

**Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0**

**Correct Marks : 1**

Question Label : Multiple Choice Question

Total Revenue for the month of April is

**Options :**

6406532266704. ✓ \$33400

6406532266705. ✗ \$12900

6406532266706. ✗ \$53900

6406532266707. ✗ \$28700

**Question Number : 85 Question Id : 640653676858 Question Type : MCQ Is Question**

**Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0**

**Correct Marks : 1**

Question Label : Multiple Choice Question

Net profit for the month of April is:

**Options :**

6406532266708. ✓ \$5920

6406532266709. ✗ \$5275

6406532266710. ✗ \$5322

6406532266711. ✗ \$5800

**RL**

**Section Id :**

64065345317

**Section Number :**

5

<b>Section type :</b>	Online
<b>Mandatory or Optional :</b>	Mandatory
<b>Number of Questions :</b>	12
<b>Number of Questions to be attempted :</b>	12
<b>Section Marks :</b>	40
<b>Display Number Panel :</b>	Yes
<b>Section Negative Marks :</b>	0
<b>Group All Questions :</b>	No
<b>Enable Mark as Answered Mark for Review and Clear Response :</b>	Yes
<b>Maximum Instruction Time :</b>	0
<b>Sub-Section Number :</b>	1
<b>Sub-Section Id :</b>	64065396840
<b>Question Shuffling Allowed :</b>	No
<b>Is Section Default? :</b>	null

**Question Number : 86 Question Id : 640653676861 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0**

**Correct Marks : 0**

Question Label : Multiple Choice Question

**THIS IS QUESTION PAPER FOR THE SUBJECT "DEGREE LEVEL : REINFORCEMENT LEARNING (COMPUTER BASED EXAM)"**

**ARE YOU SURE YOU HAVE TO WRITE EXAM FOR THIS SUBJECT?**

**CROSS CHECK YOUR HALL TICKET TO CONFIRM THE SUBJECTS TO BE WRITTEN.**

**(IF IT IS NOT THE CORRECT SUBJECT, PLS CHECK THE SECTION AT THE TOP FOR THE SUBJECTS REGISTERED BY YOU)**

**Options :**

6406532266720. ✓ YES

6406532266721. ✖ NO

**Question Number : 87 Question Id : 640653676862 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0**

**Correct Marks : 0**

Question Label : Multiple Choice Question

Note:

For numerical answer type questions, always enter your answer correct upto two decimal places without rounding up or off.

**Options :**

6406532266722. ✔ Useful Data has been mentioned above.

6406532266723. ✖ This data attachment is just for a reference & not for an evaluation.

<b>Sub-Section Number :</b>	2
<b>Sub-Section Id :</b>	64065396841
<b>Question Shuffling Allowed :</b>	Yes
<b>Is Section Default? :</b>	null

**Question Number : 88 Question Id : 640653676863 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0**

**Correct Marks : 2 Max. Selectable Options : 0**

Question Label : Multiple Select Question

Select correct statements about UCB1 algorithm:

**Options :**

6406532266724. ✖ Each arm is selected atleast twice before any arm is selected third time.

6406532266725. ✔ (Each arm is selected atleast once before any arm is selected second time.)

6406532266726. ✔ (Constant c controls degree of exploration, higher its value, higher the exploration.)

6406532266727. ✖ Constant  $c$  controls degree of exploration, higher its value, lower the exploration.

Question Number : 89 Question Id : 640653676882 Question Type : MSQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2 Max. Selectable Options : 0

Question Label : Multiple Select Question

For the (value iteration algorithm) *directly from assignments* which of the following statements are correct?

Options :

6406532266768. ✖ For a state  $s$ , as soon as  $v_{\pi}(s)$  is updated,  $\pi(s)$  is also updated. ✖

6406532266769. ✓  $\pi(s) \forall s$ , is updated only once  $v_{\pi}(s) \forall s$  changes by less than  $\theta$  (i.e. a small number).

