Networked Life ***(CSE 112)***

Prof. Michael Kearns

Final Examination

May 3, 2006

The final exam is closed-book; you should have no materials present other than the exam and a pen or pencil.

NAME: ___________________________________

PENN ID: __________________________________

Exam Score:

Problem 1:__________/15
Problem 2:__________/10
Problem 3:__________/10
Problem 4:__________/10
Problem 5:__________/10
Problem 6:__________/10
Problem 7:__________/10
Problem 8:__________/10
Problem 9:__________/7
Problem 10:_________/8

TOTAL:____________/100

1. (15 points) For each item on the left, write down the number of the item on the right that most closely corresponds to it.

(a) Differential Pricing ______	(1) Hubs & authorities
(b) Quality of Service ______	(2) No incentive for unilateral change
(c) Routers ______	(3) Multi-dimensional search
(d) Nash Equilibrium ______	(4) Price equilibrium
(e) Prisoner's Dilemma ______	(5) An online social network
(f) Arrow & Debreu ______	(6) Preferential attachment
(g) Kleinberg ______	(7) Personal valuation of goods
(h) Utility function ______	(8) Switches that receive & forward packets
(i) ping ______	(9) Guarantee of service on the Internet
(j) Watts, Dodds & Newman ______	(10) A non zero-sum game
(k) Power law ______	(11) Gladwell's Law of the Few
(l) Rich gets richer ______	(12) Tool to check reachability over IP network
(m) Orkut ______	(13) Linear log-log plot
(n) Market equilibrium ______	(14) Clearance of goods
(o) Connectors, Mavens & Salesmen ______	(15) Premium/first class service in airlines

2. (10 points) Schelling discusses the phenomenon in which individual people choose whether or not to send holiday cards to each of their friends. In this setting there is a cost associated with sending each card, but a higher cost associated with the embarrassment of not sending a card to someone who has sent a card to you.

a) One equilibrium in this setting is the situation in which everybody sends a card to everybody else. Explain why this scenario is an equilibrium.

b) The equilibrium described above is bad in the sense that everyone must pay the costs associated with sending cards to everyone else. Are there other equilibria? If not, explain why others can't exist. If so, describe the equilibrium that is socially optimal, i.e. the equilibrium with the lowest total cost to all participants.

3. (10 points) Consider the hypothetical spread of disease on a network. Here, each node represents an individual person who may or may not decide to receive a vaccination. Each edge represents potential contact between people. A person can contract the disease from any of his neighbors on the graph who are infected, but cannot contract the disease if none of his neighbors are infected.

If an individual chooses to receive the vaccination, he pays a fixed cost V and runs no risk of obtaining the disease. If a person chooses not to receive the vaccination, he pays a cost equal to N*X where N is the number of his neighbors who are not vaccinated and X is a fixed cost associated with the chance of contracting the disease. Assume that X is much larger than V.

a) Suppose everyone chooses to receive the vaccination. Is this an equilibrium? Why or why not?

b) Suppose the government decides to subsidize half of the population by providing them with the vaccination for free (i.e. this half of the population pays 0 and runs no risk of obtaining the disease). Will the other half of the population be more likely to purchase the vaccination now? Briefly explain why or why not.

c) In class we discussed the interdependent security game in which each node represents an airline and an edge between airlines A and B implies that customers may transfer between flights on A and B without having their baggage rescreened. In this game, each airline must decide whether or not to invest in a new security device for screening baggage. Briefly compare the effects of subsidization in the vaccination game with the effects of subsidization in the airline interdependent security game.

