Q3. [8 pts] The Value of Games

Pacman is the model of rationality and seeks to maximize his expected utility, but that doesn't mean he never plays games.

(a) [4 pts] **A Costly Game.** Pacman is now stuck playing a new game with only costs and no payoff. Instead of maximizing expected utility $V(s)$, he has to minimize expected costs $J(s)$. In place of a reward function, there is a cost function $C(s, a, s')$ for transitions from s to s' by action a. We denote the discount factor by $\gamma \in (0, 1)$. $J^*(s)$ is the expected cost incurred by the optimal policy. Which one of the following equations is satisfied by J^*?

- ◯ $J^*(s) = \min_a \sum_{s'} [C(s, a, s') + \gamma \max_{a'} T(s, a', s') * J^*(s')]$
- ◯ $J^*(s) = \min_{s'} \sum_a T(s, a, s')[C(s, a, s') + \gamma * J^*(s')]$
- ◯ $J^*(s) = \min_a \sum_{s'} T(s, a, s')[C(s, a, s') + \gamma * \max_{s'} J^*(s')]$
- ◯ $J^*(s) = \min_{s'} \sum_a T(s, a, s')[C(s, a, s') + \gamma * \max_{s'} J^*(s')]$
- ◯ $J^*(s) = \min_a \sum_{s'} T(s, a, s')[C(s, a, s') + \gamma * J^*(s')]$
- ◯ $J^*(s) = \min_{s'} \sum_a [C(s, a, s') + \gamma * J^*(s')]$

(b) [4 pts] **It's a conspiracy again!** The ghosts have rigged the costly game so that once Pacman takes an action they can pick the outcome from all states $s' \in S'(s, a)$, the set of all s' with non-zero probability according to $T(s, a, s')$. Choose the correct Bellman-style equation for Pacman against the adversarial ghosts.

- ◯ $J^*(s) = \min_a \max_{s'} T(s, a, s')[C(s, a, s') + \gamma * J^*(s')]$
- ◯ $J^*(s) = \min_{s'} \sum_a T(s, a, s')[\max_{s'} C(s, a, s') + \gamma * J^*(s')]$
- ◯ $J^*(s) = \min_a \min_{s'} [C(s, a, s') + \gamma * \max_{s'} J^*(s')]$
- ◯ $J^*(s) = \min_a \max_{s'} [C(s, a, s') + \gamma * J^*(s')]$
- ◯ $J^*(s) = \min_{s'} \sum_a T(s, a, s')[\max_{s'} C(s, a, s') + \gamma * \max_{s'} J^*(s')]$
- ◯ $J^*(s) = \min_a \min_{s'} T(s, a, s')[C(s, a, s') + \gamma * J^*(s')]$