SC4003 Exam 2016-2017 Semester 2

1. (a) Explain the TWO aspects that make the Multi-Agent System field different from the Artificial Intelligence field.

(4 marks)

Answer

- Two different aspect of Multi-Agent System (MAS) and Artificial Intelligence (AI)
  1. Decentralized Control and Autonomy
     - In MAS, multiple autonomous agents interact, each with their own goals, knowledge, and decision-making capabilities.

     - In AI, the focus is typically on a single intelligent agent making decisions in isolation

  2. Interaction and Coordination
     - In MAS, multiple autonomous agents interact and coordinate with each other to achieve individual or collective goals.

     - AI often deals with individual problem-solving

1. (b) Environments can be characterized as *static* vs. *dynamic*. Briefly describe these environment properties, and indicate the nice property an intelligent agent needs to achieve in a dynamic environment.

(4 marks)

Answer

- Static Environment
  - A static environment is one where the environment remains unchanged while the agent is performing its tasks.
  - The agent's actions are the only factors affecting the state of the environment.

- Dynamic Environment
  - A dynamic environment is one where the environment can change over time, independently of the agent's actions.
  - Changes in the environment can occur due to factors beyond the agent's control.

- The nice property an intelligent agent needs to achieve in a dynamic environment is reactive
  - An intelligent agent must be reactive to promptly respond to changes in its environment.

  - This ensures the agent can handle real-time events without delays, improving reliability.

1. (c) Figure **Q1** shows a simple decision network for a decision of whether Tom should take an umbrella when he goes out.

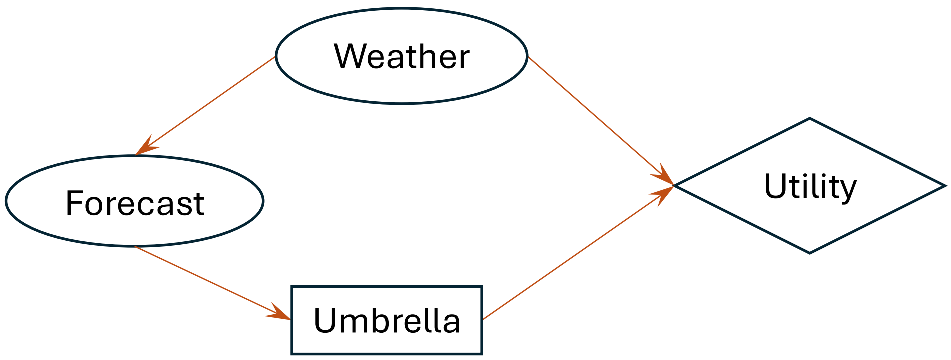

Figure Q1

The node **Forecast** indicates the weather forecast in the morning (sunny, cloudy or rainy), and the node **Weather** indicates whether or not it actually rains during the day (rain or no_rain). Tom estimates that the probability of raining during the day is 0.3. The probability of weather forecast given the actual weather, P(Forecast | Weather), is given below:

P(sunny | no_rain) = 0.7
P(cloudy | no_rain) = 0.2
P(sunny | rain) = 0.15
P(cloudy | rain) = 0.25

Tom is the most happy when it is not raining and he does not take an umbrella (utility = 100), next most happy when it is raining and he takes an umbrella (utility = 70). Tom hates carrying an umbrella when it is not raining (utility = 20), but is the most unhappy if it is raining and he does not have an umbrella (utility = 0).

1. (c) (i) For each node in the decision network, indicate its type.

(4 marks)

Answer

- Weather : Chance node
- Forecast : Chance node
- Umbrella : Decision node
- Utility : Utility node

1. (c) (ii) Given the forecast that it will be cloudy, compute the expected utility of taking an umbrella and of not taking an umbrella. Should Tom take an umbrella? Note that you may need to apply the Bayes' theorem:

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)}$$

(11 marks)

Answer

- Given:
  - P(rain) = 0.3
  - P(no_rain) = 0.7
  - P(cloudy | no_rain) = 0.2
  - P(cloudy | rain) = 0.25
  - U(no_rain, no_umbrella) = 100
  - U(rain, umbrella) = 70
  - U(no_rain, umbrella) = 20
  - U(rain, no_umbrella) = 0

- Applying Bayes' theorem:

