1

Question Number : 207 Question Id : 640653829791 Question Type : SA
Correct Marks : 2
Question Label : Short Answer Question
How many cops are necessary and sufficient to catch the robber on a cycle?
Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
2

Question Number : 208 Question Id : 640653829792 Question Type : MCQ
Correct Marks : 2
Question Label : Multiple Choice Question
How many cops are definitely enough to catch the robber on a graph of treewidth k?
Options :
6406532786907. ✖ $k/2$

6406532786908. ✖ $\sqrt{k}$

6406532786909. ✖ $k-1$

6406532786910. ✔ $k+1$

RL

Section Id :	64065359436
Section Number :	13
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	8
Number of Questions to be attempted :	8
Section Marks :	40
Display Number Panel :	Yes
Section Negative Marks :	0
Group All Questions :	No

Enable Mark as Answered Mark for Review and Clear Response :	No
Section Maximum Duration :	0
Section Minimum Duration :	0
Section Time In :	Minutes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	640653124161
Question Shuffling Allowed :	No

Question Number : 209 Question Id : 640653829793 Question Type : MCQ

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 :

6406532786911. ✔ YES

6406532786912. ✖ NO

Question Number : 210 Question Id : 640653829794 Question Type : MCQ

Correct Marks : 0

Question Label : Multiple Choice Question

Note:

1. For numerical answer type questions, enter your answer correct upto two decimal places without rounding up or off unless stated otherwise.

Options :

6406532786913. ✔ Instructions has been mentioned above.

6406532786914. ✖ This Instructions is just for a reference & not for an evaluation.

Sub-Section Number :	2
Sub-Section Id :	640653124162
Question Shuffling Allowed :	No

Question Id : 640653829795 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix

Question Numbers : (211 to 214)

Question Label : Comprehension

The following table captures the arms pulled and corresponding rewards, with increasing timestamps. The bandit has only 3 arms. ϵ-greedy policy is used to select an arm.

Assume that the reward distribution is stationary and $\epsilon = 0.3$.

Timestamp (t)	Arm (A_i)	Reward ($r_{i,t+1}$)
0	A_1	2
1	A_2	1
2	A_3	4
3	A_3	2
4	A_2	2
5	A_1	2

$A_1 = 2$
$A_2 = 1.5$
$A_3 = 3$
} avg. rewards

Best arm $\Rightarrow A_3$

$\varepsilon = 0.3 \rightarrow \frac{1}{3}$ with arm A_1

$P(\text{Explore}) = 0.3 \times \frac{1}{3} = 0.1$

$1 - \varepsilon \rightarrow$ Best arm is chosen

$0.7 \rightarrow$ arm A_3

$P(A_1) = 0$ because it is not the best arm.

$0.1 + 0 = 0.1$

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 211 Question Id : 640653829796 Question Type : SA

Correct Marks : 2

Question Label : Short Answer Question

What is the probability of choosing arm A_1 at timestamp $t = 6$?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.095 to 0.105

Question Number : 212 Question Id : 640653829797 Question Type : SA

Correct Marks : 2

Question Label : Short Answer Question

What is the probability of choosing arm A_3 at timestamp $t = 6$?

Explore $P(A_3) \rightarrow 0.3 \times \frac{1}{3} = 0.1$

Best arm $\rightarrow 0.7$ for choosing A_3

$0.1 + 0.7 = 0.8$

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.795 to 0.805

Question Number : 213 Question Id : 640653829798 Question Type : MCQ
Correct Marks : 2

Question Label : Multiple Choice Question

At timestamp $t = 6$, the arm with least estimate so far (i.e. for $0 \leq t \leq 5$) is pulled and the reward is 3. Which of the following is correct after timestamp $t = 6$?

$A_1 = 2$
$A_2 = 1.5$
$A_3 = 3$

$A_2 = \frac{1+2+3}{3} = 2$

New, avg R.
$A_1 = 2$
$A_2 = 2$
$A_3 = 3$

Options :

6406532786917. ✖ Arm A_1 is the optimal arm.

6406532786918. ✖ Arm A_2 is the optimal arm.

6406532786919. ✔ Arm A_3 is the optimal arm.

6406532786920. ✖ An optimal arm can not be determined.

6406532786921. ✖ There is a tie for the optimal arm.

Question Number : 214 Question Id : 640653829799 Question Type : MCQ
Correct Marks : 2

Question Label : Multiple Choice Question

At time stamp $t = 7$, what is the minimum reward an arm A_i $(i \in \{1, 2, 3\})$ has to provide, given A_i is selected, in order to become the optimal arm at next timestamp $t = 8$?

