1. (a) List the FIVE ongoing trends that have led to the emergence of the multiagent systems field.

(5 marks)

Answer

- Five ongoing trends:
  i. ubiquity
  ii. interconnection
  iii. intelligence
  iv. delegation
  v. human-orientation

1. (b) What are the two situations where individual agents are required? Provide an example of such agents for each situation.

(6 marks)

Answer

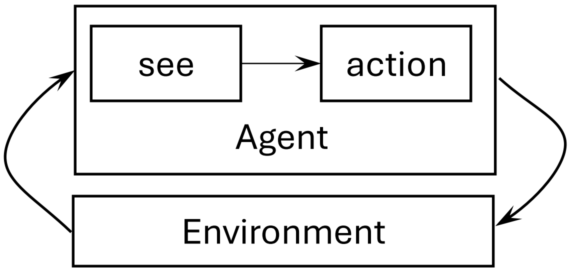
- Situation 1:
  - When autonomous action is required in high-risk or inaccessible environments
  - Example:
    - Spacecraft control agents, such as autonomous space probes like NASA's missions, that make real-time decisions without ground control.

- Situation 2:
  - When personal assistance and proactive user support are needed
  - Example:
    - Interface agents, such as email assistants, web browsing helpers, or news readers, that operate independently to enhance user experience.

1. (c) A purely reactive agent can be defined using two functions: *see: E* → *Per* and *action: Per** → *A*. With the aid of a diagram, explain how these functions define the behaviour of an agent.

(6 marks)

Answer



- The *see* function is the agent's ability to observe its environment, whereas the *action* function represents the agent's decision-making process

- *Output* of the *see* function is a *percept*:

$$see : E \rightarrow Per$$

- which maps environment states to percepts, and *action* is now a function

$$action : Per^* \rightarrow Ac$$

- which maps sequences of percepts to actions

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

(8 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. A person visits the doctor because he believes that he has the flu. At this particular time of the year, the doctor estimates that 1 out of 1000 persons suffers from the flu. The first thing the doctor examines is whether the person appears to have the standard symptoms of the flu. If the person suffers from the flu, then he will exhibit the symptoms with probability of 0.9, but if he does not have the flu, he may still have these symptoms with probability of 0.05. Then, the doctor can decide to administer a drug that has probability of 0.6 to shorten the sickness period if the person suffers from the flu (if he does not have the flu, the drug has no effect). The cost of administering the drug is $100. If the sickness period is shortened, the doctor estimates that this is worth $1000. If the sickness period is not shortened, it is worth nothing.

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

(3 marks)

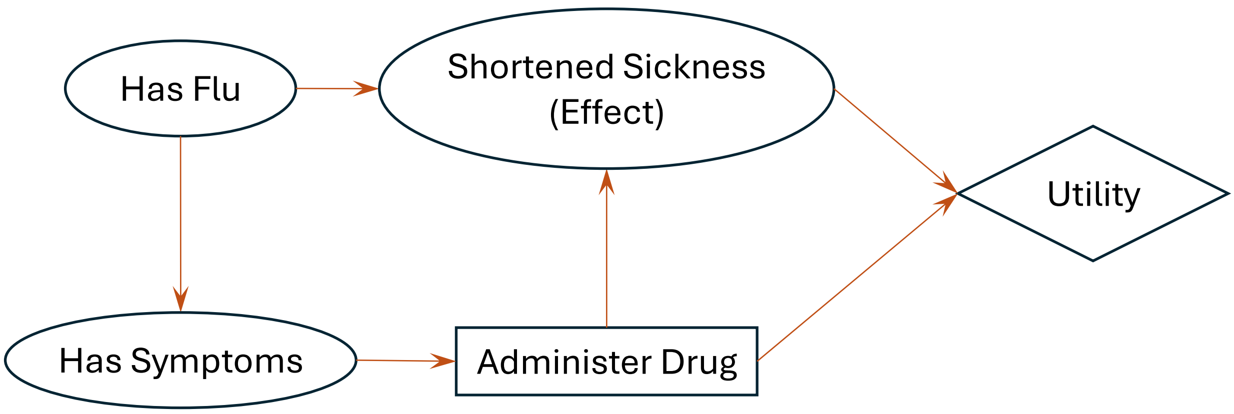
Answer

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

2. (b) Draw the decision network for this problem.