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)}$$

$$P(\text{rain} \mid \text{cloudy}) = \frac{P(\text{cloudy} \mid \text{rain}) * P(\text{rain})}{P(\text{cloudy})}$$

- Probability rule:
  $P(A) = P(A \mid B) \cdot P(B) + P(A \mid \neg B) \cdot P(\neg B)$
  P(cloudy) = P(cloudy | rain) · P(rain) + P(cloudy | no_rain) · P(no_rain)
  = (0.25 * 0.3) + (0.2 * 0.7) = 0.075 + 0.14 = 0.215

$$P(\text{rain} \mid \text{cloudy}) = \frac{P(\text{cloudy} \mid \text{rain}) * P(\text{rain})}{P(\text{cloudy})} = \frac{0.25 * 0.3}{0.215} = 0.3488$$

P(no_rain | cloudy) = 1 – 0.3488 = 0.6512

Expected utility of taking umbrella:
EU(umbrella)
= P(rain | cloudy) * U(rain, umbrella) + P(no_rain | cloudy) * U(no_rain, umbrella)
= (0.3488 * 70) + (0.6512 * 20) = 24.416 + 13.024 = 37.44

Expected utility of no umbrella:
EU(no_umbrella)
= P(rain | cloudy) * U(rain, no_umbrella) + P(no_rain | cloudy) * U(no_rain, no_umbrella)
= (0.3488 * 0) + (0.6512 * 100) = 0+ 13.024 = 65.12

Summary of EU

- EU(umbrella) = 37.44
- EU(no_umbrella) = 65.12

Decision

Since EU(no_umbrella) = 65.12 > EU(umbrella) = 37.44, Tom should not take umbrella

2. An agent inhabits a world with two states, namely, *S* and *S'*. It can perform one of the two actions, *a* and *b*. Action *a* does nothing, and action *b* flips from one state to the other state. Consider this world as a an MDP. The reward of each state is as follows: *R(S) = 3*, and *R(S') = 2*. The discount factor $\gamma = 0.5$.

   Apply the policy iteration algorithm to determine the optimal policy. Assume that the initial policy has action *a* in both states.

2. (a) What are the two main steps involved in the policy iteration algorithm? Elaborate the purpose of each step.

(6 marks)

Answer

- Policy iteration starts with some initial policy π_0 and alternates between the two steps:

  i. Policy Evaluation
     - Given a fixed policy π, calculate $U^{\pi}(s)$ for every state s
     - Determine the expected return (value) of each state under the current policy
     - The value function is updated iteratively using the Bellman expectation equation
     - Continues to the next step

  ii. Policy Improvement
     - Calculate a new policy π^{i+1} based on the updated utilities
     - For each state *s*, determine the action *a* that maximizes the expected return
     - Update the policy to π^{i+1}
     - If the policy does not change, it means the current policy is optimal, and the algorithm terminates
     - Otherwise, repeat the Policy Evaluation step with the new policy

- The process repeats until convergence to optimal policy

2. (b) Show each step in full during each iteration. What is the optimal policy?

(19 marks)

Answer

- Given:
  - States: S, S'
  - Actions: a (does nothing), b (flips state)
  - Rewards: R(S) = 3, R(S') = 2
  - Discount Factor: $\gamma = 0.5$

- Iteration 0
  - Initial Policy:
    - $\pi_0(S) = a$
    - $\pi_0(S') = a$

  - Policy Evaluation for π_0:
    - For S:

$$U^{\pi_0}(S) = R(S) + \gamma[U^{\pi_0}(S)]$$
$$U^{\pi_0}(S) = 3 + 0.5\big(U^{\pi_0}(S)\big)$$
$$0.5[U^{\pi_0}(S)] = 3$$
$$U^{\pi_0}(S) = 6$$

    - For S':

$$U^{\pi_0}(S') = R(S') + \gamma\big(U^{\pi_0}(S')\big)$$
$$U^{\pi_0}(S') = 2 + 0.5\big(U^{\pi_0}(S')\big)$$
$$0.5[U^{\pi_0}(S')] = 2$$
$$U^{\pi_0}(S') = 4$$

