498/598 Project Proposal

Audrow Nash (audrow)
Thomas Huang (thomaseh)

November 8, 2017

Paper and overview

We are interested in the paper *Curiosity-Driven Exploration by Bootstrapping Features*. This paper explores intrinsic rewards in reinforcement learning, that is alternatives to rewards when explicit rewards are infrequent, difficult to specify, or costly to evaluate. Specifically, the paper introduces an exploration method Curiosity by Bootstrapping Features (CBF), and found it to give promising results in Super Mario Bros and VizDoom. The paper proposes and evaluates several alternative notions of reward using OpenAI environments and DQN and PPO baseline algorithms.

How this paper relates

- Extends the notion of reward to sparse problems
- Introduces an exploration strategy
- Uses two of the algorithms discussed in class, DQN and PPO
- Uses and modifies OpenAI environments

Computationally reasonable?

We believe so. This paper follows the DQN implementation and performs additional computations in computing the rewards. It uses $50e6$ steps for each training environment. Running OpenAI's implementation of DQN for $50e6$ steps on the carpole environment takes less than 5 minutes on a 2015 Macbook Air with an i7 processor and 8GB of RAM. If the most computationally expensive part of DQN is updating the neural network, running the algorithm proposed in the paper should not take significantly longer than DQN. Should our implementation be significantly slower, we can use simpler environments than the paper or simulate for fewer time steps.