Project Plan

Matthew Alger
August 20, 2015

Inverse reinforcement learning (IRL) is the task of finding a reward function that an agent is trying to maximise, given the agent’s policy or trajectories sampled from the agent’s policy. A recent paper by Wulfmeier et al. extends existing approaches to use a deep-learned reward function, thus avoiding prior assumptions of a reward function linear in state space. This project aims to reimplement key previous works in IRL, and validate and extend the results found by Wulfmeier et al.

1. Read and summarise existing literature.

2. Implement the model-based IRL approaches described in Ng & Russell, 2000.
   - Finite, small state space approach from policy.
   - Large state space approximation from policy.
   - Large state space approximation from trajectories.

3. Implement the maximum entropy IRL approach described in Ziebart et al., 2008.
   - Implement gradient descent algorithm over \( \theta \), finding \( D_i \) with Ziebart’s DP algorithm.
   - Compare to Ng & Russell approach.
   - Test different environments and state representations.

4. Implement the deep IRL approach described in Wulfmeier et al., 2015.
   - Directly implement neural network with gradients from Wulfmeier.
   - Try deriving Wulfmeier gradients with Theano, Keras.
   - Compare to Ng & Russell and Ziebart and ideally reproduce results of paper.

5. Extend existing deep approach.
   - Experiment with parameters and deep structure on the Wulfmeier model.
   - Learn state representation as part of the IRL process.
   - Stretch: Try value-based deep IRL.

6. Write final report and prepare artefact for submission.