CS 440/ECE448 Homework 3

Tanishq Dubey (tdubey3)

May 7, 2016

Problem 1

1. Set of states:
   This problem can easily be defined if we assign each location a coordinate value as such:

(0,0)	(1,0)	(2,0)
(0,1)	(1,1)	(2,1)

   In addition, we can define our robot as having a pallet as P. Thus we can list our states as such:

   - $s_0 = ((0,0), P)$
   - $s_2 = ((1,0), P)$
   - $s_4 = ((2,0), P)$
   - $s_7 = ((1,1), P)$
   - $s_9 = ((2,1), P)$
   - $s_1((0,0), \neg P)$
   - $s_3 = ((1,0), \neg P)$
   - $s_5 = ((2,0), \neg P)$
   - $s_8 = ((1,1), \neg P)$
   - $s_10 = ((2,1), \neg P)$
   - $s_6 = ((0,1), P)$

   The last item must always be P since the robot must have a pallet at the loading dock. Thus eleven states are defined for our robot.

2. Set of actions:
   To define our set of actions, it can be said that our set of actions $A = a$ where:

   $$A = North(0), South(1), East(2), West(3), Stay(4)$$

   Thus we have 5 actions.

3. Initial distribution over states:
   The initial distribution over states is:

$$S_0 = s_6 = ((0,1)P)$$

4. Transition model:
   Our transition model is as such:

$$T(s_0, 0, s_0) = 1$$
$$T(s_0, 1, s_2) = 1$$
$$T(s_0, 2, s_6) = 1$$
$$T(s_0, 3, s_0) = 1$$
$$T(s_0, 4, s_0) = 1$$

$$T(s_1, 0, s_1) = 1$$
$$T(s_1, 1, s_3) = 1$$
$$T(s_1, 2, s_6) = 1$$
$$T(s_1, 3, s_1) = 1$$
$$T(s_1, 4, s_1) = 1$$

$$T(s_2, 0, s_2) = 1$$
$$T(s_2, 1, s_4) = 1$$
$$T(s_2, 2, s_7) = 1$$
$$T(s_2, 3, s_0) = 1$$
$$T(s_2, 4, s_2) = 1$$

$$T(s_3, 0, s_3) = 1$$
$$T(s_3, 1, s_8) = 1$$
$$T(s_3, 2, s_1) = 1$$
$$T(s_3, 3, s_5) = 1$$
$$T(s_3, 4, s_3) = 1$$

$$T(s_4, 0, s_9) = 1$$

$$T(s_4, 1, s_4) = 1$$
$$T(s_4, 2, s_4) = 1$$
$$T(s_4, 3, s_4) = 1$$
$$T(s_4, 4, s_2) = 1$$

$$T(s_5, 0, s_5) = 1$$
$$T(s_5, 1, s_5) = 1$$
$$T(s_5, 2, s_5) = 1$$
$$T(s_5, 3, s_3) = 1$$
$$T(s_5, 4, s_10) = 1$$

$$T(s_6, 0, s_0) = 1$$
$$T(s_6, 1, s_7) = 1$$
$$T(s_6, 2, s_6) = 1$$
$$T(s_6, 3, s_6) = 1$$
$$T(s_6, 4, s_6) = 1$$

$$T(s_7, 0, s_2) = 1$$
$$T(s_7, 1, s_6) = 1$$
$$T(s_7, 2, s_7) = 1$$
$$T(s_7, 3, s_7) = 1$$
$$T(s_7, 4, s_9) = 1$$

$$T(s_8, 0, s_3) = 1$$
$$T(s_8, 1, s_6) = 1$$
$$T(s_8, 2, s_8) = 1$$
$$T(s_8, 3, s_8) = 1$$
$$T(s_8, 4, s_10) = 1$$

$$T(s_9, 0, s_4) = 1$$

$$T(s_9, 1, s_7) = 1$$
$$T(s_9, 2, s_9) = 1$$
$$T(s_9, 3, s_9) = 1$$
$$T(s_9, 4, s_9) = 1$$

$$T(s_{10}, 0, s_5) = 1$$
$$T(s_{10}, 1, s_8) = 1$$
$$T(s_{10}, 2, s_10) = 1$$
$$T(s_{10}, 3, s_10) = 1$$
$$T(s_{10}, 4, s_10) = 1$$

