Defining Solving RL Environments(Checkpoint)

Vamsi Krishna Velivela
University at Buffalo
UB No:- 50318486
vvelivel@buffalo.edu

1 Deterministic and Stochastic Environments

Set of Actions :- Left,Right,Down,UP
States :- 36
Rewards :- 3 set of Rewards
at (5,5) - 3,at (3,3) - 1,at (4,3) - 2,at (4,5) - 2
Objective :- The main Objective is to reach the end corner position (5,5)

2 Visualizations

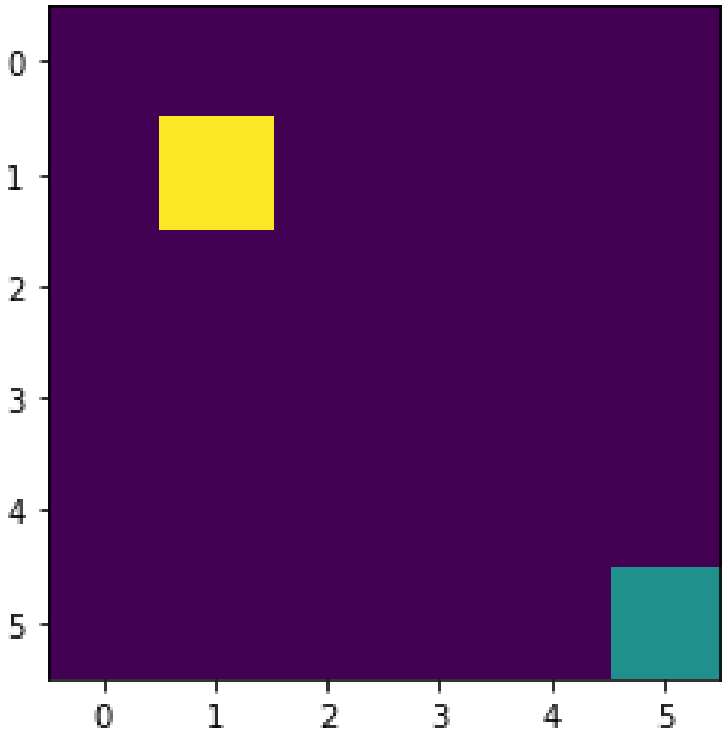

Figure 1: Deterministic Environment after moving agent from 0,0 to 1,1

3 Stochastic Environment Implementation

For all the four actions

- **Left** :- The agent goes to Left with a probability of 0.2 and stays back in it's own position with a probability of 0.8 with in bounds

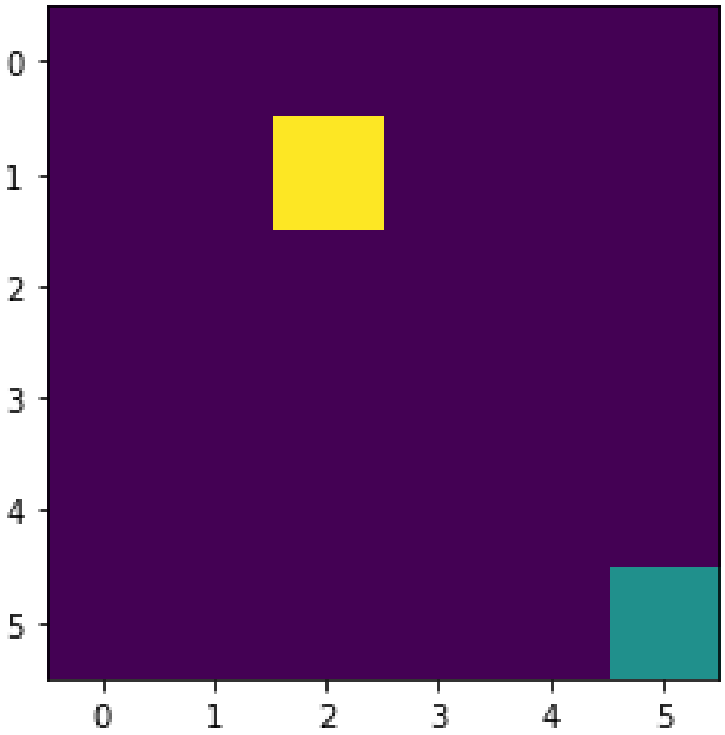

Figure 2: Stochastic Environment after moving agent from 0,0 to 1,2

- **Right** :- The agent goes to Right with a probability of 1 with in bounds
- **Down** :- The agent goes to Down with a probability of 1 with in bounds
- **Up** :- The agent goes to Up with a probability of 0.2 and stays back in it's own position with a probability of 0.8 with in bounds

4 Deterministic vs Stochastic Environments

4.1 Deterministic Policy

In deterministic policy, the action is mapped to the state. The agent follows the action and comes to a state without considering probability. In a way, the actions take deterministic outcomes without any uncertainty. $\pi(s) = a$

4.2 Stochastic Policy

In Stochastic policy, the action results to many possible states. The agent follows the action and comes to a state basing on the probability of action. $\pi(a—s) = P\pi[A = a \text{ and } S = s]$

5 Safety in AI:

The agent may get stuck in an continuous loop. To avoid that, we keep track of the number of steps taken and end the process when a max limit of number of steps is reached. In this way, we are ensuring the safety of our system.