Graphs and Networks Lecture 0

Course Description and Mechanics

Daniel A. Spielman August 21, 2013

0.1 Where and When

Tuesday and Thursday, 2:30 - 3:45, in Dunham Labs (DL) 220.

0.2 What this course will be like

The best way to learn what this course will be like is to examine the web page from when I last taught it, in 2010. The web page of this course is

`https://sites.google.com/a/yale.edu/462-562-graphs-and-networks/`.

It contains a link to the 2010 version.

The most important differences from 2010 will be:

1. There will be more problem sets (and more work overall) this time.
2. I will not divide the class into experimental and theoretical tracks. Rather, the problem sets will have both theoretical (theorem proving) and experimental parts.
3. There is a book this time: "Networks: an introduction" by Mark Newman. I will still produce lecture notes, as we will not follow the book too closely. You can get on-line access to this book through the Yale library.

0.3 Required background

To take this class, one should have had courses in discrete mathematics (at the level of Math 244 or CPSC 365) probability (at the level of MATH/STAT 241 or CPSC 365), and linear algebra (MATH 222, 225, or 230). Some more advanced courses in mathematics, statistics, or computer science would help. You should be comfortable with proofs.

0.4 Work for the course

There will be 6 or 7 problem sets during the semester. Most problem sets will have both theoretical and experimental parts. But, they will be mostly theoretical.

It should be possible to do the experiments in Matlab, or one of a number of graph analysis software packages. If you have a graph that you are interested in studying, you can use it. If not, I will supply you with a graph.

There will be no tests or exams.

The homework policy reads:

> "You may discuss the problems with other students. But, you must write your solutions independently, drawing on your own understanding. You should cite any sources that you use on the problem sets other than the textbook, TA and instructor. This means that you should list your collaborators.
>
> If you are doing the experimental problems, it goes without saying that you should write your own code.
>
> You **may not** search the web for solutions to similar problems given out in other classes. If you think this policy needs any clarification, please let me know."

0.5 Overview

The purpose of this class is to introduce many types of graphs that are studied in the sciences and engineering, to examine some elementary processes that take place on those graphs, and to introduce standard analyses of graphs. "Network" is just another word for "graph".

This course will combine mathematical and experimental approaches. Some of the lectures will consist of the presentation of definitions, theorems and proofs. Others will present observations and experiments. The best lectures will combine both.

The course will be divided into three sections. In the first section, we will learn what graphs look like. We will examine graphs from many different sources, including social networks (who knows who), the web graph (which page links to which), the internet graph (which router links to which), citation graphs (who references whose papers), protein interaction networks (which proteins interact with each other), planar graphs (which country is next to which), well-shaped meshes (pretty pictures with triangles), geometric graphs (who is near who), random graphs (whichever), random power-law graphs, and algebraically defined graphs (links determined by mathematician). We will make fundamental measurements of the real-world graphs, and will use these to compare them to abstract models of graphs. We will see that all models are bad, but that some are worse than others. We will also see that some measurements of real-world graphs are actually artifacts of the way these graphs were obtained.

In many applications, people are interested in processes that occur on the vertices or edges of a graph. Others consider processes on graphs as a means of understanding the structure of a graph.

In the second section of the course, we will study many nice properties on graphs, as well as a few counter-intuitive ones. We will focus on random walks (a toddler exploring a graph), diffusion (you spill coffee on your graph), contagion and coordination (spread of diseases on graphs), electrical flow (treating edges as resistors), gossip (spreading rumors), traffic flow (and the price of anarchy), flocking (coordination on changing graphs), and the evolution of graphs.

In the final section of the course, we will examine standard analyses of graphs. The most famous is PageRank, a means of evaluating the importance of vertices in a graph. We will also examine measures of the similarity and distance between vertices in graphs, the problem of detecting communities or clusters of vertices, hierarchical decompositions of graphs, graph drawing, the problem of predicting which edges may appear later in a graph, the problem of detecting incorrect edges or non-edges in noisy graph data, and learning and regression on graphs. We will take three approaches to justifying these analyses: telling a good story, proving that they satisfy desirable axioms, and demonstrating that they predict something meaningful.

0.6 Motivation

I believe that Graph Theory is the new Calculus. Researchers in many fields of science, engineering, and social science have spent much of the last decade actively applying graph theory. My goal in this class is to introduce you to some of these applications, and to give you the background you will need to understand the rest. There are more applications than I can possibly cover, so I'm just going to pick those I understand best.

Networks are very trendy these days, and accordingly inspire bad research as well as good. They provide very useful analogies and analyses, but have also been used to justify many bad ideas and hypothesis. This course should help students figure out which research is which.

This class is designed in opposition to standard classes in graph theory, which are typically based on a book originally written in the 1970's, and which makes no reference to uses of graph theory outside of combinatorics and theoretical computer science.

0.7 Warnings

I will be teaching a lot of material that I do not fully understand. I will probably change the content of the course from what I have advertised. This course is a work in progress, and is not nearly as clean as my other course, CPSC 365.

I use the words "vertex" and "node" interchangeably.