5. Reward Function:
   The reward function is as such:
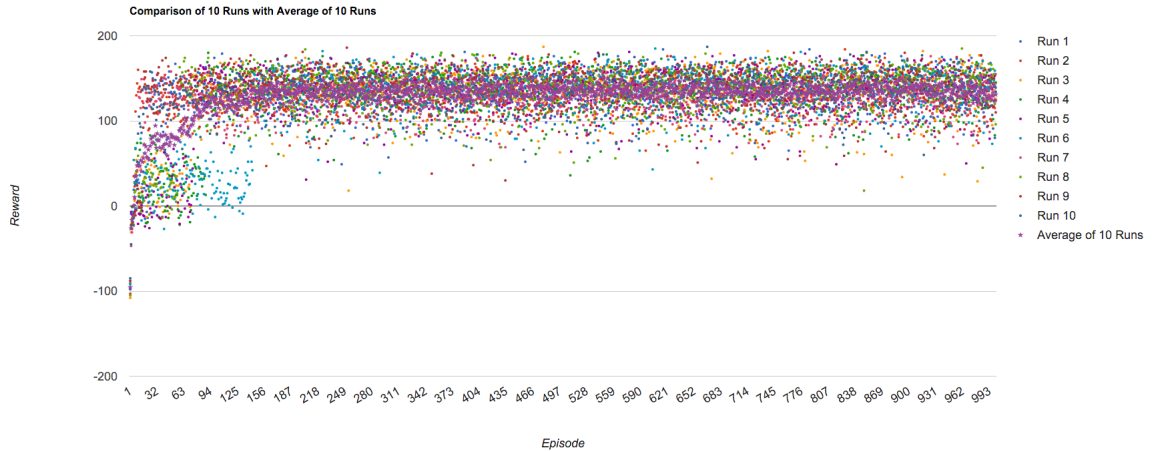$$R(s_4) = r_s$$
$$R(s_0) = r_l$$
$$R(s_1) = r_l$$
$$R(s_9) = r_l$$
$$R(s_{10}) = r_l$$

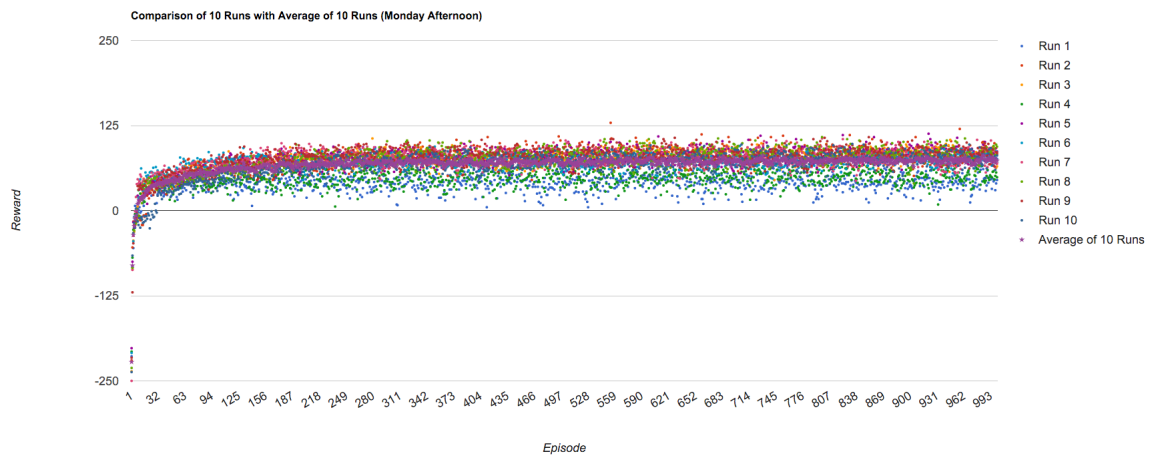Where the rewards at states s_0 and s_1 happen with a probability of p_0.