6406532266770. ✖ Terminal states are initialized with non-zero value.

6406532266771. ✓ Terminal states are initialized with zero value.

Sub-Section Number : 3

Sub-Section Id : 64065396842

Question Shuffling Allowed : Yes

Is Section Default? : null

Question Number : 90 Question Id : 640653676864 Question Type : MCQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Choose the correct statement(s) regarding explore-exploit dilemma:

Options :

6406532266728. ✓ Always exploring is not optimal.

6406532266729. ✗ Always exploiting is optimal.

6406532266730. ✗ None of these.

nothing is optimal

Question Number : 91 Question Id : 640653676870 Question Type : MCQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Which of the following is the correct Bellman equation for deterministic transitions? The symbols have the usual meaning.

Options :

6406532266739. ✗  $v_{\pi}(s) = \sum_{s'} \sum_r p(s', r | s, a) [r + \gamma v_{\pi}(s')]$

6406532266740. ✓  $v_{\pi}(s) = \sum_a \pi(a|s) [r + \gamma v_{\pi}(s')]$   $q_{\pi}(a|s)$

6406532266741. ✗  $v_{\pi}(s) = \sum_a \pi(a|s) \sum_r p(s', r | s, a) [r + \gamma q(s', a')]$

6406532266742. ✗  $v_{\pi}(s) = \sum_a \pi(a|s) \sum_{s'} [r + \gamma v_{\pi}(s')]$

Question Number : 92 Question Id : 640653676881 Question Type : MCQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Consider the following statements and select the correct option.

**Assertion:** Monte Carlo value function approximation methods do not need knowledge of the model to be implemented. ✓

**Reason:** Monte Carlo value function approximation methods require only a way to sample trajectories from the environment and aggregate the results.

**Options :**

6406532266764. ✓ Assertion and Reason are both true and Reason is a correct explanation of Assertion.

6406532266765. ✗ Assertion and Reason are both true and Reason is not a correct explanation of Assertion.

6406532266766. ✗ Assertion is true but Reason is false.

6406532266767. ✗ Assertion is false but Reason is true.

Sub-Section Number :	4
Sub-Section Id :	64065396843
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653676865 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0  
Question Numbers : (93 to 94)  
Question Label : Comprehension

The following table captures the arms pulled and corresponding rewards, with increasing timestamps. The bandit has only 3 arms.

Softmax policy is used to select an arm. Assume that the reward distribution is stationary and temperature parameter  $\tau = 1$ .

Timestamp ( $t$ )	Arm ( $A_i$ )	Reward ( $r_i$ )
0	$A_1$	$\ln(2)$ 0.69
1	$A_2$	$\ln(1)$ 0
2	$A_3$	$\ln(4)$ 1.39
3	$A_3$	$\ln(16)$ 2.77
4	$A_2$	$\ln(1)$ 0
5	$A_1$	$\ln(2)$ 0.69

Aug. Reward

$$A_1 = \ln(2) = 0.69$$

$$A_2 = 0 = \ln(1)$$

$$A_3 = 2.08$$

Based on the above data, answer the given subquestions.

### Sub questions

Question Number : 93 Question Id : 640653676866 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Short Answer Question

What is the probability of choosing arm  $A_2$  at timestamp  $t = 6$ ?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.08 to 0.10



$$p_{\mu} = \frac{e^{0.69}}{e^{0.69} + e^0 + e^{2.08}}$$

$$= \frac{1}{e^{0.69} + 1 + 8.00} = \frac{1}{11} = 0.09$$

Question Number : 94 Question Id : 640653676867 Question Type : MCQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

At time stamp  $t = 6$ , suppose the arm with least estimate so far is pulled and the reward is  $\ln(4)$ .

Which of the following is correct after timestamp  $t = 6$ ?

