1. (a) Explain why an intelligent agent should be *reactive*, *pro-active* and *social*.

(6 marks)

Answer

- 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.

- Pro-active
  - An intelligent agent should be pro-active by taking initiative to achieve its goals.

  - This leads to more efficient and goal-driven behaviour.

- 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.

1. (b) 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

1. (c) Give two benefits of reactive architectures.

(2 marks)

Answer

- Fast response time
- Simplicity

1. (d) Consider the environment $Env_1 = \langle E, e_0, \tau \rangle$ defined as follows:

$E = \{e_0, e_1, e_2, e_3, e_4, e_5, e_6\}$
$\tau(e_0, \alpha_0) = \{e_1, e_2\}$
$\tau(e_0, \alpha_1) = \{e_2, e_3\}$
$\tau(e_1, \alpha_2) = \{e_3\}$
$\tau(e_2, \alpha_3) = \{e_4, e_5\}$
$\tau(e_3, \alpha_4) = \{e_2, e_6\}$

$\tau(e_i, \alpha_i)$ defines the state transition for the environment in state e_i, given action α_i. Assume that there are two possible agents for this environment Ag_1 and Ag_2, which are defined as:

$Ag_1(e_0) = \alpha_0$
$Ag_1(e_1) = \alpha_2$
$Ag_1(e_2) = \alpha_3$

$Ag_2(e_0) = \alpha_1$
$Ag_2(e_2) = \alpha_3$
$Ag_2(e_3) = \alpha_4$

Assume $|r|$ gives the total number of states in a particular run r. For example, $|r_1| = 3$ if $r_1 = (e_0, \alpha_0, e_3, \alpha_2, e_5 \mid Ag_1, Env_1)$. The probability P and utility U of each run are given as follows:

$P(r \mid Ag_1, Env_1) = 4/(|r| + |r|)$
$U(r) = 10 - |r|$
$P(r \mid Ag_2, Env_1) = |r|/9$

(i) Write down all possible runs for Ag_1 and Ag_2.

(6 marks)

Answer

*Ag*1

- $Ag_1(e_0) = \alpha_0 \rightarrow \tau(e_0, \alpha_0) = \{e_1, e_2\}$
- $Ag_1(e_1) = \alpha_2 \rightarrow \tau(e_1, \alpha_2) = \{e_3\}$
- $Ag_1(e_2) = \alpha_3 \rightarrow \tau(e_2, \alpha_3) = \{e_4, e_5\}$

$r_1 = (e_0, \alpha_0, e_1, \alpha_2, e_3)$
$r_2 = (e_0, \alpha_0, e_2, \alpha_3, e_4)$
$r_3 = (e_0, \alpha_0, e_2, \alpha_3, e_5)$

*Ag*2

- $Ag_2(e_0) = \alpha_1 \rightarrow \tau(e_0, \alpha_1) = \{e_2, e_3\}$
- $Ag_2(e_2) = \alpha_3 \rightarrow \tau(e_2, \alpha_3) = \{e_4, e_5\}$
- $Ag_2(e_3) = \alpha_4 \rightarrow \tau(e_3, \alpha_4) = \{e_2, e_6\}$

$r_4 = (e_0, \alpha_1, e_2, \alpha_3, e_4)$
$r_5 = (e_0, \alpha_1, e_2, \alpha_3, e_5)$
$r_6 = (e_0, \alpha_1, e_3, \alpha_4, e_2, \alpha_3, e_4)$
$r_7 = (e_0, \alpha_1, e_3, \alpha_4, e_2, \alpha_3, e_5)$
$r_8 = (e_0, \alpha_1, e_3, \alpha_4, e_6)$

1. (d) cont

(ii) Calculate the expected utility for Agi and Ag2. Which one of the two agents is optimal with respect to Env_1 and U?