  - Policy Improvement for π_0:
    - For S:

$$Q(S,a) = R(S) + \gamma\big(U^{\pi_0}(S)\big) = 3 + 0.5 * 6 = 6$$
$$Q(S,b) = R(S) + \gamma\big(U^{\pi_0}(S')\big) = 3 + 0.5 * 4 = 5$$

Since Q(S, a) = 6 > Q(S, b) = 5, new policy $\pi_1(S) = a$

    - For S':

$$Q(S',a) = R(S') + \gamma\big(U^{\pi_0}(S')\big) = 2 + 0.5 * 4 = 4$$
$$Q(S',b) = R(S') + \gamma\big(U^{\pi_0}(S)\big) = 2 + 0.5 * 6 = 5$$

Since Q(S', b) = 5 > Q(S', a) = 4, new policy $\pi_1(S') = b$

  - Updated policy
    - $\pi_1(S) = a$
    - $\pi_1(S') = b$

  - since $\pi_1 \neq \pi_0$, so continue

2. (b) cont

Answer

- Iteration 1
  - Policy π_1:
    - $\pi_1(S) = a \rightarrow$ no change
    - $\pi_1(S') = b \rightarrow$ flips to $\pi_1(S)$

  - Policy Evaluation for π_1:
    - For S:

$$U^{\pi_1}(S) = R(S) + \gamma[U^{\pi_1}(S)]$$
$$U^{\pi_1}(S) = 3 + 0.5[U^{\pi_1}(S)]$$
$$0.5[U^{\pi_1}(S)] = 3$$
$$U^{\pi_1}(S) = 6$$

    - For S':

$$U^{\pi_1}(S') = R(S') + \gamma[U^{\pi_1}(S)] \quad \leftarrow \text{flips}$$
$$U^{\pi_1}(S') = 2 + 0.5 * 6 \quad \leftarrow \text{from above}$$
$$U^{\pi_0}(S') = 5$$

  - Policy Improvement for π_1:
    - For S:

$$Q(S,a) = R(S) + \gamma[U^{\pi_1}(S)] = 3 + 0.5 * 6 = 6 \quad \leftarrow \text{no change}$$
$$Q(S,b) = R(S) + \gamma[U^{\pi_1}(S')] = 3 + 0.5 * 5 = 5.5 \quad \leftarrow \text{flips}$$

      Since Q(S, a) = 6 > Q(S, b) = 5.5, new policy $\pi_2(S) = a$

    - For S':

$$Q(S',a) = R(S') + \gamma[U^{\pi_1}(S')] = 2 + 0.5 * 5 = 4.5 \quad \leftarrow \text{no change}$$
$$Q(S',b) = R(S') + \gamma[U^{\pi_1}(S)] = 2 + 0.5 * 6 = 5 \quad \leftarrow \text{flips}$$

      Since Q(S', b) = 5 > Q(S', a) = 4.5, new policy $\pi_2(S') = b$

  - Updated policy
    - $\pi_2(S) = a$
    - $\pi_2(S') = b$

  - since $\pi_2 = \pi_1$, the policy has converged!

- Optimal Policy:
  - $\pi(S) = a \rightarrow$ In state S, take action a (do nothing)
  - $\pi(S') = b \rightarrow$ In state S′, take action b (flips to state S)

3. (a) When building a team of fighter jet agents for fighting with another team of fighter jet agents, what are the additional issues we have to consider compared with designing a single fighter jet agent?

(5 marks)

Answer

i. Coordination
   - The agents must coordinate their actions to avoid interference, ensure optimal positioning, and maximize combat effectiveness.

ii. Communication
   - Agents need to share information such as enemy locations or attack strategies with effective communication.

iii. Task Allocation
   - Each agent may need to take on different roles (e.g., attacker, defender, scout) dynamically depending on the situation.

iv. Autonomy
   - Each agent must be capable of making independent decisions while also considering the overall team strategy.

v. Learning and Adaptation
   - Ensuring the team can adapt to changing environments and enemy tactics.

3. (b) Briefly describe the five stages included in the CONTRACT NET protocol.

