**Vivekanand Education Society’s Institute of Technology**

**Department of AIDS Engineering**



**Subject: Reinforcement Learning**

**Class: D16AD**

| **ROLL NO: 30** | **NAME: SUHANEE KANDALKAR** | | |
| --- | --- | --- | --- |
| **EXP NO: 10** | T**ITLE:**   **Monte Carlo methods** | | |
| **DOP:** |  | **DOS: 29/03/25** |  |
| **GRADES:** | **LOs MAPPED:** | | **SIGNATURE:** |

## 

## Aim**:**

Monte Carlo methods

## Theory**:**

**What are Monte Carlo Methods?**

Monte Carlo (MC) methods are a class of algorithms that rely on random sampling to obtain numerical results. The core idea is to use randomness to solve problems that might be deterministic in principle.

In Reinforcement Learning (RL), MC methods are used to estimate the value functions by averaging returns (total rewards) from sample episodes of experience.

**Key Concepts**

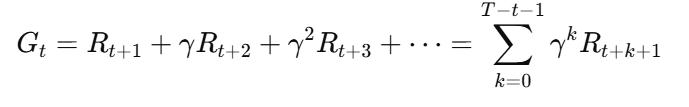
1. Episode

An episode is a sequence of states, actions, and rewards, ending in a terminal state.

1. Return (Gₜ)

The total accumulated reward from time step t onwards.

For episodic tasks:



Where:

* γ is the discount factor (0 ≤ γ ≤ 1)
* Rt​ is the reward at time step t

#### Value Function Estimation

#### There are two primary value functions:

#### State Value Function V(s)V(s)V(s) – Expected return from state s.

#### Action Value Function Q(s,a)Q(s,a)Q(s,a) – Expected return from state s taking action a.

#### Monte Carlo Prediction

#### Estimate V(s) or Q(s,a) from sampled episodes:

#### First-Visit MC: Average returns following the first time a state (or state-action) is visited.

#### Every-Visit MC: Average returns every time a state is visited in an episode.

#### Monte Carlo Control

#### Used for learning optimal policies:

* Policy Evaluation: Estimate value functions under a given policy.
* Policy Improvement: Make the policy greedy w.r.t. the estimated value function.

A common algorithm: ε-greedy policy for balancing exploration and exploitation.

[RL\_10\_30](https://colab.research.google.com/drive/1nuBJ01bIUStpNEX-fpd8d9zkRPpfvTx0)

# Conclusion :

Monte Carlo methods provide a foundational approach in Reinforcement Learning for estimating value functions based on experience. By leveraging random sampling of episodes and averaging returns, MC methods help evaluate and improve policies without requiring a model of the environment. Their episodic nature and variants like first-visit and every-visit MC offer flexibility in both value prediction and control, ultimately guiding agents toward optimal decision-making in stochastic environments.