(6 marks)

Answer

$P(r \mid Ag_1, Env_1) = 4/(|r| + |r|)$ $U(r) = 10 - |r|$

$P(r \mid Ag_2, Env_1) = |r|/9$

- $|r_1|$ = 3 states
- $|r_2|$ = 3 states
- $|r_3|$ = 3 states

- $|r_4|$ = 3 states
- $|r_5|$ = 3 states
- $|r_6|$ = 4 states
- $|r_7|$ = 4 states
- $|r_8|$ = 3 states

$EU(Ag_1) = U(r) * P(r \mid Ag_1, Env_1)$

$$= \left[(10 - |r_1|) * \left(\frac{4}{|r_1| + |r_1|}\right)\right] + \left[(10 - |r_2|) * \left(\frac{4}{|r_2| + |r_2|}\right)\right] + \left[(10 - |r_3|) * \left(\frac{4}{|r_3| + |r_3|}\right)\right]$$

$$= \left[(10 - 3) * \left(\frac{4}{3 + 3}\right)\right] + \left[(10 - 3) * \left(\frac{4}{3 + 3}\right)\right] + \left[(10 - 3) * \left(\frac{4}{3 + 3}\right)\right]$$

$$= \left(7 * \frac{2}{3}\right) + \left(7 * \frac{2}{3}\right) + \left(7 * \frac{2}{3}\right)$$

$$= 14$$

$EU(Ag_2)$

$= U(r) * P(r \mid Ag_2, Env_1)$

$$= \left[(10 - |r_4|) * \left(\frac{|r_4|}{9}\right)\right] + \left[(10 - |r_5|) * \left(\frac{|r_5|}{9}\right)\right] + \left[(10 - |r_6|) * \left(\frac{|r_6|}{9}\right)\right] + \left[(10 - |r_7|) * \left(\frac{|r_7|}{9}\right)\right] + \left[(10 - |r_8|) * \left(\frac{|r_8|}{9}\right)\right]$$

$$= \left[(10 - 3) * \left(\frac{3}{9}\right)\right] + \left[(10 - 3) * \left(\frac{3}{9}\right)\right] + \left[(10 - 4) * \left(\frac{4}{9}\right)\right] + \left[(10 - 4) * \left(\frac{4}{9}\right)\right] + \left[(10 - 3) * \left(\frac{3}{9}\right)\right]$$

$$= \left(7 * \frac{1}{3}\right) + \left(7 * \frac{1}{3}\right) + \left(6 * \frac{4}{9}\right) + \left(6 * \frac{4}{9}\right) + \left(7 * \frac{1}{3}\right)$$

$$= \frac{37}{3}$$

$$= 12.33$$

Summary of expected utility

$EU(Ag_1) = 14$

$EU(Ag_2) = 12.33$

Since $EU(Ag_1) = 14 > EU(Ag_2) = 12.33$, Ag_1 is the optimal agent.

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)

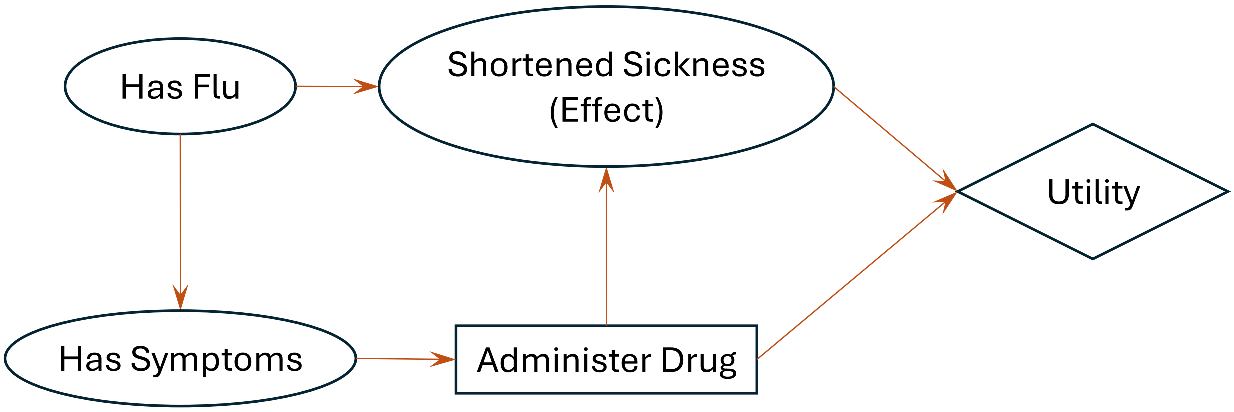
Answer

- 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:

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

(12 marks)

Answer

- Given probabilities:
  - 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:

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

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

- 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

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

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

- Given:
  - U(Drug) = -100
  - U(¬Drug) = 0
  - U(Shortened) = 1000
  - U(¬Shortened) = 0
  - U(No 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(No 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

<u>Summary of EU</u>

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

3. (a) Use a real world example to explain the need to design a coordination or cooperation protocol while building a multi-agent system.

