Example Midterm Questions II
1. **Mechanism Design**

Consider the following *multicast routing problem*. A server at $A$ can transmit content to some subset of users at nodes $D, E, F$ and $G$. To receive service, the edges between a user and the server must be activated. The mechanism outcome defines the set of activated edges, and the payments by each user.

Each edge has an associated cost, and each user has a value $w_i$ for service. The edge costs and topology are known by the mechanism designer, but the values of users are private.

The network operator uses the **VCG mechanism** to elicit values from users, and selects the set of edges to maximize total reported value minus total cost of activated edges.

(a) Briefly, is truthful reporting a dominant strategy equilibrium for users? Why or why not?

(b) What is the outcome of the VCG mechanism when users are truthful?

(c) Do the total payments cover the total edge cost incurred by the network?

(d) Modify the value of the user at $D$ so that the total payments cover the costs when users are truthful.
(e) Can users 3 and 4 collude and both submit high bids, receive service, and make no payment? Justify your answer.
2. Sponsored Search Auctions

(a) Consider a GSP auction with two positions, three bidders, and position effect 0.3 and 0.2 in positions 1 and 2, respectively (assume quality effect 1.) Provide an example of a valuation profile for which truthful bidding is not a Nash equilibrium. Justify your answer.

(b) What are the per-click VCG payments in your example in part (a)? Compare them to the GSP prices.

(c) Give three reasons why GSP and not VCG is being used for sponsored search.
3. Combinatorial auctions

(a) “For any set of OR bids, there is a cyclic order, one per atomic bid, such that the cyclic property is satisfied. Therefore the winner-determination problem can be solved in polynomial time.”

The argument is false. Where did the argument go wrong?

(b) From the perspective of winner determination, why can we assume OR bids?

(c) Express the valuation “I want at most three of 100 cars that are for sale” as a succinct OR-of-XOR or XOR-of-OR bid. Just write the bid in terms of packages, don’t worry about values. (Of course you don’t need to write out the expression completely! Cars can be denoted as items \(C_1, C_2, C_3, \ldots\) etc.).

(d) Express the valuation “I’m only interested in particular pairs of bikes, but I’d take any number of the pairs I’m interested in” as a succinct OR or XOR bid. Just write the bid in terms of packages, don’t worry about values. (Bikes can be denoted \(B_1, B_2, B_3, \ldots\). Just show how you’d do this by making up some interesting pairs.)
(e) Would you use the OR or XOR operator to combine your answers to (c) and (d) and express the valuation for someone interested in cars and bikes, but not a mixture?

(f) A bidder wants to win the wireless spectrum package “NYC license + BOS license.” Briefly, state what could go wrong in an auction in which she bids on NYC and BOS separately, and how is this addressed in a combinatorial auction?
4. **Paired-kidney exchange**

A patient can receive a kidney if the patient’s blood includes all proteins in the blood of the donor (e.g., an A-type can receive O and A, but not from B or AB).

(a) Suppose there are 4 patient-donor pairs, with (patient-donor) blood types \{B-A, AB-B, AB-A, A-A\}. Draw an **undirected** graph to indicate the feasible swaps.

(b) What is the maximum matching?

(c) How would the graph representation be modified in order to allow for matches that use 3-cycles, in addition to swaps?

(d) How can **altruistic donors** be represented in this modified graph, and why can altruistic donors have a significant impact on matching (more than just because of one additional transplant opportunity)?

(e) What is a new computational challenge when allowing for matchings with altruistic donors?
5. Network formation

In the *stay-connected game*, link formation is unilateral, the cost of each edge is “sponsored” by one agent (cost $c > 0$) and the network must be connected.

An agent’s cost is its total distance to others, plus the cost of any sponsored edges edges. For example, if an agent sponsors one edge and has a path of length 2 to one agent and a path of length 1 to another, then her total cost is $2 + 1 + c = 3 + c$.

Consider the following two graphs:

(a) Is either graph formed in a Nash equilibrium for some edge cost $c < 1$? Justify your answer.

(b) Is either graph formed in a Nash equilibrium for some edge cost $c > 1$? Justify your answer.

