

# Reinforcement Learning Assignment 5

Maida Cao (TA), Prof. Zou

2021-05-13

## 1 Introduction

Actor-Critic method reduces the variance in Monte-Carlo policy gradient by directly estimating the action-value function. The goal of this assignment is to do experiment with Asynchronous Advantage Actor-Critic (A3C). Due to the inherent advantage of PG method, AC method is able to tackle the environment with continuous action space. However, a naive application of AC method with neural network approximation is unstable for challenging problem. In this assignment, you're required to train the agent with continuous action space and have some fun in some classical RL continuous control scenarios.

## 2 Actor-Critic Algorithm

In original policy gradient  $\nabla_{\theta} \log \pi_{\theta}(s_t, a_t) v_t$ , return  $v_t$  is the unbiased estimation of expected long-term value  $Q^{\pi_{\theta}}(s, a)$  following a policy  $\pi_{\theta}(s)$  (Actor). However, original policy gradient suffers from high variance. Actor-Critic Algorithm uses Q value function  $Q_w(s, a)$ , named Critic, to estimate  $Q^{\pi_{\theta}}(s, a)$ .

In A3C, we maintain several instances of local agent and a global agent. Instead of experience replay, we asynchronously execute all the local agents in parallel. The parameter of the global agent is updated by all the local experience.

The pseudo code is listed at the end of the article. For more details of the algorithm, you can refer to the original paper in the following.

- Mnih V, Badia A P, Mirza M, et al. Asynchronous methods for deep reinforcement learning[C]//International conference on machine learning. 2016: 1928-1937.

## 3 Experiment Description

- Programming language: python3
- You are required to implement A3C algorithm.
- You should test your agent in a classical RL control environment–Pendulum. OPENAI gym provides this environment, which is implemented with python (<https://github.com/openai/gym/wiki/Pendulum-v0>).

## 4 Report and Submission

- Your report and source code should be compressed and named after “studentID+name+assignment5”.
- The files should be submitted on Canvas before June 21, 2021.