(5 marks)

Answer

- Example: A fleet of autonomous delivery drones delivering packages.

- Need for coordination and cooperation protocol
  - Without proper coordination, multiple drones might choose the same flight path, leading to collisions or delays.

  - If one drone's battery is low, by cooperation, it could request help from a nearby drone to take over its delivery task.

  - A coordination protocol ensures that drones share information, such as not attempt to deliver the same order, wasting resources.

  - It also allows prioritization of urgent deliveries and optimizes drone usage to avoid idle time.

- Designing a coordination protocol ensures smooth, conflict-free operations in shared environments.

3. (b) The CONTRACT NET protocol is one of the most widely used cooperation protocol in the multi-agent systems world. Briefly describe the operation of the 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 recognises it has a problem it wants help with.

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
   - Agents that receive the announcement decide for themselves whether they wish to bid for the task

4. Awarding
   - Agent that sent task announcement must choose between bids & decide who to "award the contract" to.

5. Expediting
   - The successful contractor then expedites the task

3. (c) Explain why "incentive compatibility" is an important property of an auction mechanism.

(5 marks)

Answer

- Reasons
  - When an auction is incentive compatible, agents are motivated to be honest, which leads to more efficient and fair outcomes.
  - Without incentive compatibility, bidders might lie about their valuations to gain an unfair advantage, which can lead to suboptimal allocation of resources.
  - Agents do not need to strategize about others' bids. They can simply bid truthfully, which reduces complexity and overhead in multi-agent systems.
  - Trust in the auction process increases when participants know that being honest is the best strategy for everyone.
  - Truthful information helps the auctioneer allocate resources to those who value them the most, maximizing overall system efficiency or social welfare.

3. (d) Explain Pareto property in voting theory.

(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) Consider the coalitional game with agents Ag={1,2} and characteristic function v defined by v({1})=5, v({2})=5, v({1,2})=20. With reference to this example, explain the meaning of the core of a coalitional game.

(5 marks)

Answer

- Given Ag = {1,2}
  - v({1}) = 5
  - v({2}) = 5
  - v({1, 2}) = 20

- The core of a coalitional game is the set of payoff allocations to agents such that:
  - Efficiency (Group Rationality)
    - $x_1 + x_2 = v(1, 2) = 20$ ------------------ (1)

  - Individual Rationality
    - $x_1 \geq v(1) = 5$ ------------------ (2)
    - $x_2 \geq v(2) = 5$ ------------------ (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), $x_2 = 20 - x_1$ ----------------- (4)
  - (3) into (4), $5 = 20 - x_1$
    $x_1 = 20 - 5 = 15$
    Allocation $5 \leq x_1 \leq 15$
  - (2) into (4), $x_2 = 20 - 5$
    $x_2 = 15$
    Allocation $5 \leq x_2 \leq 15$

  - Some possible distributions that satisfy these conditions are:
    - $x_1 = 5$ and $x_2 = 15$
    - $x_1 = 10$ and $x_2 = 10$
    - $x_1 = 15$ and $x_2 = 5$

<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	**(2, 6)**	**(1, 5)**
i cooperate	**(3, 4)**	**(2, 4)**

Payoff matrix B:

Table Q4b

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

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	**(2, 6)**	**(1, 5)**
i cooperate	**(3, 4)**	**(2, 4)**

Table Q4b

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

Payoff matrix A (Table Q4a)

- Player i (row player):
  - If j defect, i: defect = 2, cooperate = 3 → cooperate is best
  - If j cooperate, i: defect (1) vs cooperate (2) → cooperate is best
  - Player i has a dominant strategy → cooperate

- Player j (column player):
  - If i defect, j: defect = 6, cooperate = 5 → defect is best
  - If i cooperate, j: defect = 4, cooperate = 4 → defect or coop same
  - Player j has a dominant strategy → defect

Payoff matrix B (Table Q4b)

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

- Player j (Column Player):
  - If i defect, j: defect = 3, cooperate = 4 → cooperate is best
  - If i cooperate, j: defect = 5, cooperate = 4 → defect is best
  - Player j has a no dominant strategy

Summary

Matrix A: Player i has a dominant strategy → cooperate
Player j has a dominant strategy → defect
Dominant strategy pair: (i cooperate, j defect) → (3, 4)

Matrix B: Neither player has dominant strategy
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.

