Implementing SARSA Algorithm for Mountain Car Control Problem in Reinforcement Learning

Introduction

The objective of implementing the SARSA algorithm in this scenario is to train an agent to navigate a car in a 2D environment, known as the Mountain Car Control Problem, using reinforcement learning. In this problem, the car needs to reach the top of a steep hill by applying force to its engine. However, the car's engine is not powerful enough to climb the hill directly, so the agent must learn to apply forces strategically to build momentum and reach the goal.

Mountain Car Control Problem

The Mountain Car Control Problem is a classic reinforcement learning problem where an agent controls a car in a 2D environment. The car starts at the bottom of a valley between two hills and must reach the top of the right hill. However, the car's engine is not strong enough to climb the hill directly, so the agent must learn to apply forces effectively to build momentum and reach the goal.

SARSA Algorithm

SARSA (State-Action-Reward-State-Action) is an on-policy reinforcement learning algorithm used to learn the optimal policy for sequential decision-making tasks. In SARSA, the agent learns the Q-values (expected cumulative rewards) for state-action pairs and updates them based on observed rewards and transitions.

How SARSA Works

At each time step, the agent selects an action based on an exploration-exploitation strategy, such as ϵ-greedy. It takes the selected action, observes the reward, and updates the Q-value for the current state-action pair using the SARSA update rule.

SARSA Update Rule

The SARSA update rule is given by:

$$Q(s, a) \leftarrow Q(s, a) + \alpha \cdot (r + \gamma \cdot Q(s', a') - Q(s, a))$$

where:

- $Q(s, a)$ is the Q-value of taking action a in state s.
- α is the learning rate, controlling the step size of updates.
- r is the observed reward after taking action a in state s.
- γ is the discount factor, representing the importance of future rewards.
- s' is the next state after taking action a in state s.
- a' is the action selected in the next state s'.

Result

Implementing the SARSA algorithm for the Mountain Car Control Problem in reinforcement learning provides a powerful framework for training agents to navigate complex environments. By iteratively updating Q-values based on observed rewards and transitions, SARSA enables agents to learn optimal action-selection policies that maximize cumulative rewards. This approach contributes to the development of intelligent systems capable of autonomous decision-making in challenging real-world scenarios.