CSE 518A: Assignment 2

Due: Midnight, Feb 19 (Tuesday), 2019

Notes:

- Please submit your assignments using Gradescope.
- The assignment is due **by 11:59 PM on the due date.** Remember that you may not use more than 2 late days on any one homework, and you only have a budget of 3 in total.
- Please keep in mind the collaboration policy as specified in the course syllabus. You can (and are encouraged to) discuss with other students, however, you **must write down the solutions on your own**. You must also write, in the beginning of the submission, the names of students you discuss the questions with and any external sources you used in a significant manner in solving the problem.

Assignment Description:

1. (30 points) **Cooperation and Repeated Prisoner's Dilemma.** Consider the prisoner's dilemma as illustrated in Table 1, in which each player can choose to cooperate or defect. As discussed in class, in this game, both players choosing to defect is a Nash equilibrium (in fact, it's a dominant strategy equilibrium.) However, in practice, humans sometimes might choose to cooperate when facing a similar scenario, and there have been various lines of open discussion on why humans cooperate (which is also a topic of interest in crowdsourcing research). In this question, we will look at one of the potential explanations.

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	(2,2)	(0,3)
	Defect	(3,0)	(1,1)

Table 1: Prisoner's Dilemma

Repeated prisoner's dilemma. Suppose that two players play the prisoner's dilemma repeatedly from $t = 1, \ldots, T$. After each play, the players' actions are revealed, and they obtain the payoff based on their actions for that round. We also consider the setting in which players discount their future payoff by a factor of $\delta \in (0, 1)$. In particular, if a user expects to receive u_t at time t, her total utility can be written as

$$U = \sum_{t=1}^{T} \delta^{t-1} u_t$$

(a) Show that for any finite T, both players choosing to defect for every time step is the only subgame perfect equilibrium. (Hint: Use backward induction.)

(b) Consider the "tit-for-tat" strategy as follows:

- At time $t = 1$, the player chooses to cooperate.
- At time $t > 1$, her strategy is the same as the strategy her opponent takes at $t - 1$.

Assume $T \to \infty$. Show that both players playing "tit-for-tat" is a subgame perfect equilibrium when $\delta > 1/2$ (Hint: There is a useful result called *one-deviation property* that could come in handy for this question.)

Note: If we assume there exist altruistic players and players don't know whether they are interacting with self-interested players or altruistic players, it is possible to enable cooperation even in finite repeated prisoner's dilemma [1, 2].

2. (30 points) **Incentivizing Effort with Scoring Rules.** Suppose we hire an expert to predict whether it will be rainy tomorrow. Assume our expert receives one of two signals, R (Rainy) or N (Not rainy), according to the following distribution. We also assume uniform priors, that is, the prior probabilities (before any signals) of rain and not rain are the same.

	Signal	
	R	N
Rainy	80%	20%
Not Rainy	30%	70%

This table means, if it will be rainy tomorrow, the expert has 80% chance of obtaining the R signal and 20% chance of obtaining the N signal.

(a) How likely is it going to rain tomorrow if the expert receives the R signal.

As discussed in class, if we use a strictly proper scoring rule to pay the expert, the expert would has an incentive to accurately reveal her information (assuming the expert also forms her beliefs as in (a)). Suppose we use the logarithmic scoring rule

$$S(\vec{p}, i) = \log p_i + b,$$

where b is a constant to ensure positive payments, $i \in \{1, 2\}$ denotes {rainy, not rainy}, $\vec{p}$ is a two dimensional vector (p_1, p_2) denoting the expert's beliefs on the probability of rain and no rain tomorrow.

Here we explore whether strictly proper scoring rules may also incentivize the expert to *become more accurate*. Assume the expert can pay a cost c to obtain an additional signal, drawn independently from the same distribution as the first signal.

(b) How likely is it going to rain tomorrow if the expert receives two independent signals R and N?

(c) Suppose the expert is paid using the logarithmic scoring rule. Assume the expert is rational in the sense that she aims to maximize her payment minus cost. If the expert initially receives a R signal (the first signal costs 0), at what value of c will the expert choose to pay to obtain another signal? What if her initial signal is N?

(d) (Optional) Now consider the scoring rule $S(\vec{p}, i) = a \log p_i + b$. What is the minimum value of a (as a function of c) that ensures the expert always chooses to obtain another signal?

3. (40 points) **Peer Grading and Peer Prediction.** Suppose we want to design an incentive mechanism for peer grading in MOOCs (Massive Open Online Courses). In peer grading, every student assignment is randomly given to some number of other students (graders) to grade. Note that we cannot directly apply proper scoring rules, since the ground truth (i.e., whether the assignment is good or bad) will not be revealed in the future, and the goal of peer grading is to utilize graders' reports to estimate the quality of the assignment.

   To simplify the discussion, we consider one single assignment which is given to n graders. Assume an assignment is either good or bad. Let $Pr(good)/Pr(bad)$ be the prior of the assignment to be good/bad. Let $Pr(good) = 0.8$ and $Pr(bad) = 0.2$. The prior is known to all graders. Similar to the above question, assume graders obtain signals about the assignment quality with the following distribution.

	Signal	
	G	B
Good	80%	20%
Bad	40%	60%

   Graders are asked to report their signals (G or B). The goal of the mechanism is to incentivize graders to truthfully report their signals. Assume graders are rational and aim to maximize the expected number of *points* we gave them in the mechanisms described below.

   (a) Let $Pr(G)$ and $Pr(B)$ be the probability for a grader to receive signal G and B for a random assignment. Calculate the two probabilities.

   (b) Suppose we use the output agreement mechanism: for each grader report, we randomly draw another report for the same assignment. If the two reports are the same, both graders obtain 10 bonus points for their grades. If the two reports are different, both graders obtain 0 bonus points.

   Show that truthfully reporting is not a Nash equilibrium. (Hint: Condition on a grader receiving signal B, what's the probability that another grader receives signal B? Then, assume all other graders truthfully report, what should she report when receiving signal B?)

   (c) Assume $Pr(G)$ and $Pr(B)$ are known to the mechanism. We modified the output agreement slightly as below. For each grader report, we randomly draw another report for the same assignment. If the two reports are both G, both graders obtain $10/Pr(G)$ bonus points for their grades. If the two reports are both B, both graders obtain $10/Pr(B)$ bonus points for their grades. If the two reports are different, both graders obtain 0 bonus points.

   Show that truthfully reporting in this mechanism is a Nash equilibrium.

   (d) Show that in both mechanisms at (b) and (c), all graders reporting G no matter what their signals are is a Nash equilibrium. (This is one of the main obstacles peer prediction often faces, as in most settings, there exist naive uninformative equilibria.)

   Note: This is a very simplified setting. If you are interested, you can check out `https://sites.northwestern.edu/hartline/eecs-497-peer-grading/`, which is the website of a course offered by Jason Hartline at Northwestern. It contains references to this line of research.

References

[1] David Kreps, Paul Milgrom, John Roberts, and Robert Wilson. Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, 27(2):245–252, 1982.

[2] Andrew Mao, Lili Dworkin, Siddharth Suri, and Duncan J. Watts. Resilient cooperators stabilize long-run cooperation in the finitely repeated Prisoner's Dilemma. *Nature Communications*, 8(13800), January 2017.