Reinforcement Learning Assignment 1

Tianyang Yan (TA), Prof. Zou

2023-03-10

1 Introduction

The goal of this assignment is to do experiment with Dynamic Programming(DP), including iterative policy evaluation, policy iteration and value iteration. Your goal is to implement DP methods and test them in the small gridworld mentioned in the slides of Lecture 3.

2 Small Gridworld

0	1	2	3	4	5
6	7	8	9	10	11
12	13	14	15	16	17
18	19	20	21	22	23
24	25	26	27	28	29
30	31	32	33	34	35

Figure 1: Gridworld

As shown in Fig.1, each grid in the Gridworld represents a certain state. Let s_t denotes the state at grid t. Hence the state space can be denoted as $S = \{s_t | t \in 0, .., 35\}$. S_1 and S_{35} are terminal states, where the others are non-terminal states and can move one grid to north, east, south and west. Hence the action space is $A = \{n, e, s, w\}$. Note that actions leading out of the Gridworld leave state unchanged. Each movement get a reward of -1 until the terminal state is reached.

A good policy should be able to find the shortest way to the terminal state randomly given an initial non-terminal state.

3 Experiment Requirments

- Programming language: python3
- You should build the Gridworld environment and implement iterative policy evaluation, policy iteration and value iteration methods. Then run the two methods to evaluate and improve an uniform random policy $\pi(n|\cdot) = \pi(e|\cdot) = \pi(s|\cdot) = \pi(w|\cdot) = 0.25$

4 Report and Submission

- Your report and source code should be compressed and named after "studentID+name".
- The file should be submitted on Canvas on Mar. 16, 2023.