SC4003 Exam 2023-2024 Semester 2

1. (a) Clearly list the 5 steps of the Agent Control Loop.

(10 marks)

Answer

Agent Control Loop

1. Agent starts in some initial internal state i_0
2. Observes its environment state *e*, and generates a percept *see*(*e*)
3. Internal state of the agent is then updated via *next* function, becoming *next*(i_0, *see*(*e*))
4. The action selected by the agent is *action*(*next*(i_0, *see*(*e*)))
5. Goto 2

1. (b) When calculating the utility of a state sequence in Markov Decision Processes (MDP), why is it necessary to include a discount factor? What will be the effect if the discount factor is set higher or lower, respectively?

(7 marks)

Answer

- Without discount factor, the utility may have infinite state sequences.
- Discount factor:
  - Makes sure the total utility stays bounded
  - Helps algorithms converge

- The individual state rewards are discounted by a factor γ between 0 and 1
- If the discount factor is set higher (γ closer to 1):
  - Emphasize on future rewards
  - Focus on long-term planning (e.g. in games like Chess or Go)

- If the discount factor is set lower (γ closer to 0):
  - Emphasize on immediate rewards
  - Focus on short-term benefits (e.g. quick wins in robot soccer match)

1. (c) One of the properties that an intelligent agent needs to have is "social". What does this property mean? Which of the five trends in the history of computing has made this property possible?

(4 marks)

Answer

- Social
  - An intelligent agent must be social to interact and collaborate with other agents or humans.
  - This enables coordination, negotiation, and information sharing, leading to better problem-solving.
- The trend that has made the "social" property possible is interconnection. *Computer systems today no longer stand alone, but are networked into large distributed systems*

1. (d) The question 'Isn't it all just Social Science?' is an objection to the multi-agent system field. Provide arguments against this objection.

(4 marks)

Answer

- Social science studies human behaviour, norms, and societies through observation and theory.
- Multi-agent systems (MAS) focus on algorithmic design, formal models, and computational mechanisms for artificial agents.
- Example:
  In game theory, social scientists analyse human strategies, while MAS researchers implement algorithms like Nash equilibrium solvers or auction protocols in code.
- MAS is not "just" social science, it is computational, constructive, and engineering-focused.

2. John needs to decide whether or not to get coffee for his friend Jenny. Because the floor is slippery, there is a probability of 0.5 that John will drop the coffee when he gets the coffee. If John has dropped the coffee, he has the option to clean up the floor. The utilities of different cases are given in Table Q2 below.

Table Q2

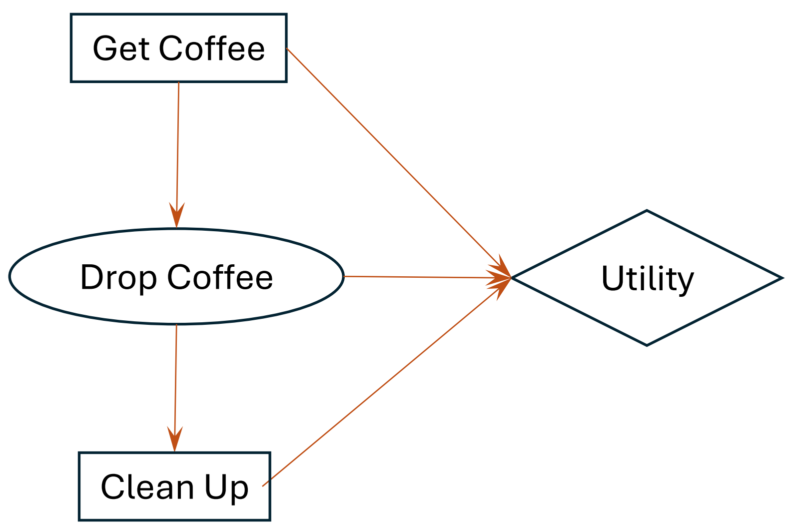
Whether to get coffee	**Whether coffee is dropped**	**Whether to clean up**	**Utility**
No	No	No	0
No	No	Yes	-20
Yes	No	No	80
Yes	No	Yes	-30
Yes	Yes	No	-100
Yes	Yes	Yes	-120

2. (a) What are the three types of nodes in a decision network for this problem? Draw the decision network.