Options :

$$\begin{aligned} A_1 &= \text{same} = 0.69 \\ A_2 &= 0.46 \\ A_3 &= 2.08 \end{aligned}$$

6406532266732. ✖ Arm  $A_1$  is optimal arm.

6406532266733. ✖ Arm  $A_2$  is optimal arm.

6406532266734. ✔ Arm  $A_3$  is optimal arm.

6406532266735. ✖ An optimal arm can not be determined.

6406532266736. ✖ There is a tie for optimal arm.

Sub-Section Number :

5

Sub-Section Id :

64065396844

Question Shuffling Allowed :

Yes

Is Section Default? :

null

Question Number : 95 Question Id : 640653676868 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Short Answer Question

A game of tic-tac-toe ends, if the board is full or either of the player wins. A player wins the game if he has 3 of its symbols in a row or a column or a diagonal. The length of the episode for following board position of tic-tac-toe is 9:

X	O	X
X	X	O
O	X	O

Consider following state of a tic-tac-toe board:

X		X
O	X	
O		

$5 + 2 \rightarrow$  anywhere  
0 puts (6<sup>th</sup>) an  
7<sup>th</sup> mark, X wins

Assuming that both players play optimally from above board position, what will be the length of this particular episode?

Response Type : Numeric



Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

7 ✓

Sub-Section Number :

6

Sub-Section Id :

64065396845

Question Shuffling Allowed :

Yes

Is Section Default? :

null

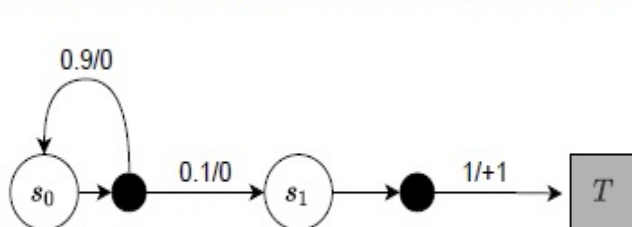
Question Number : 96 Question Id : 640653676869 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

Consider following MDP, transition probability and rewards are duly mentioned:



$$\begin{aligned} U(T) &= 0 \\ U(s_i) &= p_{\pi}(a + U(T)) \\ &= 1(1 + 0) = 1 \end{aligned}$$

Suppose  $\gamma = 1.0$  and  $T$  is terminal state. What is the value of the return  $G_0$ ?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

1 ✓✓

Sub-Section Number :

7

Sub-Section Id :

64065396846

$$\begin{aligned} U(s_0) &= 0.9(0 + U(s_0)) + 0.1(0 + 1) \\ &= 0.9 U(s_0) + 0.1 \\ U(s_0) &= 1 \end{aligned}$$

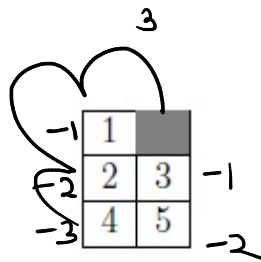
Question Shuffling Allowed : No  
Is Section Default? : null

Question Id : 640653676871 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (97 to 101)

Question Label : Comprehension

Consider following grid world:



$$\begin{aligned} q_{\pi^*}(3, \text{down}) &= R + \gamma(v_{\pi^*}(5)) \\ &= -1 + \gamma(-2) \\ &= -3 \end{aligned}$$

$u_{\pi}$	$s$
-1	1, 3
-2	2, 5
-3	4

- All transitions cost -1 reward.
- The agent can take 4 actions i.e.  $\{left, right, up, down\}$ . An action that takes the agent outside of the grid world, leaves the state unchanged.
- All transitions are deterministic.
- Gray cell represents terminal state.
- Discounting factor  $\gamma = 1$
- $\pi^*$ ,  $v_{\pi^*}(s)$  and  $q_{\pi^*}(s, a)$  represent optimal policy and corresponding state and action value functions, respectively.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 97 Question Id : 640653676872 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2 Max. Selectable Options : 0

Question Label : Multiple Select Question

Consider the following deterministic policies in table (1) :

→	
↑	↑
↑	↑

✓

(a)  $\pi_1$ 

→	
↑	↑
→	↑

✓

(b)  $\pi_2$ 

→	
→	↑
↑	↑

✓

(c)  $\pi_3$ 

→	
↑	←
↑	↑

✗

(d)  $\pi_4$ 

1	2
2	3
4	5

→

↑

↑

Table 1: Value function for different policies

Which of the following are optimal policies?

