1 Culminating Project Proposal - due April 1st

The project gives you the opportunity to delve into a class topic in a manner that suits you. Projects are meant to be instructive and fun - I encourage and expect you to apply yourself. What is important is that you use your creativity, your knowledge, your intuition and your energy to grapple with an aspect of network theory. You will not be punished for reaching high and failing.

1.1 Format

The culminating project may take one of two forms:

1. A theoretical reaction paper, in which you read at least three topic-related papers, at least one of which is not on the required reading list. You should then write approximately 5-8 pages in which you address the following points:

   (a) What is main technical content of the papers?
   (b) Why is it interesting in relation to one or more course topics?
   (c) What are the weaknesses of the papers, and how could they be improved?
   (d) What are some promising further research questions in the direction of the papers, and how could they be pursued?

   The discussion should go beyond a mere summary of the papers. For instance, you might discuss some weaknesses in the papers, and suggest alternative approaches; formulate different hypotheses than the papers and suggest experiments that could be used to test the hypotheses; discuss related work from a different field that the authors were not aware of, or a potential application of the work to a different field. I encourage you strongly to delve through Csermely for ‘weak links’.

2. An empirical project, in which you experimentally evaluate an algorithm, a model, construct a program or measure on an interesting data set. You may program something, use Pajek/Matlab, or other types of suitable software. This empirical project should be if at all possible done on a dataset that has not been examined before. You will submit a report approximately 5-10 pages long.

1.2 Final Report

Details on the structure that is expected will be given after break.
1.3 Presentation

You will give at least one presentation during the semester. The presentations will be about 20 minutes long. You will give a lecture to the class, including answering all questions from the class. It will be beneficial (but it is not required) for you to give a dry run of your lecture to me before you give it to the class. See the handout Technical Presentations for advice on giving a good talk.

1.4 Proposal

One page long, due Tuesday, April 1st, mailed and in hardcopy, with your name, containing the following sections:

1. Introduction and motivation: What you want to do and why it is worthwhile doing

2. Deliverable: What will be the result of your project (research paper, survey paper, computer program, data results, step-by-step guide, etc)

3. Methodology: How are you going to produce the deliverables (implement something, design the math, do a survey, test its functionality etc)

4. Sources: Where you will find what you need to know to do this?

5. Milestones: This is your insurance policy with me if your project fails (which can happen). I would like 3 or 4 milestones (that includes the final deliverable) on the way to the end result. It is also for your benefit so that you can track your progress.

You are encouraged to work on a project that benefits some community (Wellesley College, a neighbouring town, a country, science), but it is not required.
2 Sample Proposals Excerpts from Prior Years

2.1 A network view of sitcoms

Introduction I will analyze the scenario connectedness of plots of different sitcoms. This sort of study has not been done before, and I want to find out the difference between plot sequence/path properties and plot centrality/clustering coefficients between the cartoon sitcom "The Simpsons" and comparably successful real life sitcoms such as Frasier, Friends, and Everybody Loves Raymond. We will account for different types of plot segues/edges such as characters, misunderstandings, etc. It will be interesting to see if the Simpsons are truly an unusual sitcom using network statistical analysis of its plots.

Deliverable I plan to show the sitcom networks as graphs and write a research paper with the results of the study. I will study three randomly selected episodes of each sitcom.

Methodology I will intelligently dissect episodes of sitcoms and record the plot as a diagram with scenarios as nodes and segues into the scenarios as edges (and not just sit on my rear and laugh out loud while consuming high fructose candy). I will access these episodes using the interlibrary loan and also some DVDs in my collection. I will use Pajek to visualize and study the statistics of each plot.

Sources

- Lecture Notes
- Sitcom DVDs
- Papers on Shakespeare’s plot networks
- Online Pajek Manual

Milestones

- Selecting episodes
- visualizing the plots
- documenting path length and clustering coefficients of graphs, as well as centrality
- analysis of Simpsons vs. non-cartoon sitcom plots using the network model
2.2 Pedestrian networks

Wellesley College has been my home for four years. I have had plenty of time to explore it and spend time thinking about the campus itself and how the campus is used by students, faculty, staff, and visitors. One of the important aspects of the campus is its pedestrian focus. Pedestrian paths criss-cross the campus skirting buildings, crossing meadows, and in a couple of cases wrapping up and around buildings themselves. Rarely do they mimic the straight abstract edges in a typical depiction of a network. Instead they meander and take the more "scenic route" from building to building across campus.

I would like to analyze the pedestrian network of campus, though it is a small network, to see if it shares similar properties to the real-life networks we have looked at this semester. This project should bring the study of networks "home" to me in that I will apply it to a data set that I already know a significant amount about - unlike comic book characters or the structure of the internet. It is also worth directing attention to the pedestrian aspect of campus because this perspective is not well supported. A prime example of this is the official campus map that is very obviously designed from the perspective of someone driving through campus. Pedestrian use of the campus should be preferred, though the opposite seems to be the case. Any attention paid to the pedestrian perspective is, in my opinion, time well spent.

The results of my project will be a paper. This paper will analyze the data results of the network I have constructed to represent the pedestrian structure of the campus. My source to construct this network will be my own knowledge of the campus as well as a semi-official map that is primarily used behind the scenes in the Physical Plant that depicts more accurately the pedestrian paths on campus then other campus maps. I will analyze the network using Pajek for its clustering coefficient, average path length, etc and create a random network with the same physical parts for comparison.

Milestones

- Visually depicting the campus as a network (a physical drawing)
- Creation of the Pajek network representing the drawing
- Numeric analysis of the network (and creation and analysis of random network for comparison)
- The final paper analyzing the network
2.3 A critical assessment of network models

Introduction and Motivation  I will examine several different models of small worlds. Through the criteria of high average clustering, short average path length, and power law distribution, I will examine how accurate each model is. From there, I will couple this information with my own notions to create a new model. With this model humanity will be one step closer to finding an accurate representation of small worlds. It is important with each advance in science to step back and observe how the newest research relates to its predecessors. I will examine all aspects, especially faults, of previous models to build upon what knowledge is already available and move one step closer to the ”Holy Grail” of small-world network models.

Deliverable  I will write a research paper that will consist of analysis of several scientific papers. In addition, I will formulate my own model. This resulting representation will depend more on what needs to be addressed by an accurate model rather than actual numbers, although an equation will be provided.

Methodology  To complete this project, I will investigate a wide variety of models to get a broad range of information. After careful review, I will draw from the conclusions of those before me to create my own small world model. Pajek and Matlab will allow me to test this model’s accuracy.

Sources
- Class Notes
- Articles on small network models
- Research articles for which data of small worlds was collected
- Pajek
- Matlab

Milestones
- Find models
- Scrutinize each model to find flaws and aspects that work
- Pile positive aspects and my own corrections together
- Create an abstract model and equations
- Find data for small networks
- Test to see if equation fits graph