SC4003 Exam 2021-2022 Semester 2

1. (a) What are the three types of nodes in a decision network?

(3 marks)

Answer

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

1. (b) Two most common types of tasks for agents to perform are *achievement tasks* and *maintenance tasks*. Explain what these two types of tasks are.

(4 marks)

Answer

- Achievement Tasks
  - These tasks have a specific goal or desired state that the agent must achieve.
  - Once the goal is reached, the task is considered complete.

- Maintenance Tasks
  - These tasks involve keeping certain conditions or states true over time, rather than achieving a one-time goal.
  - The agent must act continuously or repeatedly to maintain stability or performance.

1. (c) Environments can be characterized as *static* vs. *dynamic*. Describe what these environment properties are. In addition, indicate the key property that an intelligent agent needs to achieve in a dynamic environment.

(6 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. (d) Consider the environment Env1 = $<E, e_0, \tau>$ 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:

$Ag_1(e_0) = \alpha_0$ $Ag_2(e_0) = \alpha_1$
$Ag_1(e_1) = \alpha_2$ $Ag_2(e_2) = \alpha_3$
$Ag_1(e_2) = \alpha_3$ $Ag_2(e_3) = \alpha_4$

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. An agent inhabits a world with two states, *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 Markov Decision Process (MDP). The reward of each state is as follows: *R(S) = 3*, and *R(S') = 2*.

Set 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 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:

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

    - For S':

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

  - Policy Improvement for π_0:
    - For S:

$Q(S, a) = R(S) + \gamma[U^{\pi_0}(S)] = 3 + 0.5 * 6 = 6$ ← no change

$Q(S, b) = R(S) + \gamma[U^{\pi_0}(S')] = 3 + 0.5 * 4 = 5$ ← flips

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

    - For S':

$Q(S', a) = R(S') + \gamma[U^{\pi_0}(S')] = 2 + 0.5 * 4 = 4$ ← no change

$Q(S', b) = R(S') + \gamma[U^{\pi_0}(S)] = 2 + 0.5 * 6 = 5$ ← flips

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':

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

  - Policy Improvement for π_1:
    - For S:

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

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

    - For S':

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

      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) COVID-19 has disrupted supply chains around the world. Two years into the pandemic, the global supply chain continues to sputter and break down. Each day comes news of choked ports, out-of-place shipping containers, record freight rates, and other problems that cause disruption and defy easy answer. List and briefly explain a few key ideas in multi-agent systems which can be used to improve the currently disrupted global supply chain in a post-pandemic world.

(5 marks)

Answer

i. Coordination
- Agents can negotiate, share information, and collaborate to optimize logistics, like scheduling deliveries or load balancing across routes.

ii. Decentralization
- Multi-agent systems distribute control across agents instead of relying on a central authority.

iii. Autonomous Decision-Making
- Agents can make local decisions (e.g., re-routing shipments) based on real-time information to reduce delays.

iv. Real-Time Monitoring and Adaptation
- MAS can easily adapt to changing supply chain conditions, such as surges in demand or supply shortages, by re-allocating tasks among agents.

v. Scalability and Flexibility
- Agents can sense disruptions (e.g., port congestion) and adapt routes or schedules on the fly, improving resilience and reducing downtime.

3. (b) The Contract Net protocol is the most widely studied protocol in cooperative distributed problem solving. Describe the main steps of the Contract Net.

(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) Briefly explain why Vickrey auctions are truthful.

(5 marks)

Answer

- A Vickrey auction is a type of sealed-bid second-price auction where:
  - Each bidder submits a bid without knowing others' bids.
  - The highest bidder wins, but pays the second-highest bid price.

- Why it is truthful:
  - In a Vickrey auction, 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.

3. (d) Coalitional games model scenarios where agents can benefit from cooperation. When agents join coalitions, a coalition structure is formed. Explain the concept of stability and fairness of a coalition structure.

(5 marks)

Answer

- In coalitional games, agents form groups (coalitions) to cooperate and share the total reward or utility.

- Stability
  - A coalition structure is stable if no subset of agents has an incentive to leave and form a new coalition.
  - Stability ensures that no group of agents can improve their outcome by deviating.

- Fairness
  - Fairness refers to how the total payoff is distributed among the agents in a way that is considered just or equitable.
  - A well-known fairness concept is the Shapley value, which distributes payoffs based on each agent's marginal contribution to all possible coalitions.

- Summary:
  - Stability prevents agents from leaving coalitions.
  - Fairness ensures rewards are divided in a way that reflects each agent's contribution.

3. (e) Voting is widely used to make group decisions and voting is vulnerable to manipulation. It is known that we can use complexity to avoid vote manipulation. Briefly explain the idea.

(5 marks)

Answer

- In group decision-making, voting systems can be manipulated if voters misrepresent their preferences to get a more favourable outcome.

- Computational Complexity:
    - The idea is to design voting protocols where determining a beneficial manipulation is computationally difficult.
    - By making the manipulation process complex and time-consuming, it discourages voters from attempting to manipulate the outcome.
    - In some voting rules (e.g., single transferable vote), finding a manipulative vote sequence is NP-hard.
    - By ensuring that manipulation is computationally infeasible, voting systems can preserve fairness without relying solely on trust or detection.

