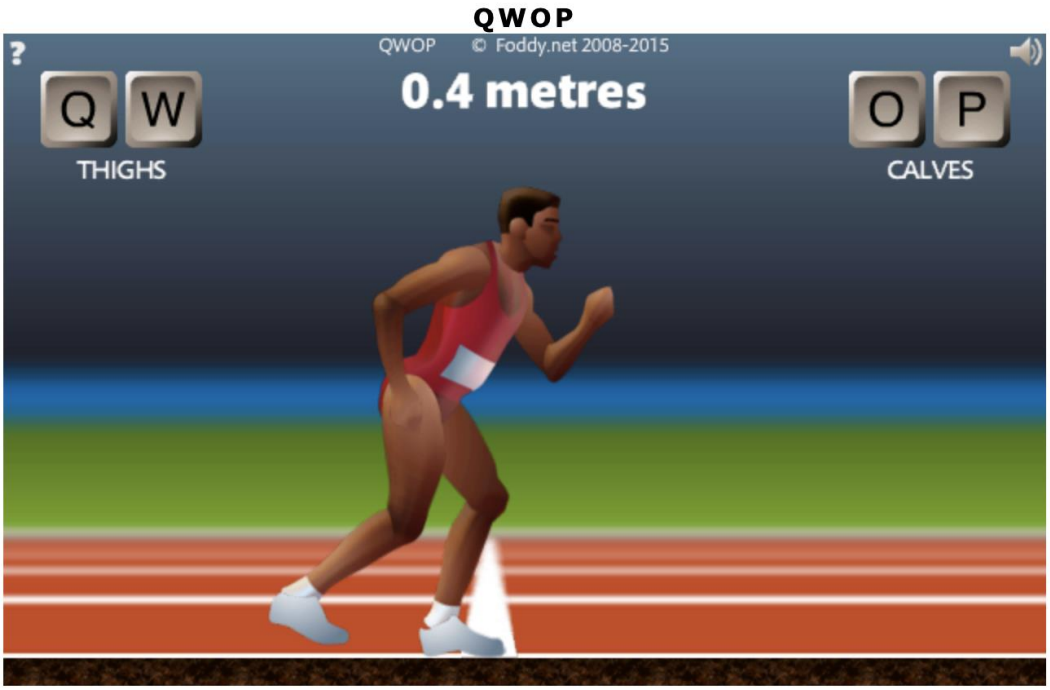
Deep Reinforcement Learning

Homework #3

Due: 2021/05/03 (Mon.) 23:59

Problem Description

1. QWOP (“`<Student_ID>_hw3_<train|test>.py`”)
   QWOP is a relatively well-known game made in Adobe Flash, available for free on the Internet. Your goal is to play as the sprinter running a 100 meter dash. You are given direct control over four muscle groups. Q, W move the runner’s left right thighs. O, P move the runner’s left right calves.



   You need to train your agent on a localhost env “`QWOP`”.

   a. How to host an environment in local, please refer to QWOP on Github.
   b. State space: 71 dimensions with limb (e.g., arm, torso, leg) coordinates, angle with y-axis, angle between adjacent limbs and current angular velocities, etc.
   c. Action space: 11 dimensions from no action to at most 2 keydowns. {‘’, ‘q’, ‘w’, ‘o’, ‘p’, ‘qw’, ‘qo’, ‘qp’, ‘wo’, ‘wp’, ‘op’}
   d. Implement your policy with any kind of DQN variants.
   e. You must write all your training and testing codes by yourself.
   f. You may store your learned results in an external file “`./<Student_ID>_hw3_data`” (with any file extension, max 50 MB), and access it with your program (during testing).
   g. After the deadline, TAs will help to compare your agent with your classmates, and then grade it.

Detailed Rules for QWOP

1. You should implement an "`act(observation)`" function in your testing code, inside the `Agent` class. See "`random_agent.py`" for an example.
2. If your program outputs invalid moves, you lose and the game ends immediately.
3. The time limit for each move is **1 second**, and the memory limit is **4 GB**. (Note that the 1-second duration may vary depending on different processors. If you use a DQN-based agent and doesn't perform additional calculations during inference, you don't need to worry about the time limit.)
4. You are allowed to access an external file for loading your learned policy. You can read the file at the following path: **"`./<Student_ID>_hw3_data`"**. (with any file extension)
5. You are not allowed to use the expert trajectories provided in QWOP.
6. You are allowed to use the following Python packages:
    a. numpy, scipy, gym, pandas, **tensorflow**, **pytorch** and the packages mentioned in the environment's repo.
    b. You are allowed to use Python's default installed packages. (e.g., sys, time, pickle, random, etc.)
    c. If you need to use other packages, state your reasons and post them on iLMS.
7. You are not allowed to use the following Python package:
    a. `stable_baselines, openai_baseline, or other RL baselines implemented by other people`

Program Submission

1. For each problem, please use Python to implement with a single source file.
2. Your files must be named as:
    a. **"`<Student_ID>_hw3_train.py`"**
    b. **"`<Student_ID>_hw3_test.py`"**
    c. **"`<Student_ID>_hw3_data`"**
    d. Please make sure that all characters of the filename are in lower case. For example, if your student id is `108062000`, the name of your program file should be `108062000_hw3_train.py` and so on.
1. Your program will be run on a GNU/Linux environment with Python 3.8: `python <Student_ID>_hw3_test.py`
2. 0 points will be given to Plagiarism. NEVER SHOW YOUR CODE to others and you must write your code by yourself. If the codes are similar to other people and you can't explain your code properly, you will be identified as plagiarism.
3. 0 points will be given if you violate the rules above.
4. If you use modularized / OOP code and want to use multiple files to keep your code structured, please upload it along with the 3 files above.

Report

1. Elaborate on how you design your agent. What advanced techniques of DQN have you used? The report is graded directly. Therefore, make sure you have included enough details and figures to help TAs understand and grade your report.
2. TAs will not refer to your code when grading your report, so make sure you have taken a screenshot of the important code snippets.
3. The report filename must be **"`<Student_ID>_hw3_report.pdf`"**. Please make sure that all characters of the filename are in lower case.

Grading Policy

1. The project accounts for 10 points (tentative) of your total grade.
2. You must submit both your source code and report. Remember the submission rules mentioned above, or your grade will be penalized. Late submission rules are specified in the Lecture 1 Slides.
3. Compress all your files directly (do not compress the folder containing your files) and upload to this Google Form before the deadline. (Total 4 files)
4. The baseline agent will not be released. Your code will be tested against them after the submission deadline.

- QWOP Environment
  - Reach the destination within a time limit of 120s, 1 meter for 0.6 points (60 %)
  - You rank compared against your classmates (faster is better) (20%)
  - Report (including all the discussions, experimental setups, experimental results, analysis, etc.) (20%)
- Extra points
  - Play the QWOP game by your hand, 1 meter for 0.05 point (maximum 10 points will be given) (5 %)
    - You only have one shot (i.e., you can't try multiple times)
    - You need to demonstrate your skill and play the game in class
    - Use your own laptop