# Economics and Computation (Spring 2026)
# Assignment #5
# — Solutions —

Due: 4/28/2026 11:59pm ET

## Problem 1: Random assignment

**[20 points]** A random assignment $P$ is *envy free* if for all $i, j \in N$ and $x \in G$,

$$\sum_{y \succeq_{\sigma_i} x} p_{iy} \geq \sum_{y \succeq_{\sigma_i} x} p_{jy}.$$

Prove that the Probabilistic Serial Mechanism produces an envy-free random assignment.

**Solution:**

First, fix any $i, j \in N$ and $x \in G$. Without loss of generality, assume that players eat throughout the time interval $[0, 1]$. Thus, as each player is eating a uniform rate, if a player eats for a duration $t$, they will accrue a total of $t$ in total probability of receiving some item (potentially across different items). Now, denote $Y = \{y : y \succeq_{\sigma_i} x\}$ to be the set of all items player $i$ weakly prefers to $x$. Further, denote $t_Y \in [0, 1]$ to be the first time that all items in $Y$ are fully eaten. At this point in the mechanism, the probabilities associated with each item $y \in Y$ have been fully allocated. Additionally, we know that $\sum_{y \succeq_{\sigma_i} x} p_{iy} = t_Y$ because agent $i$ will have spent all time $[0, t_Y]$ eating goods in $Y$ before they then begin eating the their next good. Thus, we get that

$$\sum_{y \succeq_{\sigma_i} x} p_{iy} = t_Y \geq \sum_{y \succeq_{\sigma_i} x} p_{jy}$$

Note that the inequality follows from the fact that player $j$'s total accrued probabilities at time $t_Y$ will be $t_Y$, and at this point $p_{jy}$ for all $y \in Y$ will be fully determined, and so $\sum_{y \succeq_{\sigma_i} x} p_{jy} \leq t_Y$ where equality holds only if agent $j$ has been eating goods in $Y$ for all of times $[0, t_Y]$.

## Problem 2: Cascade models

**[15 points]** In Lecture 17 we discussed the coordination game. Consider a similar game, called the *local public goods game*, which is defined using the notation used in Slide 3. The possible actions are again $a_i \in \{0, 1\}$, but here 1 corresponds to investing in a good that is useful to the neighbors of $i$, and 0 corresponds to not investing. The utility of $i$ is

$$u_i(\mathbf{a}) = \begin{cases} 1 - c & a_i = 1 \\ 1 & a_i = 0 \text{ and } n_{i,1}(\mathbf{a}_{-i}) \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

That is, $i$ gets a payoff of 1 if at least one player in their neighborhood (including themselves) invests, but investment has a cost of $c \in (0, 1)$.

Design a polynomial-time algorithm that computes a pure Nash equilibrium in a given local public goods game.

**Solution:**

Recall, a set of vertices is *independent* if no two vertices in the set are connected by an edge. An independent set is *maximal* (not to be confused with a *maximum* independent set) if every vertex in the graph is connected to at least one vertex in the independent set, so that no additional vertex can be added to the independent set without breaking interdependence.

Any action profile where players playing 1 form a maximal independent set and all other players play 0 is a Nash equilibrium. In these action profiles, the players playing 1 are only connected to players playing 0 (by independence), and each player playing 0 is connected to at least one player playing 1 (by maximality). These profiles are pure Nash equilibria because the players playing 1 will not want to switch to 0 because this will reduce their payoff from $1 - c$ to 0, and the players playing 0 will not want to switch to 1 because this will reduce their payoff from 1 to $1 - c$.

A greedy algorithm can easily compute such an independent set by iteratively adding vertices that are not connected to the previous ones. This algorithm takes $O(|V| + |E|)$.

### Problem 3: Influence maximization

**[20 pts]** Given an undirected graph $G = (V, E)$, define the following set function over subsets $S \subseteq V$:

$$f(S) = |\{(u, v) \in E : \ u \in S, v \notin S\}|.$$

Is $f$ monotone? Is it submodular? Prove or disprove each property.

