RL Crash Course + Some Thoughts

Ding (Eric) Ding

November 2022

1 Introduction

In this Project, I followed OpenAI deep reinforcement learning course, and used Spinning Up for exercises and experiments.

2 What's Omitted

Regularization, Observation normalization

3 Key Concepts

The main characters of RL are the agent and the environment. The environment is the world that the agent lives in and interacts with. At every step of interaction, the agent sees a (possibly partial) observation of the state of the world, and then decides on an action to take. The environment changes when the agent acts on it, but may also change on its own.

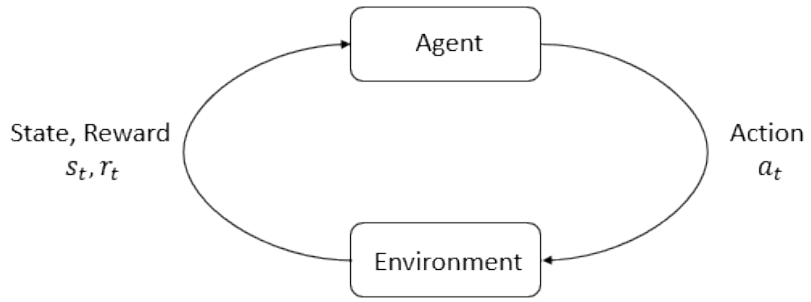

Figure 1: RL framework

The agent also perceives a reward signal from the environment, a number that tells it how good or bad the current world state is. The goal of the agent

is to maximize its cumulative reward, called return. Reinforcement learning methods are ways that the agent can learn behaviors to achieve its goal.

States and observations: represented by matrix

Action Spaces: the set of all valid actions. Discrete / continuous.

Policies: a rule used by an agent to decide what actions to take. Deterministic: $a_t = \mu_\theta(s_t)$. Stochastic: $a_t \approx \pi_\theta(\cdot|s_t)$. In deep RL, we deal with parametrized policies, with parameters that can be adjusted by optimization.

4 Ideas

4.1 Observation

Fully observed environment vs. partially observed

4.2 Action

Rule out certain actions in action space