Web Science (201500025) Exam resit Part 1

January 31, 2018, 8:45-11:45 hrs, CR 4A

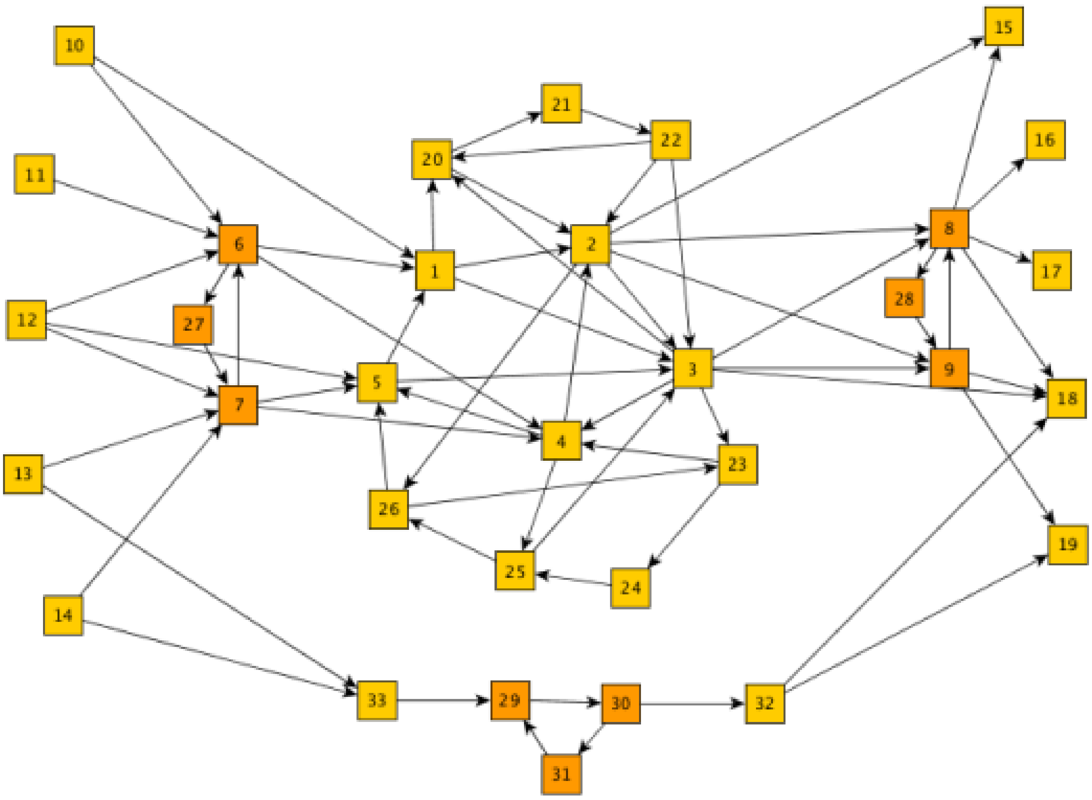
Please pay attention to the following:

- This is a closed book exam: it is not allowed to use the book or any other material.
- Use of calculators, mobile phones etc. is also not allowed.
- Please write your name and student ID on your solution pages.
- Each question mentions the related chapter/topic from the book.
- There are 24 regular sub-questions. For each regular subquestion you get 2 points if correctly answered. In addition, there are three bonus subquestions (explicitly indicated as BONUS), which are worth 2 point each.
- Calculation of the grade G:
  if $R + B < Rmax$: $G = 10(R + B)/Rmax$
  otherwise: $G = 10$

  where
  - R is the number of points from regular questions
  - B is the number of points from bonus questions
  - Rmax is the maximum score from regular questions (= 48)

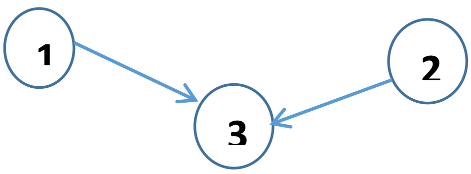
Question 1 (Structure of the Web, Chapter 13)

Identify SCC, In nodes and out nodes from the graph depicted below.



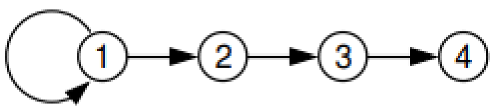
Question 2 (PageRank, Chapter 14)

Consider the graph for web pages given in the following figure.



a) Compute the page-rank for the web. Comment if the answer is intuitive or not.
b) Modify the ranking approach so that a value for the rank can be obtained.

Consider the Web graph in the following figure.



c) Apply the HITS page-ranking algorithm to the graph. Compute manually the first 3 iterations of the algorithm. Can you predict what the page ranks will be, in the limit?
d) What are the ordinal ranks (in other words, the order in which the user is shown these pages) at the end of your computation?

