IDS 564 (Social Media and Network Analytics) April 25, 2016
Prof. Ali Tafti

DO NOT OPEN THIS EXAM,
Until you are instructed to do so.

Before you begin the exam, please make sure that your exam booklet contains pages 1 through 14 (including this cover sheet, pages are double-sided). Please sign this cover sheet and the top of *every page of this exam booklet***.** You will need to do this in order to get credit for this exam. When you are done, you will need to submit the exam question booklet along with your answer sheet. You may mark up the question booklet.

The length of the exam is timed at: 90 minutes.

Use of mobile telephony devices (smartphones) or tablet computers, for any reason , are prohibited. During the exam, conversation with anyone other than proctors is prohibited.

You will need a #2 pencil to fill the scantron answer sheet.

Statement:

I acknowledge this is a *closed book, closed notes, closed computer exam;* with the exception of a single standard size page of notes with my own handwriting. In submitting this exam, I attest to be fully compliant with academic integrity policies of the UIC College of Business. I have not intentionally used or attempted to use unauthorized information, people, or study aids in any academic exercise during the exam. I have not and will not provide or receive from another person, any kind of unauthorized assistance on this examination.

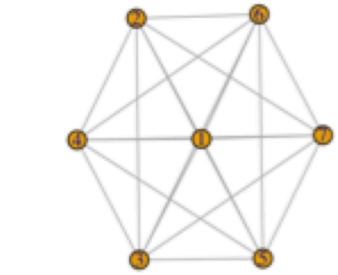
Once the exam begins, I will not be permitted to re-enter the exam room after stepping outside for any reason. Communicating with anyone other than a proctor during the exam will result in an automatic grade of zero.

Please Print Your name:

Your signature

Please sign here:

For all multiple choice questions, please select the single best answer that applies. Please read all choices carefully before answering. Questions are NOT listed in order of difficulty.



diameter(g.full)

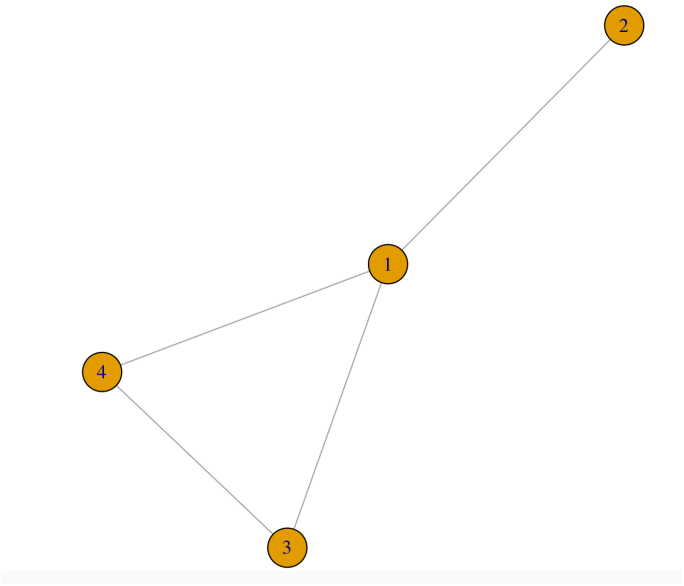
diameter(g.ring)

diameter (g.tree)

diameter(g.star)

1) The R function diameter() returns an integer representing the diameter of the graphs depicted above with their corresponding names: g.full, g.ring, g.tree, and g.star. Which of the following lists the diameter of the graphs in increasing order, from lowest to highest?
   A. g.star, g.ring, g.tree, g.full
   B. g.tree, g.ring, g.full, g.tree
   C. g.full, g.star, g.tree, g.ring
   D. g.full, g.star, g.ring, g.tree
   E. g.star, g.full, g.tree, g.ring

Please sign here:

