
Defining & Solving RL Environments

Sivashanmugam Saravanan Omprakash
Department of Mechanical & Aerospace Engineering
University at Buffalo (SUNY)
Buffalo, NY 14260
saravan3@buffalo.edu

Abstract

The goal is to acquire experience in defining and solving RL environments, following OpenAI Gym standards. First, we focus on defining deterministic & stochastic environments that are based on Markov Decision Process (MDP).

1. Defining RL Environments

Everything the *agent* interacts with, comprising everything outside the agent, is called the *environment*. These interact continually, the agent selecting actions and the environment responding to these actions and presenting new situations to the agent ^[1].

Report

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

	Deterministic Environment		Stochastic Environment	
	Number	Set	Number	Set
Action	4	{Left, Up, Right, Down}	4	{Left, Up, Right, Down}
States	9	{{(0, 0), (0, 1), (0, 2), (1, 0), (1, 2), (1, 2), (2, 0), (2, 1), (2, 2)}	9	{{(0, 0), (0, 1), (0, 2), (1, 0), (1, 2), (1, 2), (2, 0), (2, 1), (2, 2)}
Rewards	4	{-5, 1, 5, 10}	4	{-5, 1, 5, 10}

2. Provide visualizations of your environment.

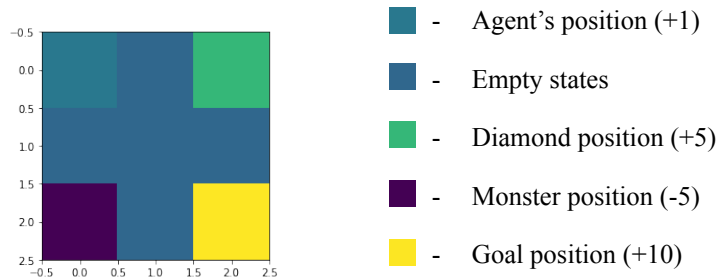


Figure 1: Default environment layout

Figure 1 shows the layout of the environment. There are 9 states, 4 actions allowed, & 4 reward signals placed in the environment.

3. *How did you define the stochastic environment?*

Stochasticity for the environment has been defined under the step function by choosing a sample from a uniform distribution between 0.0 & 1.0. If the random sample is greater than 0.75 the given action is executed, else an action from the action space is chosen at random.

4. *What is the difference the deterministic & stochastic environment?*

Deterministic Environment	Stochastic Environment
$P(s', r s, a) = \{0,1\}$	$\sum_{s',r} P(s', r s, a) = 1$
Outcome can be determined based on state	Outcome cannot be determined for certain based on state
Example: Making a move in Chess	Example: Rolling a die

5. **Safety in AI:** *Write a brief review explaining how you ensure the safety of your environments.*

In order to keep the agent within the deterministic/stochastic grid environment, the actions that makes the agent to exit the environment is clipped to the min & max limits of the grid. This means that, even if such an action is executed, the agent remains in its current state.