Project 3 – Reinforcement Learning

Rohman Sultan

December 5, 2021

1 Introduction

This project is about creating a reinforcement learning algorithm that can learn a policy for a specific activity using task-based rewards. As well as converting a continuous environment to a discrete environment to make it ideal for reinforcement learning.

The first stage involves balancing a pole with a one-dimensional movement that only moves left and right. This is done by finding the optimal values for alpha (learning rate), gamma (discount factor) that is applied to the Q learning algorithm.

The last stage involves a car with limited power to drive up a hill. The values found in the first stage are applied here. The environment values and bins needed to be set for this stage.

2 Approach

The first stage, I set the alpha to 0.5 because the learning should not generally be too high (swing back and forth between the lowest and highest points) or too low (require substantial amount of iterations). I set gamma to 0.999 as it worked for this type of problem.

The last stage, I changed alpha to 0.15 to see how it would perform and the results turned out great. I set the bins to [20, 20] and n_episodes to 1000. I found the environment values by calling the observation space method.

3 Results and discussions

Although the solutions may not be the optimized approach, I was able to solve the Cart-Pole and Car Mountain problem. I believe there other strategies to quickly learn the task without affecting much of the algorithm.