(10 marks)

Answer

- Decision network nodes
  - Decision nodes : rectangle shape
  - Chance nodes : oval shape
  - Utility nodes : diamond shape



2. (b) If John gets coffee and drops it, should he clean up? If John gets coffee but does not drop it, should he clean up? If John does not get coffee, should he clean up? (6 marks)

Answer

Whether to get coffee	Whether coffee is dropped	Whether to clean up	Utility
No	No	No	0
No	No	Yes	-20
Yes	No	No	80
Yes	No	Yes	-30
Yes	Yes	No	-100
Yes	Yes	Yes	-120

i) If John gets coffee and drops it:
- Not clean up: Utility $U(\neg clean, get, drop) = -100$
- Clean up: Utility $U(clean, get, drop) = -120$

Decision: Since utility -100 > -120, John should not clean up

ii) If John gets coffee but does not drop it:
- Not clean up: Utility $U(\neg clean, get, \neg drop) = 80$
- Clean up: Utility $U(clean, get, \neg drop) = -30$

Decision: Since utility 80 > -30, John should not clean up

iii) If John does not get coffee:
- Not clean up: Utility $U(\neg clean, \neg get) = 0$
- Clean up: Utility $U(clean, \neg get) = -20$

Decision: Since utility 0 > -20, John should not clean up

Decisions
- If John gets coffee and drops it, he should not clean up
- If John gets coffee but does not drop it, he should not clean up
- If John does not get coffee it, he should not clean up

2. (c) Should John go to get coffee for Jenny? Show clearly all the steps of calculation. (9 marks)

Answer

Whether to get coffee	**Whether coffee is dropped**	**Whether to clean up**	**Utility**
No	No	No	0
No	No	Yes	-20
Yes	No	No	80
Yes	No	Yes	-30
Yes	Yes	No	-100
Yes	Yes	Yes	-120

P(drop | get) = 0.5
P(¬drop | get) = 0.5

From 2(b),

If John gets coffee and drops it: U(get, drop) = max (-100, -120) = -100
If John gets coffee but does not drop it: U(get, ¬drop) = max (80, -30) = 80
If John does not get coffee: U(¬get) = max (0, -20) = 0

Expected utility for getting coffee:
EU(get) = P(drop | get) * U(get, drop) + P(¬drop | get) * U(get, ¬drop)
= [0.5 * (-100)] + (0.5 * 80)
= -50 + 40
= -10

Expected utility for not getting coffee:
EU(¬get) = U(¬get) = 0

<u>Summary of EU</u>

- EU(get) = -10
- EU(¬get) = 0

Since EU(¬get) = 0 > EU(get) = -10, John should not get coffee for Jenny

3. (a) Building agents with LLM (large language model) is now a very hot area in AI. In an LLM-powered autonomous agent system, LLM functions as an agent's brain, complemented by several key components such as planning, memory and tool use. List and briefly explain at least three key ideas we discussed in the course which can be used to build a successful LLM-powered multi-agent system.

(6 marks)

Answer

i) Planning and Decision-Making

- LLMs can generate plans, but require integrating formal planning techniques to ensure coherence and goal-directed behaviour.
- Integrating planning allows agents to break down complex tasks into sub-goals and decide next steps.

ii) Communication and Coordination

- Agents use structured communication to share information, goals, and actions.
- Multi-agent communication help LLM agents coordinate actions and avoid conflict.

iii) Tool Use and Environment Interaction

- Agents can use external tools (e.g., calculators, APIs) via function calling or plugins.
- Extends capabilities beyond pure language reasoning.

3. (b) A few major countries will run presidential election (voting) this year. Explain the concept of manipulation in voting.

(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. (c) Consider a third-price auction where the winner is the bidder who submits the highest bid, but he/she only pays the third highest bid. Is the third-price auction truthful? Explain why.

(5 marks)

Answer

- No, the third-price auction is not truthful.

- In a third-price auction, the winner pays the third highest bid. This creates an incentive for bidders to strategically misreport their true valuation.

- Bidders can inflate their bid to ensure winning without affecting payment.

