OffWorld, Inc.
2031 E. Foothill Blvd., Pasadena, CA 91107, USA
www.offworld.ai

OFFWORLD PRACTICAL CHALLENGE ON REINFORCEMENT LEARNING AND ROBOTICS

This document presents a set of practical assignments on the topic of reinforcement learning and robotics to evaluate the skill of training reinforcement learning agents. Apart from being a technical challenge it also serves as a good introduction to the type of work an RL engineer might be doing at OffWorld, while working on an array of projects under the general umbrella of applying machine learning and reinforcement learning techniques to robotics with the ultimate goal of developing an AI that is sufficient to facilitate self-learning real-world robots.

The challenge consists of three parts of increasing complexity:

- setting up the environment and making sure all the components of the system are ready and functional,
- training an agent to solve one of the tasks in the OffWorld Gym environment by training your favorite RL algorithm end-to-end in simulation, reporting its performance and sample complexity, and finally
- training an agent on a real physical robot, by adapting the agent pre-trained in simulation to the real environment or using human demonstrations to reduce the sample complexity of the end-to-end approach.

Please see the following sections for the detailed explanation of what you can try to do and what is expected.

As part of your submission please provide (where relevant) the code that you used to reach the solution including instructions on how to run this code, screenshots and/or videos of the critical steps and results, and a report that outlines the step that you took to solve each part of the challenge and an explanation why you opted to using one or another algorithm or approach. You are given one week to complete the exercise starting from the date of your choosing.

Note: OffWorld Gym is running in Beta mode, if you encounter issues with the library or the physical environment please let us know, we will pause the challenge clock, fix the issue and resume. Since the physical robot is only available to one person at a time, we will factor this in and provide additional time for training if necessary.

SETUP THE ENVIRONMENT AND RUN A TEST

In this introductory task you will need to get familiar with the OffWorld Gym framework (https://gym.offworld.ai and https://github.com/offworld-projects/offworld-gym) and make sure you can run it on your system.

Please follow the installation steps available at https://gym.offworld.ai/docs/installation.html and showcase the successful completion with a couple of screenshots or a video of both the simulated and the real versions of the environment.

TRY TO SOLVE AN OFFWORLD GYM ENVIRONMENT END-TO-END

At the moment there are two environments available in OffWorld Gym you can choose from: `OffWorldDockerMonolithDiscreteSim`, `OffWorldDockerMonolithContinuousSim`. Both are the versions of the same task where the robot has to approach a visual beacon in the middle of the environment by relying on visual inputs only. The episode is reset, and the robot is sent to a random location, if it hits the boundary of the environment or exceeds the maximal episode length of 100 steps. The only reward signal is a sparse reward of +1 for approaching the monolith within a certain radius. The only difference between the environments is the action space: four discrete actions or two continuous velocities respectively. Please have a look at the examples provided in the repository: https://github.com/offworld-projects/offworld-gym/tree/develop/examples.

In this task you can use any of your favorite (open access) deep and/or reinforcement learning libraries and readily available implementations of RL algorithms, with the exception of (D)DQN because Double DQN solution is already implemented in our examples (Rainbow is allowed).

The goal is to achieve intelligent behavior, report the choices you have made along the way, report the learning curves, the performance and the sample complexity and record a video of the resulting agent.

In our experience training an agent in simulation on a single workstation takes at least 5 hours, so we recommend starting early.

TRAIN AN AGENT ON A PHYSICAL ROBOT USING SIM-TO-REAL TRANSFER FROM YOUR PREVIOUS AGENT OR WITH HUMAN DEMONSTRATIONS

We highly recommend that the code you use for training in this task can tolerate interruptions, snapshotting networks weights (and the replay buffer) to be able to resume the training if the robot stalls, you lose the connection with the environment, etc.

The approach you took in simulation is bound to take much longer to train on a real robot (in our experiments at least 2 days with the `OffWorldMonolithDiscreteReal` and `OffWorldMonolithContinuousReal` environments), so for real-world applications of reinforcement learning we need to rely on some advanced techniques to make the task feasible.

One method we find useful is adapting an agent you have trained in simulation to the real physical world. We tried to make the dynamics and visual features as similar as possible, however there are significant differences as well, so closing the “reality gap” presents an interesting challenge.

Another approach would be to help the agent skip the time-consuming process of initial random exploration and use human expert inputs to generate some good initial data.

While working on this part of the challenge you can also try running the architecture you have created for the previous task on the real robot directly (and let it run in the background), measuring the learning speed and performance and then comparing it to the results you will obtain via sim-to-real or human demonstrations.

We appreciate that this challenge has many variables at play, and some insignificant detail, or an unfortunate random seed can be the cause of no intelligent behavior emerging. Please do not be discouraged from submitting if that happens, at this stage we are more interested in learning how you approach a problem and structure your thinking, reporting, and code.

Please send you solutions to ilya.kuzovkin@offworld.ai.

We hope you enjoy this exercise, good luck!