Software Agents

SNAKES AND LADDERS

Coursework – IN3016/INM426

WASSIM BEN YOUSSEF | BENJAMIN AUZANNEAU

Table of Content

1. Basic Case

1.1 Domain and Task

In this paper, we will solve a problem by implementing the Q-Learning reinforcement learning algorithm. In Q-Learning, an agent moves across states within a domain based on a predefined set of reward located in the R-Matrix. The domain in which the software agent will be deployed is a simple 5x5 snakes and ladders (Figure 1.).Each horizontal set of five states in the grid represent levels that are interconnected by snakes (in red) or ladders (in green).

The journey of our software agent begins in a lower level state (1) and its task is to find the shortest path to an upper level state (21) by navigating across levels, up the ladders, and avoiding going down the snakes. Each numbered square in Figure 1. represents a state. Ladders are two way transitions between levels and snakes are downward transitions only. To make the task of our agent a more challenging, we have created a shortcut to the goal state only accessible by voluntarily going down a snake (Figure 2.). That path also happens to be the shortest one with a total of 12 actions from start to goal state. The second shortest path is the most logical and least challenging (Figure 3.). It consists of 14 actions, all either horizontal or up ladders, avoiding all the snakes.

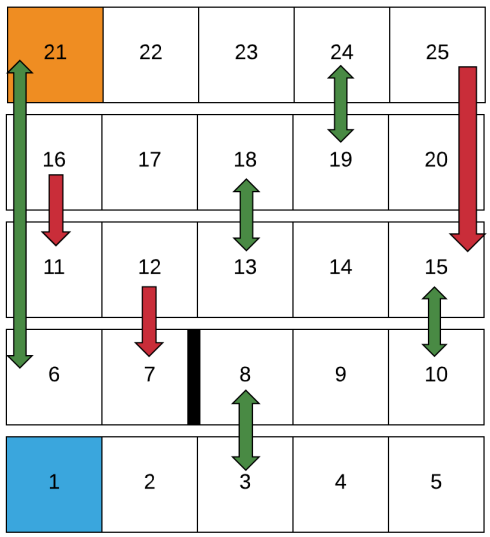


Figure 1. – Domain Graph

1.2 State Transition Function

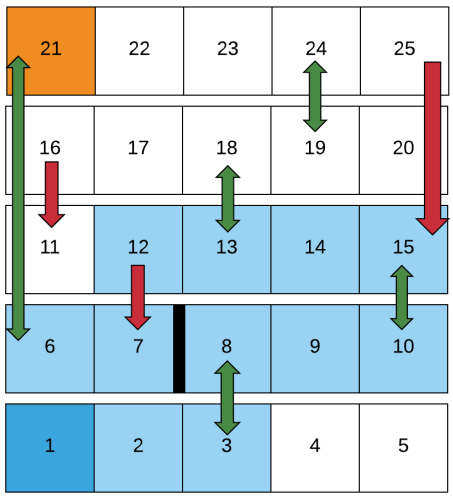


Figure 2. – Shortcut Path

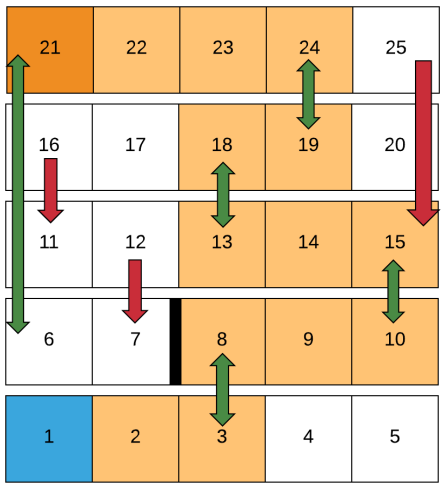


Figure 3. – Logical Path

Our state transition grid begins in State 1, where our agent will be deployed. The goal state is state 21 (Figures 1 to 3). Starting from state 1, each groups of 5 horizontal states represent levels (1-5 is level 1, 6-10 level 2, etc…). Each action represents a move from a state to another whether it be horizontally across a level, or vertically across levels up or down ladders or snakes. While our agent is free to move up or down ladders, snake transitions will force it down a level (or more). Neither does it have the possibility to transition up a snake. In the case of no snakes or ladders, the agent will have to move horizontally across the level states searching for an opportunity to transition up or down a level.