$A_3 = \frac{2+4}{2} = 3$

$A_1 = \frac{2+2+x}{3} \leq 3$

$2+2+x = 9$

$x \geq 5$

Options :

6406532786922. ✖ Arm A_1 has to provide a reward of 4 or more.

6406532786923. ✔ Arm A_2 has to provide a reward strictly more than 6.

6406532786924. ✖ Arm A_3 has to provide a reward strictly less than 0.

6406532786925. ✖ It can not be determined.

6406532786926. ✖ None of these.

$A_2 = \frac{1+2+3+x}{4} \leq 3$

$1+2+3+x \leq 12$

$x \geq 6$

if $x = 6$ then A_2, A_3 is equally best

if $x > 6$, A_2 wins

Sub-Section Number : 3
Sub-Section Id : 640653124163
Question Shuffling Allowed : Yes

Question Number : 215 Question Id : 640653829800 Question Type : SA
Correct Marks : 2

Question Label : Short Answer Question

Suppose $\gamma = 0.2$ and the reward sequence is $R_t = 2 \times t$ for $t \in \{1, 2, 3, 4, \ldots\}$. What is the value of the return G_0 ?

$G_0 = 2 + \gamma 4 + \gamma^2 6 + \gamma^3 8 \ldots$
$= 2 \cdot (1 + 2\gamma + 3\gamma + 4\gamma^3)$
$= \sum_{t=1}^{\infty} \gamma^{t-1} R_t = \sum_{t=1}^{\infty} \gamma^{t-1} (2t)$
$= 2 \sum_{t=1}^{\infty} \gamma^{t-1} t$
$= 2 \sum_{t=1}^{\infty} t (0.2)^{t-1}$

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Range

We know $\sum t x^{t-1} = \frac{1}{(1-x)^2}$

Text Areas : PlainText

Possible Answers :

3.12 to 3.13

Sub-Section Number : 4

Sub-Section Id : 640653124164

Question Shuffling Allowed : No

Question Id : 640653829801 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix

Question Numbers : (216 to 219)

Question Label : Comprehension

Consider following MDP represented in a tabular form:

s	a	s'	$p(s'\|s,a)$	r
s_0	f	s_1	0.5	2
s_0	f	s_2	0.5	2
s_1	f	s_1	0.5	0
s_1	f	s_3	0.5	0
s_2	f	s_3	0.5	1
s_2	f	T	0.5	1
s_3	f	T	1.0	1

Table 1: An MDP

The symbols have the usual meaning. There are only 4 non-terminal states, $S = \{s_0, s_0, s_1, s_2, s_3\}$ and only one action possible in each state $A = \{f\}$ (i.e. forward). T represents the terminal state.

Answer the given subquestions:

Sub questions

Question Number : 216 Question Id : 640653829802 Question Type : SA

Correct Marks : 2

Question Label : Short Answer Question

What is the expected return starting from state s_0 if $\gamma = 1$?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

3.245 to 3.255

18
25
090

Question Number : 217 Question Id : 640653829803 Question Type : SA

Correct Marks : 2

Question Label : Short Answer Question

What is the expected return starting from state s_2 if $\gamma = 0.8$?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

1.395 to 1.405

$V(T) = 0$

$V(s_3) = Pa(a + \gamma V(T)) = 1(1+0) = 1$

$V(s_2) = Pa(a + \gamma V(T)) + Pa(a + \gamma V(s_3))$

$= 0.5(1+0) + 0.5(1 + 0.8 \times 1)$

$= 0.5 + 0.5(1.8)$

$= 0.5 + 0.9 = 1.4$

Question Number : 218 Question Id : 640653829804 Question Type : SA

Correct Marks : 3

Question Label : Short Answer Question

What is the expected return starting from state s_1 if $\gamma = 0.8$?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.66 to 0.67

$V(s_1) = Pa(a + \gamma V(s_1)) + Pa(a + \gamma V(s_3))$

$= 0.5(0 + 0.8 V(s_1)) + 0.5(0 + 0.8)$

$= 0 + 0.4 V(s_1) + 0.4$

$V(s_1) = \frac{0.4}{(1-0.4)} = \frac{0.4}{0.6} = \frac{2}{3} = 0.667$

Question Number : 219 Question Id : 640653829805 Question Type : SA

Correct Marks : 4

Question Label : Short Answer Question

What is the expected return starting from state s_0 if $\gamma = 0.8$?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

