BE562: Computational Biology: Genomes Networks Evolution
Fall 2017 Final Project - Research Frontiers

Congratulations on completing the 'foundations' half of the course! By now, lectures, recitations, and problem sets should have given you the algorithmic, machine learning, and biological breadth to tackle many diverse problems in computational biology. While the lectures will continue to expose you to recent advances and current research directions, the recitations and assignments for the second half of the course are dedicated to your final project, including project brainstorming, project feedback, and project write-up and presentation advice.

(a) Project brainstorming and feedback: The first part of any research project is coming up with a good, innovative, concrete, and feasible idea. For this course, it is also important that your project topic be computational in nature. Here are characteristics of a good project for this course:

- Your project should use the computational algorithms and methods we have studied. At a minimum, your project should apply an existing computational method. But you will get many more points for innovation (see below) if you extend an existing computational method or even develop a new one given the concepts and methods you have learned from class.

- Your project should have a clear computational or biological goal. An example of a computational goal would be the development of an algorithm to perform a novel type of analysis. An example of a biological goal would be to answer a biological question or test a biological hypothesis by using computational methods. The best projects have both: you develop a novel computational method and use it to address a clear and well-defined biological question.

- Your project should be feasible and you should have a well-thought out plan for executing your project. This means that you need to ensure that you have the resources (time, data, computers, etc) required. It also means that you need to consider alternatives if your Plan A does not work out. Will you be stuck if you run into a road-block? Or do you have contingency plans?

- Your project should challenge you and the field. The best projects move the field forward. You will be rewarded for taking on more challenging projects, but of course you also need to produce results. The risk-reward balance is something you need to consider.

Given these considerations, there is no single recipe for getting a good idea for a project, and our best ideas frequently come in unexpected ways. You may already have an idea of what you would like to pursue, and may have already found a team to execute it with. But if you're still brainstorming, there are several resources available:

- During lectures, we have already discussed several research directions that can be pursued as final projects. Going back through your notes can give you more ideas.

- A brief look at the calendar of remaining lectures can also give you ideas about current research directions, and the lecture notes from last year (linked from the class website) can give you more details on most of these, that can help you build your own research projects.

- The problem sets that you have completed provide possible initial working starting points for projects that extend the algorithms and programs you have already written in new research directions.

- And of course, browsing recent publications in Nature, Science, PLoS Biology, Genome Research, Nucleic Acids Research, PNAS, PLoS Computational Biology, the Journal of Computational Biology, PubMed, Google Scholar, is a great way to get ideas of recent research ideas, datasets, results, that you can expand upon for your project.

In addition, if you have questions about your project idea, you can always schedule time to talk with either Prof. Galagan or Ariel.

(b) Writing a research proposal: DUE 11/06/17 by 8:00 AM

Project proposals should describe the foundational work in the particular area, what you intend to accomplish and the resources and data sets you intend to use. Proposals should be 3 pages (Specific Aims + Project Description) in length and contain a plan of action for your project. Please include a schedule/timeline of benchmarks for your project. Also include the significance of your work. One proposal is due per group.

For a rough guide, see NIH new guidelines for Research Plan:
http://grants.nih.gov/grants/writing_application.htm
http://www.niaid.nih.gov/researchfunding/grant/pages/appsamples.aspx

Your research proposal should including the following elements all combined into one single document. Note that *page limits are strict* (citations do not count towards the page limit).

1) **Specific Aims**.
   a) This page should start with a one paragraph summary of the background of your project, the key problem or question, and a statement of the overall goals of your project.
   b) This paragraph should then be followed by 2-3 paragraphs describing each of the specific aims of your project. These are the objectives of your research project, what you want to accomplish. The project aims should be driven by the hypothesis you set out to test. Make sure they are highly focused.

2) **Project Description**. The project description should be split into three sections
   a) **Significance**: Why is your proposal important? How will it advance the field? What is the key problem that you will address? *Why should we care about your project*? This should be no more than a page, and often half a page will do.

   b) **Innovation**: How is your approach new and unique? For this class, it is also important to highlight *how your project uses and extends computational methods in a novel way*.

   c) **Research Strategy**: What you will do and how you will do it. This should typically start with any relevant background material. Then you would typically describe your approach for each of your specific aims in turn. Make sure you provide enough detail that people can evaluate your likelihood of success. Also make sure that you have thought through problems that might arise. You should also include sections on dataset availability, and special resources you will use, and a timeline. Students working in groups should include a collaboration plan, roles of individual investigators, and the specific coordination of activities to ensure success of the group.

   Note: Students can work individually or can work in groups of up to 3 people.

3) A **maximum two-page CV or resume** for each member of your team (do not provide your GPA or other information you do not want to share with other students in the course, as these will be circulated!).

(c) Reviewing Peer Proposals: DUE Mon 11/10/17 by 8:00 PM.

Once you have written your proposal they will be peer-reviewed. For this course your peers are your fellow student colleagues. Thus you will each review projects from other groups.

You will review projects according to the following three criteria:

Significance. Does the project address an important problem or a critical barrier to progress in the field? If the aims of the project are achieved, how will scientific knowledge and/or computational methods be improved? Is the project challenging?

Innovation. Does the project use computational methods in new or innovative ways? Does it seek to develop new computational approaches? Will the application of computational methods provide novel insight into a biological question?

Approach. Are the overall strategy, methodology, and analyses well-reasoned and appropriate to accomplish the specific aims of the project? Are potential problems, alternative strategies, and benchmarks for success presented?

You will each write critiques of three submitted proposals. Please organize your reviews into the three sections above. For each section propose specific suggestions and changes and summarize associated strengths and weaknesses. While reviewers will know the identity of proposal authors, reviews will be kept anonymous.

For your own project, you should obviously continue working throughout all stages of the review process, extending your ideas, gathering data, background information, algorithmic and machine learning techniques and tools that you can use, and additional biological background, and beginning the implementation of your ideas. You can incorporate feedback as you go, but preliminary work will be crucial in your revised proposal to provide a response and/or rebuttal to reviewers.

(d) Writing a revised final proposal and preliminary results: DUE Mon 11/17/17 at 8:00 PM

You will revise your proposal based on the reviews you received. Requires two components:

- A **1 page Introduction to the Resubmission Proposal.** On this page, you will list each of the major criticisms and respond to each on in turn.

- A **Revised Project Proposal** - Provide revised aims, based on feedback, and adjusted scope based on length of the term, and adjusted timeline. You should address the perceived weaknesses by correcting them when possible, perceived points of failure by providing additional background that will ensure success, and also present your continued progress and additional results since writing the initial proposal, to ensure feasibility despite reviewer's criticisms. In some cases, you may need to change your proposal altogether, and provide a new plan, but we hope this will be rare.

(e) Writing your results in a scientific paper format: DUE Wed 12/13/17 at 8:00 PM

We will be providing a separate handout on this topic (Guidelines for the final project report and presentations).

(f) Final Project Presentations: Thurs 12/7/17 (CGS 527 1:30pm) & Fri 12/8/17 (EPC 207 10am)

Give a 10 minute presentation on your project to the class. We will post a doodle poll for you to sign up for a time slot.