(5 marks)

Answer

- The contract net includes five stages:
  1. Recognition
  2. Announcement
  3. Bidding
  4. Awarding
  5. Expediting

1. Recognition
   - In this stage, an agent who recognises that there are tasks need to be complete

2. Announcement
   - In this stage, the agent with the task sends out an announcement of the task which includes a specification of the task to be achieved

3. Bidding
   - Interested members that receive the announcement decide for themselves whether they wish to bid for the task

4. Awarding
   - As a manager agent, must choose between bids & decide who to "award the contract" to.

5. Expediting
   - The selected member then expedites the task

3. (c) Anyone who wishes to register a new vehicle in Singapore must first obtain a Certificate of Entitlement (COE). COEs are bid through the COE Open Bidding System. The number of successful bidders is limited by the COEs available for each particular COE category. Each successful bidder pays the price of the highest unsuccessful bid. Assume that each bidder wants only one COE. Is the bidding mechanism truthful? Justify your answer.

(5 marks)

Answer (version 1)

- The COE bidding mechanism is truthful.

- Why it is truthful:
  - In COE bidding system, bidding your true value is a dominant strategy.
  - If you bid more than your true value, you might win and overpay, leading to a loss.
  - If you bid less than your true value, you risk losing even when the item is worth it to you.
  - Since the price you pay does not depend on your own bid, you have no incentive to lie, your best choice is to be honest.

Answer (version 2)

- The COE bidding mechanism is not truthful.

- Why it is not truthful:
  - In this COE system, the price paid is the highest rejected bid, which depends on other bidders' behaviour.
  - Bidders observe others' bids and adjust their strategies accordingly.
  - They may bid less than their true value initially and only increases as needed.
  - Bidders may engage in last-moment bids to avoid price war.
  - Bidders may raise bids aggressively to intimidate others.
  - Bidders may strategically reduce demand (bid lower than their true willingness to pay) to lower the market price.
  - This open auction encourages strategic behaviour based on competitors' actions.

3. (d) Explain manipulation in voting theory.

(5 marks)

Answer

- Manipulation in voting refers to strategic behaviour where voters do not vote for their true preferences in order to influence the outcome in their favour.

- Voters may believe that voting truthfully will not lead to their most preferred outcome, so they vote strategically to prevent a less desirable result.

- For example, in a 3-candidate election (A, B, C), a voter who prefers A > B > C might vote for B if they believe A has no chance to win, just to prevent C from winning.

- The Gibbard-Satterthwaite theorem shows that, under certain conditions, all non-dictatorial voting systems are susceptible to manipulation.

- Manipulation undermines fairness and trust in the election process, voters can benefit by strategically misrepresenting their preferences.

3. (e) Consider the coalitional game with agents Ag={1,2} and characteristic function v defined by v({1})=9, v({2})=6, v({1,2})=17. With reference to this example, explain the meaning of the core of a coalitional game.

(5 marks)

Answer

- Given Ag = {1,2}
  - v({1}) = 9
  - v({2}) = 6
  - v({1, 2}) = 17

- The core of a coalitional game is the set of payoff allocations to agents such that:
  - Efficiency (Group Rationality)
    - $x_1 + x_2 = v(1, 2) = 17$ ------------------ (1)

  - Individual Rationality
    - $x_1 \geq v(1) = 9$ ------------------ (2)
    - $x_2 \geq v(2) = 6$ ------------------ (3)

  - Coalitional Rationality
    - No subset of agents has an incentive to break away, because they cannot do better on their own than the allocation they receive in the grand coalition.

- Finding the core
  - From (1), $x_2 = 17 - x_1$ ----------------- (4)
  - (3) into (4), $6 = 17 - x_1$
    $x_1 = 17 - 6 = 11$
    Allocation $9 \leq x_1 \leq 11$
  - (2) into (4), $x_2 = 17 - 9$
    $x_2 = 8$
    Allocation $6 \leq x_2 \leq 8$

  - Possible distributions that satisfy these conditions are:
    - $x_1 = 9$ and $x_2 = 8$
    - $x_1 = 10$ and $x_2 = 7$
    - $x_1 = 11$ and $x_2 = 6$

<u>Interpretation of the core</u>

- The core guarantees stability
  - No agent or sub-set of agents has incentive to deviate from the coalition.
  - Together, they get the maximum value and split it fairly.

4. Consider the two payoff matrices A and B in Table Q4a and Table Q4b, respectively. The first number in each entry is the payoff received by the row player *i*; while the second number is the payoff received by the column player *j*.