Options :

6406532266743. ✓  $\pi_1$

6406532266744. ✓  $\pi_2$

6406582266745. ✓  $\pi_3$

6406532266746. ✗  $\pi_4$

6406532266747. ✗ Can not be determined.

Question Number : 98 Question Id : 640653676873 Question Type : MSQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2 Max. Selectable Options : 0

Question Label : Multiple Select Question

Refer to table (1) and choose the correct statements from the following:

→	
↑	↑
↑	↑

(a)  $\pi_1$

→	
↑	↑
→	↑

(b)  $\pi_2$

→	
→	↑
↑	↑

(c)  $\pi_3$

→	
↑	←
↑	↑

(d)  $\pi_4$

Table 1: Value function for different policies

Options :

6406532266748. ✓  $\pi_1 = \pi_2$

6406532266749. ✗  $\pi_2 < \pi_3$

6406532266750. ✗  $\pi_1 \leq \pi_4$

6406532266751. ✓  $\pi_3 \geq \pi_4$

6406532266752. ✗ None of these

Question Number : 99 Question Id : 640653676874 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Short Answer Question

Compute  $v_{\pi^*}(4)$ .

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

-3

Question Number : 100 Question Id : 640653676875 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

Compute  $q_{\pi^*}(3, \text{down})$ .

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

-3 ✓

Question Number : 101 Question Id : 640653676876 Question Type : MSQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2 Max. Selectable Options : 0

Question Label : Multiple Select Question

Select the pairs of states  $(i, j), i \neq j$   
such that  $v_{\pi^*}(i) = v_{\pi^*}(j)$

Options :

6406532266755. ✗ 1 and 2.

6406532266756. ✓ 1 and 3.

6406532266757. ✗ 2 and 3.

6406532266758. ✓ 2 and 5.

6406532266759. ✗ 2 and 4.

6406532266760. ✗ 4 and 5.

Sub-Section Number :

8

Sub-Section Id :

64065396847

Question Shuffling Allowed :

No

Is Section Default? :

null

Question Id : 640653676877 Question Type : COMPREHENSION Sub Question Shuffling

Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix

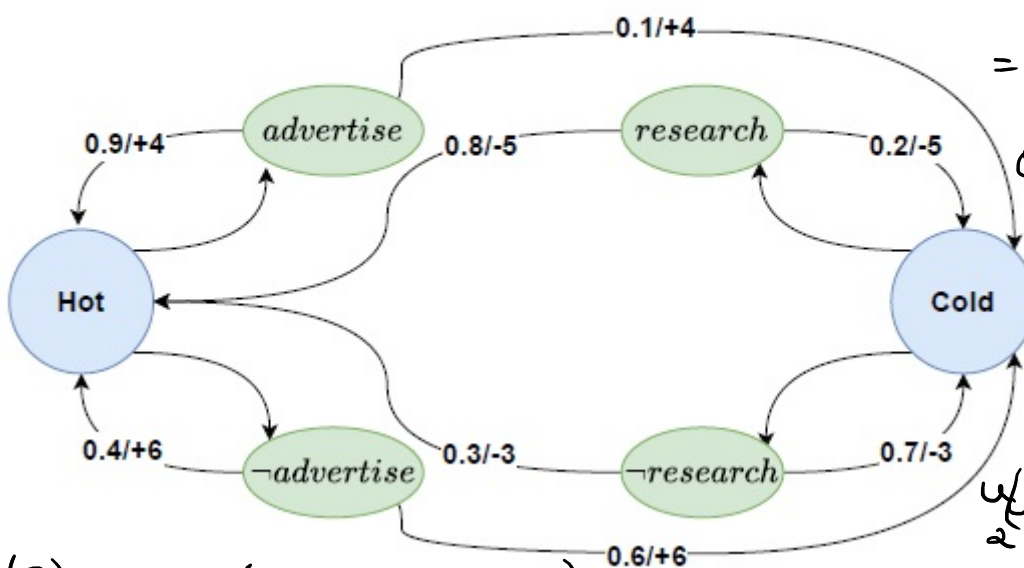
Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (102 to 104)