(c) Give an example of a Nash equilibrium graph for some edge cost $c > 1$. Justify your answer.
(d) Why is the star graph socially optimal (for any sponsoring structure) for large edge costs?
6. Network formation

In the bilateral connections games, edges are undirected, link formation is bilateral, and an edge between a pair of agents costs each agent an amount $c > 0$. An agent’s value is its total discounted distance to others, for discount factor $\delta \in (0, 1)$. For example, if an agent has a path of length 2 to one agent and a path of length 1 to another then its total value is $\delta^2 + \delta$. The utility to an agent is value minus cost.

The game is studied under **link stability**, which requires:

(i) for an edge that exists between a pair of agents, neither agent strictly prefers to remove the edge, and

(ii) for two agents without an edge between them, if one agent strictly prefers to add the edge then the other agent strictly prefers not to add the edge.

Consider the following two graphs:

(a) Explain why $c \leq \delta$ is required for graph (i) to be link stable. [Hint: consider edges that are present.]

(b) What is a tight lower-bound on $c$ (perhaps stated in terms of $\delta$) for graph (i) to be link stable? **Justify your answer.** [Hint: consider edges that are not present.]
(c) Prove that graph (ii) cannot be link stable for any values of $c$ and $\delta$. [Hint: Use case analysis on $c$ relative to $\delta$.]
7. Network Games

Consider a game on a network where each agent chooses a “0” or a “1”, and its payoff depends on the actions of its neighbors.

(a) In a local substitutes game, if no agent in the neighborhood of an agent selects “1” then the agent prefers a “1.” If one or more agents in the neighborhood select a “1” then the agent prefers “0.”

What are the possible Nash equilibria on the following network?

(b) State an optimization problem for which the set of vertices that play “1” in a Nash equilibrium of the local substitutes game is a solution.
(c) Consider a coordination game (or local complements game), where if strictly more than 1/3 of the neighbors of an agent choose “1” it prefers “1,” and otherwise it prefers “0.” List out the possible Nash equilibria in each of the following graphs. [Hint: the number of equilibria in each graph is 2, 4 and 4, respectively.]

(d) The cohesiveness of a set of vertices $S$ in an undirected graph is

$$coh(S) = \min_{i \in S} \frac{|N_i \cap S|}{|N_i|},$$

where $N_i$ is the set of neighbors of $S$ in the graph.

Why would a play of “0” by each of a set of vertices with cohesiveness greater than 2/3 block the spread of a 1-cascade on an undirected network in the context of the coordination game in part (c)?
8. **Top-trading cycle mechanism**

Consider the TTC mechanism for house allocation.

(a) Consider seven agents, with preference orders:

1: \( 2 \succ_1 3 \succ_1 1 \)
2: \( 3 \succ_2 4 \succ_2 5 \succ_2 2 \)
3: \( 1 \succ_3 3 \)
4: \( 2 \succ_4 1 \succ_4 5 \succ_4 4 \)
5: \( 4 \succ_5 5 \)
6: \( 5 \succ_6 7 \succ_6 6 \)
7: \( 5 \succ_7 6 \succ_7 7 \)

The preference orders are truncated at the point where an agent prefers the house it already owns. What is the behavior of the TTC mechanism on this example? **State the trades in each round.**

(b) Explain why the outcome of the TTC mechanism is Pareto Optimal. You can argue with respect to this particular example if you choose. [Hint: argue by contradiction.]

(c) Suppose the *random serial dictator* mechanism (RSD) was used instead, where the houses of the seven agents were first pooled. What important property of TTC is not achieved by RSD?

(d) State two differences between the 3-cycle matching problem in paired-kidney exchange and the matching problem solved in a single round of the TTC mechanism.
9. Two-sided Matching

Consider the following preference orders for four firms and four workers (most preferred to least preferred):

\begin{align*}
\text{f1:} & \quad \text{w1} \quad \text{w2} \quad \text{w3} \quad \text{w4} \\
\text{f2:} & \quad \text{w2} \quad \text{w1} \quad \text{w4} \quad \text{w3} \\
\text{f3:} & \quad \text{w3} \quad \text{w4} \quad \text{w1} \quad \text{w2} \\
\text{f4:} & \quad \text{w4} \quad \text{w3} \quad \text{w2} \quad \text{w1} \\
\text{w1:} & \quad \text{f4} \quad \text{f3} \quad \text{f2} \quad \text{f1} \\
\text{w2:} & \quad \text{f3} \quad \text{f4} \quad \text{f1} \quad \text{f2} \\
\text{w3:} & \quad \text{f2} \quad \text{f1} \quad \text{f4} \quad \text{f3} \\
\text{w4:} & \quad \text{f1} \quad \text{f2} \quad \text{f3} \quad \text{f4}
\end{align*}

