RBE595 - Week 6 Assignment

Keith Chester

Due date: February 19, 2023

In this assignment, we looked at a simple stochastic environment of 6 consecutive cells, labeled 0 through 4. Either end of the line of cells were terminal states - cell 0 with a reward of 1, cell 5 with a reward of 5. All other cells were a reward of 0 with a choice to go either left or right. There existed an 80% chance of the robot in our environment following the designated action, a 15% chance of the robot merely staying where it is, and a 5% chance for the robot to go the opposite of what was chosen.

Here we developed two Monte Carlo agents - an exploring starts agent and a first visit agent. We looked at the convergence of the agent at the end of a set number of episodes, from 1 episode to 500 episodes. For each of these episodes, we performed the test with 100 different agents. We then looked at the average $Q(s, a)$ of each state/action pair, as well as the $V(s)$ of each.

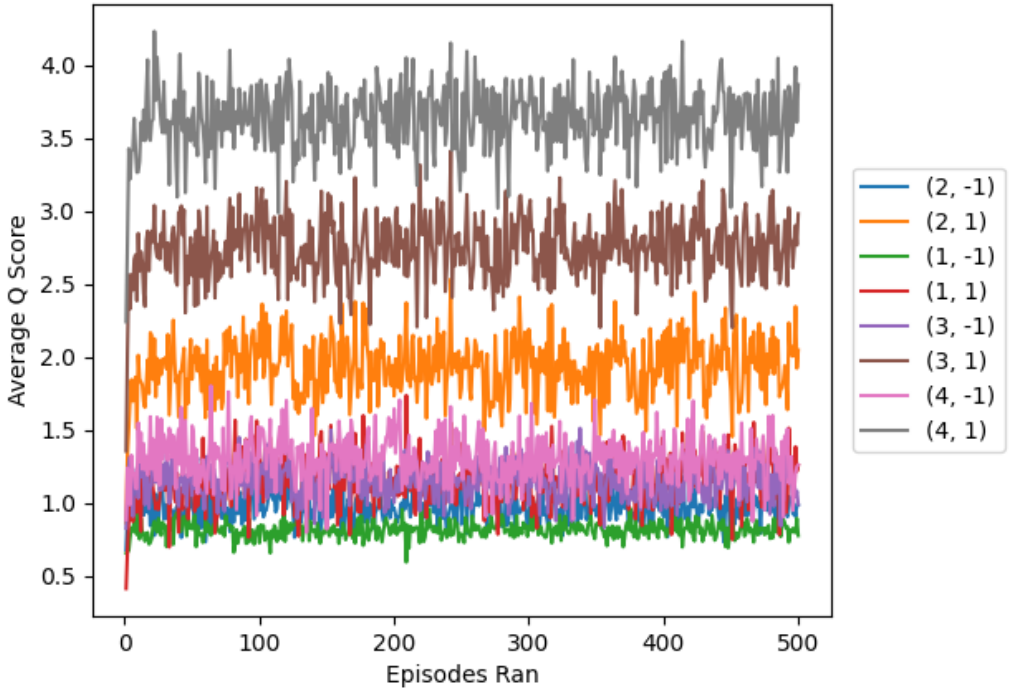

Figure 1: Q(s,a) Avg for Exploring Starts Monte Carlo Agent

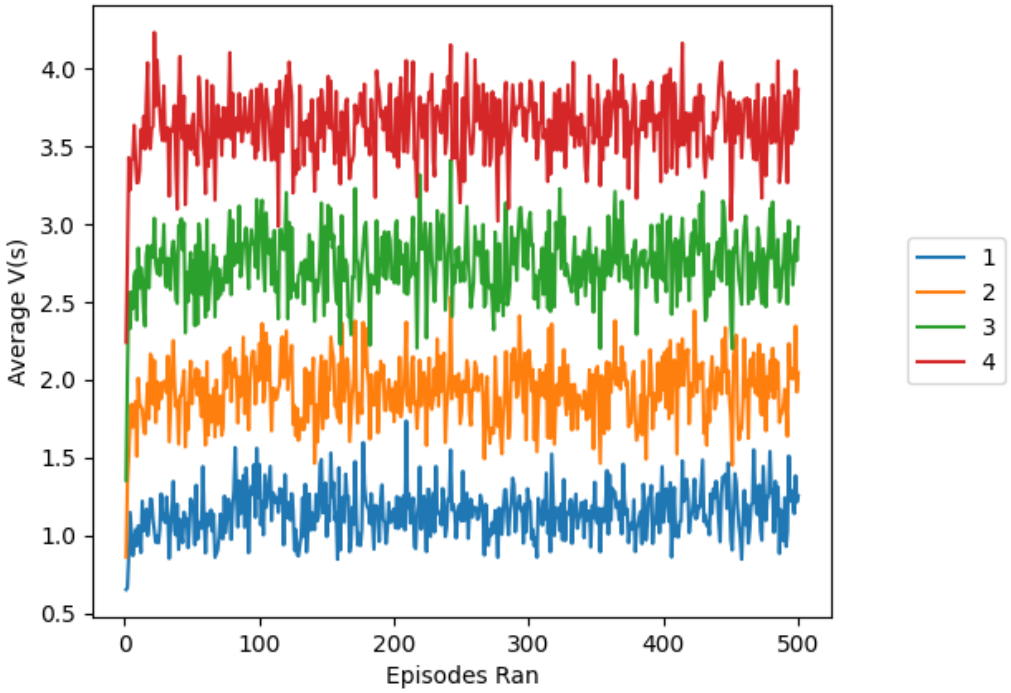

Figure 2: V(s) Avg for Exploring Starts Monte Carlo Agent

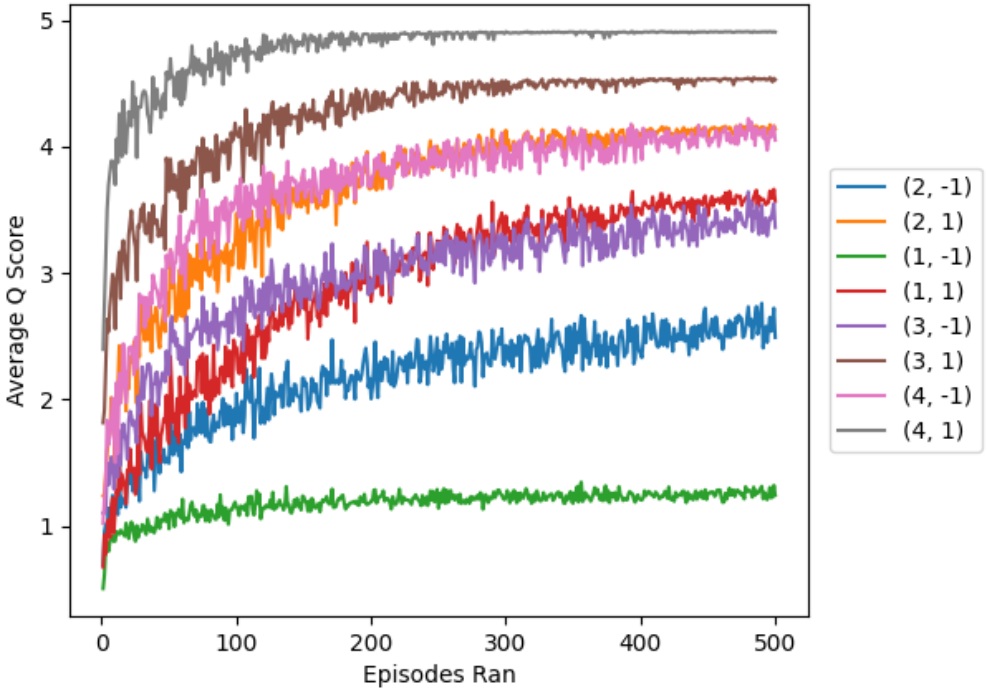

Figure 3: Q(s,a) Avg for First Visit Monte Carlo Agent

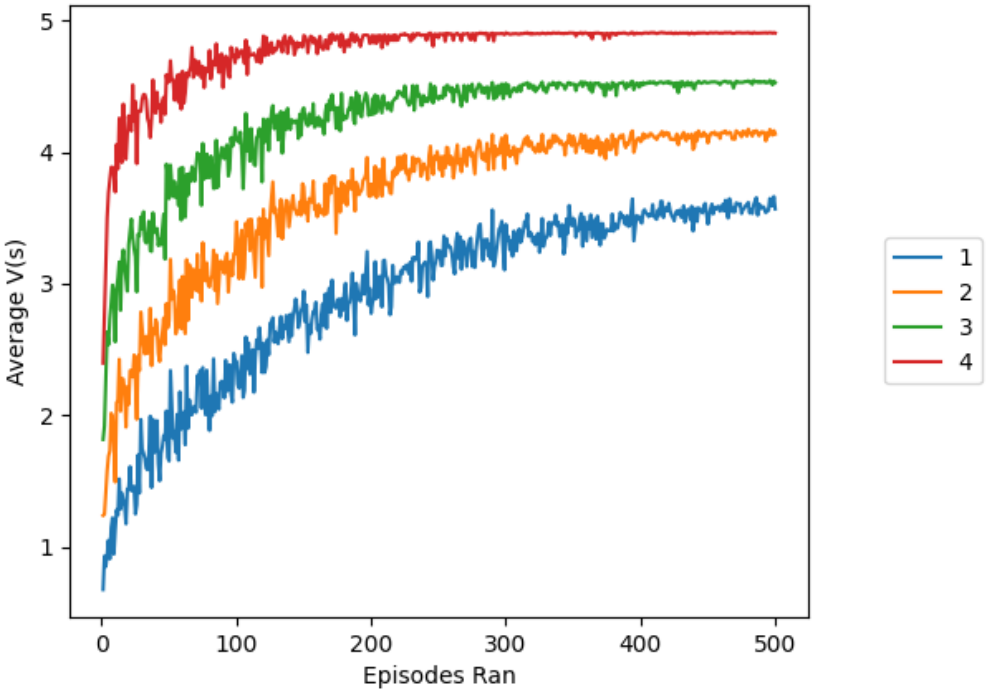

Figure 4: V(s) Avg for First Visit Monte Carlo Agent