(8 marks)

Answer



2. (c) Compute the expected utility of administering the drug and of not administering the drug, if the person appears to have standard symptoms of the flu. Note that you may need to apply Bayes' theorem: $$\mathrm{P}(\mathrm{A} \mid \mathrm{B}) = \frac{\mathrm{P}(\mathrm{B} \mid \mathrm{A})\,\mathrm{P}(\mathrm{A})}{\mathrm{P}(\mathrm{B})}$$

(12 marks)

Answer

- Pprobabilities:
  - P(Flu) = 1 /1000 = 0.001
  - P(¬Flu) = 1- 0.001 = 0.999
  - P(Symptoms | Flu) = 0.9
  - P(Symptoms | ¬Flu) = 0.05

- Applying Bayes' theorem: $$\mathrm{P}(\mathrm{A} \mid \mathrm{B}) = \frac{\mathrm{P}(\mathrm{B} \mid \mathrm{A})\,\mathrm{P}(\mathrm{A})}{\mathrm{P}(\mathrm{B})}$$

$$\mathrm{P}(\mathrm{Flu} \mid \mathrm{Symptoms}) = \frac{\mathrm{P}(\mathrm{Symptoms} \mid \mathrm{Flu}) * \mathrm{P}(\mathrm{Flu})}{\mathrm{P}(\mathrm{Symptoms})}$$

Probability rule: P(A) = P(A | B) * P(B) + P(A | ¬B) * P(¬B)

P(Symptoms) = P(Symptoms | Flu) · P(Flu) + P(Symptoms | ¬Flu) · P(¬Flu)
= (0.9 * 0.001) + (0.05 * 0.999)
= 0.05085

$$\mathrm{P}(\mathrm{Flu} \mid \mathrm{Symptoms}) = \frac{\mathrm{P}(\mathrm{Symptoms} \mid \mathrm{Flu}) * \mathrm{P}(\mathrm{Flu})}{\mathrm{P}(\mathrm{Symptoms})} = \frac{0.9 * 0.001}{0.05085} = 0.0177$$

P(¬Flu | Symptoms) = 1 – 0.0177 = 0.9823

- Given:
  - U(Drug) = -100
  - U(¬Drug) = 0
  - U(Shortened) = 1000
  - U(¬Shortened) = 0
  - U(¬ Effect) = 0
  - P(Shortened) = 0.6
  - P(¬Shortened) = 0.4

U(Effect) = P(Shortened) * U(Shortened) + P(¬Shortened) * U(¬Shortened)
= (0.6 * 1000) + (0.4 * 0)
= 600

Expected utility of administering the drug:
EU(Drug) = P(Flu | Symptoms) * U(Effect) + P(¬Flu | Symptoms) * U(¬ Effect) + U(Drug)
= (0.0177 * 600) + (0.9823 * 0) + (-100)
= 10.62 - 100
= -89.38

Expected utility of not administering the drug:
EU(¬Drug) = U(¬Drug) = 0

2. (d) Should the doctor administer the drug to the person?

(2 marks)

Answer

Summary of EU

- EU(Drug) = -89.38
- EU(¬Drug) = 0
- Since EU(¬Drug) = 0 > EU(Drug) = -89.38, the doctor should not administer the drug

3. (a) In 2019, Google released the Google Research Football Environment where agents aim to master the world's most popular sport - football. Modelled after popular football video games, the Football Environment provides a physics based 3D football simulation where agents control either one or all football players on their team, learn how to pass between them, and manage to overcome their opponent's defense in order to score goals. Lots of researchers are still working on the problem. List and briefly explain a few key ideas in multi-agent systems which can be used to build a successful team of agents for football.

(6 marks)

Answer

i. Coordination
   - Agents must work together to achieve a common goal (e.g. scoring goals or defending), synchronized movement.

ii. Communication
   - Agents need effective communication to share information (e.g., teammates positions, ball control), decides who to pass, or shoot.

iii. Role Assignment
   - Agents adopt roles (e.g., striker, defender) dynamically based on game state.
   - A defender agent may switch to a midfield role if the team loses possession.

iv. Decentralized Decision Making
   - Agents make decisions independently, yet still align with the team strategy, to maximise the chances of winning.

v. Opponent Modelling
   - Agents must predict the behaviour or strategies of opposing team, develop counter-strategies.

vi. Learning and Adaptation
   - Agents can use reinforcement learning or other learning methods to improve their performance over time, such as when to pass, shoot, or reposition, based on success in past games.

3. (b) The Contract Net protocol is the most widely studied protocol in cooperative distributed problem solving. Briefly explain sub-contracting in Contract Net and its benefit.

(5 marks)

Answer

- In the Contract Net Protocol (CNP), a manager agent announces a task.

- Bidders (other agents) submit proposals.

- The manager selects the best proposal, assigning the task to a contractor.

