length Requirements: Specified page lengths assume your document is single spaced with 11 or 12 point font. You should include figures but your document shouldn't be all figures. It will be obvious if you're padding your space with large unnecessary figures.

Where to Find Articles: The best place to find articles in the biological sciences and in computational biology is at the National Library of Medicine at www.pubmed.org. If you search from on campus you can find full text pdf access to most of the referenced articles. After clicking through to the article abstract page, you may see an icon on the right side of the screen (see image below) – in many cases this will take you to the full-text article. When it doesn't your next step should be to search the UofT electronic resource page for the journal containing your article

<http://main.library.utoronto.ca/eir/resources.cfm?T=J>

UofT may have access to your journal through another channel.



The screenshot shows the PubMed interface. At the top, there are logos for NCBI and PubMed, along with the text "A service of the National Library of Medicine and the National Institutes of Health". Below the logos is a navigation bar with tabs for "All Databases", "PubMed", "Nucleotide", "Protein", "Genome", "Structure", "OMIM", "PMC", "Journals", and "Books". A search bar contains "PubMed" and "for", with "Go" and "Clear" buttons. Below the search bar are tabs for "Limits", "Preview/Index", "History", "Clipboard", and "Details". The "Display" section shows "AbstractPlus", "Show 20", "Sort by", and "Send to". The search results show "All: 1" and "Review: 1". The first result is "1: Curr Opin Struct Biol. 2004 Aug;14(4):487-94." with a red circle around the "ELSEVIER FULLTEXT ARTICLE" link. Below the result is a box titled "Advances in computational protein design." with the authors "Park S, Yang X, Saven JG." and a brief description of the article. To the right of the result is a "Related Links" section with several links to related articles.

Project Due Date: The projects will be due by end of day on Dec 18th, 2009. This is a firm date and time – based on when grades are due to the undergraduate office. There will be no extensions. Ideally you will turn your project in before this date so you can enjoy more of your term / holiday break.