B.Sc. Engg. SWE 6th Semester — 14 May 2024

ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)

ORGANISATION OF ISLAMIC COOPERATION (OIC)

Department of Computer Science and Engineering (CSE)

SEMESTER FINAL EXAMINATION — **SUMMER SEMESTER, 2022-2023**
DURATION: 3 HOURS — **FULL MARKS: 150**

CSE 4617: Artificial Intelligence

Programmable calculators are not allowed. Do not write anything on the question paper.
Answer **all 5 (five)** questions. Figures in the right margin indicate full marks of questions with corresponding COs and POs in parentheses.

1. Consider that a game of Pacman is being played on the grid shown in Figure 1.

	D	
G	P	G

Figure 1: Pacman Grid for Question 1

Here, P indicates the position of Pacman, G indicates the position of the ghosts, and D indicates the position of a food dot. For simplicity, the ghosts remain stationary throughout the game.

To determine the policy for playing the game, our feature-based Q-learning agent, Pacman, uses two features, F_g and F_d defined as follows:

$$F_g(s,a) = F_1(s) + F_2(s,a) + F_3(s,a)$$
$$F_d(s,a) = F_4(s) + 2 \times F_5(s,a)$$

where

$F_1(s)$ = number of ghosts within 1 step of state s
$F_2(s,a)$ = number of ghosts Pacman touches after taking action a from state s
$F_3(s,a)$ = number of ghosts within 1 step of the state Pacman ends up in after taking action a
$F_4(s)$ = number of food dots within 1 step of state s
$F_5(s,a)$ = number of food dots eaten after taking action a from state s

After a few episodes of Q-learning, the weights are w_g = -10 and w_d = 100. The discount factor, $\gamma = 0.5$ and the learning rate, $\alpha = 0.5$. The action space of Pacman is $\{left, right, up, down, stay\}$. Pacman can take any actions from a state given it does not go beyond the grid.

a) Considering the actions, a that are available from the current position, s of Pacman, answer the following: 16 + 8 (CO1) (PO1)

   i. Calculate $F_g(s,a)$ and $F_d(s,a)$.

   ii. Calculate the approximate Q-value of the state-action pair (s,a), $\hat{Q}(s,a)$.

b) Recommend the optimal policy for Pacman from its current position following the $\hat{Q}(s,a)$ values that you calculated in 1.a) ii. Argue on how good the policy is considering the alternatives. 3 + 3 (CO3) (PO3)

c) From its current position, s, Pacman moves up to go to the cell, s' containing the food dot and eats it. We observe a reward, $R(s,a,s') = 250$. Considering the actions, a' that are available from s', answer the followings: 6 + 4 (CO1) (PO1)

   i. Calculate the exact Q-value of the state-action pair (s,up), $Q(s,up)$.

   ii. Update w_g and w_d.

2. a) Consider a Markov Decision Process (MDP) having three states S_1, S_2, and S_3, with rewards -1, -2, and 0 for executing an action from the state, respectively. S_3 is the terminal state where only the exit action is available that gives no reward. There is no other living reward. S_1 and S_2 each have two possible actions: a and b. 12×2 (CO3) (PO3)

The transition function is described as follows:

- In S_1, action a moves the agent to S_2 with probability 0.8 and makes the agent stay in S_1, otherwise.
- In S_2, action a moves the agent to S_1 with probability 0.8 and makes the agent stay in S_2, otherwise.
- In either S_1 or S_2, action b moves the agent to S_3 with probability 0.1 and makes the agent stay in its current state, otherwise.

Assume that the discount factor, $\gamma = 1$. With detailed steps, recommend the optimal policy for each state considering:

i. The initial policy has action b in both states.
ii. The initial policy has action a in both states.

b) Sometimes MDPs are formulated with a reward function $R(s, a)$ that depends on the current state and the action taken, or $R(s, a, s')$ that also depends on the outcome state. Modify the Bellman Equation to determine the value of a state based on these formulations. 6 (CO2) (PO2)

c) Suppose that we define the value of a state to be the maximum (as opposed to summation) reward obtained from its future states. Does this result in stationary preferences? Justify your position. $2 + 6$ (CO1) (PO1)

3. Consider the problem of solving two 8-puzzles. A sample game is shown in Figure 2:

8		6
5	4	7
2	3	1

2	8	1
	4	3
7	6	5

Figure 2: Sample 8-puzzle game for Question 3

a) Formulate the scenario as a game by identifying the start state, action, successor function, and goal test. 8 (CO2) (PO2)

b) Determine the size of the reachable state space. 3 (CO1) (PO1)

c) Suppose we make the problem adversarial by considering two players who take turns moving. Before each move, a coin is flipped to determine the puzzle on which to make a move in that turn. The winner is the first to solve one puzzle. $4 + 6$ (CO3) (PO3)

i. With proper justification, recommend an algorithm for finding the optimal move in this game.
ii. "The game will never end if both players play perfectly." - Do you agree with the statement? Justify your position.

