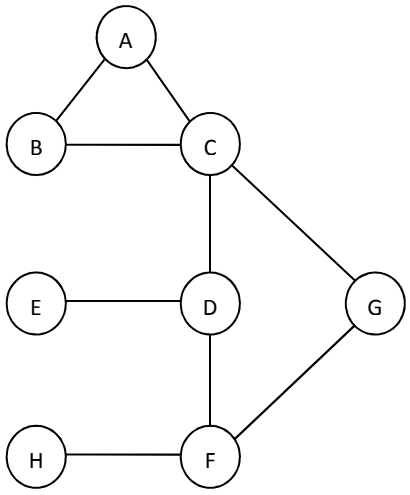
Embeddedness and Betweenness Name _____

Intro to Web Science

10 points

Use the graph on the right to answer questions 1-7 below.



1. (1 pt) Find the **clustering coefficient** of the nodes A and C.

2. (1 pt) Which edge(s) have the **smallest** neighborhood overlap?

3. (1 pt) Which edge(s) have the **largest** neighborhood overlap?

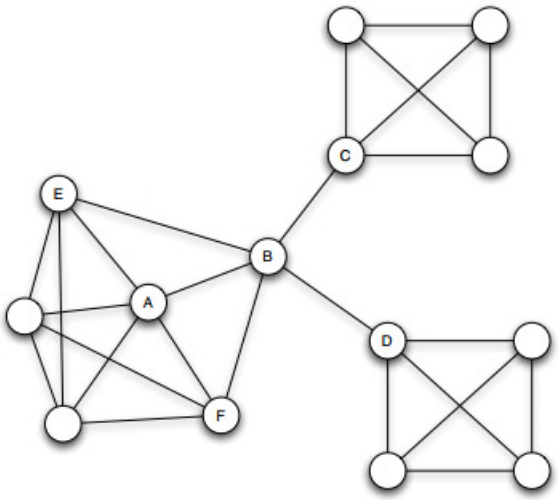
4. (1 pt) Find the **embeddedness** of edges E-D and B-C.

5. (1 pt) Which edges have an embeddedness of **zero**?

6. (1 pt) Find the **betweenness** of edges A-B and F-G.

7. (1 pt) Suppose the network on the right represents the social network of individuals who work in a business that is separated into multiple departments, each represented by the tightly knit components. B manages each of the departments, and as shown by the graph, there is currently no social interaction between the departments.

   Suppose E has just been transferred to C's department. Assuming E keeps her previous social network intact, how will the transfer affect B?



8. (3 pts) The **Girvan-Newman** algorithm needs to find the edges of highest betweenness at each step. Chapter 3 of the book *Networks, Crowds, and Markets* (available online here) describes an efficient technique to determine the flow values beginning in subsection "B. Computing Betweenness Values" of section 3.6.

   a) First, show what the graph above (used for problems 1-6) would look like when ordered by layers using a breadth-first search **beginning at node C**, just like the example in Figure 3.18.

   b) Then determine the shortest paths from C to each of the nodes, writing the number of shortest paths down next to each node as shown in Figure 3.19.

   c) Finally, determine the flow values by beginning at the bottom node and working up, and write each flow value next to the edges as shown in Figure 3.20.