In level 2 the agent does not have the possibility to transition from state 7 to 8. That separation is represented by a bold wall in Figures 1 to 3. Instead, he will have to transition up to level 3 before going back down to level 2 via a snake (12→7) . That option is both the shortest path (12 steps) and the main challenge for the agent as it will have to learn that upwards is not always the best solution. The most logical and second shortest path (14 steps) is shown in figure 3. We refer to state 13 as the Crossroad, the point at which the agent will either chose to take the Shortcut or the Logical path.

The following annotations represents the transition possibilities from each state:

1 → {2}
2 → {1,2}
3 → {2,4,8}
4 → {3,5}
5 → {4}
6 → {7,21}
7 → {6}
8 → {3,9}
9 →{8,10}
10 → {9,15}
11 → {12}
12 → {7}
13 → {12,14,18}
14 → {13,15}
15 → {10,14}
16 → {11}
17 → {16,18}
18 → {13,17,19}
19 → {18,20,24}
20 → {19}
21 → {6,**21**,22}
22 → {**21**,23}
23 → {22,24}
24 → {19,23,25}
25 → {15}

Green = Ladder Transition
Red = Snake transition
Bold = Goal state

1.3 Reward Function

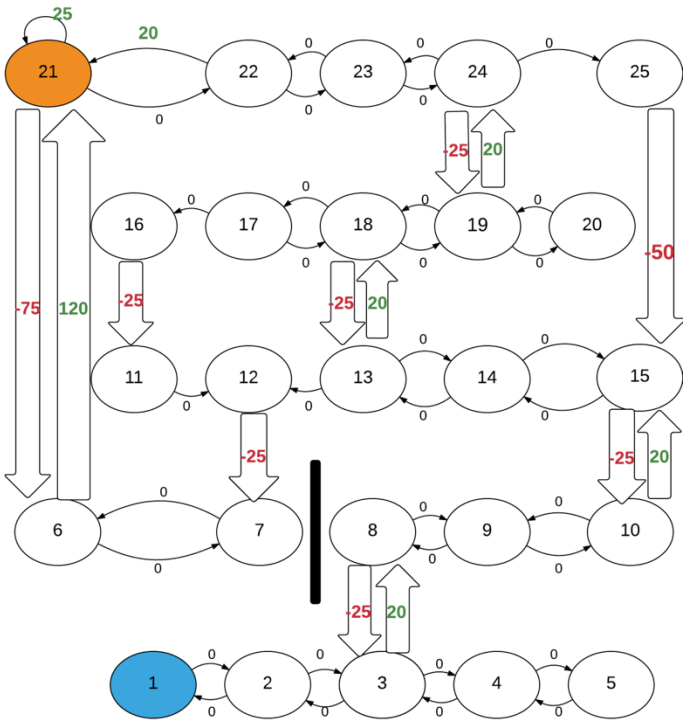


Figure 4. – Transitions and Rewards

Our initial reward function operates vertically. Positive rewards are attributed for transitioning upwards in the grid and negative rewards for transitioning downwards. There are no rewards for transitioning horizontally from state to state. Transitions up a ladder will activate a reward of +20 and going down -25. The logic behind the extra negative reward is to inform the agent that his general direction towards the goal state is upwards. In the case of a snake transitioning down two levels (25--> 15), we simply multiply the negative reward by two.

When transitioning up the three-level ladder (6⟶21) , part of the “challenge path” , our agent will receive the highest reward of our environment, +120. This should allow it to accept the prior negative reward and still identify that path as the shortest one to the goal state. Going down that ladder, the standard ladder reward is multiplied by three for a total negative reward of -75.

In order to define the state transition possibilities of the agent, we have set the reward of impossible state transitions as -1000. Within our code, we set a rule which prevents our agent to transition into those states and only across our defined environment. Our reward function is represented as an R-Matrix in Figure 5.

State	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
2	0	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
3	-1000	0	-1000	0	-1000	-1000	-1000	20	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
4	-1000	-1000	0	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
5	-1000	-1000	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
6	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	120	-1000	-1000	-1000	-1000
7	-1000	-1000	-1000	-1000	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
8	-1000	-1000	-25	-1000	-1000	-1000	-1000	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
9	-1000	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
10	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000	-1000	-1000	-1000	-1000	20	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
11	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
12	-1000	-1000	-1000	-1000	-1000	-1000	-25	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
13	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000	0	-1000	-1000	-1000	20	-1000	-1000	-1000	-1000	-1000	-1000	-1000
14	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
15	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-25	-1000	-1000	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
16	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-25	-1000	-1000	-1000	-1000	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000
17	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000	-1000
18	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-25	-1000	-1000	-1000	0	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000
19	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000	0	-1000	-1000	-1000	20	-1000
20	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000	-1000	-1000	-1000	-1000	-1000
21	-1000	-1000	-1000	-1000	-1000	-75	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	25	20	-1000	-1000	-1000
22	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	20	-1000	0	-1000	-1000
23	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000	0	-1000
24	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-25	-1000	-1000	-1000	0	-1000	0
25	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-50	-1000	-1000	-1000	-1000	-1000	-1000	-1000	-1000	0	-1000

