CS269I: Exercise Set #3

Due by 11:59 PM on Wednesday, October 19, 2016

Instructions:

(1) You can work individually or in a pair. If you work in a pair, the two of you should submit a single write-up.

(2) Submission instructions: We are using Gradescope for the homework submissions. Go to www.gradescope.com to either login or create a new account. Use the course code 9P8K49 to register for CS269I. Only one person needs to submit the assignment. When submitting, please remember to add your partner's name (if any) in Gradescope.

(3) Please type your solutions if possible. We encourage you to use the LaTeX template provided on the course home page.

(4) Write convincingly but not excessively. You should be able to fit all of your solutions into two pages, if not less.

(5) Except where otherwise noted, you may refer to the course lecture notes and the specific supplementary readings listed on the course Web page *only*.

(6) You can discuss the exercises verbally at a high level with other groups. And of course, you are encouraged to contact the course staff (via Piazza or office hours) for additional help.

(7) If you discuss solution approaches with anyone outside of your group, you must list their names on the front page of your write-up.

(8) No late assignments will be accepted, but we will drop your lowest exercise set score.

Lecture 5 Exercises

Exercise 16

Consider a two-player game in which Alice chooses either action A or action B, and Bob chooses either action C or action D. Recall that an outcome (X, Y) of the game is a *Nash equilibrium* if each player is playing a best response to the other. Equivalently, neither player can increase his or her payoff by unilaterally switching strategies.[1] Give an example of a game (with two players with two actions each) that does not have any Nash equilibria.[2]

Exercise 17

Now suppose that a game (with two players with two actions each) is played $N \geq 2$ times (where both players know N). The final payoff to a player is defined as the sum of the payoffs earned in the N "stage games." A *strategy* is now a (deterministic) function that takes as input the history-so-far (i.e., the actions taken by both players in the first i stages) and returns an action to play in the next stage. A Nash equilibrium is now defined as a strategy pair (one for Alice, one for Bob) such that each player's strategy is a best response to

[1] For simplicity, we're disallowing randomized strategies. What we're calling a Nash equilibrium is usually called a *pure-strategy* Nash equilibrium, to emphasize that each player deterministically chooses an action.

[2] If randomized strategies are allowed, then every game (with any finite number of players and strategies) has at least one Nash equilibrium. (This is what Nash proved, back in 1950.)

that of the other (i.e., no other strategy nets larger total payoff against the other's strategy). Prove that if (X, Y) is a Nash equilibrium in the stage game, then the following is a Nash equilibrium in the repeated game: Alice always plays X, and Bob always plays Y.

Exercise 18

Show by example that a repeated game can have a Nash equilibrium in which the actions chosen by the players in the first stage do not constitute a Nash equilibrium of the stage game. (Two players, two actions each, and $N = 2$ suffices.)

Exercise 19

Recall the payoff matrix from lecture for the Prisoner's Dilemma:

	Cooperate	Defect
Cooperate	$2, 2$	$-1, 3$
Defect	$3, -1$	$0, 0$

Recall the Tit-for-Tat strategy: at stage 1, cooperate; at stage i, do whatever the other player did in stage $i-1$. Prove that the Tit-for-Tat strategy never wins a head-to-head match: no matter what strategy Bob uses, if Alice uses Tit-for-Tat, then Alice's total payoff is at most that of Bob's.

Lecture 6 Exercises

Exercise 20

Consider the following extension of the model of badges discussed in lecture. Rather than one badge, there is a sequence of k badges. The ith badge is acquired after T_i successes (with $T_1 < T_2 < \cdots < T_k$).[3] We extend the model from lecture by letting $v_i \geq 0$ denote the additional value that the user gets from the ith badge, above and beyond that already conferred by the $(i-1)$th badge. (So if the user has earned the first i badges, her total value from them is $\sum_{j=1}^{i} v_j$.) All other aspects of the model remain the same (the preferred action p, the cost $h(p, q)$ of playing an action q other than p, the discount factor γ, and the discounted expected utility.) Explain how to modify the approach from lecture to solve for the optimal strategy of the user (i.e., the optimal activity level q_s for each $s \in \{0, 1, 2, \ldots, \}$). Be sure to explicitly state the recurrence that you're solving and the order in which you solve the different subproblems.

Exercise 21

Apply the method of the previous exercise to solve for the optimal user behavior with 2 badges and the following parameter values: $v_1 = 2$, $v_2 = 1$, $T_1 = 3$, $T_2 = 6$, $\gamma = .75$, $h(p, q) = (p - q)^2$, and $p = \frac{1}{2}$. For $s \in \{0, 1, 2, \ldots, 6\}$, report the discounted future expected utility U_s starting from state s (assuming optimal play) and the optimal activity level q_s.[4]

Exercise 22

Now assume that there are two different types of activities possible (e.g., answering questions and upvoting), with a different badge for each. There is one badge for having at least T_1 successes at the first activity, and a second for having at least T_2 successes at the second activity. Assume that the user has value v_1 for the first badge, v_2 for the second badge, and $v_1 + v_2$ for both badges.

The preferred action of the user is now a pair $\mathbf{p} = (p_1, p_2)$, indicating the preferred level of each activity. (Assume that $p_1, p_2 \geq 0$ and $p_1 + p_2 \leq 1$.) Similarly, at each stage the user now chooses a pair $\mathbf{q} = (q_1, q_2)$ of activity levels, with $q_1, q_2 \geq 0$ and $q_1 + q_2 \leq 1$. There is again a cost function $h(\mathbf{p}, \mathbf{q})$ (e.g., the Euclidean

[3] E.g., think of frequent flyer programs.

[4] You might find the `fmincon` function in Matlab useful for this.

distance between $\mathbf{p}$ and $\mathbf{q}$) and a discount rate γ, and we again assume that the user maximizes discounted expected utility.

Explain how to modify the approach from lecture to solve for the optimal strategy of the user (i.e., the optimal activity level $\mathbf{q}_\mathbf{s}$ for every possible state $\mathbf{s} = (s_1, s_2)$, where s_1 and s_2 denote the number of successes in each of the two activities). Be sure to explicitly state the recurrence that you're solving and the order in which you solve the different subproblems.