CMSC 858L — Fall 2009 — Class Project

The project will be done individually, although you can work together, each doing parts of some larger project.

There are 3 options:

1) Algorithm Design. Improve on either MCL, Newman’s modularity, or Graph Summarization in some new way. For example, your GS implementation is faster than the reference implementation, or gives a better compression ratio, or handles other kinds of graph patterns, or supports extra data on the nodes / edges, or compresses hypergraphs, or makes better function predictions, or has a NP-hardness proof. Be creative about what you mean by “improve.” (This is a crucial skill for an academic).

As an alternative, you can argue convincingly why one particular approach to improvement does not work. For example, if your original plan was to improve Graph Summarization by using an integer programming formulation, and you don’t obtain an improvement, even after some careful thought and experimentation, you might conclude that that idea, though reasonable, won’t work. Explain how you came to that conclusion and support it with some experimental results.

2) Programming Option. Write a useful, interesting Cytoscape plug-in that doesn’t already exist. Think carefully about some of the network analysis tasks a biologist might be interested in performing and write a plug-in to help users perform that task. You should really try to understand the task from the biologist’s viewpoint, and design your plug-in accordingly. Be careful of blindly applying a social network analysis task to biology.

3) Other. Another project related to biological networks that I approve. (I expect many people will chose this option.)

You must:

1. Send me, via email, a 1–2 paragraph proposal for the project by 5pm, Friday, November 6. If you choose option #3, your writeup should be 1 page long. The write-up should say which of the above 3 options you have chosen, and describe the project briefly. If two or more people choose a very similar project, there may be an negotiation stage where we try to minimize the overlap.

2. Present the results of your project in a ~ 15 minute presentation during the last 2 days of class (Dec 7 and Dec 9).

3. Turn in a write-up of the project (3-5 pages) by the start of the Final Exam (Dec 16).

Evaluation criteria:

1. Clarity of presentation and write-up.

2. Project of appropriate scope (not too small not too large).

3. Soundness of the results. Did you try the natural approaches? Are your “experiments” correctly done?


Advice: Remember, this is supposed to be a 5- or 6-week-long project, not a graduate thesis. Part of the project is choosing the right sized question to work on. The best way to go about this is to pick some question that is obviously too large (“Can network clustering improve labeling of metagenomic samples?”) and then choose one small part to investigate for the project. If that small part seems promising we can expand it into a full paper later.