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, -4)**	**(-3, -2)**
A: down	**(1, 1)**	**(-2, -4)**

Payoff matrix 2:

Table Q4b

	B: left	**B: right**
A: up	**(4, 0)**	**(0, 6)**
A: middle	**(2, 1)**	**(4, 1)**
A: down	**(1, 4)**	**(3, 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, -4)	(-3, -2)
A: down	(1, 1)	(-2, -4)

Table Q4b

	B: left	B: right
A: up	(4, 0)	(0, 6)
A: middle	(2, 1)	(4, 1)
A: down	(1, 4)	(3, 1)

Payoff matrix 1 (Table Q4a)

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

- Player B (column player):
  - If A plays up, B: left = -4, right = -2 → right is best
  - If A plays down, B: left = 1, right = -4 → left is best
  - Player B has no dominant strategy

Payoff matrix 2 (Table Q4b)

- Player A (row player):
  - If B plays left, A: up = 4, mid = 2, down = 1 → up is best
  - If B plays right, A: up = 0, mid = 4, down = 3 → mid is best
  - Player A has no dominant strategy

- Player B (column player):
  - If A plays up, B: left = 0, right = 6 → right is best
  - If A plays middle, B: left = 1, right = 1 → left or right same
  - If A plays down, B: left = 4, right = 1 → left is best
  - Player B has no dominant strategy

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.

Payoff matrix 1 (Table Q4a)

Table Q4a

	B: left	**B: right**
A: up	**(2, -4)**	**(-3, -2)**
A: down	**(1, 1)**	**(-2, -4)**

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

Payoff matrix 2 (Table Q4b)

Table Q4b

	B: left	**B: right**
A: up	**(4, 0)**	**(0, 6)**
A: middle	**(2, 1)**	**(4, 1)**
A: down	**(1, 4)**	**(3, 1)**

• (A: up, B: left) • A: up = 4, mid = 2, down = 1 → A prefers up • B: left = 0, right = 6 → B prefers right • Not Nash equilibrium	• (A: up, B: right) • A: up = 0, mid = 4, down = 3 → A prefers mid • B: left = 0, right = 6 → B prefers right • Not Nash equilibrium
• (A: middle, B: left) • A: up = 4, mid = 2, down = 1 → A prefers up • B: left = 1, right = 1 → left or right same • Not Nash equilibrium	• (A: middle, B: right) • A: up = 0, mid = 4, down = 3 → A prefers mid • B: left = 1, right = 1 → left or right same • Nash equilibrium
• (A: down, B: left) • A: up = 4, mid = 2, down = 1 → A prefers up • B: left = 4, right = 1 → B prefers left • Not Nash equilibrium	• (A: down, B: right) • A: up = 0, mid = 4, down = 3 → A prefers mid • B: left = 4, right = 1 → B prefers left • Not Nash equilibrium

Summary of Nash Equilibrium

Payoff matrix 1: No Nash equilibrium

Payoff matrix 2: (A: middle, B: right) → (4, 1)

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

	B: left	**B: right**
A: up	**(2, -4)**	**(-3, -2)**
A: down	**(1, 1)**	**(-2, -4)**

Payoff matrix 1 (Table Q4a)

- (A: up, B:left) → (2, -4)
  - No other better domination
  - Pareto optimal

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

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

- (A: down, B:right) → (-2, -4)
  - Dominated by (1, 1)
  - Not Pareto optimal

4. (c) cont.

Table Q4b

	B: left	B: right
A: up	(4, 0)	(0, 6)
A: middle	(2, 1)	(4, 1)
A: down	(1, 4)	(3, 1)

Payoff matrix 2 (Table Q4b)

- (A: up, B:left) → (4, 0)
  - Dominated by (4, 1)
  - Not Pareto optimal

- (A: up, B:right) → (0, 6)
  - No other better domination
  - Pareto optimal

- (A: middle, B:left) → (2, 1)
  - Dominated by (4, 1)
  - Not Pareto optimal

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

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

- (A: down, B:right) → (3, 1)
  - Dominated by (4, 1)
  - Not Pareto optimal

Summary of Pareto optimal

Payoff matrix 1:	(A: up, B:left)	→ (2, -4),
	(A: down, B:left)	→ (1, 1)
Payoff matrix 2:	(A: up, B:right)	→ (0, 6),
	(A: middle, B:right)	→ (4, 1),
	(A: down, B:left)	→ (1, 4)

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, -4)**	**(-3, -2)**
A: down	**(1, 1)**	**(-2, -4)**

Table Q4b

	B: left	**B: right**
A: up	**(4, 0)**	**(0, 6)**
A: middle	**(2, 1)**	**(4, 1)**
A: down	**(1, 4)**	**(3, 1)**

Payoff matrix 1 (Table Q4a)

- (A: up, B: left): Total payoff = 2 + (-4) = -2
- (A: up, B: right): Total payoff = -3 + (-2) = -5
- (A: down, B: left): Total payoff = 1 + 1 = 2
- (A: down, B: right): Total payoff = -2 + (-4) = -6

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

Payoff matrix 2 (Table Q4b)

- (A: up, B: left): Total payoff = 4 + 0 = 4
- (A: up, B: right): Total payoff = 0 + 6 = 6
- (A: middle, B: left): Total payoff = 2 + 1 = 3
- (A: middle, B: right): Total payoff = 4 + 1 = 5
- (A: down, B: left): Total payoff = 1 + 4 = 5
- (A: down, B: right): Total payoff = 3 + 1 = 4

- Maximize social welfare = 6
- Outcome: (A: up, B: right) → (0, 6)

Summary of maximize social welfare

Payoff matrix 1: (A: down, B: left) → (1, 1)

Payoff matrix 2: (A: up, B: right) → (0, 6)

Their outcomes achieve the highest total payoff maximizes social welfare.