Payoff matrix A:

Table Q4a

	j defect	***j cooperate***
i defect	**(7, 7)**	**(6, 7)**
i cooperate	**(7, 7)**	**(7, 8)**

Payoff matrix B:

Table Q4b

	j defect	***j cooperate***
i defect	**(2, 3)**	**(3, 3)**
i cooperate	**(1, 5)**	**(4, 2)**

4. (a) Identify which strategy pairs (if any) in these two payoff matrices are in dominant strategy equilibrium. Briefly explain your answer.

(5 marks)

Answer

Table Q4a

	j defect	***j cooperate***
i defect	**(7 , 7)**	**(6, 7)**
i cooperate	**(7, 7)**	**(7, 8)**

Table Q4b

	j defect	***j cooperate***
i defect	**(2, 3)**	**(3, 3)**
i cooperate	**(1, 5)**	**(4, 2)**

Payoff matrix A (Table Q4a)

- Player i (row player):
  - If j defect, i: defect = 7, cooperate = 7 → defect or cooperate same
  - If j cooperate, i: defect = 6, cooperate = 7 → cooperate is best
  - Player i has a dominant strategy → cooperate

- Player j (column player):
  - If i defect, j: defect = 7, cooperate = 7 → defect or cooperate same
  - If i cooperate, j: defect = 7, cooperate = 8 → cooperate is best
  - Player j has a dominant strategy → cooperate

Payoff matrix B (Table Q4b)

- Player i (row player):
  - If j defect, i: defect = 2, cooperate = 1 → defect is better
  - If j cooperate, i: defect = 3, cooperate = 4 → cooperate is better
  - Player i has no dominant strategy

- Player j (column player):
  - If i defect, j: defect = 3, cooperate = 3 → defect or cooperate same
  - If i cooperate, j: defect = 5, cooperate = 2 → defect is best
  - Player j has a dominant strategy → defect

Summary

Matrix A: Player i has a dominant strategy → cooperate
Player j has a dominant strategy → cooperate
Dominant strategy pair: (i cooperate, j cooperate) → (7, 8)

Matrix B: Player i has no dominant strategy
Player j has a dominant strategy → defect
No dominant strategy pair

- Dominant strategy is when there is no single action yields a better or equal payoff in all situations compared to the other actions.

4. (b) Identify which strategy pairs (if any) in these two payoff matrices are in Nash equilibrium. Briefly explain your answer.

(8 marks)

Answer

- To identify Nash equilibrium, we need to find strategy pairs where neither player has an incentive to change their strategy.

Table Q4a

	j defect	***j cooperate***
i defect	**(7 , 7)**	**(6, 7)**
i cooperate	**(7, 7)**	**(7, 8)**

Payoff matrix A (Table Q4a)

• (i defect, j defect) • i: defect = 7, coop = 7 → defect or coop • j: defect = 7, coop = 7 → defect or coop • Nash equilibrium	• (i defect, j cooperate) • i: defect = 6, coop = 7 → i prefers coop • j: defect = 7, coop = 7 → defect or coop • Not Nash equilibrium
• (i cooperate, j defect) • i: defect = 7, coop = 7 → defect or coop • j: defect = 7, coop = 8 → j prefers coop • Not Nash equilibrium	• (i cooperate, j cooperate) • i: defect = 6, coop = 7 → i prefers coop • j: defect = 7, coop = 8 → j prefers coop • Nash equilibrium

Table Q4b

	j defect	***j cooperate***
i defect	**(2, 3)**	**(3, 3)**
i cooperate	**(1, 5)**	**(4, 2)**

Payoff matrix b (Table Q4b)

• (i defect, j defect) • i: defect = 2, coop = 1 → i prefers defect • j: defect = 3, coop = 3 → defect or coop • Nash equilibrium	• (i defect, j cooperate) • i: defect = 3, coop = 4 → i prefers coop • j: defect = 3, coop = 3 → defect or coop • Not Nash equilibrium
• (i cooperate, j defect) • i: defect = 2, coop = 1 → i prefers defect • j: defect = 5, coop = 2 → j prefers defect • Not Nash equilibrium	• ((i cooperate, j cooperate) • i: defect = 3, coop = 4 → i prefers coop • j: defect = 5, coop = 2 → j prefers defect • Not Nash equilibrium

Summary of Nash Equilibrium

Payoff matrix A: (i defect, j defect) → (7, 7)

(i cooperate, j cooperate) → (7, 8)

Payoff matrix B: (i defect, j defect) → (2, 3)

Neither player has an incentive to change their strategy

4. (c) Identify which outcomes in these two payoff matrices are Pareto optimal. Briefly explain your answer.

