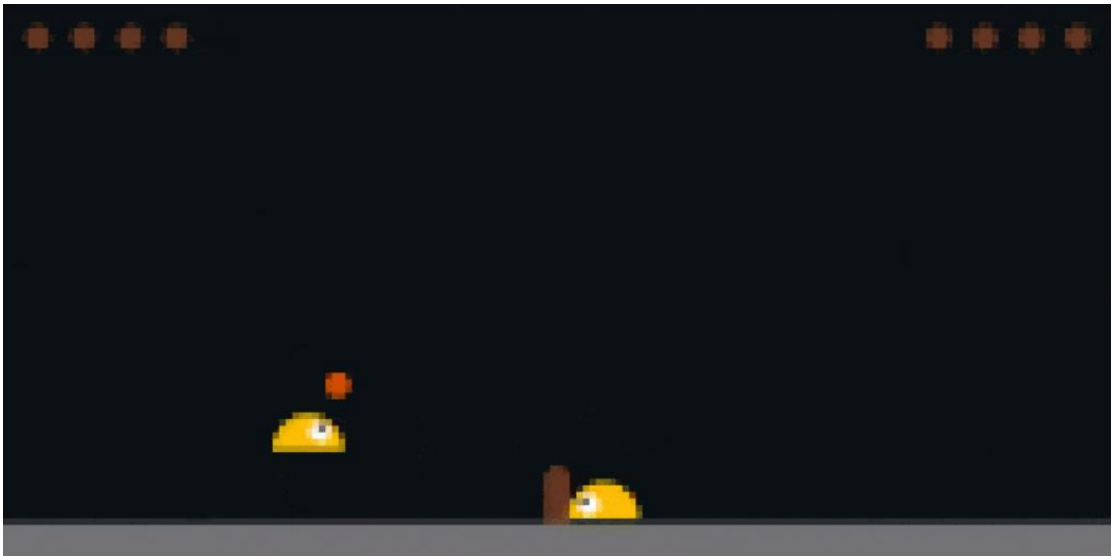
Deep Reinforcement Learning

Homework #2

Due: 2021/04/12 (Mon.) 23:59

Problem Description

1. Slime Volleyball (“`<Student_ID>_hw2_<train|test>.py`”)
   Please refer to the [Slime Volleyball Gym Environment] on GitHub:



   a. 3 types of environments:

Environment Id	Observation Space	Action Space
SlimeVolley-v0	Box(12)	MultiBinary(3)
SlimeVolleyPixel-v0	Box(84, 168, 3)	MultiBinary(3)
SlimeVolleyNoFrameskip-v0	Box(84, 168, 3)	Discrete(6)

   You need to train your agent on “`SlimeVolleyNoFrameskip-v0`”.
   b. Implement your policy with DQN.
   c. Other improvements are allowed (optional).
   d. You must write all your training and testing code by yourself.
   e. You may store your learned results in an external file “`./<Student_ID>_hw2_data`” (readonly, max 50 MB), and access it with your program (during testing).
   f. After deadline, TAs will help to compare your agent with your classmate and grade it.

Detailed Rules for Slime Volleyball

1. You should implement an "`act(observation, reward, done)`" function in your testing code. See "`random_agent.py`" for an example.
2. If your program outputs invalid moves, you lose and the game ends immediately.
3. 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>_hw2_data`".
5. You are allowed to use the following Python package:
    a. numpy, scipy, gym, pandas, **tensorflow**, **pytorch** and the packages mentioned in the environment's repo.
    b. You are allowed to use Python default installed packages. (e.g., sys, time, pickle, random, etc.)
    c. If you need to use other packages, state your reasons and post on iLMS.

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>_hw2_train.py`"
    b. "`<Student_ID>_hw2_test.py`"
    c. "`<Student_ID>_hw2_data`"
    d. and 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_hw2_train.py` and so on.
1. Your program will be ran in a GNU/Linux environment with Python 3.8: `python <Student_ID>_hw2_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 do you used? The report is graded directly. So, make sure you have included enough details and figures to help TAs 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>_hw2_report.pdf`" and please make sure that all characters of the filename are in lower case.

Grading Policy

1. The project accounts for 15 points (tentative) of your total grade.
2. You must submit both your source code and report. Remember the submission rules mentioned above, or you will be punished on your grade. 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.

- Slime Volleyball Gym Environment
  - Beat Baselines agent (25%)
  - Rank of the compared results (35%)
  - Report (Discussion) (40%)