4. (10 points) Consider the undirected graph shown in Figure 4.1. Answer following questions with respect to Figure 4.1:

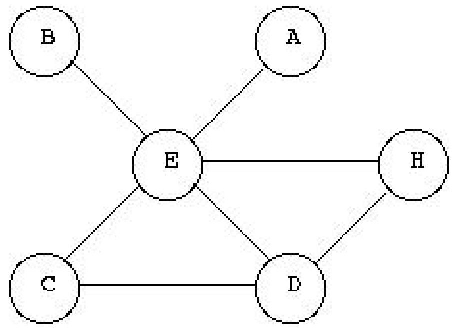


Figure 4.1

a) What is the value of worst-case diameter in this graph? Identify the vertices that are farthest apart (if more than one such pair is possible, then list all of them).

b) What is the size of the largest clique in the graph? List the vertices in each clique of this size.

c) What is the minimum number of edges that could be removed to make the graph unconnected? List these edges. (If more than one answer is possible, then list all of them.)

d) What is the clustering coefficient of node E?

e) What is the minimum number of edges whose addition to the existing graph would increase the size of the largest clique by 1 (compared to the size of the largest clique in the existing graph)? List these edges.

5. (10 points) Consider a network of exchange in which each node represents either a buyer or a seller. Buyers start with one dollar and value only wheat; sellers start with one unit of wheat and value only dollars. All edges on the network connect a buyer with a seller and represent the fact that the buyer and the seller are allowed to trade with each other. Assume that each seller must sell all of her wheat at a single price, and each buyer will purchase wheat only from the seller(s) in his neighborhood with the lowest price.

a) Suppose the network contains ten buyers and six sellers and is connected. Will the average price of wheat be higher or lower than one dollar? (No explanation necessary.)

b) Suppose Seller A is connected to three buyers and that two of these buyers have no other neighbors. What is the minimum price that Seller A will be able to demand at equilibrium? (Again, no explanation necessary.)

c) Is it possible in a buyer-seller network for two sellers of the same degree to charge different prices at equilibrium? If so, draw a network in which this is the case, clearly labeling the buyers and sellers and marking the two sellers of the same degree. If not, explain why sellers of the same degree must charge the same price.

6. (10 points) Consider a repeated game of Prisoner's Dilemma. The game matrix for a single round of this game is given below.

	Cooperate	Defect
Cooperate	-1, -1	-10, -0.25
Defect	-0.25, -10	-8, -8

a) If the game will be played for a known number of rounds R, there is only one equilibrium. State what this equilibrium is and describe why no other can exist.

b) If the game will be played for R rounds where R is *unknown* it is possible for both players to receive higher payoffs at equilibrium. Describe a strategy that can be followed by both players to lead to these higher payoffs. Why does this strategy result in an equilibrium only when R is unknown?

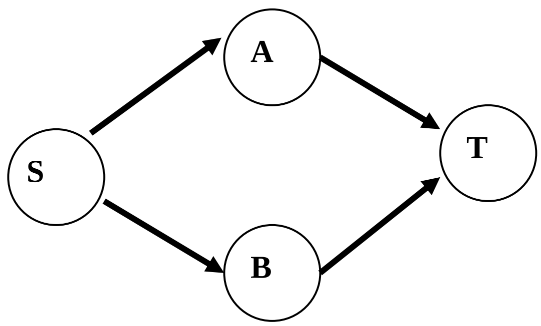
7. (10 points) Some internet services, such as email, can tolerate a network lag of 1 or 2 seconds without any noticeable degradation of service. Others, such as internet telephony, would be basically unusable with such a lag. In class we discussed a possible economic-based approach to handling such differences in demand for quality of service that was based on an analogy to the Paris subway system.

a) Briefly describe this approach and its appeal, including properties such as self-regulation.

b) Give an example of a real-world domain in which a similar technique has worked.

8. (10 points) Consider the following network where *S* is the source node and *T* is the termination node. Traffic (*e.g.* Internet packets, drivers on a freeway etc.) flows from *S* to *T*. The latency function $L_e(x)$ of an edge *e* specifies the latency or delay that traffic on that edge will suffer when a fraction *x* of the total traffic is traveling on that edge. Assume that each packet or driver is selfish and tries to minimize its own total latency.