- If the contractor cannot complete the task alone or more efficiently by delegation, it can sub-contract the task using the same CNP mechanism, becoming a manager for that subtask.

- Benefit:
  - Large or complex tasks can be distributed efficiently across multiple or specialized agents, preventing bottlenecks and reducing total execution time.

3. (c) Briefly explain whether we can apply Vickrey auctions to auctions for bundles of goods (i.e., combinatorial auctions).

(5 marks)

Answer

- A Vickrey auction is a sealed-bid auction where the highest bidder wins but pays the second-highest bid.

- In combinatorial auctions, bidders place bids on bundles of goods, not just individual items.

- Yes, Vickrey's principles can apply to combinatorial auctions via the Vickrey-Clarke-Grove (VCG) mechanism

- The VCG mechanism:
  - Each bidder submits their valuation for any combination of items.
  - The auctioneer determines an allocation maximizes total social welfare.
  - Each winning bidder pays the cost to others by their presence in the auction.
  - It is incentive compatible for the same reason as the Vickrey auctions.
  - Bidders are incentivized to reveal their true valuations for bundles.
  - This maximizes total welfare across all bidders.

3. (d) Briefly explain the difference between core and Shapley value.

(5 marks)

Answer

- The Core and the Shapley Value are both solution concepts in cooperative game theory but they differ in their approach and the type of solution they provide.

- There is no direct relationship between them.

- Differences:
  - Stability vs. Fairness:
    - The core focuses on stability, ensuring no coalition has an incentive to deviate.
    - The Shapley value focuses on fairness, distributing payoffs based on contributions.

  - Multiplicity vs. Uniqueness:
    - The core can contain multiple allocations or be empty.
    - The Shapley value provides a unique allocation for each player.

3. (e) Briefly explain the difference between plurality voting procedure and Borda count voting procedure.

(5 marks)

Answer

- Differences:
  - Voting Method:
    - Plurality Voting: Voters select one candidate.
    - Borda Count: Voters rank all candidates.

  - Winner Determination:
    - Plurality Voting: The candidate with the most votes wins.
    - Borda Count: The candidate with the highest total points wins.

  - Consideration of Preferences:
    - Plurality Voting: Only the top preference of each voter is considered.
    - Borda Count: All preferences are considered, providing a more comprehensive view of voter preferences.

Plurality Voting Example:
Candidate A gets 40 votes
Candidate B gets 35 votes
Candidate C gets 25 votes
Candidate A wins,
despite 60% preferring others

Borda Voting Example:
(1^{st} place 3 pt, 2^{nd} place 2 pt, 3^{rd} place 1 pt)
Voter 1: A > B > C
Voter 2: B > C > A
Voter 3: C > B > A
A gets 5 pt, B gets 7 pt, C gets 6 pt
Candidate B wins

4. Consider the payoff matrix in Table Q4. 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**.

Table Q4

	B: left	**B: middle**	**B: right**
A: up	**1, -1**	**-1, 0**	**2, 0**
A: middle	**3, 1**	**0, 2**	**1, 1**
A: down	**2, 2**	**-1, 1**	**1, -1**

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

(5 marks)

Answer

- Player A (row player):
  - If B plays left, A: up = 1, mid = 3, down = 2 → middle is best
  - If B plays middle, A: up = -1, mid = 0, down = -1 → middle is best
  - If B plays right, A: up = 2, mid = 1, down = 1 → up is best
  - Player A has no dominant strategy

- Player B (column player):
  - If A plays up, B: left = -1, mid = 0, right = 0 → mid or right is best
  - If A plays middle, B: left = 1, mid = 2, right = 1 → middle is best
  - If A plays down, B: left = 2, mid = 1, right = -1 → left is best
  - Player B has no dominant strategy

- 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 the payoff matrix 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.

	B: left	B: middle	B: right
A: up	1, -1	-1, 0	2, 0
A: middle	3, 1	0, 2	1, 1
A: down	2, 2	-1, 1	1, -1

1. (A: up, B: left) → (1, -1)
- A against B: left → up = 1, mid = 3, down = 2 → A prefers middle
- B against A: up → left = -1, mid = 0, right = 0 → B prefers mid or right
- Not Nash equilibrium

2. (A: up, B: middle) → (-1, 0)
- A against B: middle → up = -1, mid = 0, down = -1 → A prefers middle
- B against A: up → left = -1, mid = 0, right = 0 → B prefers mid or right
- Not Nash equilibrium

3. (A: up, B: right) → (2, 0)
- A against B: right → up = 2, mid = 1, down = 1 → A prefers up
- B against A: up → left = -1, mid = 0, right = 0 → B prefers mid or right
- Nash equilibrium