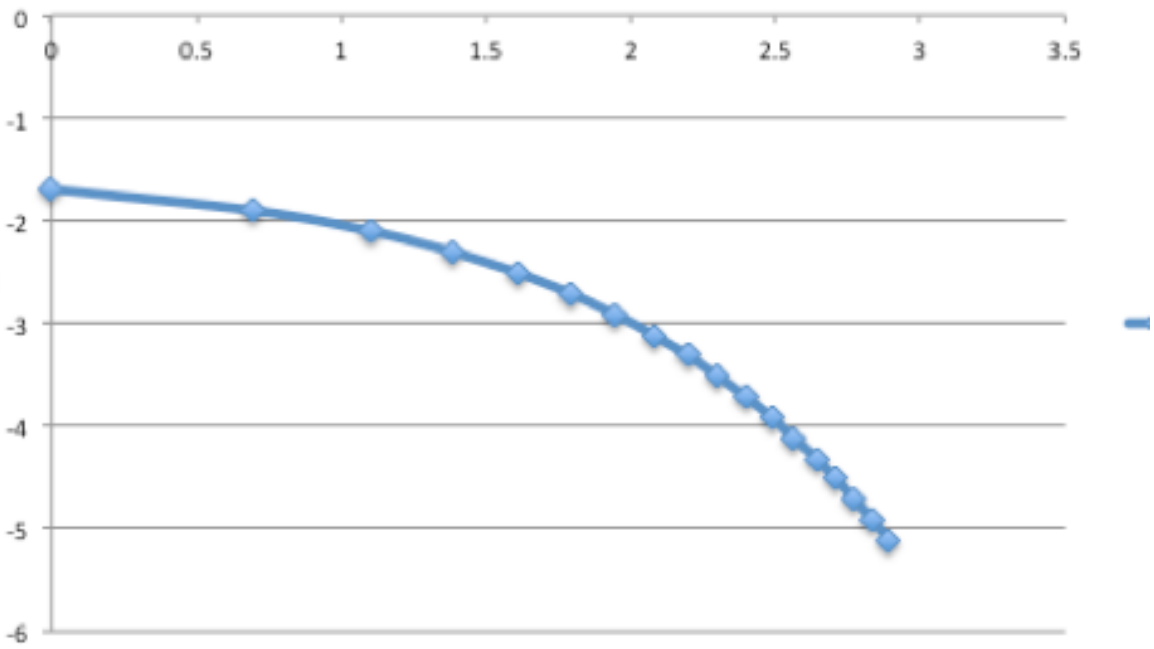


2) Which of the following is true of the above graph?
   A. The (global) clustering coefficient for the graph is the same as the average clustering coefficient of the nodes.
   B. The clustering coefficient for node 2 is equal to zero.
   C. The clustering coefficient for nodes 3 and 4 is equal to 1.
   D. The clustering coefficient for node 1 is equal to ½.
   E. A and C only.

3) What is the overall (global) clustering coefficient for the above graph?
   A. 1/3
   B. 1/2
   C. 2/5
   D. 3/5
   E. None of the above.

4) Which of the following best describes the figure below?
   A. A log-log plot of the cumulative distribution for an Erdos-Renyi random graph.
   B. A log-log plot of the probability density function an Erdos-Renyi random graph.
   C. A log-log plot of the cumulative density function a Barabasi scale free network.
   D. A log-log plot of the probability density function a Barabasi scale free network.
   E. A and B only

Please sign here:

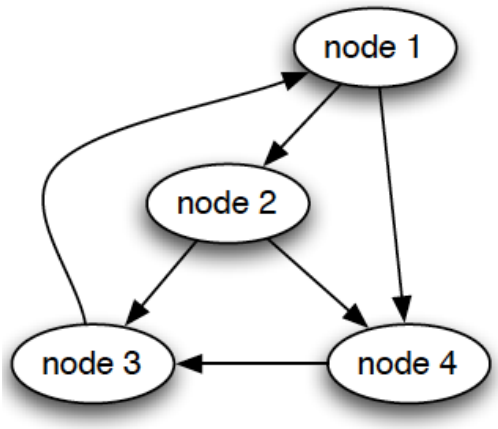


5) Which of the following is true in the graph below?
   A. The edge connecting H and B is a local bridge of span 3.
   B. The edge connecting H and B is a local bridge of span 4.
   C. Node I is a local bridge of span 4.
   D. Node I has a clustering coefficient of 1/3.
   E. B and D only.

Please sign here:

6) Which of the following describes a scenario in which the graph above violates the strong triadic closure property?
   A. B maintains only weak ties with its neighbors. H maintains strong ties with all of its neighbors with the exception of its tie with B, which is a weak tie.
   B. H maintains only strong ties with its neighbors; while B maintains weak ties with all of its neighbors with the exception of its tie with H, which is a strong tie.
   C. Except for the weak tie that connects nodes H and B, both nodes maintain strong ties with all of their other neighbors.
   D. All of the above
   E. B and C only.

7) Consider the directed network below.



