MDP and Bellman Equations

Jadhav Rudranil Sandiip
Roll Number: 24095043
CSoC 2025 – Intelligence Guild

June 23, 2025

1 Introduction

This report provides detailed information regarding my solutions to the COPS CSOC Reinforcement Learning tasks.

2 File: `ProjectTasks.pdf`

This file contains all the theoretical questions assigned in this week's task. I primarily referred to Sutton and Barto's textbook for learning the concepts. As a result, many of the answers may closely resemble the wording from the book.

3 File: `Chapter3.pdf`

This file includes my solutions to Chapter 3 of Sutton and Barto. I have solved and written explanations for approximately 75% of the problems from this chapter.

4 File: `Chapter4.pdf`

I was able to solve only the first three problems from Chapter 4. The following sections describe the corresponding code files:

4.1 File: `Ex4-1.py`

This question required knowledge about state values of given environment. At the time, I was unaware that the book provides these values in the next section. Therefore, I implemented a policy evaluation method in code. Since this was my first experience with RL code, I referred to a similar solution online to help guide my implementation.

4.2 File: `Ex4-2.py`

This problem required a few modifications to the code from `Ex4-1.py` due to the introduction of a new state. While the answer could be calculated using results from `Ex4-1.py` it wasn't intuitive, hence i had to rewrite the `next_state` function to get the correct answer.

5 Original Frozen Lake

Value iteration and policy iteration algorithms were implemented using the standard Frozen Lake environment from the Gymnasium library.

6 Frozen Lake 20x20 Custom Map

The primary change for the custom 20x20 map was the map size, as the Gymnasium environment consists only 4x4 and 8x8 maps. The custom map was generated using the library's random map generator function. This approach helped keep the code cleaner and reduced potential errors compared to manually defining the map.

7 Frozen Lake with 8 Actions

The original goal of this task was to create a custom environment and register it using Gymnasium. However, adding multiple actions to the default Frozen Lake environment caused several issues. I was unable to debug my code hence I created a new environment named `FrozenLake8Action` from scratch. This implementation was inspired by the environment I built in `Ex4-1.py`.

8 Policy Iteration Vs Value Iteration

Note : Since the next weeks task is comparing all 5 methods , I have skipped this comparison in this report and will add it in the next weeks report to maintain clarity and ensure a more insightful and well-supported analysis.