Assignment 1 Part 2

ELEN6885 Reinforcement Learning

October 2018

1 Introduction

1.1 Background

Winter is here. You and your friends were tossing around a frisbee at the park when you made a wild throw that left the frisbee out in the middle of the lake. The water is mostly frozen, but there are a few holes where the ice has melted. If you step into one of those holes, you'll fall into the freezing water. At this time, there's an international frisbee shortage, so it's absolutely imperative that you navigate across the lake and retrieve the disc. **However, the ice is slippery, so you won't always move in the direction you intend.** The episode ends when you reach the goal or fall in a hole.[3]

1.2 Mathematical Model

The states form a 4 * 4 grid. There are 4 kinds of states. "S" is the safe starting point, "F" represents frozen surface, which is safe as well. "H" represents a hole. You "fall to your doom" if you enter "H" states. "G" is your goal where the frisbee is located.

S	F	F	F
F	H	F	H
F	F	F	H
H	F	F	G

The indices of states are shown below.

00	01	02	03
04	05	06	07
08	09	10	11
12	13	14	15

At each step, you can take 4 actions: "LEFT", "DOWN", "RIGHT", "UP" represented by indices 0 to 3 respectively. Your next state is then given by the environment. The episode ends when you reach the goal or fall in a hole. You receive a reward of 1 if you reach the goal, and 0 otherwise.[1]

2 Task

In this assignment, you need to implement the following algorithms: Epsilon-Greedy, Policy Iteration, Value Iteration, Q-Learning and SARSA. Open Assignment-1-Part-2.ipynb with Jupyter Notebook and follow the detailed instructions step by step. We also have gym-tutorial.ipynb for you to get familiar with Gym package.

3 Environment Setup

3.1 Download Anaconda

Anaconda[2] is a free and open source distribution of the Python and other languages for data science and machine learning related applications. It also includes a package management system that can help you easily deal with your packages. It is suitable for Windows, Linux and MacOS X. Download Python 3.6 version of Anaconda Navigator. We only provide support and test your code on Python 3.6. Python 3.5 should work. If you have already installed Anaconda 2, head to Section 3.3 to create a Python 3.6 environment.

3.2 Install Anaconda

Follow the installation guide to install Anaconda Navigator.

Anaconda installs a new copy of Python 3.6 independent from your system Python. To verify your installation, you can type the following commands in console (command line for windows).

```
python -V
conda list
```

"python -V" outputs your current default Python version. "conda list" outputs a list of packages already installed.

3.3 Create a new virtual environment

A virtual environment can isolate package management from different projects. By using virtual environment, people can install different packages for different projects without affecting each other.

Select "Environments" in the sidebar. Click "create" at the bottom. Give your new environment a name ("RL" as an example), choose "3.6" for Python version and create. Once the creation is done, open your terminal / CMD, type in the following commands to install necessary packages for this assignment.

```
source activate <the name of the new environment>
pip install --upgrade pip
pip install --upgrade gym
pip install --upgrade scipy
pip install --upgrade matplotlib
```

Please don't use third party libraries other than Numpy, Gym and Matplotlib. Otherwise, we may get an error when executing your code.

3.4 Jupyter Notebook with Anaconda Navigator

Jupyter Notebook is installed for the base (root) environment. But you need to manually install Jupyter Notebook for every environment you create. Click "Home" in the sidebar, switch to the new environment and install Jupyter Notebook.

After installation, click "Launch", Jupyter Notebook is open in a browser.

4 Submission

After you finish your coding, open Assignment-1-Part-2.ipynb. Restart the kernel and run all the cells to get all of the outputs. (Comment all the cells that play the game). Save this notebook as a pdf file. Then restart the kernel again and clear all of the outputs. Zip the notebook, the generated pdf and the RLalgs folder into a package. Rename it as [Assignment1-FirstName-LastName-UNI.zip] (e.g. Assignment1-Minhui-Li-ml4026.zip) and upload it on Courseworks.

5 Bug Report

Minhui Li ml4026@columbia.edu

6 Acknowledgement

This assignment referenced some parts of the Assignment 1 in the year 2017, designed by Prof. Chong Li, Chen-Yu Yen, Lingyu Zhang, Xing Yuan and Qing Lan.

References

[1] OpenAI Gym. Frozenlake-v0 source. `https://github.com/openai/gym/blob/master/gym/envs/toy_text/frozen_lake.py`.

[2] Anaconda Inc. Anaconda. `https://www.anaconda.com/`.

[3] OpenAI. Frozenlake-v0. `https://gym.openai.com/envs/FrozenLake-v0/`.