Study Notes of RL: Implementation

Elynn Chen[1]

[1]ORFE, Princeton University
[1]OpenAI

2019-03-01

Abstract

This article summarizes implementation of RL algorithms.

Key words: Reinforcement learning, Multi-Armed Bandits (MAB), Finite Markov Decision Processes (MDP), dynamic programming, Monte Carlo methods, Temporal-Difference (TD) learning, value function, optimal policy, Generalized Policy Iteration (GPI), sample updates, expected updates, bootstrapping, on-policy and off-policy, .

1 AWS Setup

To set up AWS, follow the tutorial by FastAI. Once an instance is created and started, use the following codes to start virtual environment and launch Jupyter notebook.

Listing 1: Connect to AWS from Local Machine

```
ssh -L localhost:8888:localhost:8888 ubuntu@<your instance IP>
source activate conda_tensorflow_p36
jupyter notebook
```

The interactive productivity booster for the AWS CLI provides an useful tool for using AWS.

2 Google CoLab

This is good for play around with free computing resources. However, you only have 12 hours for each session and no storage on the server. You need to install necessary packages every time you start. Use the following in .ipynb and then play around with gym.

Listing 2: Google CoLab

```
# install necessary packages
!pip install gym
!git clone https://github.com/openai/spinningup.git
!pip install -e spinningup
!pip install tensorflow==1.12.0
!pip install gym pyvirtualdisplay > /dev/null 2>&1
!apt-get install -y xvfb python-opengl ffmpeg > /dev/null 2>&1

# import packages
from spinup import ppo
import tensorflow as tf
import gym
```

3 Environment Setup

This environment setup follows the instruction here.

3.1 Local

Installation of system packages on Mac requires Homebrew.

Listing 3: Setup SpinningUp Environment

```
conda create -n spinningup python=3.6
```

```
# you need to activate the environment before your start
source activate spinningup

# install openMPI, need first install Homebrew.
brew install openmpi

# installing spinningup
git clone https://github.com/openai/spinningup.git
cd spinningup
pip install -e .

# check install
# try running PPO in the LunarLander-v2 environment with
python -m spinup.run ppo --hid "[32,32]" --env LunarLander-v2 --exp_name instal

# After it finishes training, watch a video of the trained policy with
python -m spinup.run test_policy data/installtest/installtest_s0

# And plot the results with
python -m spinup.run plot data/installtest/installtest_s0

# The last step somehow went wrong with error, will check back later.
```

3.2   Cloud (AWS)

Basically the same procedure once you log on to the VM through terminal. The only difference is the operating system is Ubuntu. To install "OpenMPI", you will need to use

Listing 4: Installing OpenMPI on Ubuntu

```
sudo apt-get update && sudo apt-get install libopenmpi-dev
```

4   Implementation Plans for Next Week

4.1   Mar 4th – Mar 8th.

Next week, I am going to delve deeper into the implementation of deep RL. This include the following algorithms.

- Deep Q-Network (DQN)

- Deep Deterministic Policy Gradients (DDPG)

To serve as a guideline, The following algorithms are implemented in the Spinning Up package.

- Vanilla Policy Gradient (VPG)

- Trust Region Policy Optimization (TRPO)

- Proximal Policy Optimization (PPO)

- Deep Deterministic Policy Gradient (DDPG)

- Twin Delayed DDPG (TD3)

- Soft Actor-Critic (SAC)

5   Useful Resources

OpenAI Curriculum

OpenAI Spinning Up in Deep RL

SpinningUp GitHub Repo

OpenAI Spinning Up

References

Sutton, R. S. and A. G. Barto (2018). *Reinforcement learning: An introduction.* MIT press.

Appendix A Proof

Appendix B Table and Plot