BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI,
K K BIRLA GOA CAMPUS
Open Book, Open Laptop, Internet NOT allowed

Subject Name: CS F407 - Artificial Intelligence — Date: 10 May 2023
Examiner Name: Sravan Danda — Marks: 40
Duration: 3 hours (10 AM - 1 PM)

Begin your answer to each question on a new page. Attempt all questions. Marks corresponding to each question is highlighted in bold within square braces at the end of the question. You are allowed to carry any material on your laptops. However, you would not be allowed to use the internet. In case of ambiguities in any of the questions, clearly state your assumptions and attempt the question(s).

1. A used-car buyer is deciding whether to buy car c_1. Assume that there is time to carry out at most one test, and that t_1 is the test of c_1 and costs \$$x$ ($x > 0$). A car can be in good shape (quality $Q = q^+$) or bad shape (quality $Q = q^-$), and the test might help indicate what shape the car is in. $t_1(c_1)$ denotes whether or not the car c_1 has passed the test and $Q(c_1)$ denotes the quality of the car. Car c_1 costs \$1,500, and its market value is \$2,000 if it is in good shape; if not, \$700 in repairs will be needed to make it in good shape. The buyer's estimate is that c_1 has a 60% chance of being in good shape (prior distribution). Tests can be described by the probability that the car will pass or fail the test given that the car is in good or bad shape. We have the following information: $P\{t_1(c_1) = pass | Q(c_1) = q^+\} = 0.8$, $P\{t_1(c_1) = pass | Q(c_1) = q^-\} = 0.1$.

   (a) Calculate the expected net gain (in \$) from buying c_1, given no test.

   (b) Calculate the posterior probabilities that the car is in good shape given each possible test outcome.

   (c) Calculate the optimal decisions given either a pass or a fail, and their expected net gains as a function of x.

   (d) Calculate the value of information (in \$) of the test. What is the maximum value of x beyond which the optimal plan for the buyer does not involve the test t_1?

   [1+(1+1)+(1+1+1+1)+(1+1)]

2. Suppose each of two players announces a non-negative integer equal to at most 100. If $x_1 + x_2 \leq 100$ where x_i is the number announced by player i then each player receives a payoff of x_i. If $x_1 + x_2 > 100$ and $x_i < x_j$, then player i gets x_i and player j gets $100 - x_i$. If $x_1 + x_2 > 100$ and $x_i = x_j$ then each player receives 50.

   (a) What are the maximum and minimum payoffs for each player (including irrational moves of the players)?

   (b) Is there a dominant strategy for each player? Justify your answer.

   (c) Find a pure-strategy Nash equilibrium? Justify your answer.

   [2+(1+3)+(1+1)]

3. Fig 1 shows a rectangular gridworld representation of a simple finite MDP. The cells of the grid correspond to the states of the environment. At each cell, four actions are possible: **left, right, up, and down**, which deterministically cause the agent to move one cell in the respective direction on the grid. Actions that would take the agent off the grid leave its location unchanged, but also result in a reward of -1. Other actions result in a reward of 0, except those that move the agent out of the special states A and B. From state A, all four actions yield a reward of $+10$ and take the agent to A'. From state B, all actions yield a reward of $+5$ and take the agent to B'. Assume that the discounting factor γ satisfies $0 < \gamma < 1$. There are no terminal states.

   (a) Suppose a constant c is added to all the one-step transition rewards (including the actions that attempt to take the agent off the grid), argue that the expected utilities of all the states (of a given policy π) in the modified MDP are translated by a constant v_c. Find the relation between v_c, c and γ.

   (b) Using (a) or otherwise argue that the relative order of the expected utility values of the states (of a given policy) π remain the same in the modified MDP with translated one-step rewards?

   [(3+2)+1]

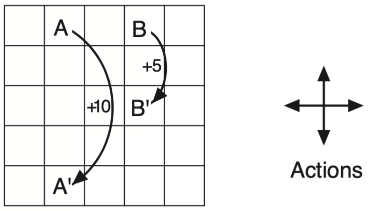

Figure 1: Figure for Problem 3

4. Let $v_1, \cdots, v_n$ denote the estimated values of the bid of all n agents participating in the bid. Assume that all the agents act rationally. Consider a descending-bid auction given as follows: The auctioneer starts at a price p (with $p > 2max_{1\leq i\leq n}v_i$) and lowers the price in integer multiples of d (fixed and known) until some buyer is willing to accept that price. (If multiple bidders accept the price, one is arbitrarily chosen as the winner.) For example, the starting ask is p and if no one bids, the next ask is $p - d$ and if no one bids, the next ask is $p - 2d$ and so on. Let b_{max} denote the value of the winning bid. Assume that $0 < d < max_{1\leq i\leq n}v_i - min_{1\leq i\leq n}v_i$.
   (a) Does the descending-bid always result in the bidder with the highest value for the item obtaining the item? Justify.
   (b) Compute the revenue to the auctioneer for the descending-bid mechanism in terms of the quantities $v_1, \cdots, v_n, p, d$.
   (c) Consider a modified descending-bid where except that at the end, the winning bidder, the one who bid b_{max}, pays only $\dfrac{b_{max}}{2}$ rather than b_{max}. Compute the revenue to the auctioneer for the modified descending-bid mechanism in terms of the quantities $v_1, \cdots, v_n, p, d$.
   (d) Is the revenue to the auctioneer for the modified descending-bid mechanism always at least as much as the revenue to the auctioneer for the descending-bid mechanism? Justify.

[2+2+2+(1+1)]

5. A gambler has the opportunity to make bets on the outcomes of a sequence of coin flips. If the coin comes up heads, he wins as many dollars as he has staked on that flip; if it is tails, he loses his stake. The game ends when the gambler wins by reaching his goal of 100, or loses by running out of money. On each flip, the gambler must decide what portion of his capital to stake, in integer numbers of dollars. This problem can be formulated as an undiscounted, episodic, finite MDP. The state is the gambler's capital, $s \in \{1, \cdots, 99\}$ and the actions are stakes, $a \in \{1, \cdots, min\{s, 100-s\}\}$. The reward is zero on all transitions except those on which the gambler reaches his goal, when it is $+1$. A policy is a mapping from levels of capital to stakes. Let p denote the probability of the coin coming up heads. Assume $p < 0.5$. Let π^* denote an optimal policy and for each $s \in \{0, 1, \cdots, 99, 100\}$, $U^{\pi^*}(s)$ denotes the expected utility of the state s under the policy π^*.
   (a) Argue that $U^{\pi^*}(1) < U^{\pi^*}(2)$ holds. (Hint: Check the number of possible actions at state 1).
   (b) Suppose $0 < s < 100$, does $U^{\pi^*}(s) < pU^{\pi^*}(s+1) + (1-p)U^{\pi^*}(s-1)$ hold? Justify.
   (c) Argue that the utilities of the states of an optimal policy π^* are strictly increasing i.e. they satisfy $U^{\pi^*}(s_1) < U^{\pi^*}(s_2)$ whenever $0 \leq s_1 < s_2 \leq 100$ as follows - Let $k = max\{s : U^{\pi^*}(s) > U^{\pi^*}(s-1),\ s \in \{1, \cdots, 100\}\}$, it is easy to see from (a) that $k \geq 2$. Show the following:
      i. If $k < 100$, then show that $U^{\pi^*}(k) \geq U^{\pi^*}(k+1) \geq U^{\pi^*}(k+2) \geq \cdots \geq U^{\pi^*}(100)$.
      ii. If $k < 100$, are any of the Bellman equations violated? Justify and complete the proof.

[1+(1+1)+(3+3)]