Assignment #2: Markov Decision Processes

Course: Artificial Intelligence
Instructor: Zuhaib Hussain
Student: Muhammad Zaid
ID: 211370214
Project: Web-Based Visualization of Value Iteration and Policy Iteration

Introduction

This project focuses on the implementation and visualization of Markov Decision Processes (MDPs) using a Grid-World environment. The objective is to help users understand how optimal policies emerge through Value Iteration and Policy Iteration algorithms.

MDP Formulation

The Grid-World environment is modeled as an MDP where each cell represents a state. The agent can move in four directions: up, down, left, and right. Transitions are stochastic, with an 80% chance of moving in the intended direction and a 20% chance of random movement. The reward structure consists of +10 for reaching the goal state, -10 for entering a negative terminal state, and -0.1 for each step taken. The discount factor (γ) is user-controlled through the web interface.

Algorithm Implementation

Value Iteration computes the optimal value function by iteratively applying the Bellman optimality equation until convergence. Policy Iteration alternates between policy evaluation and policy improvement steps to determine the optimal policy. Both algorithms were implemented and visualized using a Streamlit-based web interface.

Web Application

The web application allows users to select between Value Iteration and Policy Iteration, adjust the discount factor using a slider, and observe step-by-step convergence. The value function is displayed numerically and the policy is visualized using directional arrows.

Observations and Results

Value Iteration converges steadily but may require more iterations, while Policy Iteration typically converges faster due to direct policy updates. The discount factor significantly affects agent behavior, with higher values encouraging long-term planning. Both algorithms successfully produced optimal policies.

Conclusion

This assignment successfully demonstrates the practical implementation of MDPs and reinforcement learning algorithms. The interactive visualization enhances conceptual understanding of policy formation and convergence behavior.

GitHub Repository: https://github.com/Zaid-code123/GridWorld-MDP-Visualization