

Homework #1

COSE490(02) Reinforcement Learning

Goal: The goal of the first homework is (i) to let the student familiar with the simulator environment, called Open AI Gym, and (ii) to understand the basic programming structure for the RL.

Tasks: Use the provided file, and run the program in there. Understand how one can implement policy evaluation, policy improvement, policy iteration, and value iteration.

- Task 1: Construct program environment to run the provided program (e.g., python, jupyter notebook, and programs for RL and Open AI Gym). The provided file is in the format of “jupyter notebook”
- Task 2: Make the (provided) program run without error.

Submission:

- After running the program (all the cells), print the results as a PDF file.
- Submit the PDF file through BlackBoard (BB).
- Submission is due on **March 24 (before the class hour)**. *Feb.*

Some useful tips:

- Too recent version of python may not work. For example, I had troubles with Python 3.9.2, while Python 3.8 works fine (Win10).
- If you get stuck and desperate for a help, please contact the TA through email (antaegun20@korea.ac.kr). Please start the email title as “[COSE490(02)]”

The below is some helpful information, if you are not familiar with RL programming.

0. What is Open AI Gym?

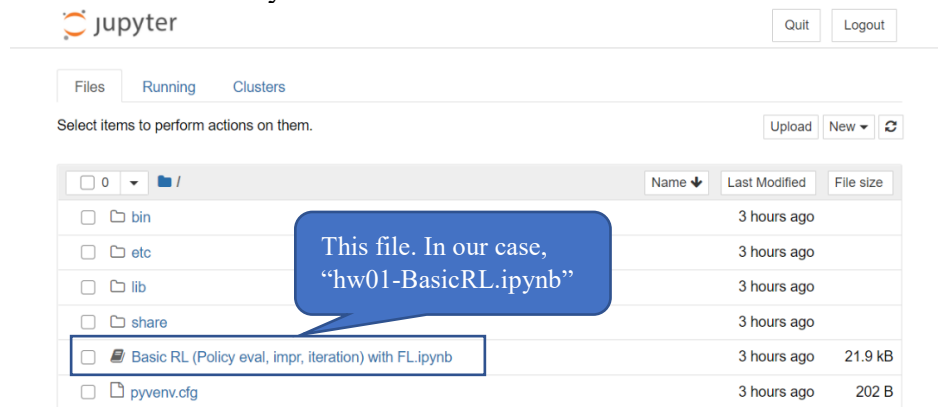
Open AI Gym is a toolkit for developing and comparing reinforcement learning algorithms. There are many docs that you can find at <https://gym.openai.com/docs/>

1. Basic installation guide

A. Ubuntu (20.04LTS)

- Python requirements
 - Python 3.5+ (check your python version with `python -V` command)
- Installation using pip
 - `pip install gym`
 - `pip install git+https://github.com/mimoralea/gym-walk#egg=gym-walk`
- Installation Python using source
 - `git clone https://github.com/openai/gym`
 - `cd gym`
 - `pip install -e .`
`pip install -e .[all]`: this command is for full installation; will require more dependencies, such as cmake or pip version

- iv. Install jupyter notebook
 - pip install jupyter notebook
 - Download the provided homework file and move it to virtual environment directory.
 - Type “jupyter notebook” in console or terminal → This will run the jupyter notebook and show you some screen like below.



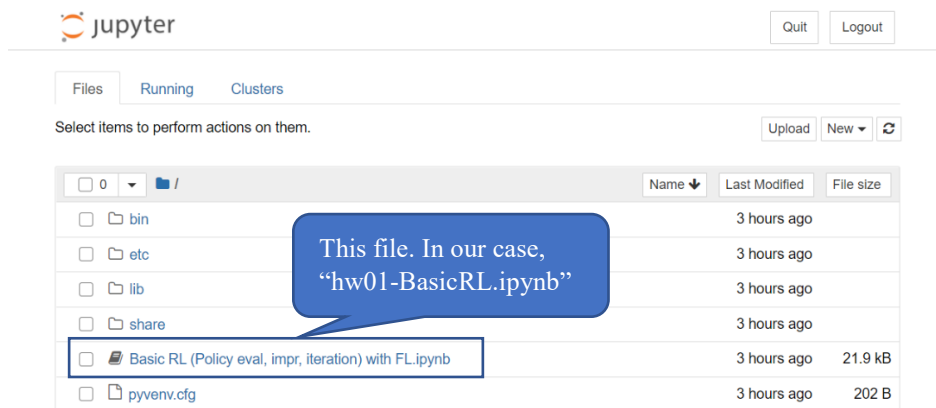
- Use the Internet, and figure out how to use “jupyter notebook” to create / edit / run programs.

B. Windows: Use ANACONDA

- i. Search anaconda in google – download 64bit Graphical installer – install & open anaconda prompt
- ii. Making virtual environment
 - conda create -n [env_name]
- iii. Activating virtual environment
 - conda activate [env_name]
 - will change (base) -> (env_name)
 - deactivate
 - will change (base) -> (env_name)
 - can check env names using ‘conda env list’
- iv. Installation using conda
 - conda install gym
 - conda install git+https://github.com/mimoralea/gym-walk#egg=gym-walk
 - conda install <package_name> (e.g. numpy)
- v. Install jupyter notebook
 - conda install jupyter notebook
 - Download homework file and move it to virtual environment directory.
 - Open a console by type “cmd” in the search space below.



- Type “jupyter notebook” in console or terminal → This will run the jupyter notebook and show you some screen like below.



- Use the Internet, and figure out how to use “jupyter notebook” to create / edit / run programs.

2. Installation (for program libraries for RL)

To run a basic RL program, there are several libraries that you need to install, e.g., “numpy”, etc. Once you try to run the provided program in the jupyter notebook, you may meet many errors for the missing components. Please use the power of Google, and find how to install those libraries. In most case, typing a few commands is all what you need.

3. A bit of information about the Environment for the games in the provided file

A. FrozenLake-v0

- Reference: [Gym \(openai.com\)](https://gym.openai.com/)
- Deterministic environment
 - Board (example)

```
SFFF      (S: starting point, safe)
FHFH      (F: frozen surface, safe)
FFFH      (H: hole, fall to your doom)
HFFG      (G: goal, where the frisbee is located)
```

- State: your position on board
- Action: move ($\uparrow \downarrow \leftarrow \rightarrow$)
- Reward: **reach G and get reward 1**
- iii. Policy Iteration
 - With current policy (Random or Go-get), play on map and get experience.
- iv. Policy Evaluation
 - Get state-value function from experience in policy iteration.
- v. Policy Improvement
 - With achieved state-values function, update action-value Q

B. SlippyWalkFive-v0

- Reference: <https://github.com/mimoralea/gdrl>
- Stochastic Environment
 - Board

Terminal	1	2	3	Terminal
----------	---	---	---	----------

- Agent starts at .2
- Action: move ($\leftarrow \rightarrow$),
stochastic (success: $\frac{1}{2}$, stay: $\frac{1}{3}$, backward $\frac{1}{6}$)
- Reward: When you reach the right Terminal, and you get reward 1.