4. (A: middle, B: left) → (3, 1)
- A against B: left → up = 1, mid = 3, down = 2 → A prefers middle
- B against A: middle → left = 1, mid = 2, right = 1 → B prefers middle
- Not Nash equilibrium

5. (A: middle, B: middle) → (0, 2)
- A against B: middle → up = -1, mid = 0, down = -1 → A prefers middle
- B against A: middle → left = 1, mid = 2, right = 1 → B prefers middle
- Nash equilibrium

6. (A: middle, B: right) → (1, 1)
- A against B: right → up = 2, mid = 1, down = 1 → A prefers left
- B against A: middle → left = 1, mid = 2, right = 1 → B prefers middle
- Not Nash equilibrium

4. (b) cont

Answer

	B: left	B: middle	B: right
A: up	**1, -1**	**-1, 0**	**2, 0**
A: middle	**3, 1**	**0, 2**	**1, 1**
A: down	**2, 2**	**-1, 1**	**1, -1**

7. (A: down, B: left) → (2, 2)

- A against B: left → up = 1, mid = 3, down = 2 → A prefers middle
- B against A: down → left = 2, mid = 1, right = -1 → B prefers left
- Not Nash equilibrium

8. (A: down, B: middle) → (-1, 1)

- A against B: middle → up = -1, mid = 0, down = -1 → A prefers middle
- B against A: down → left = 2, mid = 1, right = -1 → B prefers left
- Not Nash equilibrium

9. (A: down, B: right) → (1, -1)

- A against B: right → up = 2, mid = 1, down = 1 → A prefers left
- B against A: down → left = 2, mid = 1, right = -1 → B prefers left
- Not Nash equilibrium

<u>Summary of Nash Equilibrium</u>

- (A: up, B: right) → (2, 0)
- (A: middle, B: middle) → (0, 2)
- Neither player has an incentive to change their strategy

4. (c) Identify which outcomes in the payoff matrix 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.

	B: left	**B: middle**	**B: right**
A: up	**1, -1**	**-1, 0**	**2, 0**
A: middle	**3, 1**	**0, 2**	**1, 1**
A: down	**2, 2**	**-1, 1**	**1, -1**

1. (A: up, B: left) → (1, -1)
   - Dominated by (2, 0) & others
   - Not Pareto optimal

2. (A: up, B: middle) → (-1, 0)
   - Dominated by (3, 1) & others
   - Not Pareto optimal

3. (A: up, B: right) → (2, 0)
   - Dominated by (3, 1) & others
   - Not Pareto optimal

4. (A: middle, B: left) → (3, 1)
   - No other better domination
   - Pareto optimal

5. (A: middle, B: middle) → (0, 2)
   - Dominated by (2, 2)
   - Not Pareto optimal

4. (c) cont

Answer

	B: left	B: middle	B: right
A: up	**1, -1**	**-1, 0**	**2, 0**
A: middle	**3, 1**	**0, 2**	**1, 1**
A: down	**2, 2**	**-1, 1**	**1, -1**

6. (A: middle, B: right) → (1, 1)
   - Dominated by (3, 1) & others
   - Not Pareto optimal

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

8. (A: down, B: middle) → (-1, 1)
   - Dominated by (0, 2) & others
   - Not Pareto optimal

9. (A: down, B: right) → (1, -1)
   - Dominated by (2, 0) & others
   - Not Pareto optimal

Summary of Pareto optimal

- (A: middle, B: left) → (3, 1)
- (A: down, B: left) → (2, 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.

	B: left	**B: middle**	**B: right**
A: up	**1, -1**	**-1, 0**	**2, 0**
A: middle	**3, 1**	**0, 2**	**1, 1**
A: down	**2, 2**	**-1, 1**	**1, -1**

- (A: up, B: left): Total payoff = 1 + (-1) = 0
- (A: up, B: middle): Total payoff = -1 + 0 = -1
- (A: up, B: right): Total payoff = 2 + 0 = 2
- (A: middle, B: left): Total payoff = 3 + 1 = 4
- (A: middle, B: middle): Total payoff = 0 + 2 = 2
- (A: middle, B: right): Total payoff = 1 + 1 = 2
- (A: down, B: left): Total payoff = 2 + 2 = 4
- (A: down, B: middle): Total payoff = -1 + 1 = 0
- (A: down, B: right): Total payoff = 1 + (-1) = 0

- Maximize social welfare = 4

Summary of maximize social welfare

- (A: middle, B: left) → (3, 1)
- (A: down, B: left) → (2, 2)
- Both outcomes achieve the highest total payoff maximize social welfare.