In class, we defined cost of a flow as the mean of all latencies incurred by traffic in that flow. We also defined optimal flow as the flow that minimizes this cost. On the other hand, a flow is at Nash equilibrium if no traffic (*e.g.* packet or driver) can improve its latency by unilaterally taking an alternative route.



Latency functions for each edge in the above network are as follows:

$$L_{SA}(x) = 0.5x + 0.4, L_{SB}(x) = 1, L_{AT}(x) = 1, L_{BT}(x) = 0.6x + 0.3$$

a) In the above network, one unit of traffic needs to travel from *S* to *T*. The traffic can be divided fractionally along the two possible paths, *S-A-T* and *S-B-T*. What is the cost of Nash equilibrium flow of this one unit of traffic from *S* to *T*? Show your work.

b) Without explicitly computing the value, what can you say about the cost of optimal (minimum cost) flow of this one unit of traffic from *S* to *T* in this network?

9. (7 points)

a) In Figure 9.1, Watts' α parameter is plotted along the horizontal axis. Values of path length (L) and clustering coefficient (C) are plotted along the vertical axis.

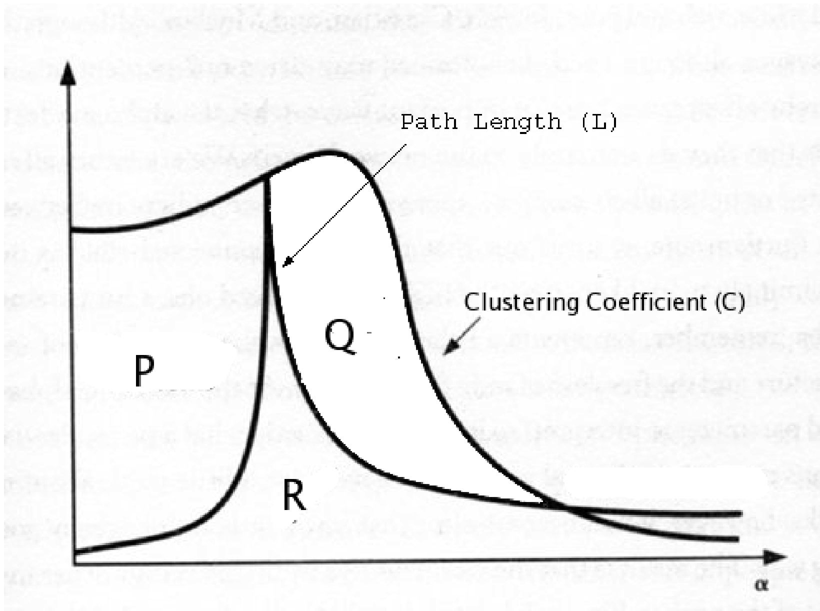


Figure: 9.1

(i) In Figure 9.1, identify the region corresponding to small world networks. Indicate your choice by selecting one of the three regions below:

P **Q** **R**

(ii) State two characteristic features (which can be inferred from Figure 9.1) of such small world networks.

b) In one or two words, answer the following:

(i) State the network formation model with the following properties: small diameter, with heavy tailed degree distribution, but without high clustering.

(ii) State the network formation model with the following properties: small diameter without high clustering and without heavy tailed degree distribution.

10. (8 points) Kleinberg proposed the HITS algorithm for ranking web pages as either hubs or authorities. The intuition behind this algorithm is that a good hub should point to many good authorities, and a good authority should be pointed to by many good hubs.

a) How might one manipulate the HITS algorithm in order to artificially increase the authority ranking of a page? Is this same manipulation technique possible under the PageRank algorithm? Briefly explain your answer.

b) The HITS algorithm makes strong assumptions about the nature of the internet. Aside from manipulability, what are some problems we might run into if we tried to use HITS as a ranking tool for search over the full internet?