Assignment 1 - Defining & Solving RL Environments

Arjun Ramesh Kaushik

UB ID - 50413327

Department of Computer Science & Engineering

kaushik3@buffalo.edu

https://github.com/arjunrkaushik/Reinforcement_Learning_Fall22

Abstract

The goal of the assignment is to acquire experience in defining and solving reinforcement learning environments, following OpenAI Gym standards. The assignment consists of two parts. The first focuses on defining deterministic and stochastic environments that are based on Markov decision process. In the second part we will apply two tabular methods to solve environments that were previously defined.

1 Defining RL Environments

1.1 Describe the deterministic and stochastic environments, which were defined (set of actions/states/rewards,main objective, etc).

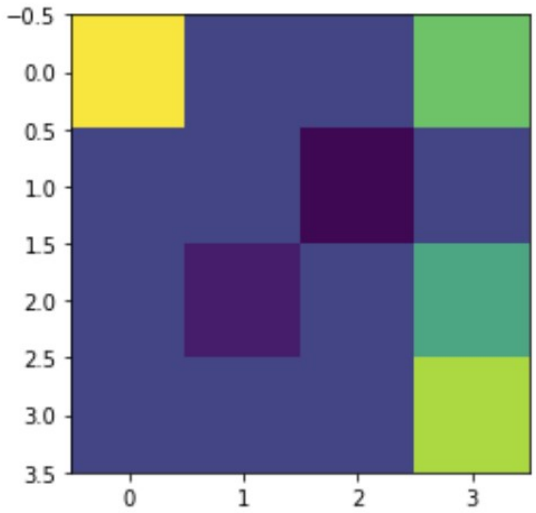
We have used the same set of States, Actions and Rewards for both Deterministic and Stochastic Environments.

States(S) = { (0,0), (0,1), (0,2), (0,3), (1,0), (1,1), (1,2), (1,3), (2,0), (2,1), (2,2), (2,3), (3,0), (3,1), (3,2), (3,3) }

Actions(A) = { Up, Down, Left, Right }

Rewards(R) = { -5, -1, 3, 6, 10 }

1.2 Provide visualizations of your environments



(0,0) - Agent start position - Zero Reward
(0,3) - Diamond position - Positive Reward
(2,3) - Coin position - Positive Reward
(1,2) - Fire position - Negative Reward
(2,1) - Wall position - Negative Reward
(3,3) - End position - Positive Reward

1.3 How did you define the stochastic environment?

At each timestep, a random action is picked. We have another variable - **temp** - defined inside the environment that picks a random value between 0 and 1. If **temp** is below **0.3**, then the agent sticks with the random action that was chosen above. Else, the agent performs any action that is picked randomly from the **Actions Set excluding the previously picked action**. This stochasticity is only present in 3 of 16 possible states of the agent, at (0,3), (2,2), and (3,0).

1.4 What is the difference between the deterministic and stochastic environments?

Deterministic Environment	**Stochastic Environment**
The next state of the agent, given a state and action, is always certain at any point in time	For a given action and state, there's no certainty regarding the agent's next state
A real life example would be - traffic signal. There is a certainty with respect to the action taken by the agent. Red signals stop. Green signals go.	A real life example would be - playing soccer.The next action and thereby the next state of a soccer player is unpredictable.
$P(s', r\|s, a) = \{0, 1\}$	$\sum_{s',r} P(s', r\|s, a) = 1$

1.5 Safety in AI: Write a brief review (5 sentences) explaining how you ensure the safety of your environments. E.g. how do you ensure that agent choose only actions that are allowed, that agent is navigating within defined state-space, etc

We ensure that the agent is navigating within the defined state-space by using the **np.clip()** function. In our environment, its used as - **np.clip(agent_pos,0,3)**. This keeps the row and column values of the agent between 0 and 3.

Our Action set consists of only 4 actions - {Left, Right, Up, Down} - which are denoted by integers between 0 and 3. The agent is made to pick a random action(integer) by using the **np.random.choice()** function. As an example, **np.random.choice(5,3)** picks 3 integers between 0 and 5.