David Tandetnik
CS4100
3/8/20

Project Proposal: Robot Hand Manipulation with OpenAI Gym

INTRODUCTION

For this project I'd like to work on the problem of manipulating a block using a robotic hand, as provided by the OpenAI Gym environment *HandManipulateBlock*. The goal of this problem is to identify a series of movements with the robot arm's actuators that orient a block, held in the robot's hand, into an arbitrary position. More formally, the input of this problem is an initial hand position and box orientation, and the output is a series of movements that result in a goal box orientation. This problem is interesting because of how it directly maps to real-world scenarios of robot automation; one can imagine that the ability to have a robot hand move a box to an arbitrary position would be useful in industries such as shipping and manufacturing.

APPROACH

The approach to solving this problem will be based on reinforcement learning, as it is a useful technique for teaching robots how to interact with physical environments. The simulation of this environment will be provided by OpenAI's Gym toolkit in the form of one of their recently released robot environments.

I intend to start with the RL techniques we have learned in class, such as Q-Learning and SARSA. However, given the size of the state space, I anticipate that without massive simplifications of the problem scope, these two approaches will not work particularly well. In their research, OpenAI researchers did attempt reduced goal definitions at first, such as only having the robot rotate the block along one axis, before progressing to more difficult goal definitions (1). (The subsequent goals are naturally achieving proper rotation in two axes, then three axes, and finally proper rotation in three axes as well as proper position in 3D space.) I will likely also progress in this order.

OpenAI researchers also used the Deep Q Networks and Deep Deterministic Policy Gradient algorithms in their approach (1). They additionally proposed and used a new technique called Hindsight Experience Replay in tandem with DQN and DDPG (2) (3). I anticipate that I will need to make use of some of these algorithms in order to achieve meaningful results, but as mentioned, I will only attempt them after observing a baseline with simple Q-Learning.

RESULTS

Ideally, the outcome of this project will be an RL agent that can manipulate the robot arm and block into a desired orientation with a reasonable success rate. I would like to at minimum show the simulation achieving this under the simplified goal definition described above, but I will move on to more difficult goal states as time permits. Additionally, while an exact success rate is difficult to predict given time and computation restrictions, I would like to at least explore some of the more advanced algorithms detailed in OpenAI's research paper and show their relative performance to the baseline algorithms we've discussed in class.

REFERENCES

1) Matthias, et al. "Multi-Goal Reinforcement Learning: Challenging Robotics Environments and Request for Research." *ArXiv.org*, 10 Mar. 2018, arxiv.org/abs/1802.09464.
2) Andrychowicz, et al. "Hindsight Experience Replay." *ArXiv.org*, 23 Feb. 2018, arxiv.org/abs/1707.01495.
3) Rivlin. "Reinforcement Learning with Hindsight Experience Replay." *Medium*, Towards Data Science, 28 May 2019, towardsdatascience.com/reinforcement-learning-with-hindsight-experience-replay-1fee5704f2f8.