Question Label : Comprehension

Consider following transition diagram of an MDP:

Assume  $\gamma = 1$ .



$U_2(\text{hot})$  2<sup>nd</sup> iteration

$$Q(\text{ad}) = R + \gamma(P_{\text{ad}} \times U_1(\text{hot}) + P_{\text{ad}} \times U_1(\text{cold}))$$
$$= 4 + 1(0.9 \times 6 + 0.1 \times (-3))$$
$$= 4 + (5.4 + (-0.3)) = 5.1 + 4 = 9.1$$
$$Q(\text{-ad}) = 6 + (0.4 \times 6 + 0.6 \times (-3))$$
$$= 6 + (2.4 + (-1.8))$$
$$= 6 + 0.6 = 6.6$$

$U_2(\text{hot}) = 9.1$

$U_2(\text{cold})$

$$Q(R) = -5 + (0.8 \times 6 + 0.2 \times (-3)) = -5 + (4.8 + (-0.6)) = -5 + 4.2 = -0.8$$

$$Q(-R) = -3 + (0.3 \times 6 + 0.7 \times (-3)) = -3 + (1.8 - 2.1) = -3 - 0.3 = -3.3$$

Based on the above data, answer the given subquestions.

Sub questions

1 round of iteration  $\Rightarrow U(\text{hot}) = 6$

$U(\text{cold}) = -3$

Question Number : 102 Question Id : 640653676878 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

What will be the value of  $v(\text{hot})$  after one round of value iteration? Assuming  $v(\text{hot})$  and  $v(\text{cold})$  are initialized with 0. Note the value function is updated synchronously.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

6



**Question Number :** 103 **Question Id :** 640653676879 **Question Type :** SA **Calculator :** None

**Response Time :** N.A **Think Time :** N.A **Minimum Instruction Time :** 0

**Correct Marks :** 3

**Question Label :** Short Answer Question

What will be the value of  $v(cold)$  after one round of value iteration? Assuming  $v(hot)$  and  $v(cold)$  are initialized with 0. Note the value function is updated synchronously.

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

-3



**Question Number :** 104 **Question Id :** 640653676880 **Question Type :** SA **Calculator :** None

**Response Time :** N.A **Think Time :** N.A **Minimum Instruction Time :** 0

**Correct Marks :** 4

**Question Label :** Short Answer Question

What will be the value of  $v(hot) + v(cold)$  after **two** rounds of value iteration? Assuming  $v(hot)$  and  $v(cold)$  are initialized with 0. Note the value function is updated synchronously.

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

$$9.1 + (-0.8) = \underline{\underline{8.3}}$$

Possible Answers :

10.75 to 10.85

*this is wrong*

## Industry 4.0

Section Id :	64065345318
Section Number :	6
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	11
Number of Questions to be attempted :	11
Section Marks :	20
Display Number Panel :	Yes
Section Negative Marks :	0
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	Yes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	64065396848
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 105 Question Id : 640653676883 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 0

Question Label : Multiple Choice Question

THIS IS QUESTION PAPER FOR THE SUBJECT "DEGREE LEVEL : INDUSTRY 4.0 ( COMPUTER BASED EXAM )"