Networks and Markets
Homework 1

Instructor: Rafael Pass   TA: Yoav Ben Shimon

Part 1: Game Theory

1. For each of the following three two-player games, find (i) all strictly *dominant* strategies, (ii) the action profiles which survive iterative removal of strictly *dominated* strategies (ISD), and (iii) all pure-strategy Nash equilibria. Give a *brief justification* for each part.

(a)

	$(*, L)$	$(*, R)$
$(U, *)$	(5, 4)	(4, 5)
$(D, *)$	(4, 4)	(0, 0)

(b)

	$(*, L)$	$(*, R)$
$(U, *)$	(2, 2)	(2, 1)
$(D, *)$	(3, 2)	(0, 3)

(c)

	$(*, L)$	$(*, R)$
$(U, *)$	(6, 5)	(4, 5)
$(D, *)$	(5, 4)	(2, 2)

2. Consider the two-player game given by the following payoff matrix:

	$(*, L)$	$(*, M)$	$(*, R)$
$(t, *)$	(-1, 2)	(5, 1)	(0, 0)
$(m, *)$	(1, 2)	(-1, 0)	(6, 2)
$(b, *)$	(4, 1)	(3, 1)	(2, 0)

   (a) Does either player have a strictly dominant strategy? If so, which player, what strategy, and why? If not, what is the smallest number of entries in the payoff matrix which would need to be changed so that some player did have a strictly dominant strategy? Justify why this is the minimum, i.e. there is no smaller value that works.

   (b) What are player 1's and player 2's best-response sets given the action profile (m, L)? Explain in words why these are the best responses.

   (c) Find all pure-strategy Nash equilibria for this game. (Argue why all that you wrote are PNEs and why there are no others.) Is there a single action profile from which best-response dynamics (BRD) might converge to all pure-strategy Nash equilibria? If yes, describe the action profile and the choices by which the BRD process converge to each equilibria. If no, explain why this is not possible for any action profile.

3. (a) Prove the following: If player 1 in a two-person game has a dominant strategy s_1, then there is a pure-strategy Nash equilibrium in which player 1 plays s_1 and player 2 plays a best response to s_1.
   (b) Prove/disprove: The equilibrium from part (a) is necessarily a *unique* pure-strategy Nash equilibrium.
   (c) Prove/disprove: There exists a pure-strategy Nash equilibrium where player 1 does not play s_1.
   (d) Assume now s_1 is instead a *strictly dominant* strategy for player 1 (not just dominant). How do the answers to (a)-(c) change? Justify your answer.

4. Formulate a normal-form game (as a payoff matrix) that has a unique pure-strategy Nash equilibrium, but for which best-response dynamics does not always converge (i.e. there are possible starting states and choices in the BRD process for which BRD will not converge). Justify your answer. (*Hint:* Try modifying a game where BRD does not converge to give it a unique PNE.)

5. Consider the following two-person game between a kicker (row player) and a goalie (column player). Before the penalty kick, the kicker can decide to either kick to the left, middle, or right, and the goalie can decide to either guard left, middle, or right. Based on what each player decides, they have the following payoffs:

	(*,GL)	(*,GM)	(*,GR)
(KL,*)	(-1, 1)	(-0.5, 0.5)	(1,-1)
(KM,*)	(0.5, -0.5)	(-1,1)	(0.5, -0.5)
(KR,*)	(1,-1)	(-0.5, 0.5)	(-1,1)

Intuitively, this says that the goalie blocks the kick if they both choose the same direction. The kicker scores a goal if they choose opposite directions. The goalie has a slight advantage to guard the middle (it can better react if the kicker chooses left or right). And the kicker has a slight advantage if it kicks middle and the goalie commits to left or right.

   (a) Prove that there is no pure-strategy Nash equilibrium in this game.
   (b) Suppose each player chooses each action uniformly at random (so 1/3 probability for each available choice). Compute the expected utility for each player.
   (c) The kicker notices that kicking to the middle is not working out very well for her, so she decides to randomly kick to the left or right instead (each with probability 1/2). Compute the goalie's strategy that maximizes its expected utility. Specifically, provide probabilities x_L, x_M, x_R (such that $x_L + x_M + x_R = 1$) where x_L is the probability the goalie chooses GL, x_M for GM, and x_R for GR. Then compute the expected utility for the goalie under those probabilities. Justify why the value you computed is maximal, i.e. is a best response.
   (d) **Bonus** (and a hint for the previous question): Prove there will always exist a "pure" best response to *any* mixed strategy of the kicker, where the goalie plays one of the actions with probability 1.

(e) Prove that the mixed action profile from part (c) is a Nash equilibria. Namely, neither player has a (strictly) profitable deviation.

Part 2: Graph Theory

Note: Unless stated otherwise, please assume for any problem involving graphs that we refer to *undirected* and *unweighted* graphs.

6. Given a graph, we call a node x in this graph *pivotal* for some pair of nodes y and z if x (not equal to y or z) lies on every shortest path between y and z.

   (a) Give an example of a graph in which every node is pivotal for at least one pair of nodes. Explain your answer.

   (b) For any integer $c \geq 1$, construct a graph where every node is pivotal for at least c different pairs of nodes. That is, if I give you any value for $c \geq 1$, your explanation should tell me how to construct such a graph for that c. Explain your answer.