(7 marks)

Answer

- An outcome is said to be Pareto optimal if there is no other outcome that makes one agent better off without making another agent worse off. (win-win)

- If an outcome is Pareto optimal, then at least one agent will be reluctant to move away from it (because this agent will be worse off).

- If an outcome ω is not Pareto optimal, then there is another outcome ω' that makes everyone as happy, if not happier, than ω.

Table Q4a

	j defect	***j cooperate***
i defect	**(7 , 7)**	**(6, 7)**
i cooperate	**(7, 7)**	**(7, 8)**

Payoff matrix A (Table Q4a)

- (i defect, j defect) ⟶ (7, 7)
  - Dominated by (7, 8)
  - Not Pareto optimal

- (i defect, j cooperate) ⟶ (6, 7)
  - Dominated by (7, 8) & others
  - Not Pareto optimal

- (i cooperate, j defect) ⟶ (7, 7)
  - Dominated by (7, 8)
  - Not Pareto optimal

- (i cooperate, j cooperate) ⟶ (7, 8)
  - No other better domination
  - Pareto optimal

4. (c) cont.

Table Q4b

	j defect	***j cooperate***
i defect	**(2, 3)**	**(3, 3)**
i cooperate	**(1, 5)**	**(4, 2)**

Payoff matrix B (Table Q4b)

- (i defect, j defect) → (2, 3)
  - Dominated by (3, 3)
  - Not Pareto optimal

- (i defect, j cooperate) → (3, 3)
  - No other better domination
  - Pareto optimal

- (i cooperate, j defect) → (1, 5)
  - No other better domination
  - Pareto optimal

- (i cooperate, j cooperate) → (4, 2)
  - No other better domination
  - Pareto optimal

Summary of Pareto optimal

Payoff matrix A: (i cooperate, j cooperate) → (7, 8)

Payoff matrix B: (i defect, j cooperate) → (3, 3)
(i cooperate, j defect) → (1, 5)
(i cooperate, j cooperate) → (4, 2)

No other outcome makes one player better off without making the other worse off

4. (d) Identify which outcomes in these two payoff matrices maximize social welfare. Briefly explain your answer.

(5 marks)

Answer

- Social welfare is measured as the sum of all players' payoffs. The outcome with the highest total payoff maximizes social welfare.

Table Q4a

	j defect	***j cooperate***
i defect	**(7 , 7)**	**(6, 7)**
i cooperate	**(7, 7)**	**(7, 8)**

Table Q4b

	j defect	***j cooperate***
i defect	**(2, 3)**	**(3, 3)**
i cooperate	**(1, 5)**	**(4, 2)**

Payoff matrix A (Table Q4a)

- (i defect, j defect): Total payoff = 7 + 7 = 14
- (i defect, j cooperate): Total payoff = 6 + 2 = 13
- (i cooperate, j defect): Total payoff = 7 + 7 = 14
- (i cooperate, j cooperate): Total payoff = 7 + 8 = 15

- Maximize social welfare = 15
- Outcome: (i cooperate, j cooperate) → (7, 8)

Payoff matrix B (Table Q4b)

- (i defect, j defect): Total payoff = 2 + 3 = 5
- (i defect, j cooperate): Total payoff = 3 + 3 = 6
- (i cooperate, j defect): Total payoff = 1 + 5 = 6
- (i cooperate, j cooperate): Total payoff = 4 + 2 = 6

- Maximize social welfare = 6
- Outcome: (i defect, j cooperate) → (3, 3)
  (i cooperate, j defect) → (1, 5)
  (i cooperate, j cooperate) → (4, 2)

Summary of maximize social welfare

Payoff matrix A: (i cooperate, j cooperate) → (7, 8)

Payoff matrix B: (i defect, j cooperate) → (3, 3)
(i cooperate, j defect) → (1, 5)
(i cooperate, j cooperate) → (4, 2)

Their outcomes achieve the highest total payoff maximizes social welfare.