Project 3: RLDM - CS 7642

GID: mmendiola3

Presentation: https://youtu.be/

Overview

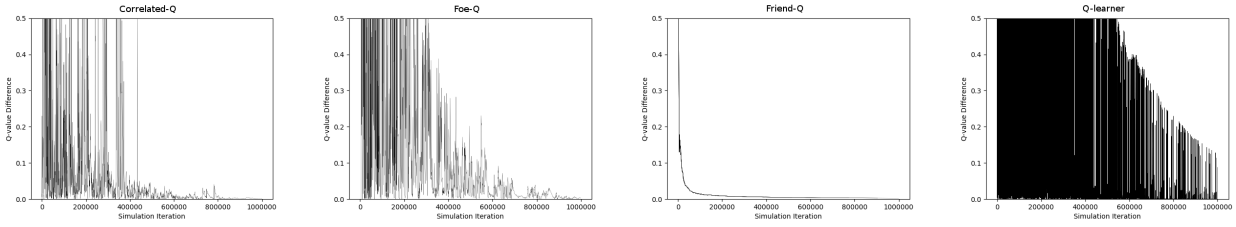
This paper will review the reproduction of experiments done in "Correlated Q-Learning" (Greenwald and Hall 2003). We will explore the agents designed to learn policies for the multi-agent soccer environment outlined in the original paper, and compare the result. Finally, we will discuss the challenges and assumptions made in order to reproduce the graphs in Figure 3.

Soccer Environment

The soccer experiment rules were reproduced with a basic environment class built with an interface similar to those in OpenAI gym environments. It maintains state and models the actions and rewards outlined in the original paper. Experiments with various agents are able to make calls to env.step(actions) and receive back next state, rewards, and the done flag. Actions, state, and rewards includes data on both players in the soccer environment.

Experiment Setup

Using the soccer environment model, experiments were run with four learning agents (uCE-Q, Foe-Q, Friend-Q, and Q-learning). Each experiment was run over 10e5 steps (env steps, not episodes). The change in Q values for a specific state, action(s) pair was recorded over these trials. This state is illustrated in Figure 4 of the original paper and the action(s) were A=South, B=Stick.



Layers – I tested many combinations of layers and nodes to find the simplest network capable of approximating Q values with enough granularity to guide the lander towards the goal. I tested single hidden layers with nodes between 16 and 1024 and two hidden layer networks with the second layer either equal or half that of the fist layer. Each of these experiments resulted in worse or similar initial performance or increased training time.

References

Greenwald, Amy, and Keith Hall. 2003. "Correlated-Q Learning." In *Proceedings of the Twentieth International Conference on International Conference on Machine Learning*, 242–49. ICML'03. Washington, DC, USA: AAAI Press. http://dl.acm.org/citation.cfm?id=3041838.3041869.