**Solution:** $f$ is not monotone. Proof by counterexample. An easy counterexample is by considering $S = V$ and so $f(S) = 0$ because there are no edges of $G$ that cross the boundary out of $S$. Making this clearer, consider $V = \{1, 2, 3\}$ and $E = \{(1, 2), (2, 3), (3, 1)\}$ (a triangle). Then, consider $S_1 = \{1\}$ and $S_2 = \{1, 2, 3\} = V$. Then, we have that $S_1 \subseteq S_2$ but

$$f(S_1) = |\{(1, 2), (1, 3)\}| = 2$$

$$f(S_2) = |\emptyset| = 0$$

and so $f(S_1) > f(S_2)$, contradicting monotonicity.

$f$ is submodular. We wish to show that for any $X \subseteq Y \subseteq V$ and any $z \notin Y$ then $f(X \cup \{z\}) - f(X) \geq f(Y \cup \{z\}) - f(Y)$. We first note that

$$f(X \cup \{z\}) - f(X) = |\{(z, v) \in E : v \notin X \cup \{z\}\}| - |\{(v, z) \in E : v \in X\}|$$

because we are adding in the edges connected to $z$ and the rest of $V$, but are getting rid of the edges that are connected to $z$ from vertices inside $X$. Similarly,

$$f(Y \cup \{z\}) - f(Y) = |\{(z, v) \in E : v \notin Y \cup \{z\}\}| - |\{(v, z) \in E : v \in Y\}|$$

Now, since $X \subseteq Y$, it is clear that

$$|(z, v) \in E : v \notin X \cup \{z\}\}| \geq |(z, v) \in E : v \notin Y \cup \{z\}\}|$$
$$|(v, z) \in E : v \in X\}| \leq |(v, z) \in E : v \in Y\}|$$

and thus we conclude that

$$f(X \cup \{z\}) - f(X) \geq f(Y \cup \{z\}) - f(Y)$$

## Problem 4: No-regret learning

**[20 pts]** Consider the analysis of (deterministic) weighted majority in Lecture 19, Slides 11–14. Assume that there is a perfect expert that never makes mistakes. Show that there is a value of $\epsilon$ such that the modified weighted majority algorithm (Slide 14) makes at most $\log_2 n$ mistakes, where $n$ is the number of experts.

**Solution:** Inspired by the construction in lecture, we define $M$ to be the total number of mistakes made so far by the algorithm and $W$ be the total weight, which starts at $n$ as each of $n$ experts contributes 1 to the weight at the start.

Every time the algorithm makes a mistake, this means that a weighted majority of the experts, comprising at least $\frac{1}{2}$ of the current total weight, were wrong. An $\epsilon$ fraction of this weight is taken away from these experts. This means that after each mistake, W drops by at least $\frac{\epsilon}{2}$, so after $M$ mistakes, $W \leq n(1 - \frac{\epsilon}{2})^M$. Since the perfect expert never makes mistakes, its weight remains forever at 1.

This means that at any given point,

$$1 \leq n\left(1 - \frac{\epsilon}{2}\right)^M,$$

since the sum of all the weights has to be at least as large as the highest weight that contributes to this sum.

Taking the $\log_2$ of both sides, we have

$$\log_2\left(\frac{1}{n}\right) \leq M \log_2\left(1 - \frac{\epsilon}{2}\right)$$

After some arithmetic with the logarithms, we have

$$\log_2(n) \geq M \log_2\left(\frac{2}{2-\epsilon}\right)$$

If we set $\epsilon = 1$, then $\log_2(n) \geq M \log_2(2)$, and it follows that $\log_2(n) \geq M$. Thus, there is a value of $\epsilon$ such that the modified weighted majority algorithm makes at most $\log_2 n$ mistakes.

### Problem 5: Feature attribution

**[25 points]** The Shapley value is hard to compute, but it is easy to estimate accurately using a Monte Carlo algorithm. Specifically, given a player $i$ whose Shapley value we wish to estimate, consider the following algorithm: For $t = 1, \ldots, m$, sample a random permutation $\pi_t$ and compute $v(S^i_{\pi_t} \cup \{i\}) - v(S^i_{\pi_t})$; then return the marginal contribution of $i$ averaged across the $m$ samples.

