

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2023

B.Tech-VIII Semester (CSE/IT)

COURSE CODE (CREDITS): 19B1WCI837 (3)

MAX. MARKS: 35

COURSE NAME: REINFORCEMENT LEARNING

COURSE INSTRUCTORS: DHA

MAX. TIME: 2 Hours

Note: All questions are compulsory. Marks are indicated against each question in square brackets.

Q1. a) Explain the importance of *Gamma* (γ) and *Lambda* (λ) in temporal differencing algorithms? **[CO-3, Marks: 2+2+3]**

b) Give the advantages of using Temporal Difference Learning over Monte Carlo Learning.

c) Show comparison between the above two by simulating the Black Jack game.

Q2. Explain in detail Monte Carlo Backup, Temporal Difference Learning Backup and Dynamic Programming Backup in terms of strength of their exhaustive search along with Backup diagrams. **[CO-3, Marks: 3]**

Q3. Explain TD (λ) and λ -return in detail. **[CO-3, Marks: 3]**

Q4. a) How is On-Policy Learning different from Off-Policy Learning?

[CO-3, Marks: 3+2+3+3]

b) Is one better than the other in above question? Justify your answer.

c) Explain On-Policy Control with Sarsa Algorithm. Illustrate with Windy Gridworld example.

d) Give algorithm for Sarsa Algorithm for On-Policy Control.

Q5. a) What is Monte Carlo (MC) Control? **[CO-3,4, Marks: 2+2+4]**

b) What are the steps involved in MC Control?

c) Explain ϵ -Greedy Exploration. What is the advantage of using it over standard Monte Carlo Exploration?

Q6. a) How does the Monte Carlo prediction method compute the Value Function?

[CO-2,3, Marks: 2+1]

b) Why do we use Dynamic Programming in Planning problem in Reinforcement Learning?