   (c) Give an example of a graph having at least four nodes in which there is a single node x which is pivotal for every pair of nodes not including x. Explain your answer.

7. Consider a graph G on n nodes.

   (a) What is the fewest number of edges such that G is connected? Give an example with that many edges, and argue why any fewer edges must result in a graph G which is disconnected.

   (b) What is the fewest number of edges such that any two nodes in G have a shortest path length of 1? Again, prove that this is the minimum by arguing that no fewer is possible and that the number you give is attainable.

   (c) Repeat part (b) for a shortest path length of at most 2.

8. Given some connected graph G, let the *diameter* of the graph $D(G)$ be the maximum distance (i.e. shortest path length) between any two nodes. Let the *average distance* of the graph $A(G)$ be the expected shortest path length between a randomly selected pair of distinct nodes.

   Formally, the average distance is defined as $A(G) = \sum_{k=1}^{D(G)} k \cdot \Pr[\text{distance is } k]$ where $\Pr[\text{distance is } k]$ is the number of pairs of distinct nodes with distance k divided by the total number of pairs of distinct nodes.

   (a) Let G_n be a path graph on n nodes. i.e., a graph with nodes $V = \{1, \dots, n\}$ and edges $E = \{\{i, i+1\} \ : \ 1 \leq i < n\}$. Compute $A(G_n)$. The answer should not include summation (Σ).

   (b) Let G be a graph with m nodes. What is the minimal possible value of $\frac{D(G)}{A(G)}$? Provide an example graph G with m nodes such that the minimum is attained, and prove that no smaller value is possible for any graph G.

   (c) Let $c \geq 1$ be some constant. Prove there exists a graph G s.t. $\frac{D(G)}{A(G)} \geq c$.
   Partial points will be given for providing a candidate graph, and full points for a proof it satisfies the requirement.
   Hint: You can use the union bound to analyze the average distance.

(d) **Bonus (hard):** Let $a \geq 1$ and assume there exists a graph G_a with $A(G_a) = a$. Give a graph G such that $\frac{D(G)}{A(G)} \geq c$ and $A(G)$ is approximately a.

(e) Discuss what the diameter and average distance of a social network (given as a graph) might represent. What might it mean if the diameter is very similar to the average distance? What might it mean if the diameter is much greater?

Part 3: Coding: Shortest Paths

9. For this problem you will submit both `hw1.py`, as well as written/graphical answers in your main pdf submission. Please follow instructions carefully. Please don't hesitate to contact us with any questions.

    (a) In hw1.py, implement the `UndirectedGraph` class, and implement a function `create_graph(n, p)` that produces an `UndirectedGraph` with n nodes where each pair of nodes is connected by an edge with probability p.

    (b) Implement a general shortest-path algorithm for graphs, as described in lecture, that works on your graph. In hw1.py, include a function `shortest_path(G, i, j)` that outputs the length of the shortest path from node i to j in your graph G. Make sure to handle the case where the graph is disconnected (i.e. no shortest path exists) by outputting -1.

    (c) In hw1.py, implement `avg_shortest_path(G, num_samples=1000)`, to estimate the average shortest path between a random pair of two (connected) nodes in the graph, by taking the average over `num_samples` random pairs of (connected) nodes.

    Next, construct a graph for $n = 1000$ and $p = 0.1$. Run `avg_shortest_path` on your graph, taking `num_samples=1000`, and include your estimate (a number) in your written pdf submission. Does it seem reasonable? Briefly justify your answer.

    (d) For $n = 1000$, run your average shortest-path algorithm for many values of p. Specifically, choose p from $p = 0.01$ to $p = 0.04$ using .01 increments, and then $p = 0.05$ to $p = 0.5$ using .05 increments). Plot the average shortest path as a function of p and include the graph or image in your main .pdf file.

    Note: For $p = 0.01$ there is actually a small but reasonable chance (around 4%) to produce a disconnected graph. If this occurs, resample and produce a connected graph for the purposes of gathering data.

    (e) Intuitively explain the behavior of the data you found; specifically, as p increases (in particular, look at the larger values, e.g. 0.3 and above), what function $f(p)$ does the average shortest path length asymptotically approach? Justify why it behaves this way.

10. Now run your code on Facebook social network data:

    (a) In `hw1.py`, implement `create_fb_graph`. In particular, you should refer to the file "facebook_combined.txt"; the data is formatted as a list of undirected edges between 4,039 nodes, numbered 0 through 4038. You will need to parse this data as part of your code; knowing how to do this will be useful for subsequent assignments!

    You do not need to submit 'facebook_combined.txt' on Gradescope; the autograder has a copy. (Submit `hw1.py` only.)

(b) Repeat the same analysis as in part 9(c) (i.e. run your algorithm on 1000 random pairs of nodes and determine the average shortest path length). Include your code in `hw1.py`. Include your estimate (a number) in your written pdf submission. Does it seem reasonable? Briefly justify your answer.

(c) For the Facebook data, estimate the probability p that two random nodes are connected by an edge. Explain how you computed p.

(d) Is the average shortest path length of the Facebook data greater than, equal to, or less than you would expect it to be if it were a random graph with the same number of nodes and value of p? (To answer this, you may wish to run your code from question (8c) using the p you determined in part (9b) and 4039 nodes.) Explain why you think this is the case.