Figure 5. – R-Matrix

1.4 Action Selection Policy

We chose to give our agent an Epsilon-greedy policy to explore the snakes and ladder environment. After initializing the ε value between 0 and 1, a random number between or equal to 0 and 1 is generated. Given that number is equal or higher than ε, the agent will exploit the environment by choosing the action with the highest reward value. On the other hand, the agent will explore by choosing a random action if the number is smaller than ε.

For our agent to move from an exploring to an exploiting mind-set as time passes, we implement a reduction factor by multiplying the ε value after each action. ε is to be multiplied by 0.99999 if it is equal or higher than 0.5 and by 0.9999 if lower than 0.5. The ε-greedy policy is particularly interesting in our domain as there are two paths to explore. The most logical one is not the optimal path to the goal state. In turn, it is key to let the agent explore the environment before slowly moving towards a more aggressive action selection policy.

Epsilon value(ε): 0.8

We initialize our ε-value at 0.8 to ensure our agent prioritizes exploration at the beginning of the learning process. Different ε-values will be tested and investigated in the later parts of the paper.

1.5 Q-Learning initialization

There are two key parameters to Q-Learning that affect the attitude of the agent. The learning rate, α, determines how much importance is to be given to new information over old information. As such, the higher the learning rate, the faster the agent will learn. The second parameter, discount factor γ, determines the importance given to future rewards. If the value is close to 1, the agent will work towards a long-term high reward. On the other hand, the agent will not consider future rewards if the value is 0.

Learning Rate(α): 0.5
Discount Factor(γ): 0.8

We initialize our Learning Rate value at 0.5 as a trade-off between efficiency and effectiveness. The Discount Factor value is set at 0.8 to ensure our agent works towards a high long-term reward. Different values of (α) and (γ) will be tested and investigated in the second part of the paper.

1.6 Q-Learning Algorithm

In this section, we run the Q-Learning algorithm by hand until the first ladder transition, the 3rd action. We update the Q-Matrix (Figure 6.) from its initial state of zero everywhere. We also update our Epsilon values according to our implemented reduction factor.

Here is the algorithm used to apply Q-learning algorithm using Greedy policy:

1. Initialize the values of γ, α, ε, the first state and the goal state
2. Initialize the Reward matrix R and initialize the Q-matrix by a matrix containing only 0s with the size of the R matrix
3. For each episode :
    a. While the goal state is not reached Do :
        i. Set a random variable B between 0 and 1
        ii. If B >= ε (exploit):
            1. The next state will be the one with the highest q value among the possible states :
                a. If the q-values of the different available states are equal, we select randomly the next state
                b. Else we select the next state as the one with the highest Q-value
            2. Set the value of the state as the next state
            3. Update the Q-matrix using the formula :

$$Q_{new}(S_t, a_t) = Q_{old}(s_t, a_t) + \alpha(r_{t+1} + \gamma max_a Q(s_{t+1}, a) - Q_{old}(s_t, a_t))$$

        iii. Else (if B < ε)(explore):
            1. Randomly chose the next state among the possible states
            2. Update the Q-matrix using the same formula
            3. Set the value of the state as the next state
        iv. Update the value of ε :
            1. if ε >=0.5: ε <- ε *0.99999

2. else : ε <- ε *0.9999

End Do

End For

Now we apply this algorithm for 3 steps:
First we initialize the values :

- γ = 0.8, α = 0.5, ε = 0.8
-the first state is state 1 and the final state is state 21
-the R matrix is the one given before and the Q is a 25x25 matrix full of 0s

First step:

We select randomly B = 0.1

B < ε so we select randomly the next state among the possible states, but here the only possible state is 2 so we set the next state at 2