Assume that $v(S) \in [0,1]$ for all $S \subseteq N$ and $v$ is monotonic. Show that, given $\epsilon, \delta > 0$ and $m = O(\ln(1/\delta)/\epsilon^2)$, the above algorithm outputs an estimate $\hat{\sigma}_i$ of the Shapley value of $i$ such that $|\sigma_i - \hat{\sigma}_i| < \epsilon$ with probability at least $1 - \delta$.

**Guidance:** Each sample is a random variable; what can you say about their expectations? Plug these random variables into Hoeffding's Inequality: Let $X_1, \ldots, X_m$ be i.i.d. random variables bounded in $[0,1]$ with $\mathbb{E}[X_j] = \mu$ for $j = 1, \ldots, m$, then

$$\Pr\left[\left|\frac{1}{m}\sum_{t=1}^{m} X_t - \mu\right| \geq \epsilon\right] \leq 2 \cdot e^{-2m\epsilon^2}.$$

**Solution:** As given, we use the following Monte Carlo algorithm to estimate the Shapley value of player $i$ by sampling random permutations and averaging the marginal contributions.

For each sample $t = 1, ..., m$, we:

1. Randomly generate a permutation $\pi_t$ of all players
2. Find $S^i_{\pi_t}$, the set of all players preceding $i$ in permutation $\pi_t$
3. Calculate $X_t = v(S^i_{\pi_t} \cup \{i\}) - v(S^i_{\pi_t})$, which represents $i$'s marginal contribution
4. After $m$ samples, return $\hat{\sigma}_i = \frac{1}{m}\sum_{t=1}^{m} X_t$ as our estimate

Each sample $X_j$ is an independent random variable bounded in $[0,1]$ (since $v(S) \in [0,1]$ for all coalitions $S$ and it is monotonic). Furthermore, $\mathbb{E}[X_j] = \sigma_i$ because the Shapley value is precisely the expected marginal contribution across all permutations. Since our algorithm samples permutations uniformly at random, each $X_j$ provides an unbiased estimate of the true Shapley value $\sigma_i$.

Applying Hoeffding's inequality to these i.i.d. bounded random variables:

$$\Pr\left[\left|\frac{1}{m}\sum_{t=1}^{m} X_t - \sigma_i\right| \geq \epsilon\right] \leq 2 \cdot e^{-2m\epsilon^2}$$

Taking the complement and using our estimator notation:

$$\Pr\left[|\hat{\sigma_i} - \sigma_i| < \epsilon\right] \geq 1 - 2 \cdot e^{-2m\epsilon^2},$$

which is equivalent to

$$\Pr\left[|\sigma_i - \hat{\sigma_i}| < \epsilon\right] \geq 1 - 2 \cdot e^{-2m\epsilon^2}.$$

We thus introduce $\delta$ such that

$$\Pr\left[|\sigma_i - \hat{\sigma_i}| < \epsilon\right] \geq 1 - 2 \cdot e^{-2m\epsilon^2} \geq 1 - \delta$$

We focus on:

$$1 - 2 \cdot e^{-2m\epsilon^2} \geq 1 - \delta$$

This is equivalent to saying that

$$\delta \geq 2 \cdot e^{-2m\epsilon^2}$$

$$\ln \delta \geq \ln(2) - 2m\epsilon^2$$

Solving for $m$, we need at least:

$$m \geq \frac{\ln(2) - \ln(\delta)}{2\epsilon^2} = \frac{\ln(2) + \ln(1/\delta)}{2\epsilon^2}$$

Therefore, given $\epsilon, \delta > 0$ and $m = O(\ln(1/\delta)/\epsilon^2)$ samples, the algorithm with a polynomial number of samples outputs an estimate $\hat{\sigma}_i$ of the Shapley value of $i$ such that $|\sigma_i - \hat{\sigma}_i| < \epsilon$ with probability at least $1 - \delta$.

This shows we can approximate Shapley values to arbitrary precision with high probability using only a polynomial number of samples, making an approximation tractable even for large games.