Homework 3: This assignment is due on 04/14/2016 11:00 AM. Please hand in your homework in class, NO emails!

1: The Structure of the Web (40%)

As new links are created and old ones are removed among an existing set of Web pages, the pages move between different parts of the bow-tie structure of the Web.

(a) Describe an example of a graph where removing a single edge can reduce the size of the largest strongly connected component by at least 1000 nodes. (Clearly you shouldn’t attempt to draw the full graph; rather, you can describe it in words, and also draw a schematic picture.)

(b) Describe an example of a graph where adding a single edge can reduce the size of the set OUT by at least 1000 nodes. (Again, you should describe the graph rather than actually drawing it.)

2: Power Laws & Rich-Get-Richer Phenomena (30%)

Suppose that some researchers studying educational institutions decide to collect data to address the following two questions.

- As a function of $k$, what fraction of UMD classes have $k$ students enrolled?
- As a function of $k$, what fraction of 3rd-grade elementary school classrooms in Maryland have $k$ pupils?

Which one of these would you expect to more closely follow a power-law distribution as a function of $k$? Give a brief explanation for your answer.

3: Link Analysis & Web search (30%)

Consider the limiting values that result from the Basic PageRank Update Rule (i.e. the version where we do not use a scaling factor $s$). Such limiting values are described as capturing an equilibrium based on direct endorsement. This gives a way to check whether an assignment of numbers to a set of Web pages forms an equilibrium set of PageRank values as follows: The numbers should add up to 1 “and” they should remain unchanged when we apply the Basic PageRank Update Rule again.

(a) Does the assignment of numbers to the nodes in Figure 1 form an equilibrium set of PageRank values for this network of Web pages? Give an explanation for your answer.

(b) Does the assignment of numbers to the nodes in Figure 2 form an equilibrium set of PageRank values for this network of Web pages? Give an explanation for your answer.
Figure 1: Graph of Question 3(a)

Figure 2: Graph of Question 3(b)