The procedure for computing PageRank with this scaling factor of 0.8 is as follows. Start with the adjacency matrix of the graph, divide each entry by the number of outgoing edges for the node corresponding to the row, multiply each entry by the scaling factor and then add the ratio of (1 – scaling factor) over the number of nodes. This results in the following matrix N.

$$\begin{bmatrix} .05 & .45 & .05 & .45 \\ .05 & .05 & .45 & .45 \\ .85 & .05 & .05 & .05 \\ .05 & .05 & .85 & .05 \end{bmatrix}$$

Please sign here:

The transpose of the above matrix N^T is a new matrix whose rows are the columns of N. By definition, PageRank r* is a measure of eigenvector centrality in which r* = $N^T r^*$.

Let $PR(1)_n$, $PR(2)_n$, $PR(3)_n$, and $PR(4)_n$ represent the computed PageRank values of nodes 1 through 4 respectively, after some number of iterations n. Which of the following formulas below specifies how to update the PageRank value for node 3 in the next iteration n+1, with a scaling factor of 0.8:

A. $PR(3)_{n+1} = \frac{1}{2} PR(2)_n + PR(4)_n$
B. $PR(3)_{n+1} = (\frac{1}{2} PR(2)_n + PR(4)_n) \text{ X } 0.8 + 0.05$
C. $PR(3)_{n+1} = (\frac{1}{2} PR(2)_n + PR(4)_n) \text{ X } 0.8$
D. $PR(3)_{n+1} = (\frac{1}{2} PR(2)_n + PR(4)_n) \text{ X } 0.8 - PR(1)_n$
E. $PR(3)_{n+1} = (0.85\ PR(1)_n) \text{ X } 0.8 + 0.05$

8) Referring to the same diagram, which of the following represents a listing of equilibrium PageRank values for nodes 1 through 4, in order, if the scaling factor is 0.8 (rounded to two digits)?

A. 0.29, 0.17, 0.30, 0.23
B. 0.85, 0.16, 0.11, 0.20
C. 0.31, 0.15, 0.31, 0.23
D. 0.17, 0.30, 0.30, 0.22
E. 0.16, 0.11, 0.20, 0.85

9) Referring to the same diagram, which of the following represents a listing of equilibrium PageRank values for nodes 1 through 4, in order, as the scaling factor gets very close to 1 (rounded to two digits)?

A. 0.29, 0.17, 0.30, 0.23
B. 0.85, 0.16, 0.11, 0.20
C. 0.31, 0.15, 0.31, 0.23
D. 0.17, 0.30, 0.30, 0.22
E. 0.16, 0.11, 0.20, 0.85

Please sign here:

10) Referring to the same diagram, suppose that the network is modified so that the edge from node 3 to node 1 is removed, and instead, an edge from node 3 to node 4 is added. Hence, there is now a 2-node cycle involving nodes 3 and 4. Which of the following statements is accurate?
    A. As the scaling factor approaches 0, the PageRank values of nodes 3 and 4 each approach 0.5, while the PageRank values of the other nodes approach zero.
    B. As the scaling factor approaches 0, the PageRank values of nodes 3 and 4 each approach 1, while the PageRank values of the other nodes approach zero.
    C. As the scaling factor approaches 1, the PageRank values of nodes 3 and 4 each approach 0.5, while the PageRank values of the other nodes approach zero.
    D. As the scaling factor approaches 1, the PageRank values of nodes 3 and 4 each approach 1, while the PageRank values of the other nodes approach zero.
    E. B and C only.

11) Which of the following is true of networks that are generated through a preferential attachment process, where higher degree nodes are more likely to attract subsequent connections?
    A. They exhibit a degree distribution with exponential decay.
    B. They exhibit long tails, with a greater density of both very low-degree and very high-degree nodes than networks generated randomly.
    C. They exhibit a scale-free degree distribution.
    D. All of the above.
    E. B and C only.

12) Consider a model of network growth where, in each round, a new node joins the network by first choosing a node randomly among the existing set of nodes in the network. Rather than forming a link with that node, the new node instead picks randomly among the neighbors linked to that node. What kind of network does this process generate?
    A. A preferential attachment network, exhibiting an exponentially decreasing degree distribution.
    B. A preferential attachment network, exhibiting a scale-free degree distribution.
    C. A random growth network, exhibiting an exponentially decreasing degree distribution.
    D. A random growth network, exhibiting a Poisson degree distribution.
    E. A random growth network, exhibiting a scale-free degree distribution.