Payoff matrix A (Table Q4a)

Table Q4a

	j defect	***j cooperate***
i defect	**(2, 6)**	**(1, 5)**
i cooperate	**(3, 4)**	**(2, 4)**

• (i defect, j defect) • i: defect = 2, coop = 3 → i prefers coop • j: defect = 6, coop = 4 → j prefers defect • Not Nash equilibrium	• (i defect, j cooperate) • i: defect = 1, coop = 2 → i prefers coop • j: defect = 6, coop = 4 → j prefers defect • Not Nash equilibrium
• (i cooperate, j defect) • i: defect = 2, coop = 3 → i prefers coop • j: defect = 4, coop = 4 → defect or coop • Nash equilibrium	• (i cooperate, j cooperate) • i: defect = 1, coop = 2 → i prefers coop • j: defect = 4, coop = 4 → defect or coop • Nash equilibrium

Payoff matrix b (Table Q4b)

Table Q4b

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

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

Summary of Nash Equilibrium

Payoff matrix A: (i cooperate, j defect) → (3, 4)

(i cooperate, j cooperate) → (2, 4)

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

(i cooperate, j defect) → (3, 5)

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	**(2, 6)**	**(1, 5)**
i cooperate	**(3, 4)**	**(2, 4)**

Payoff matrix A (Table Q4a)

- (i defect, j defect) → (2, 6)
  - No other better domination
  - Pareto optimal

- (i defect, j cooperate) → (1, 5)
  - Dominated by (2, 6)
  - Not Pareto optimal

- (i cooperate, j defect) → (3, 4)
  - No other better domination

- Pareto optimal (i cooperate, j cooperate) → (2, 4)
  - Dominated by (2, 6)
  - Not Pareto optimal
  - Pareto optimal

4. (c) cont.

Table Q4b

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

Payoff matrix B (Table Q4b)

- (i defect, j defect) → (2, 3)
  - Dominated by (3, 4) & others
  - Not Pareto optimal

- (i defect, j cooperate) → (3, 4)
  - Dominated by (3, 5)
  - Not Pareto optimal

- (i cooperate, j defect) → (3, 5)
  - No other better domination
  - Pareto optimal

- (i cooperate, j cooperate) → (2, 4)
  - Dominated by (3, 4) & others
  - Not Pareto optimal

Summary of Pareto optimal

Payoff matrix A: (i defect, j defect) → (2, 6),
(i cooperate, j defect) → (3, 4)

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

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	**(2, 6)**	**(1, 5)**
i cooperate	**(3, 4)**	**(2, 4)**

Table Q4b

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

Payoff matrix A (Table Q4a)

- (i defect, j defect): Total payoff = 2 + 6 = 8
- (i defect, j cooperate): Total payoff = 1 + 5 = 7
- (i cooperate, j defect): Total payoff = 3 + 4 = 7
- (i cooperate, j cooperate): Total payoff = 2 + 4 = 6

- Maximize social welfare = 8
- Outcome: (i defect, j defect) → (2, 6)

Payoff matrix B (Table Q4b)

- (i defect, j defect): Total payoff = 2 + 3 = 5
- (i defect, j cooperate): Total payoff = 3 + 4 = 7
- (i cooperate, j defect): Total payoff = 3 + 5 = 8
- (i cooperate, j cooperate): Total payoff = 2 + 4 = 6

- Maximize social welfare = 8
- Outcome: (i cooperate, j defect) → (3, 5)

Summary of maximize social welfare

Payoff matrix A: (i defect, j defect) → (2, 6)

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

Their outcomes achieve the highest total payoff maximizes social welfare.