- Example:
  - Suppose there are three bidders A, B, C with true valuations of $10, $8, and $6, respectively.
  - If all bidders bid truthfully, A wins the bid and pays $6.
  - If A overbid significantly, say $100, A would still win and still pay $6.
  - In this case, it demonstrates that A bid does not influence the price he pays, so long as he remains the highest bidder.

- Due to bidder's payment is not directly tied to their own bid (but to the third-highest bid), they do not face the same pressure to reveal their true valuation.

3. (d) One desirable property of a good voting mechanism is Pareto property. Explain the meaning of Pareto property.

(5 marks)

Answer

- A voting mechanism satisfies the Pareto property if every voter prefers candidate A over candidate B, then B should not be chosen as the winner.

- Example:
  - Suppose there are three candidates (A, B, and C) and ten voters.
  - If all voters prefer candidate A over candidate B and C, then candidate A must win.
  - Otherwise, it violates the Pareto property.

- The Pareto property ensures that the voting mechanism respects unanimous preferences, leading to outcomes that are collectively beneficial and fair.

3. (e) Data pricing is to price data as an asset to promote the healthy development of data sharing, exchange, and reuse. Can Shapley value be applied for data pricing? Explain why.

(4 marks)

Answer

- Yes, Shapley value can be applied for data pricing.

- The Shapley value, derived from cooperative game theory, fairly distributes value among contributors based on their marginal contribution.

- Shapley value offers a principled and fair way to determine the price of data by quantifying the contribution of each dataset to the collective value generated through sharing and reuse.

- Shapley value satisfies several desirable properties for fair allocation, including:
  - Efficiency
  - Symmetry
  - Additivity
  - Dummy agent (null agent)

4. Consider the two payoff matrices 1 and 2 in Table Q4a and Table Q4b, respectively. The first number in each entry is the payoff received by the row player **A** while the second number is the payoff received by the column player **B**.

Payoff matrix 1:

Table Q4a

	B: left	**B: right**
A: up	**(-2, -1)**	**(2, 0)**
A: down	**(-1, 0)**	**(1, -1)**

Payoff matrix 2:

Table Q4b

	B: left	**B: right**
A: up	**(0, -2)**	**(-1, 0)**
A: down	**(-2, 1)**	**(0, -1)**

4. (a) Identify the dominant strategies (if any) of each player in these two payoff matrices. Briefly explain your answer.

(5 marks)

Answer

Table Q4a

	B: left	B: right
A: up	(-2, -1)	(2, 0)
A: down	(-1, 0)	(1, -1)

Table Q4b

	B: left	B: right
A: up	(0, -2)	(-1, 0)
A: down	(-2, 1)	(0, -1)

Payoff matrix 1 (Table Q4a)

- Player A (row player):
  - If B plays left, A: up = -2, down = 2 → down is best
  - If B plays right, A: up = 2, down = 1 → up is best
  - Player A has no domination

- Player B (column Payer):
  - If A plays up, B: left = -1, right = 0 → right is best
  - If A plays down, B: left = 0, right = -1 → left is best
  - Player B has no domination

Payoff matrix 2 (Table Q4b)

- Player A (Row Player):
  - If B plays left, A: up = 0, down = -2 → up is best
  - If B plays right, A: up= -1, down = 0 → down is best
  - Player A has no domination

- Player B (Column Player):
  - If A plays up, B: left = -2, right = 0 → right is best
  - If A plays down, B: left = 1, right = -1 → left is best
  - Player B has no domination

Summary

- In both matrices, neither player has a dominant strategy
- This is because 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

	B: left	B: right
A: up	(-2, -1)	(2, 0)
A: down	(-1, 0)	(1, -1)

Payoff matrix 1 (Table Q4a)

• (A: up, B: left) → (-2, -1) • A: up = -2, down = -1 → A prefers down • B: left (-1) vs. right (0) → B prefers right • Not Nash equilibrium	• (A: up, B: right) → (2, 0) • A: up = 2, down = 1 → A prefers up • B: left = -1, right = 0 → B prefers right • Nash equilibrium
• (A: down, B: left) → (-1, 0) • A: up = -2, down = -1 → A prefers down • B: left = 0, right = -1 → B prefers left • Nash equilibrium	• (A: down, B: right) → (1, -1) • A: up = 2, down = 1 → A prefers down • B: left = 0, right = -1 → B prefers left • Not Nash equilibrium

