

Jacob Tucker  
CS 482 Artificial Intelligence  
Project 3  
Reinforcement Learning

Task: To implement a reinforcement learning algorithm given a python program. The models must accomplish the tasks of balancing a pole on a cart for at least 200 ticks and move a car up a hill to a flag.

Approach:

In cart.py, the q learning algorithm is implemented on line 142.

$$Q[s, action] += \alpha * (\text{reward} + (\gamma * \text{predicted\_value}) - Q[s, action])$$

Everything is precalculated except for alpha and gamma, which refer to the learning rate and discount of future rewards. Alpha was set to 0.05, a very low learning rate, and gamma was set to 0.99, meaning future rewards are almost as desirable as immediate rewards. These numbers were reached through many iterations of trial and error until it became clear that the very low alpha and very high gamma were proving successful consistently.

For car.py, the alpha was set to 0.7 and the gamma was set to 0.95. While testing with the given 50001 episodes, it became clear that many values of alpha would create successful models, however, through trial and error, it appears that roughly 0.7 was producing successful models the fastest. It is likely that these models could be trained with much less than 50001 episodes. During testing, some models were producing successful tests after 5001 episodes. The models in this assignment were trained with the given 50001 episodes to ensure they would be successful.

Results:

The resulting models consistently accomplish the task. The cartpole model has a very low learning rate, so it requires a large number of episodes to become successful in accomplishing the task. However, the car has a much higher learning rate and could likely produce successful models in less episodes. Both trained models have not failed the task in any of the tests I have done.