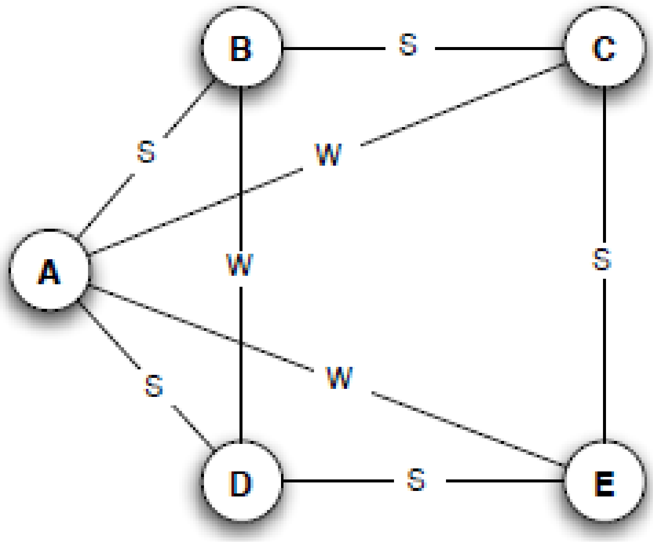
Question 3 (Strong and Weak Ties, Chapters 3 + 4)

(a) Explain (informally) briefly what *triadic closure* is and how it plays a role in the formation of social networks. Draw a schematic picture.

(b) Provide the formal definition of the *Strong Triadic Closure* (STC) property and of a *local bridge*. Draw a picture if needed.

(c) Assume that a network satisfies the STC property. Prove that if a node in the network has at least two strong ties, and any local bridge involving this node, is a weak tie.

(d) In the social network depicted below, in which edge is labeled as either a strong or a weak tie, which nodes satisfy the STC property, and which do not? Explain your answer.



(e) Which edge(s) should be added to make the graph above satisfy the STC property? Give all possible label(s) of those edge(s).

Question 4 (Information Cascades, Chapter 16)

Introduction. In the notation of Section 16, consider two states of the world G and its complement $\bar{G}$. There is a *high* signal (denoted H) suggesting that the state of the world is G, and a low signal (denoted L) suggesting that the state of the world is $\bar{G}$. Assume $P(G) = P(\bar{G}) = 1/2$. Assume that $P(H|G) = 3/5$, $P(L|G) = 2/5$, also $P(H|\bar{G}) = 2/5$ and $P(L|\bar{G}) = 3/5$. Each person chooses between Accept (A) (indicating G is true), or Reject (R) (indicating $\bar{G}$ is true). Motivate your answers using Bayes' rule when appropriate.

(a) Formulate Bayes' rule.

(b) Consider the introduction. Suppose you are the first person and you observe H? Will you choose A or R, and why?

(c) Consider the introduction. Suppose the two people before you choose A, and you observe a low signal L. Will you choose A or R, and why?

(d) Consider the introduction. Suppose that you are the 10-th person to make a choice, and you observed that everyone before you has chosen A, that is, we are in an A-cascade. What is the probability that the cascade is incorrect, so that the state is $\bar{G}$ given that we are in this A-cascade?

(e) (**BONUS**) Bayes rule can also be written as (you do not have to show this):

$$P(G|S) = \frac{P(G)P(S|G)}{P(S|G)P(G) + P(S|\bar{G})P(\bar{G})},$$

Here $\bar{G}$ denotes the complement of G, and S denotes a signal. Assume $P(G) = P(\bar{G})$, show that if $P(S|\bar{G}) < P(S|G)$ then $P(G|S) > P(G)$, and give an interpretation.

Question 5 (Chapter 17, network effects)

Consider the economic model in Chapter 17 where consumers occupy the interval $[0, 1]$ and their reservation price without network effects is given by $r(x) = 1 - x$. The function $f(z)$ measures the benefit to each customer from having a fraction z of the population using the good. When proportion z of the population buys the product, then the new reservation price of customer x with network effects is $r(x)f(z)$, where $f(z) = \sqrt{z}$.