(a) Consider the following two stable matchings

\begin{align*}
\text{matching 1:} & \quad (\text{f1, w2}) \quad (\text{f2, w1}) \quad (\text{f3, w3}) \quad (\text{f4, w4}) \\
\text{matching 2:} & \quad (\text{f1, w1}) \quad (\text{f2, w2}) \quad (\text{f3, w4}) \quad (\text{f4, w3})
\end{align*}

What is the “join” of these two matchings (where each firm gets its most-preferred worker across both matchings)?

(b) What do you notice about this matching?

Consider the following preference orders for three firms and three workers (most preferred to least preferred):

\begin{align*}
\text{f1:} & \quad \text{w1} \quad \text{w2} \quad \text{w3} \\
\text{f2:} & \quad \text{w1} \quad \text{w3} \quad \text{w2} \\
\text{f3:} & \quad \text{w1} \quad \text{w2} \quad \text{w3} \\
\text{w1:} & \quad \text{f1} \quad \text{f2} \quad \text{f3} \\
\text{w2:} & \quad \text{f1} \quad \text{f2} \quad \text{f3} \\
\text{w3:} & \quad \text{f1} \quad \text{f3} \quad \text{f2}
\end{align*}

(c) There’s something about the firms’ preferences that seems to make this matching instance difficult. What is it?

(d) How does stability address this concern in general, and in this example in particular?
10. **Price of Anarchy**

There is an atomic routing game with two agents, each with one unit to route from $s$ to $u$. Each edge in the figure is associated with a cost function. Each agent selects a path and incurs cost equal to the sum of cost on edges it selects, and wants to minimize her own cost in equilibrium.

(a) What is a Nash equilibrium of this game? Justify your answer, and state the total cost.

(b) What is a socially optimal (efficient) solution?

(c) Define the POA for a cost-minimization game $\Gamma$. Use $c_i(a)$ for the cost to $i \in N$ for action profile $a$, total cost $C(a) = \sum_{i \in N} c_i(a)$. 
(d) The (5/3,1/3)-smoothness of this game states

\[ \sum_{i \in N} c_i(a_i^{opt}, a_{-i}) \leq (5/3)C(a^{opt}) + (1/3)C(a), \]

for all \( a \in A \), and provides \( POA(\Gamma) \leq 5/2 \). Verify the smoothness property where profile \( a \) is the Nash flow.
11. **Privacy**

(a) What is the sensitivity of the following queries in regard to hospital admissions data? State any assumptions.

(i) A count of the number of admitted people by medical condition, when each person can have multiple conditions.

(ii) A count of the number of admitted people by medical condition, when each person can have at most one condition.

(iii) The fraction of admitted people with heart failure, where you should assume the total number of admitted people is not known without a query.

(iv) A count of the number of patients admitted to a hospital in Boston on January 1, 2013.

(v) A count of the number of patients receiving emergency room treatment in each of a hospital in San Francisco and a hospital in Boston on Monday mornings in January.

(vi) A count of the total number of patients treated in the emergency room during January in a hospital in Boston.
(vii) A count of the total number of people who were re-admitted to at least one hospital in Boston during 2012.

(b) Briefly, what role does sensitivity play in the Laplace perturbation, when applied to a query that returns a count of the number of records with some property?

(c) Briefly, what property of differential privacy tells us that we can execute differentially-private queries of a database and then publish the results with compromising privacy, allowing for any other analysis to be done on these published results?
12. Quick Fire

(a) A regularity of real world networks is a heavy-tailed degree distribution. For each log-log plot, does it suggest a degree distribution that is power law, heavy-tailed, or neither?

(b) True or False. Bitcoin solves the money printing problem by using a peer-to-peer protocol that prevents anyone from creating new money.

(c) The edge density of a graph is the fraction of edges over all possible edges, and the clustering coefficient is the fraction of neighbors with edges over all possible edges between neighbors, averaged over all vertices.

What is the clustering coefficient and edge density of these two graphs? Do you consider them to have high edge clustering?

(d) True of False. Randomization is necessary to achieve differential privacy for useful algorithms.