

CAS CS-103 Midterm Solutions

Many problems have more than one good answer. Below are example answers receiving full credit for each problem.

1. (4 points) Explain the difference between "voyager" and "deep thought" questions, and give an example of each.

Voyager is a general question - so general not
even sure what the question is (ie. Information on Solar
System). Deep thought is more focused than voyager,
but still complex (what was economy like in 1905 + why?)

2. (4 points) The Web 101 textbook names four types of web resources for answering questions, and we added a couple more types in lecture. List any four of these types.

Resource #1 Clearinghouse

Resource #3 Peer to Peer

Resource #2 General Search Engine

Resource #4 Subject Tree

3. (4 points) For each of the four types of web resources you listed above, choose **one** of the following network concepts and explain why that concept is most crucial to that particular type of web resource. You may use the same network concept for more than one type of web resource.

1. Indegree centrality

4. Tree

2. Outdegree centrality

5. Clustering

3. Betweenness centrality

6. Affiliation (bipartite)

Resource #1 uses network concept # 2 because Clearinghouses have high outdegree centrality because usually made by a guru that links to relevant sources.

Resource #2 uses network concept # 2 because Search engines like google rate by amount of citations inward for author credibility.

Resource #3 uses network concept # 6 because Peers answer the questions and submit the questions so there's a bipartite relationship.

Resource #4 uses network concept # 4 because Places like about.com use a tree hierarchy for search in which you choose topics to narrow the query.

4. (2 pts) Write a formula for the density of an undirected graph with vertex set V and edge set E :

$$D = \frac{\text{Actual Edges}}{\text{possible edges}} = \frac{|E|}{\frac{1}{2}|V|(|V|-1)}$$

5. (2 pts) Explain in a few words what is the degree of a vertex: the degree of a vertex x

is the number of nodes adjacent to x

8. (2 pts) Suppose you have a graph with 10 vertices and a density of 0.01. If the edges are distributed randomly across the graph, approximately how many vertices would you expect the largest connected component to have?

Answer: 2

$$.01 = \frac{E}{5(n)}$$

$$\frac{|E|}{|V|} = .45 = \frac{.45}{10}$$

9. (2 pts) Same as #8 above but for a graph with 1000 vertices: If the graph has a density of 0.01 and the edges are distributed randomly across the graph, approximately how many vertices would you expect the largest connected component to have?

Answer: 1000

$$E = 4995$$

$$\frac{|E|}{|V|} = 4.995$$

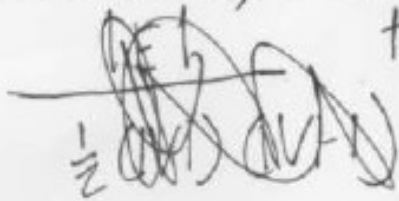
10. (2 pts) What are two defining traits of a tree? (Do not use "hierarchical" or "roots")

a) All nodes connected.

b) No cycles.

11. (3 pts) Suppose you add up the degree of every vertex in a graph and divide by $|V|$. The result is the average degree of a vertex. Explain why this is equal to $2*|E|/|V|$

Because each edge increases the degree of two different nodes
by one. Therefore, $2*|E|$ gives you the sum of ~~all~~ the ~~degrees~~ degrees of all
the vertices. Dividing by $|V|$ averages that sum across
all vertices.

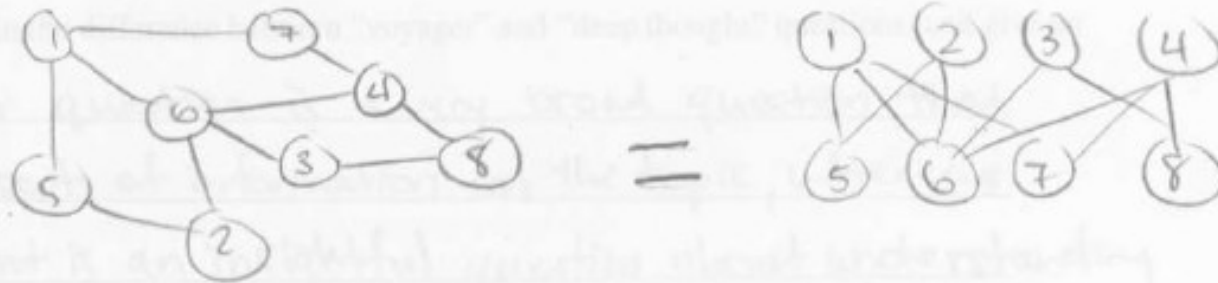


Consider the following graph written in explicit set notation:

$$V = \{1, 2, 3, 4, 5, 6, 7, 8\}$$

$$E = \{(1,5), (1,6), (1,7), (2,5), (2,6), (3,6), (3,8), (4,6), (4,7), (4,8)\}$$

12. (3 pts) Draw the graph neatly as a map of nodes and edges, labeling each node clearly.



13. (3 pts) Construct an adjacency matrix for the graph using the following template. Only write in the "1"s and leave the "0"s blank.

	1	2	3	4	5	6	7	8
1					1	1	1	
2					1	1		
3						1		1
4						1	1	1
5								
6								
7								
8								

14. (2 pts) What special kind of network did you draw in #12 and #13 above? (Hint: The answer has to do with which vertices are adjacent and which vertices are not adjacent to each other.)

Affiliation / Bipartite Network

15. (3 pts) Draw a network with one node labeled "x" such that node x has low degree centrality ($\text{degree}(x)=2$) and node x has high betweenness centrality. Your network must be connected and have at least ten nodes.

