

Homework 3

Ted Xiao

March 8, 2017

1 Basic Q-Learning Performance

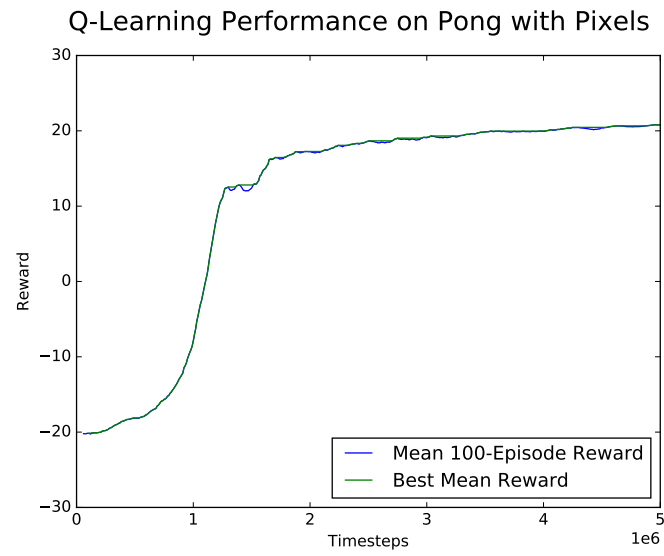


Figure 1: These results correspond to the default hyperparameters. It was run for 5 million timesteps on GPU.

2 Experimenting with Hyperparameters

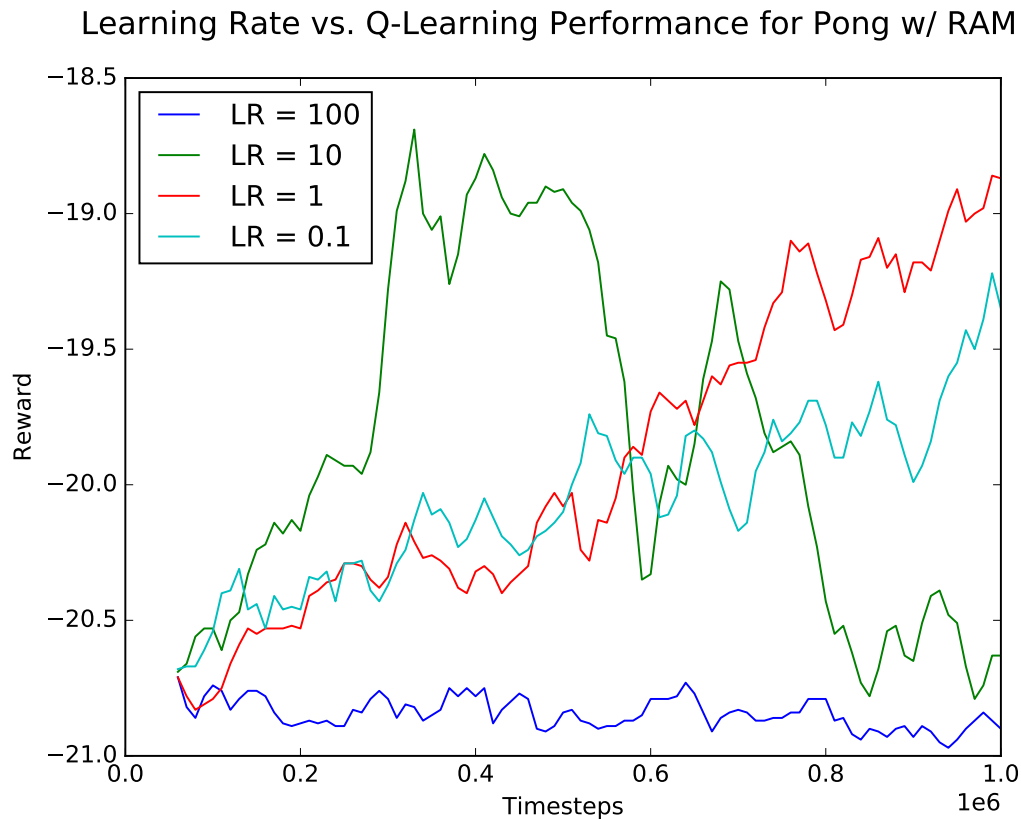


Figure 2: The plot above shows the effect of learning rate on various trials of the Q-Learning model on Pong with RAM for 1 million steps. I chose to experiment with tuning learning rate (learning rate multiplier) because I was interested in comparing the sensitivity of exploration on Q-Learning with the sensitivity of learning rate on other tasks (Q-Value Iteration, Behavior Cloning). My intuition was that high learning rates would perform badly, and that seems to be the case. I was surprised that a LR of 10 was able to learn very quickly at first, which suggests that perhaps a scheduler that starts very high and decays fast might perform well.