Please sign here:

13) Consider a model of network growth where, in each round, a new node joins the network by choosing ***m*** nodes randomly among the ***t*** nodes that already exist in the beginning of the round. Thus, each pre-existing node has a probability of ***m/t*** of getting a new link in each round. Which of the following represents the expected degree at time ***t*** of a node born in round ***i*****:**

A. m + m/(i+1) + m/(i+2) + … + m/t
B. m(1 + exp(t/i))
C. m/(i+1)
D. m/(i-1) + m/(i - 2) + ... + m /(i – t)
E. A and B only

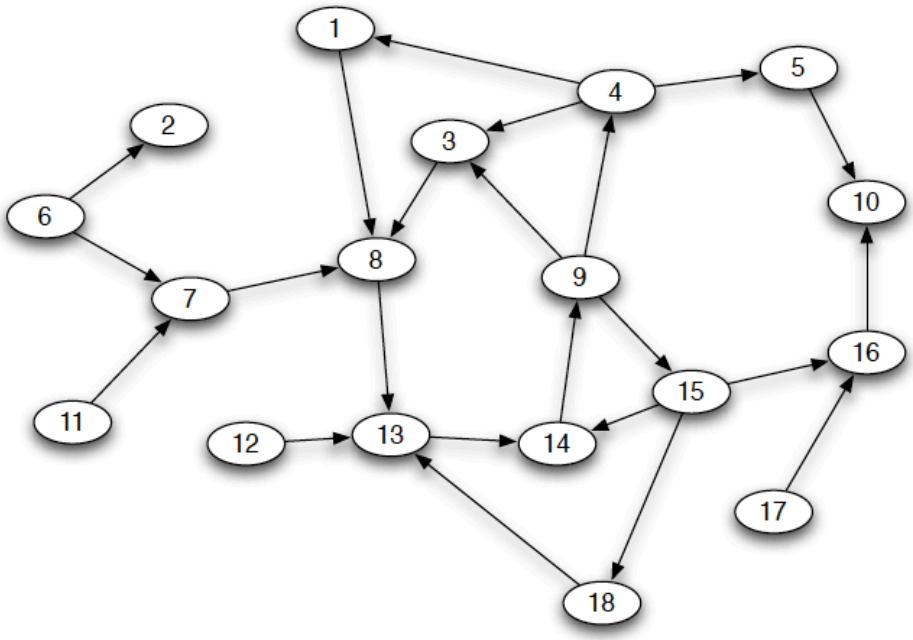
14) Suppose that the operators of a news site are considering changing the way that links are sorted on the front web page of the site. Which of the following ways of sorting links on the front page will result in the popularity distribution of articles to most closely follow a power-law distribution?

A. Sort the links to articles based upon a number of personalized parameters such as the location, gender and age of the user.
B. **Sort the links to articles in order of decreasing popularity, with the most-viewed articles listed first.**
C. Sort the links to articles with the most recent articles appearing first.
D. Sort the articles in random order.
E. Sort the articles in decreasing order of length of the article text.

15) In the network below, which of the following is a list of nodes within the strongest connected component?

A. 1, 6, 10 and 9
B. 14, 16, 10 and 4
C. 3, 4, 9 and 10
D. 1, 4, 8 and 7
E. 1, 3, 14, and 18

Please sign here:



16) Chapter 4 of the Easley and Kleinberg (2010) textbook discusses the Schelling model of segregation. Chapter 5 discusses structural balance in network. Which of the following statements characterizes both of these two theoretical models?

A. Both theories are based upon dynamic simulations and randomized parameters that test how a network evolves over time.
B. A property of local structure or neighborhood preference determines the overall organization of the network.
C. Rules that govern the overall organization of the network determine the actions and preferences of individuals in the system.
D. All of the above
E. B and C only

17) Which of the following is a correct statement about the *strong* form of structural balance in networks?

A. No triangles can have two positive edges and one negative edge.
B. Any triangles with three edges must have exactly one or three positive edges.
C. No cycle in the network can have an odd number of negative edges.
D. Nodes in the network can be separated into no more than two different factions, where each pair of nodes within the same faction are connected by friendship (with a positive edge), and any

Please sign here:

pair of nodes across different factions are enemies (with a negative edge).

E. All of the above.

18) Which of the following is a correct statement about the *weak* form of structural balance in networks?

A. No triangles can have two positive edges and one negative edge.

B. Any triangles with three edges must have exactly one or three positive edges.

C. No cycle in the network can have an odd number of negative edges.

D. Nodes in the network can be separated into no more than two different factions, where each pair of nodes within the same faction are connected by friendship (with a positive edge), and any pair of nodes across different factions are enemies (with a negative edge).

E. All of the above.