

**BACHELOR OF TECHNOLOGY (CBCS) (2021-COURSE)**  
**B. Tech. Sem - VIII Computer Science & Engineering AI & ML : SUMMER : 2025**  
**SUBJECT: ITC-VI: REINFORCEMENT LEARNING**

Day : Monday  
Date : 02/06/2025

S-23997-2025

Time : 02:30 PM-05:30 PM  
Max. Marks : 60

**N.B.**

- 1) All questions are **COMPULSORY**.
- 2) Figures to the **RIGHT** indicate **FULL** marks.
- 3) Assume suitable data, if necessary.
- 4) Draw neat and labeled diagrams, **wherever** necessary.

**Q.1** Imagine you are tasked with building a recommendation system for a streaming service. How would you approach the problem using machine learning techniques? Discuss the types of algorithms you might consider and the challenges you anticipate. (10)

**OR**

**Q.1** Explain in brief exploration-exploitation dilemma in RL, with the help of example. (10)

**Q.2** Describe the key differences between a Markov Chain and a standard directed graph. Provide an example of a real-world scenario that can be modeled as a Markov Chain. (10)

**OR**

**Q.2** Describe the intuitive meaning behind the Bellman equation's recursive structure. Why is it useful to express the value function in this way? (10)

**Q.3** Explain the core idea behind Temporal Difference (TD) learning. How does it differ from Monte Carlo methods in terms of updating value estimates? (10)

**OR**

**Q.3** Describe TD(0) in detail. What are its strengths and weaknesses, and how does it update value estimates? (10)

**Q.4** State and explain Bellman's principle of optimality. How does this principle form the foundation for DP algorithms in MDPs? (10)

**OR**

**Q.4** Describe the process of iterative policy evaluation. How does it use the Bellman expectation equation to estimate the value function of a given policy? (10)

**Q.5** Explain the core idea behind Gradient Monte Carlo methods for function approximation. How do they differ from tabular MC methods? (10)

**OR**

**Q.5** Describe the Semi-gradient TD(0) algorithm. Why is it called 'semi-gradient,' and what are its advantages and limitations? (10)

**Q.6** Explain the Naive REINFORCE algorithm. What are its strengths and weaknesses? Why is it considered a high-variance algorithm? (10)

**OR**

**Q.6** Discuss the concepts of bias and variance in the context of reinforcement learning. How do they affect the learning process, and how do they relate to policy gradient methods? (10)

\* \* \* \* \*