4. Consider that you live in the city of Townsville. The city is quite safe having the chance of a burglary at 0.1% and the chance of an earthquake at 0.2%. Your home is equipped with a burglar alarm that is reliable in detecting burglary, but it may be triggered by minor earthquakes. Specifically, the alarm goes off 94% of the time when there is a burglary, and this increases to 95% when an earthquake occurs simultaneously. Additionally, the alarm is triggered 29% of the time during an earthquake, even if there is no burglary. Without either an earthquake or burglary, there is a 0.1% chance that the alarm will go off randomly.

   Your neighbors, Jamal and Munira, have agreed to notify you at work if they hear the alarm. Jamal calls 90% of the time when he hears the alarm, but occasionally (5%) mistakes the telephone ring for the alarm. Munira, due to nearby construction noise, might miss the alarm, but calls you 70% of the time when the alarm rings. She also calls 1% of the time even when the alarm does not go off.

   Assume that the probabilities summarize a potentially infinite set of circumstances that are not mentioned here.

   a) Formulate the scenario as a Bayesian Network considering causal relations and construct the associated Conditional Probability Tables. 20 (CO2) (PO2)

   b) Identify and prove at least two conditional independence from your Bayesian Network. 6 (CO3) (PO3)

   c) Find the probability that the alarm has sounded, but neither a burglary nor an earthquake has occurred, and both Jamal and Munira called you. 2 (CO1) (PO1)

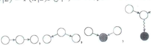
5. a) Imagine a game between two players Pascal and Fermat where each turn consists of the roll of a fair 6-sided dice. On each turn, the dice is rolled and the face that lands as the top of the dice is considered as outcome of that turn. Pascal gets a point when the outcome is even and Fermat gets a point when the outcome is odd. The first player to get T points wins the game and is rewarded M USD. Suppose the game is interrupted with Pascal leading by $X-Y$, where $0 \leq Y < X < T$. Considering the game must end prematurely, determine the general formula on how a rational agent should fairly split the reward. 7 (CO1) (PO1)

   b) Assume that you are building a rational diagnosis robot for a hospital that screens patients and assigns them doctors and/or tests based on observed symptoms. Based on the past history of the hospital, you have observed that 1 out of every 10000 patients admitted in the hospital have meningitis. You know that the disease meningitis causes a patient to have a stiff neck 80% of the time. Again, a person not having meningitis can have a stiff neck 1% of the time. A test for meningitis requires the collection of cerebrospinal fluid via a spinal tap, which is expensive. $9+7$ (CO3) (PO3)

      i. Suppose that a patient reports that they have a stiff neck. Should the robot send them for the test? Justify your choice.

      ii. Is there any reason to disagree with your choice in Question 5.b)i? Provide brief arguments in support of your decision.

Appendix

- $V^*(s) = \max_a \sum_{s'} T(s,a,s')\,[R(s,a,s') + \gamma V^*(s')]$
- $V_{k+1}(s) = \max_a \sum_{s'} T(s,a,s')\,[R(s,a,s') + \gamma V_k(s')]$, where $V_0(s) = 0$
- $V^\pi(s) = \sum_{s'} T(s,\pi(s),s')\,[R(s,\pi(s),s') + \gamma V^\pi(s')]$
- $V^\pi_{k+1}(s) = \sum_{s'} T(s,\pi(s),s')\,[R(s,\pi(s),s') + \gamma V^\pi(s')_k]$, where $V^\pi_0(s) = 0$
- $\pi^*(s) = \operatorname{argmax}_a \sum_{s'} T(s,a,s')\,[R(s,a,s') + \gamma V^*(s')]$
- $\pi_{i+1}(s) = \operatorname{argmax}_a \sum_{s'} T(s,a,s')\,[R(s,a,s') + \gamma V^{\pi_i}(s')]$
- $\pi^*(s) = \operatorname{argmax}_a Q^*(s,a)$
- $V^\pi(s) = (1-\alpha)V^\pi(s) + \alpha \times \text{sample}$, where $\text{sample} = R(s,\pi(s),s') + \gamma V^\pi(s')$
- $Q_{k+1}(s,a) = \sum_{s'} T(s,a,s')\left[R(s,a,s') + \gamma \max_{a'} Q_k(s',a')\right]$
- $Q(s,a) = R(s,a,s') + \gamma \max_{a'} f(Q(s',a'), N(s',a'))$
- $w_i = w_i + \alpha[\text{difference}]f_i(s,a)$, where $\text{difference} = \left[r + \gamma \max_{a'} Q(s',a')\right] - \hat{Q}(s,a)$
- $P(x|y) = \frac{P(x,y)}{P(y)}$
- $P(x_1, x_2, \ldots, x_n) = \prod_i P(x_i|x_1, \ldots, x_{i-1})$
- $P(x|y) = \frac{P(y|x)}{P(y)}P(x)$
- $\forall x,y,z : P(x,y|z) = P(x|z)P(y|z)$ or $\forall x,y,z : P(x|z,y) = P(x|z)$, given $X \perp\!\!\!\perp Y|Z$
- Active Triples: 
- Inactive Triples: