1 Clustering Coefficient and Neighborhood Overlap (10%)

Provide example graphs that contain the node pair (a, b) for which

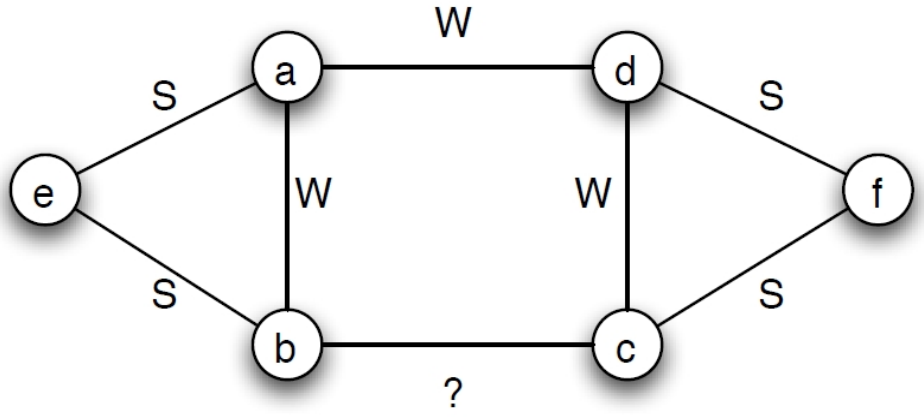
1. both a and b have *high* clustering coefficient but (a, b) has *low* neighborhood overlap.
2. both a and b have *low* clustering coefficient but (a, b) has *high* neighborhood overlap.

2 Strong and Weak Ties (10%)

Consider the following definition and network, in which each edge, except the edge connecting nodes b and c, is labeled as strong (S) or weak (W).

Strong triadic closure assumption: if the node x has strong ties to nodes y and z, then there must be a tie, either weak or strong, between y and z.

1. According to our discussion about strong and weak ties, with the strong triadic closure assumption mentioned above, what is the expected label for the edge between b and c? Give a brief explanation for your answer.
2. What are the differences between weak and strong ties in social networks?. Use the above comparison to explain the strength of weak ties in social networks.



3 The Structure of the Web (30%)

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.

1. 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.)
2. 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.)

4 Power Laws (25%)

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

1. As a function of k, what fraction of UML classes have k students enrolled?
2. As a function of k, what fraction of 3rd-grade elementary school classrooms in Massachusetts have k pupils?

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

5 Power Laws & Rich-Get-Richer Phenomenon (25%)

Erdős and Rényi (1960) studied a model of growth for graphs in which, at each step, two nodes are chosen uniformly at random and a link is inserted between them. Do you think power laws and the rich-get-richer phenomena are likely to be observed in these random graphs. Provide a brief explanation for your answer.