We calculate the new value of Q(1,2) :

$$Q_{new}(1,2) = Q_{old}(1,2) + \alpha(r_{1,2} + \gamma max\ (Q(2,1), Q(2,3)) - Q_{old}(1,2))$$

$$Q_{new}(1,2) = 0 + 0.5(0 + 0.8 max\ (0,0) - 0)$$

$$Q_{new}(1,2) = 0$$

So the Q-matrix stays a 25x25 matrix full of 0s

ε takes the value 0.8*0.99999 = 0.799992

Second step:

We select randomly B = 0.6

B < ε so we select randomly the next state among the possible states, we set the next state at 3 randomly

We calculate the new value of Q(2,3):

$$Q_{new}(2,3) = Q_{old}(2,3) + \alpha(r_{2,3} + \gamma max\ (Q(3,4), Q(3,8), Q(3,2)) - Q_{old}(2,3))$$

$$Q_{new}(2,3) = 0 + 0.5(0 + 0.799992 max\ (0,0,0) - 0)$$

$$Q_{new}(2,3) = 0$$

So the Q-matrix stays a 25x25 matrix full of 0s

ε takes the value 0.799992*0.99999 = 0.799984

Third step:

We select randomly B = 0.8

B > ε so we select the next state as the one with the highest reward value among the possible states. Here all the next states have the same Q-value, so we select randomly the next state. Let us assume the next state randomly chosen is 8.

We calculate the new value of Q(3,8):

$$Q_{new}(3,8) = Q_{old}(3,8) + \alpha(r_{3,8} + \gamma max\ (Q(8,3), Q(8,9)) - Q_{old}(3,8))$$

$$Q_{new}(3,8) = 0 + 0.5(20 + 0.799984 max\ (0,0) - 0)$$

$$Q_{new}(3,8) = 10$$

ε takes the value 0.799992*0.99999 = 0.799984

New Q-Matrix (The 10 is located in 3rd row and 8th column. All other values are 0.):

$$\begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & & 10 & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