Table Q4b

	B: left	B: right
A: up	(0, -2)	(-1, 0)
A: down	(-2, 1)	(0, -1)

Payoff matrix 2 (Table Q4b)

• (A: up, B: left) → (0, -2) • A: up = 0, down = -2 → A prefers up • B: left = -2, right = 0 → B prefers right • Not Nash equilibrium	• (A: up, B: right) → (-1, 0) • A: up = -1, down = 0 → A prefers down • B: left = -2, right = 0 → B prefers right • Not Nash equilibrium
• (A: down, B: left) → (-2, 1) • A: up = 0, down = -2 → A prefers up • B: left = 1, right = -1 → B prefers left • Not Nash equilibrium	• (A: down, B: right) → (0, -1) • A: up = -1, down = 0 → A prefers down • B: left = -1, right = -1 → B prefers left • Not Nash equilibrium

Summary of Nash Equilibrium

Payoff matrix 1: (A: up, B: right) → (2, 0)
(A: down, B: left) → (-1, 0)
Neither player has an incentive to change their strategy

Payoff matrix 2: No Nash equilibrium

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

	B: left	**B: right**
A: up	**(-2, -1)**	**(2, 0)**
A: down	**(-1, 0)**	**(1, -1)**

Payoff matrix 1 (Table Q4a)

- (A: up, B:left) → (-2, -1)
  - Dominated by (2,0)
  - Not Pareto optimal because can improve to (2, 0)

- (A: up, B:right) → (2, 0)
  - No other better domination
  - Pareto optimal because no better outcome for both players

- (A: down, B:left) → (-1, 0)
  - Dominated by (2,0)
  - Not Pareto optimal because can improve to (2, 0)

- (A: down, B:right) → (1, -1)
  - Dominated by (2,0)
  - Not Pareto optimal because can improve to (2, 0)

4. (c) cont.

Table Q4b

	B: left	**B: right**
A: up	**(0, -2)**	**(-1, 0)**
A: down	**(-2, 1)**	**(0, -1)**

Payoff matrix 2 (Table Q4b)

- (A: up, B:left) → (0, -2)
  - Dominated by (0, -1)
  - Not Pareto optimal because can improve to (0, -1)

- (A: up, B:right) → (-1, 0)
  - No other better domination
  - Pareto optimal

- (A: down, B:left) → (-2, 1)
  - No other better domination
  - Pareto optimal

- (A: down, B:right) → (0, -1)
  - No other better domination
  - Pareto optimal

Summary of Pareto optimal

Payoff matrix 1: (A: up, B:right) → (2, 0)
Payoff matrix 2: (A: up, B:right) → (-1, 0),
(A: down, B:left) → (-2, 1),
(A: down, B:right) → (0, -1)

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

	B: left	B: right
A: up	(-2, -1)	(2, 0)
A: down	(-1, 0)	(1, -1)

Table Q4b

	B: left	B: right
A: up	(0, -2)	(-1, 0)
A: down	(-2, 1)	(0, -1)

Payoff matrix 1 (Table Q4a)

- (A: up, B: left): Total payoff = -2 + (-1) = -3
- (A: up, B: right): Total payoff = 2 + 0 = 2
- (A: down, B: left): Total payoff = -1 + 0 = -1
- (A: down, B: right): Total payoff = 1 + (-1) = 0

- Maximize social welfare = 2
- Outcome: (A: up, B: right) → (2, 0)

Payoff matrix 2 (Table Q4b)

- (A: up, B: left): Total payoff = 0 + (-2) = -2
- (A: up, B: right): Total payoff = -1 + 0 = -1
- (A: down, B: left): Total payoff = -2 + 1 = -1
- (A: down, B: right): Total payoff = 0 + (-1) = -1

- Maximize social welfare = -1
- Outcome: (A: up, B: right) → (-1, 0)
  (A: down, B: left) → (-2, 1)
  (A: down, B: right) → (0, -1)

Summary of maximize social welfare

Payoff matrix 1: (A: up, B: right) → (2, 0)

Payoff matrix 2: (A: up, B:right) → (-1, 0),
(A: down, B:left) → (-2, 1),
(A: down, B:right) → (0, -1)

Their outcomes achieve the highest total payoff maximizes social welfare.