(170)

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATIONS- 2026

B.Tech-VI/VIII Semester (CSE/IT/BT/ECE/CE)

COURSE CODE (CREDITS): ~~25B1WCI640~~/19B1WCI837 MAX MARKS: 25

COURSE NAME: Reinforcement Learning

COURSE INSTRUCTOR: Saurav Singh MAX. TIME: 1 Hour 30 Min

Note: (a) All questions are compulsory.

(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems

(c) Use of calculator is allowed

Q.No	Question	CO	Marks
Q1	Consider a Markov reward process with the state space {1,2,3} , and initial state Probabilities are $\mu(1)=1/3$, $\mu(2)=1/3$, $\mu(3)=1/3$. The one step probability transition matrix is given below $P = \begin{array}{\|c\|c\|c\|} \hline 0.1 & 0.2 & 0.7 \\ \hline 0.4 & 0 & 0.6 \\ \hline 0.5 & 0.1 & 0.4 \\ \hline \end{array}$ Find the following: (a) $P(X_2=3 \mid X_1=2)$ (b) $P(X_3=1 \mid X_1=2)$ (c) $P(X_2=2)$ (d) $P(X_9=2 \mid X_7=2, X_5=1, X_0=3)$ (e)Consider the reward function R(1,1)= 0, R(1,2)=0, R(1,3)=0 , R(2,1)= 3, R(2,2)=3, R(2,3)=5 and R(3,1)=2, R(3,2)=2,R(3,3)=1. Now find the expected γ-discounted cumulative reward for $\gamma = 0.1$.	CO3	1+2+2 +2+3
Q2	In the tic-tac-toe game, consider the following: (a) If we award zero reward for every transition irrespective of action, then what is the optimal policy? Justify your answer. (b)Now consider the following: We award one unit reward for every transition irrespective of action. And for every action, all valid states are equally likely. Then what will be the optimal policy? Justify your answer.	CO3	2+3

<table>
<tr><td>Q3</td><td>In a Markov Decision Process (MDP), the reward function is defined as a function of the current state, the action taken, and the next state. Derive the corresponding value function and explain each term involved in the formulation.</td><td>CO3</td><td>5</td></tr>
<tr><td>Q4</td><td>(a) Two fair dice are rolled. Let X be the sum of outcomes. Let event A: "sum is greater than 8". Compute E[X|A]

(b) The joint pmf of X and Y is:
<table><tr><th>X\Y</th><th>1</th><th>2</th></tr><tr><td>1</td><td>0.1</td><td>0.2</td></tr><tr><td>2</td><td>0.3</td><td>0.4</td></tr></table>
Find : E[X|Y=2] E[Y|X=1]</td><td>CO2</td><td>2+3</td></tr>
</table>