Eric Johnson
erjohns3
ECE 448

HW3

Problem 1

States:
There are 11 states, the first 5 are for each of the square the the robot could be besides the stack while it has the pallet, the 6th state is when the robot is on the stack square and has stacked the pallet, and the last 5 are all the squares besides the stack while the robot doesn't have the pallet.

Actions
There are 5 actions, one for each directions the robot could move in and one for just not moving at all.

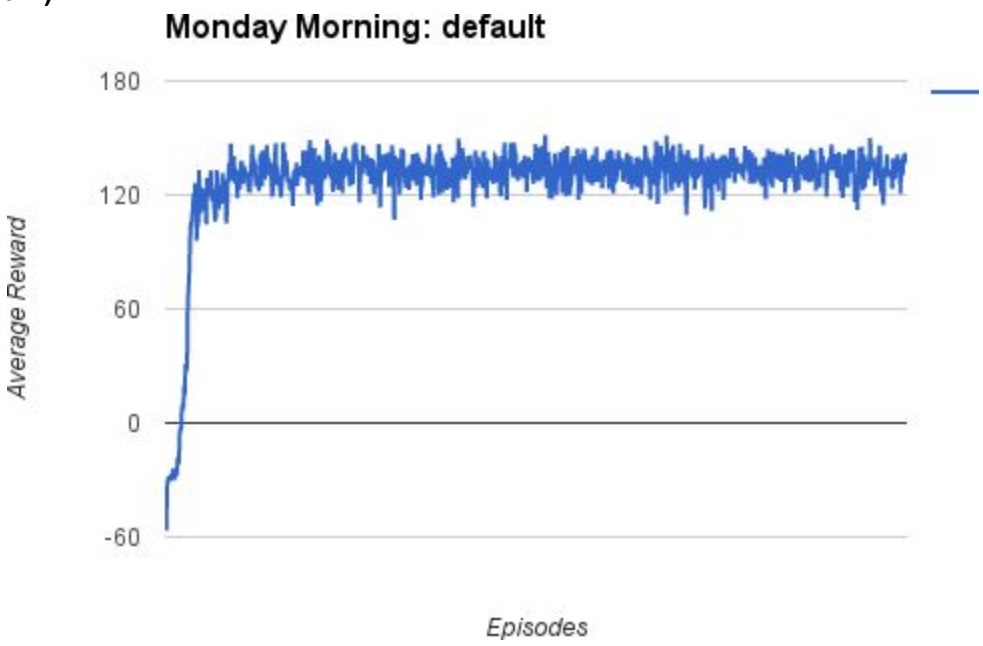
Initial Distribution
The initial distribution is that the robot has a probability of 1 of starting at position (1,0), a forklift has a probability of 1 of starting at position (1,2), an obstacle has a probability of 1 of starting at position (0,0), a stack has a probability of 1 of starting at position (0,2) , and the loading dock has a probability of 1 of starting at position (1,0).

Transition
The initial distribution over the states is that when the robot picks an action for a certain direction, it has a 100% chance of going there. If the robot is at the edge of the world though, and it attempts to move in the direction of that edge has 100% chance of not moving.

Rewards
The reward for going to the stack is r_s , the reward for going to the obstacle is $-r_l$, and the reward for going to the forklift is $-p_0 r_l$.

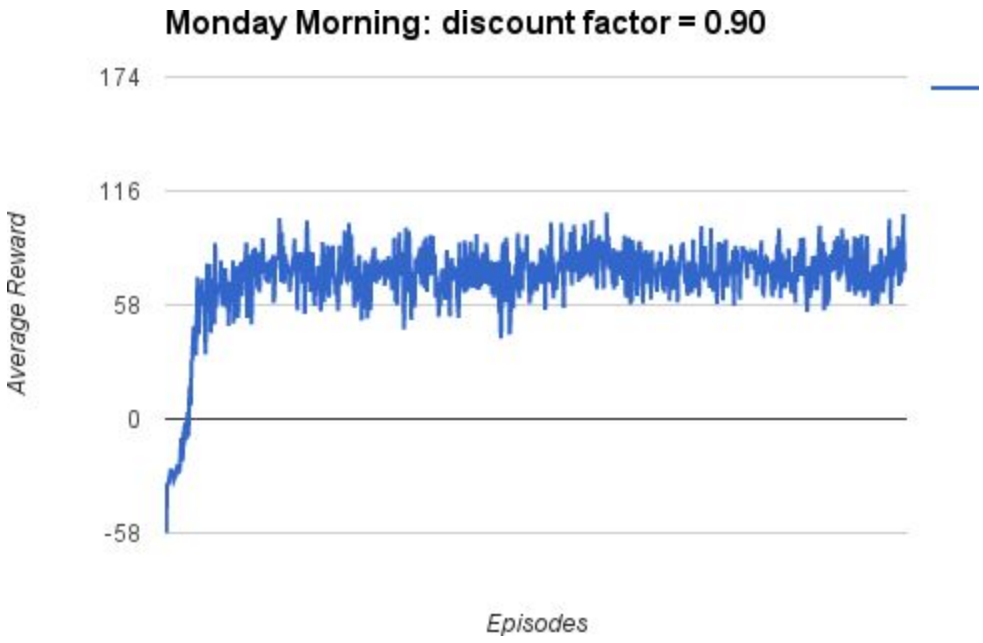
Problem 2

a i.)



a ii.)