2.82 to 2.83

$V(s_0) = Pa(a + \gamma(Vs_1)) + Pa(a + \gamma(Vs_2))$

$= 0.5(2 + 0.8 \times 0.667) + 0.5(2 + 0.8 \times 1.4)$

$= 0.5 \times 2.5336 + 0.5(3.12)$

$= 1.2668 + 1.56$

1.5600
1.2668
2.8268

Sub-Section Number : 5
Sub-Section Id : 640653124165
Question Shuffling Allowed : Yes

Question Number : 220 Question Id : 640653829806 Question Type : MCQ
Correct Marks : 2
Question Label : Multiple Choice Question
Which of the following is the correct Bellman equation for stochastic transitions, stochastic policy and stochastic rewards? The symbols have the usual meaning.
Options :

6406532786932. ✖ $v_\pi(s) = \sum_{s'} \sum_r p(s', r|s, a)[r + \gamma v_\pi(s')]$

6406532786933. ✔ $v_\pi(s) = \sum_a \pi(a|s) \sum_{s'} \sum_r p(s', r|s, a)[r + \gamma v_\pi(s')]$

policy transition probab. rewards

6406532786934. ✖ $v_\pi(s) = \sum_a \pi(a|s)[r + \gamma v_\pi(s')]$

6406532786935. ✖ $v_\pi(s) = \sum_a \pi(a|s) \sum_r p(s', r|s, a)[r + \gamma q(s', a')]$

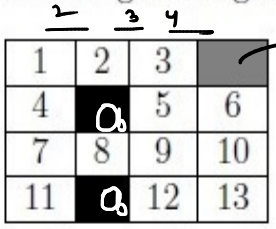
6406532786936. ✖ $v_\pi(s) = \sum_a \pi(a|s) \sum_{s'}[r + \gamma v_\pi(s')]$

6406532786937. ✖ None of these

Sub-Section Number : 6
Sub-Section Id : 640653124166
Question Shuffling Allowed : No

Question Id : 640653829807 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix
Question Numbers : (221 to 223)
Question Label : Comprehension

Consider the following 4 × 4 grid world:



- 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 or in an obstacle, leaves the state unchanged.
- All transitions are deterministic.
- Gray cell represents terminal state. Top right corner is terminal state.
- Black cells represent obstacles.
- Each visitable cell/state is numbered.
- Discounting factor $\gamma = 1$
- π^*, $v_{\pi^*}(s)$ 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 : 221 Question Id : 640653829808 Question Type : SA
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 :
-4

Question Number : 222 Question Id : 640653829809 Question Type : SA
Correct Marks : 2
Question Label : Short Answer Question

Compute $q_{\pi^*}(7, down)$.

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

-7

Question Number : 223 Question Id : 640653829810 Question Type : MSQ

Correct Marks : 3 Max. Selectable Options : 0

Question Label : Multiple Select Question

$v_{\pi^*}(i) = v_{\pi^*}(j)$

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

Options :

6406532786940. ✖ 1 and 2.

6406532786941. ✖ 1 and 3.

6406532786942. ✖ 2 and 3.

6406532786943. ✔ 2 and 5.

6406532786944. ✔ 13 and 9

6406532786945. ✔ 4 and 8.

6406532786946. ✖ None of these

Check my table above

Sub-Section Number :	7
Sub-Section Id :	640653124167
Question Shuffling Allowed :	No

Question Id : 640653829811 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Question Numbers : (224 to 227)

Question Label : Comprehension

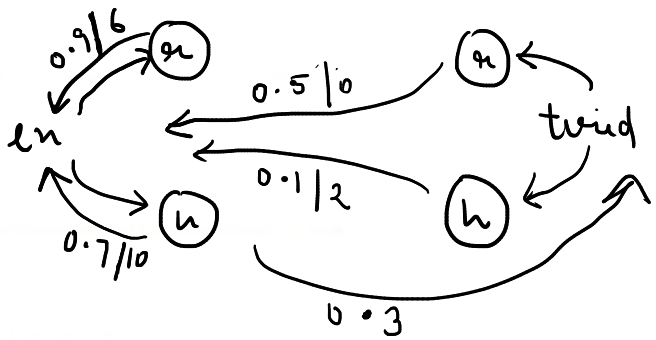
Suppose Akash wanted to make an informed decision about whether to hike or relax over the weekend. Akash prefers to hike, but is worried about getting tired. Such a problem can be modeled as an MDP with two states, *energetic* and *tired*, and two actions, *relax* and *hike*. Thus $S = \{energetic, tired\}, A = \{relax, hike\}$. Based on experience, Akash estimates that the dynamics $p(s'|s, a)$ is given by following table:

