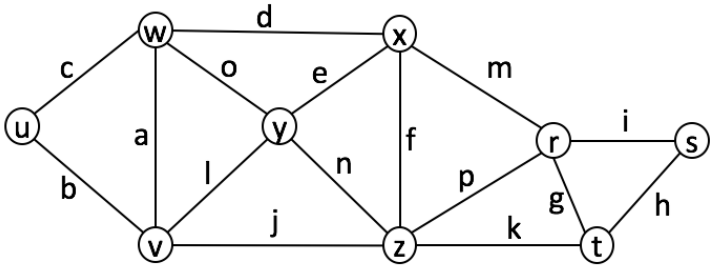
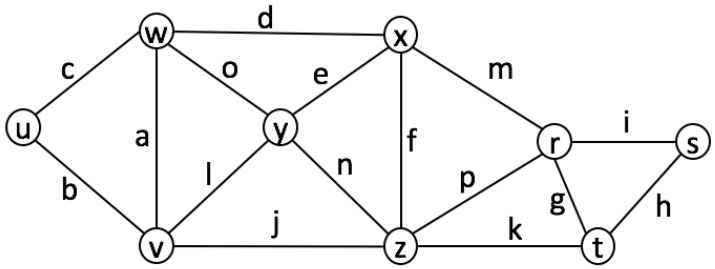
BCB 570 HW2
Spring 2018

1) For the network shown, compute the following plots by hand. Show your work (10 points).



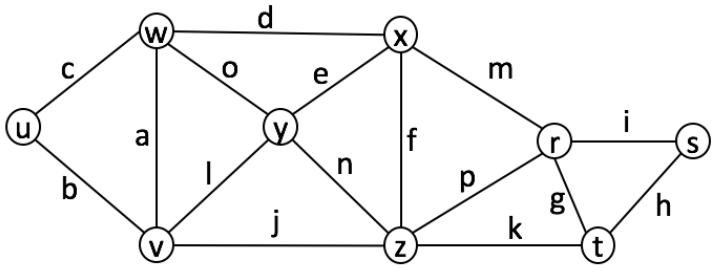
a) Plot the connectivity distribution of k vs. P(k)
b) Plot the distribution of the clustering coefficient, k vs. C(k)

2) In the problem below, draw the depth first tree that results when the DFS algorithm is applied starting at the specified vertex. Include the df_numbers (the order where nodes are searched. Use the a) lexicographic order as the default priority and b) the reverse lexicographic order. (10 points)



- Vertex w
- Vertex s

3) In the problem below, draw the breadth first tree that results when the BFS algorithm is applied starting at the specified vertex. Include the bf_numbers (the order where nodes are searched. Use the a) lexicographic order as the default priority and b) the reverse lexicographic order. (10 points)



- Vertex w
- Vertex s

BCB 570 HW2
Spring 2018

For the next problems, you will be learning about using *iGraph* to create and analyze networks. Three graph files have been made available to you in edge list format: graphA.txt, graphB.txt, and graphC.txt. The edge list format is one of the many standard formats for representing a network, and in it each line represents the two vertices that make up an edge. You may assume that all edges are undirected. Also keep in mind that one or more of these graphs may not have all the properties you expect it to, which may lead to errors. In such a case, errors are acceptable as long as you explain why they arise. All scripts and results files generated in this homework should be zipped together and turned in as one file on Canvas. Also, try to use either a Python Notebook or R markdown for the assignment.

4) Use iGraph (it is recommended that you do this work in Python or R) to build a graph of a fully connected 5-point star with numbered vertices. Show the graph in your script. (10 pts)

5) For each given graph A-C answer the following questions:

i. Is the graph dense, sparse, or somewhere in-between? Justify your answer. (5 pts)
ii. What is the diameter and the radius of the graph? (5 pts)
iii. What are the average and global clustering coefficients? What's the difference between these measures? (10 pts)
iv. What is the average shortest path length? (5 pts)
v. Use two different centrality measures to determine the central nodes of the graph. Describe what the measures are. How many nodes are the same amongst both measures? Discuss why those nodes were the ones both measures found to be central. (15 pts)

6) How would you go about determining whether these graphs follow a power law distribution? Provide a step-by-step plan for doing this and assessing the goodness of fit. Hint: the iGraph package has a routime called *fit_power_law* in R and *power_law_fit* in Python. (10 pts)

7) How would you go about determining whether these graphs are scale-free? Provide a step-by-step plan for doing this. (10 pts)