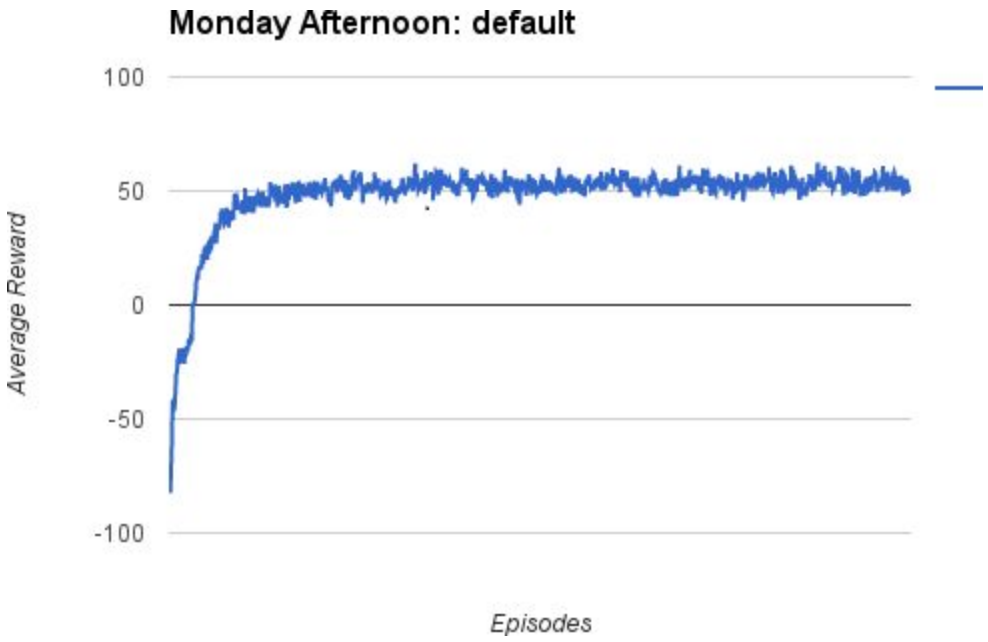
The lower discount factor causes the reward to not reach as high and also not converge towards a single value as much.



Problem 2

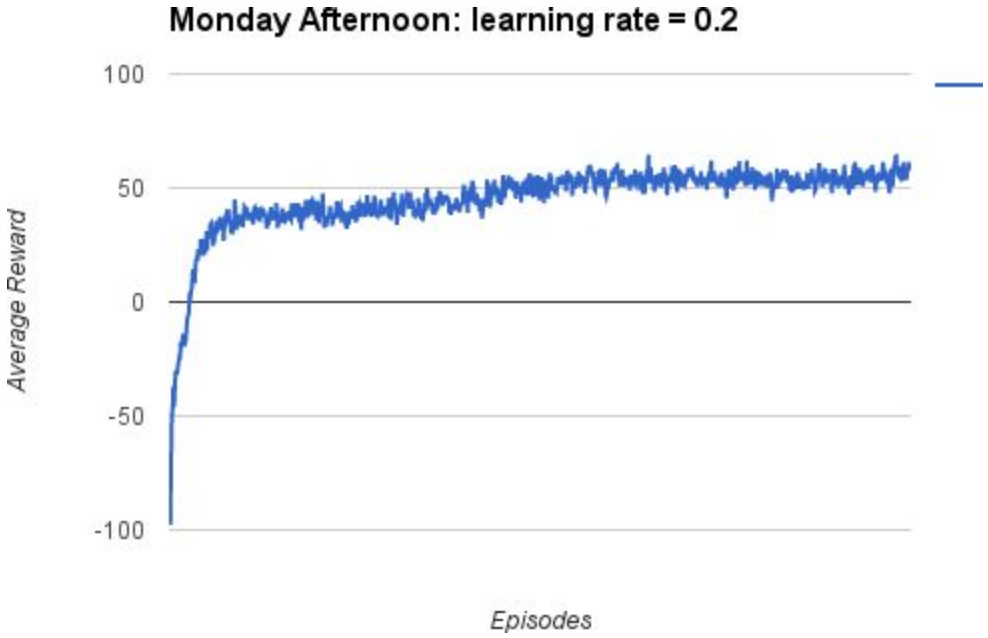
b i.)

The reward in the second scenario reach as high, but they do converge more than the default.



b ii.)

The high learning rate causes the rewards to take longer to reach their maximum, but it might have a slightly higher maximum than the default.



Problem 2

b iii.)

The lower exploration value causes the rewards to converge on a higher amount but has a slightly lower slope than the default.