# CSE 518A: Assignment 3

Due: Midnight, October 12 (Wednesday), 2022

**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 4 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, both players choosing to defect is a Nash equilibrium (in fact, it's a dominant strategy equilibrium.) However, in practice, human behavior sometimes does not follow this prediction: humans sometimes might choose to cooperate when facing a similar scenario.

   There have been various lines of open discussions on why humans would choose to cooperate in this setting. 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 $\delta \geq 1/2$ (Hint: There is a useful result called *one-shot deviation principle* that could come in handy for this question.)

Note: This simplified setting implies that users might choose to cooperate if they care enough about their future payoffs. There have also been works addressing the finite repeated prisoner's dilemma [2, 3]. Roughly speaking, 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.

2. (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. The goal of peer grading is to collect truthful opinion from graders, so we can later utilize graders' reports to estimate the quality of the assignment. Note that we cannot directly apply proper scoring rules to ensure truthful reports, since the ground truth (i.e., whether the assignment is good or bad) will not be revealed.

To simplify the discussion, let us consider one single assignment which is given to multiple graders. Assume an assignment is either good or bad. Let $Pr(Good) = 0.8$ and $Pr(Bad) = 0.2$ be the prior of the assignment to be good and bad. Assume the prior is known to all graders. [1]

After a grader reviews the assignment, assume she obtains a signal about the assignment quality with the following distribution. You can interpret the "signal" as the grader's noisy assessment of the assignment quality. Here the assignment quality is denoted as {Good, Bad}, and the grader signal is denoted as {G, B}.

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

This table means, if the assignment is good, the grader has 80% chance of obtaining signal "G" and 20% chance of obtaining signal "B". Similarly, if the assignment is bad, the grader has 40% chance of obtaining signal "G" and 60% chance of obtaining signal "B".

Graders are asked to report their signals (G or B). Our goal is to design a mechanism that offers "bonus points" to incentivize graders to truthfully report their signals (as reports are correlated with the quality, if we collect enough truthful reports, we can estimate the assignment quality using ideas similar to our label aggregation lectures). We assume graders are rational and aim to maximize the expected number of bonus points obtained from the mechanism.

Please answer the following questions.

[1] To interpret the prior, you can think about the scenario in which the grader knows that she is getting an assignment randomly drawn from a pool in which 80% of them are good.

(a) Let $Pr(G)$ and $Pr(B)$ be the probabilities for a grader to receive signal $G$ and $B$ for an assignment randomly drawn from all assignments. Calculate the two probabilities.

(b) Suppose we use the output agreement mechanism: for each grader report for an assignment, we randomly draw another grader report for the same assignment. If the two reports are the same, both graders obtain 10 bonus points. 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 $B$ for the same assignment? 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. If the two reports are both $B$, both graders obtain $10/Pr(B)$ bonus points. 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 in this line of research, 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.

3. (30 points) **Information Design with Bayesian Persuasion.** Suppose a crowdsourcing company wants to recruit interns from students in our class. The company only wants to recruit the very best students. Assume among students who apply for the position, 30% of the students meet their requirement (we say a student is *good* if she/he meets the company's requirement). This 30% ratio (the prior) is common knowledge and known to everyone. For each student applying for the position, the company only knows that there is a 30% chance the student is good but does not know whether a particular student is good. Also assume that I have the perfect knowledge and know whether each student is good or not in the class.

When a student applies for the position, the company asks me for a recommendation letter (a binary signal on whether the student is good). The company needs to make the hiring decision solely based on the prior (30%) and my recommendation letter.

Assume the company obtains an utility of $1$ for hiring a good student, an utility of $-1$ for hiring a not-good student, and an utility of $0$ for not hiring. There is no budget on how many students the company can hire. The goal of the company is to make hiring decisions to maximize the expected utility. For a particular student applying for the position, when the expected utility for hiring and the expected utility for not hiring are the same, the company will choose to hire the student (i.e., when there is a tie in utility, the decision favors hiring).

My objective is to have as many students hired as possible, i.e., I obtain an utility of $1$ if a student is hired, and $0$ otherwise. I can decide on a strategy on how to write a letter. My strategy is represented by a conditional distribution: conditional on that the student is (or is not) good, what's the probability I should say that the student is good. I need to decide on a strategy to write letters for all students, and that the company is aware of my strategy (e.g., from past experience).

Now let's consider my strategies for writing the recommendation letters.

(a) (Everyone is good) Consider the strategy that, for every student applying for the position, I provide a letter saying that the student is good. What is the expected ratio of students who will get hired by the company? (Hint: Recall that the company knows my strategy. Given the common prior and the signal, i.e., my letter, what is the company's posterior belief for the applicant to be good? Given the company's utility function, how should they make the decision based on the posterior?)

(b) (Honest) Consider the strategy that, for every student applying for the position, I provide an honest letter that says the student is (or is not) good based on my knowledge (which is assumed to be perfectly correct). What is the expected ratio of students who will get hired by the company?

(c) There exists an optimal strategy that 60% of students will get hired in expectation. Please characterize that optimal strategy, i.e., what is the probability I should say the student is good when the student is indeed good, and what is the probability I should say the student is good when the student is not?

Note: This is a simplified setup in Bayesian persuasion [1], where a *sender* (me) wants to design information (recommendation letters) to persuade a *receiver* (the company) to take certain actions. This information design question has been everywhere, e.g., online retailers might decide to show different product features to persuade customers to make purchase, recommendation systems might decide whether to include other users' ratings to persuade users to take recommendations, public health officials might want to provide vaccine information to persuade people to take vaccines, etc.

To simplify discussion, we allow *lying* in the information strategy in the above question. In fact, we don't need to explicitly lie to achieve the same effect. Selectively disclosing a subset of information could lead to the same effect, even under the condition each piece of information is true information (e.g., providing evidence showing that the student is good but not mentioning the counter-evidence). The advantageous party that has the power to choose what information to present actually has a lot of leverage to influence people's decisions, even under the condition that people know exactly the strategy of the advantageous party.

## References

[1] Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

[2] 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.

[3] 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.