Problem 2

a) Scenario 1: Monday Morning

Comparison of 10 Runs with Average of 10 Runs

i
   It can be seen that with learning, the agent drastically improves it performance. As shown with

5

the "Average" plot, the robot is increasing in performance and remaining in the positive reward zone as well.
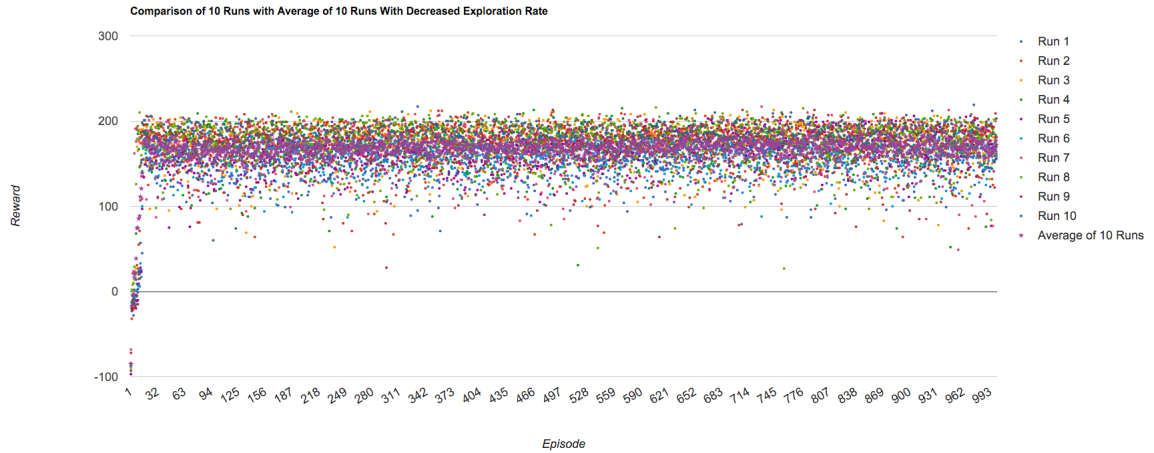

The afternoon simulation has a much smaller steady state reward and in addition has a much smaller spread in the steady state reward when compared to the default Monday morning simulation. For all following simulations, similar results were witnessed in both the morning and afternoon simulations.
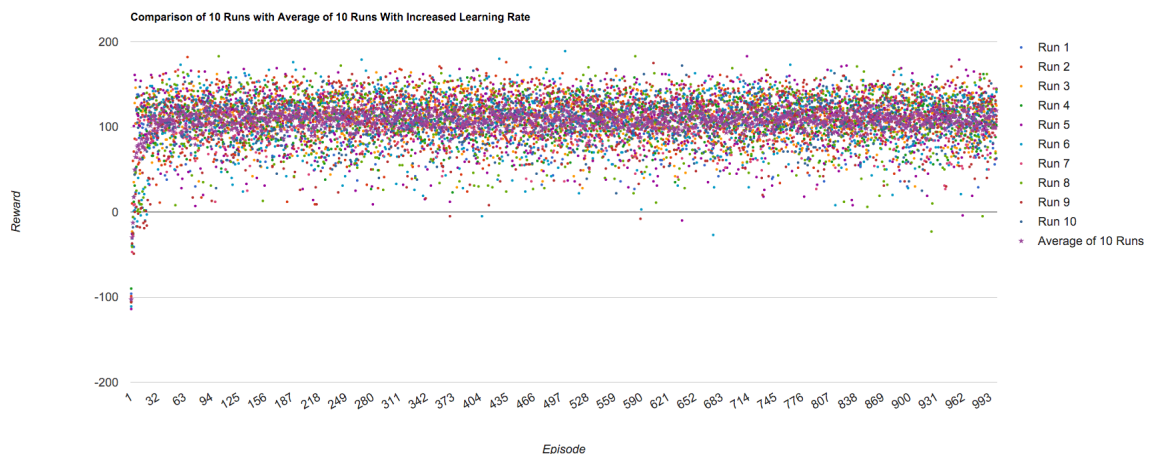
ii

By increasing the learning rate, it can be seen that the initial learning slope is greatly increased, causing the agent to learn faster than the initial learning value.

Comparison of 10 Runs with Average of 10 Runs With Decreased Exploration Rate

Reward

Episode

Run 1
Run 2
Run 3
Run 4
Run 5
Run 6
Run 7
Run 8
Run 9
Run 10
Average of 10 Runs

iii

The decrease in learning rate does not cause a significant change between the increase in learning rate plot. The only noticeable change is that, compared to the increased learning rate graph, this plot has a much more linear initial slope, perhaps due to the fact that the agent is more likely to choose the best option for movement. However, the initial learning slope for the agent is much steeper than the default values for the agent. ∞