CS391R Project Proposal

Lucas Kabela lak2258

September 17th 2020

The problem I will be investigating is how different settings of experience replay affect robot learning. **This problem is interesting** because proper settings for experience replay improve sample efficiency in training robotics with reinforcement learning (RL) which is key to progress in the field. It can also provide a number of benefits such as faster, more stable training and convergence in sparse reward settings. Some **papers I will be reading** which will inform my approach are Fedus et al. [2020], Wan and Xu [2018] Zhang and Sutton [2017], Schaul et al. [2015], and Andrychowicz et al. [2017]. The first three papers evaluate experience replay, similar to what this work will do, while the second two papers are Prioritized Experience Replay and Hindsight Experience Replay, two experience replay techniques which have defined and driven state of the art deep RL. For data, I will be **collecting new data** using a number of RL environments: the classic pole cart, mountain car, and robotic manipulation task(s) - likely a peg grasping/placing problem and/or pushing. The first two will use OpenAI Gym, and the grasping task(s) will use robosuite. The **algorithms** I will be comparing (varying hyperparameters and combinations) are Prioritized Experience Replay, Hindsight Experience Replay, and several other replay techniques with a DQN like that of Rainbow Hessel et al. [2017]. I will further look for opportunities to augment or explore shortcomings of the experience replay techniques in this work. I will also use libraries for the network when appropriate. **Results will be evaluated** along several axes - the number of episodes needed to reach a certain threshold/succeed for each problem, the maximum return reached by each technique in a certain number of steps, and the ratio of learning updates to experience collected (replay ratio). This can be visualized with tables and mostly line graphs comparing learning vs number of episodes per each technique.

References

William Fedus, Prajit Ramachandran, Rishabh Agarwal, Yoshua Bengio, Hugo Larochelle, Mark Rowland, and Will Dabney. Revisiting fundamentals of experience replay, 2020.

Tracy Wan and Neil Xu. Advances in experience replay, 2018.

Shangtong Zhang and Richard S. Sutton. A deeper look at experience replay. *CoRR*, abs/1712.01275, 2017. URL `http://arxiv.org/abs/1712.01275`.

Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. Prioritized experience replay, 2015.

Marcin Andrychowicz, Filip Wolski, Alex Ray, Jonas Schneider, Rachel Fong, Peter Welinder, Bob McGrew, Josh Tobin, Pieter Abbeel, and Wojciech Zaremba. Hindsight experience replay, 2017.

Matteo Hessel, Joseph Modayil, Hado van Hasselt, Tom Schaul, Georg Ostrovski, Will Dabney, Dan Horgan, Bilal Piot, Mohammad Azar, and David Silver. Rainbow: Combining improvements in deep reinforcement learning, 2017.