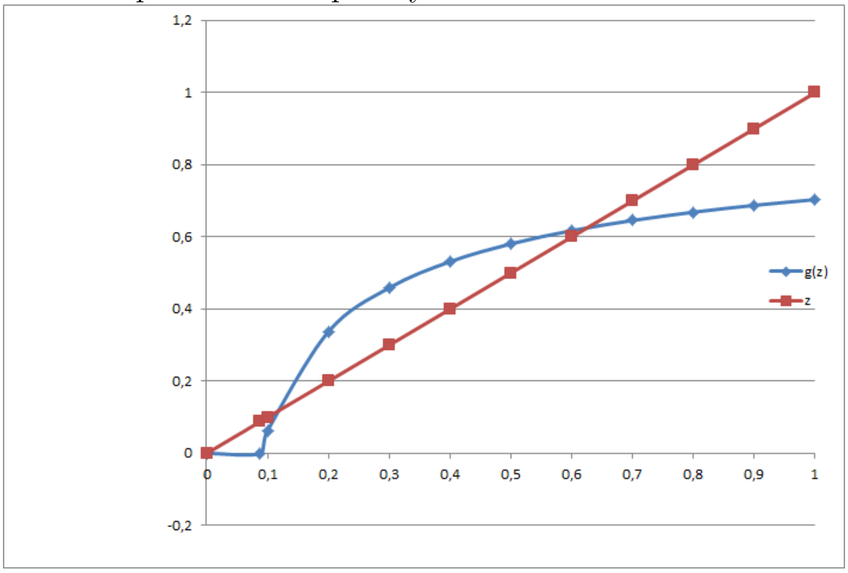
(a) First consider the situation *without* network effects. For each value $p^* > 0$ define the fraction of customers that buy the product. Explain why this point is a stable market equilibrium.

(b) Now consider the situation *with* network effects. When the shared expectation is $z \geq 0$, the fraction of people $\hat{z}$ that will actually buy the product can be denoted as : $\hat{z} = g(z)$.

(c) Provide an expression for $g(z)$, based on z and p^*.

(d) Assume $z = 1/9$. Calculate the exact value of the price p^* for which this value of z is an equilibrium.

(e) Consider the graph of the functions $g(z)$ shown below. Is $z = 1/9$ a stable or a non-stable equilibrium? Explain your answer.

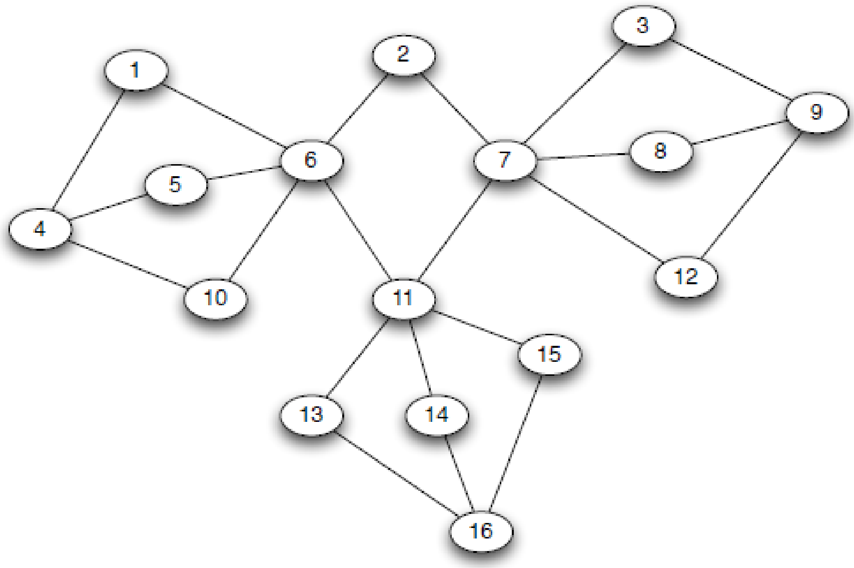


Question 6 (Chapter 19) Extensions of the cascade model

Consider the model from Chapter 19 for the diffusion of a new behaviour through a social network. Everyone starts with behavior B and any node will switch to A if at least a fraction q of its neighbors has adopted A. Such a model can be derived from a networked coordination game in which each node has a choice between two possible behaviours, A and B. Two nodes connected by an edge both receive a payoff of a if they both select A, a payoff of b if they both select B, and a payoff of 0 otherwise.

(a) Derive an expression for the threshold q in terms of a and b.

(b) Suppose a and b are such that $q = 1/2$. Consider the network depicted below. Assume node 1, node 5 and node 10 start with behaviour A. Indicate how the process of adopting A will proceed. Which nodes will eventually adopt A?



(c) Find three clusters in the network, each with density greater than 1/2, with the property that no node belongs to more than one of these clusters.

(d) Find a set of three initial adopters from A (different from those listed under (b)) that could cause the entire network to adopt A.

(e) Can there be a set of two nodes in the network that could cause a complete cascade? Either indicate such a set or explain why this is not possible.

(f) (**BONUS**) As described above, consider a set of initial adopters of A with a threshold of q. Give a detailed proof that a complete cascade can only occur if and only if there is no cluster of density greater than $1 - q$.

(g) (**BONUS**) Consider an infinite network in which the degree of each node is k, with threshold $q > 1/2$. Consider a cascade starting with m arbitrary nodes. Provide an upper bound on the number of nodes (in terms of k,m,and possibly q) that will adopt A.