CSE 573 - Project Proposal

Nicholas Andrews
Joseph David

May 6, 2022

1   Motivation

Reinforcement learning is one of the three basic machine learning paradigms, alongside supervised and unsupervised learning. The main characteristic of reinforcement learning as opposed to the other types of learning is that no prior information or training data is given to the agent. Instead, the agent simply learns through experience. The agent is assumed to know the state of its environment at any given point and seeks to maximize its own cumulative reward.

OpenAI gym is a collection of simulations designed for learning agents to solve as benchmarks in order to compare to other reinforcement models. The simulations are designed such that an intelligent agent should learn to complete any of the simulated tasks. For this project we will solve various OpenAI gym environment problems, such as Cart-Pole, by implementing the REINFORCE and A3C algorithms for reinforcement learning.

The REINFORCE algorithm is responsible for updating our policy network after the agent has completed (or failed) the task and received rewards. The A3C algorithm introduces a valuation network that is used by multiple learning agents (in the same environment) to update their policy networks [1]. These learning agents share their experiences in a global network, which they all have access to.

2   Simulation

We will use the Jupyter Notebook provided by pjreddie, aka Joseph Redmon, which includes a code skeleton for getting an agent up and running on OpenAI gym, but leaves out the important algorithm implementations for reinforcement learning [2]. We will have access to a GPU through Google Colab, which alleviates the hardware constraints since an agent can require many hours or days of training. As we are using a Jupyter Notebook, our algorithms will be implemented in Python. We will use the library Pytorch in order to create the policy's neural network and for sampling actions from the policy.

3   Timeline

Our project milestones and timeline are shown in the table below.

Task	Sub-task	Start	Due
Project Proposal	-	-	5/6
Jupyter Notebook	-	5/6	6/1
-	Actor	5/6	5/13
-	REINFORCE	5/13	5/18
-	Critic	5/18	5/22
-	A3C	5/22	5/29
-	Cartpole Testing	5/29	6/1
Lunar Lander	-	6/1	6/4
Project Presentation	-	6/4	6/8
Project Report	-	6/4	6/10

References

1.   Mnih, V. *et al.* Asynchronous Methods for Deep Reinforcement Learning. `https://arxiv.org/abs/1602.01783` (2016).

2.   Redmon, J. *Reinforcement Learning Homework* `https://github.com/pjreddie/rl-hw`. 2018.