s	a	$s' =$ energetic
energetic	relax	0.9
energetic	hike	0.7
tired	relax	0.5
tired	hike	0.1

energetic → relax, hike
tired → relax, hike

So, if Akash is energetic and hikes, there is a 30% chance of becoming tired. If Akash is energetic and relaxes, Akash will more likely remain energetic. If Akash is tired and relaxes, there is a 50% chance of becoming energetic. If Akash is tired and hikes, there is only a 10% chance of becoming energetic. Akash estimates the (immediate) rewards to be:

s	a	r
energetic	relax	6
energetic	hike	10
tired	relax	0
tired	hike	2



Thus, Akash always enjoys hiking more than relaxing. However, Akash feels much better overall when energetic, and hiking results in being tired more than relaxing does. Assuming $v(energetic)$ and $v(tired)$ are initialized with 0. Using value iteration, answer the subquestions. Note the value function is updated synchronously. Assume $\gamma = 0.8$.

Sub questions

Question Number : 224 Question Id : 640653829812 Question Type : SA
Correct Marks : 2
Question Label : Short Answer Question

What will be the value of $v(energetic)$ after one round of value iteration?

Response Type : Numeric

v_0(energy) = 0
v_0(tired) = 0
max of energetic reward = 10

v_1 Energetic $Q(Relax) = R + \gamma (P_{en} \times v_0(energy) + P_{en} \times v_0(tired))$
$= 6 + 0.8(0.9 \times 0 + 0.1 \times 0) = 6$

similarly = $Q(hike) = 10$

$v_{energetic} = 10$

Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
10

Question Number : 225 Question Id : 640653829813 Question Type : SA
Correct Marks : 2
Question Label : Short Answer Question

What will be the value of $v(tired)$ after one round of value iteration?

max of tired reward = 2

$v(tired) = 2$

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
2

Question Number : 226 Question Id : 640653829814 Question Type : SA
Correct Marks : 3
Question Label : Short Answer Question

What will be the value of $v(energetic)$ after *two* rounds of value iteration?

$v_2(energetic)$

$Q(Relax) = R + 0.8\,(Pr \times v_1(energy) + Pr \times v_1(tired))$

$= 6 + 0.8(0.9 \times 10 + 0.1 \times 2)$

$= 6 + 0.8(9 + 0.2)$

$= 6 + 0.8(9.2) = 13.36$

this is max

$Q(hike) = 10 + 0.8(0.7 \times 10 + 0.3 \times 2)$

$= 10 + 0.8(7 + 0.6) = 10 + 0.8(7.6) = 16.08$

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Range
Text Areas : PlainText
Possible Answers :
16.03 to 16.13

Question Number : 227 Question Id : 640653829815 Question Type : SA
Correct Marks : 3
Question Label : Short Answer Question

What will be the value of $v(tired)$ after *two* rounds of value iteration?

$v_2(tired) = 4.8$

$Q(Relax) = 0 + \gamma(0.5 \times 10 + 0.5 \times 2)$

$= 0 + 0.8(5 + 1) = 0.8 \times 6 = 4.8$

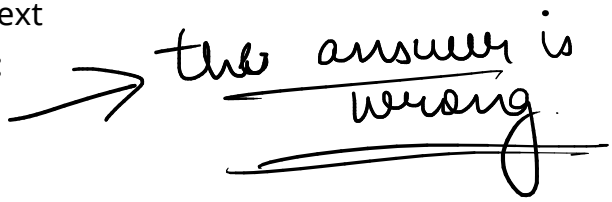
$Q(hike) = 2 + 0.8(0.1 \times 10 + 0.9 \times 2)$

$= 2 + 0.8(1 + 1.8) = 2 + 0.8 \times 2.8$

$= (4.24)$

Response Type : Numeric
Evaluation Required For SA : Yes

Show Word Count : Yes
Answers Type : Range
Text Areas : PlainText
Possible Answers :
12.23 to 12.25



Game Theory

Section Id :	64065359437
Section Number :	14
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	6
Number of Questions to be attempted :	6
Section Marks :	25
Display Number Panel :	Yes
Section Negative Marks :	0
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	No
Section Maximum Duration :	0
Section Minimum Duration :	0
Section Time In :	Minutes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	640653124168
Question Shuffling Allowed :	No

Question Number : 228 Question Id : 640653829816 Question Type : MCQ
Correct Marks : 0
Question Label : Multiple Choice Question
THIS IS QUESTION PAPER FOR THE SUBJECT "DEGREE LEVEL : GAME THEORY (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 :
6406532786951. ✔ YES
6406532786952. ✖ NO