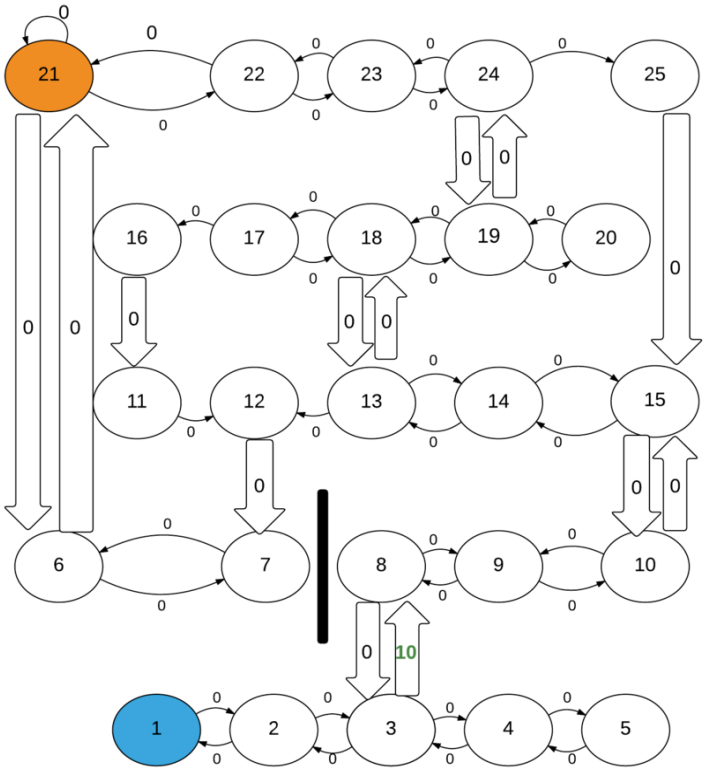


Figure 6. Updated Q-Values

1.7 Analysis of results

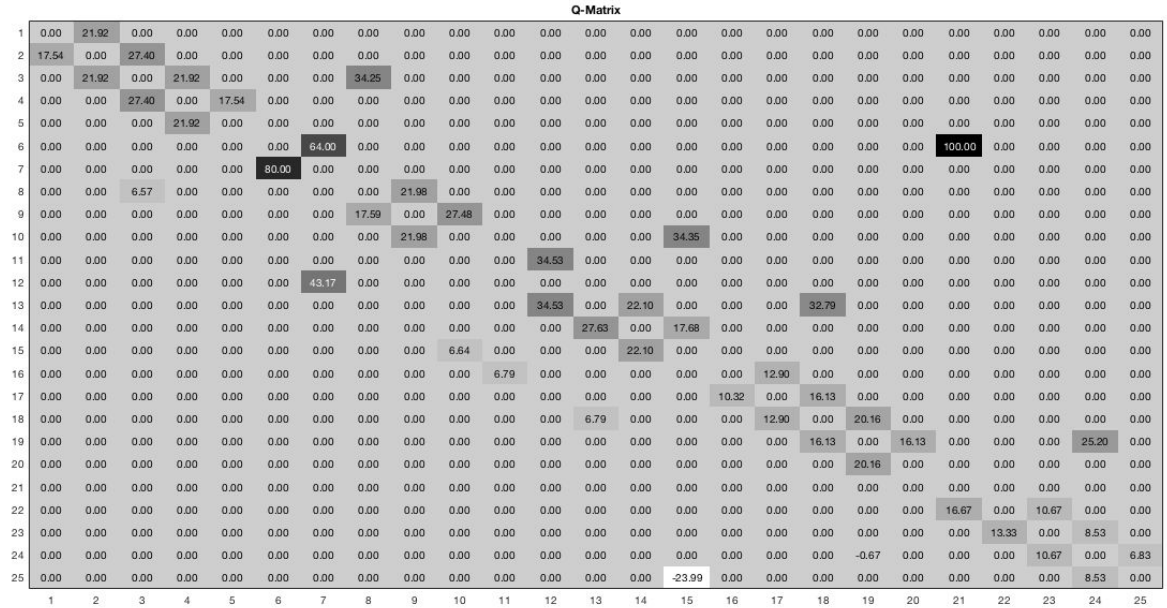


Figure 7. – Updated Q-Matrix

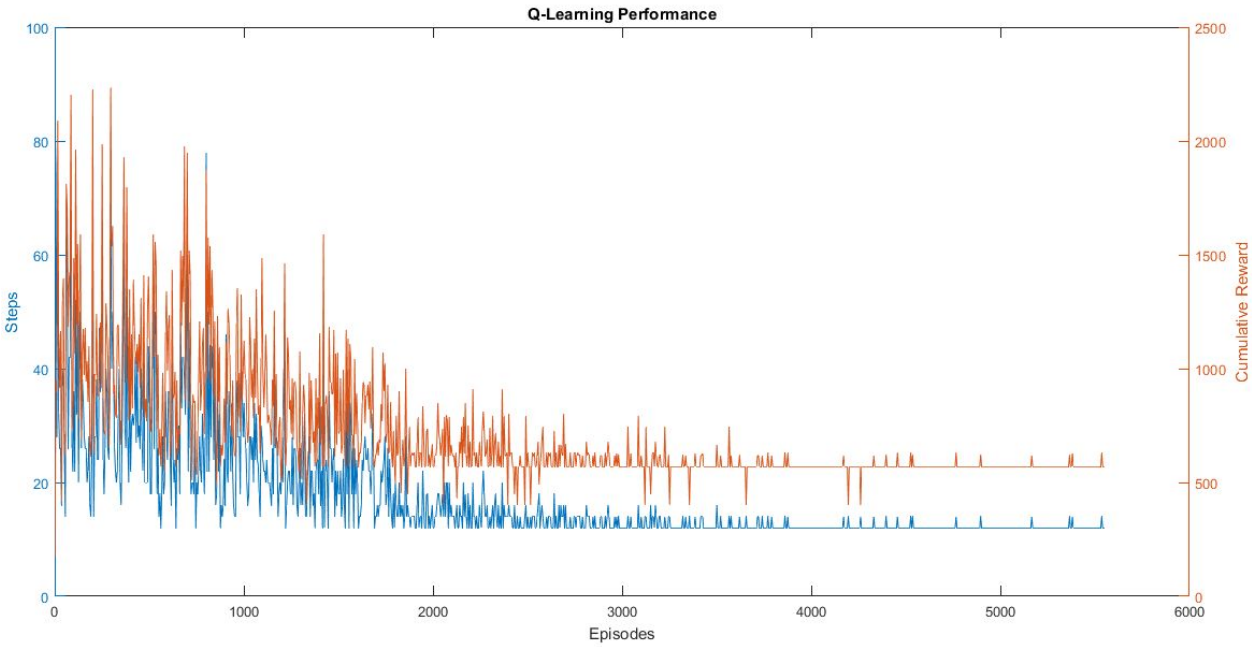


Figure 8. – Cumulative Reward and Steps vs Episodes

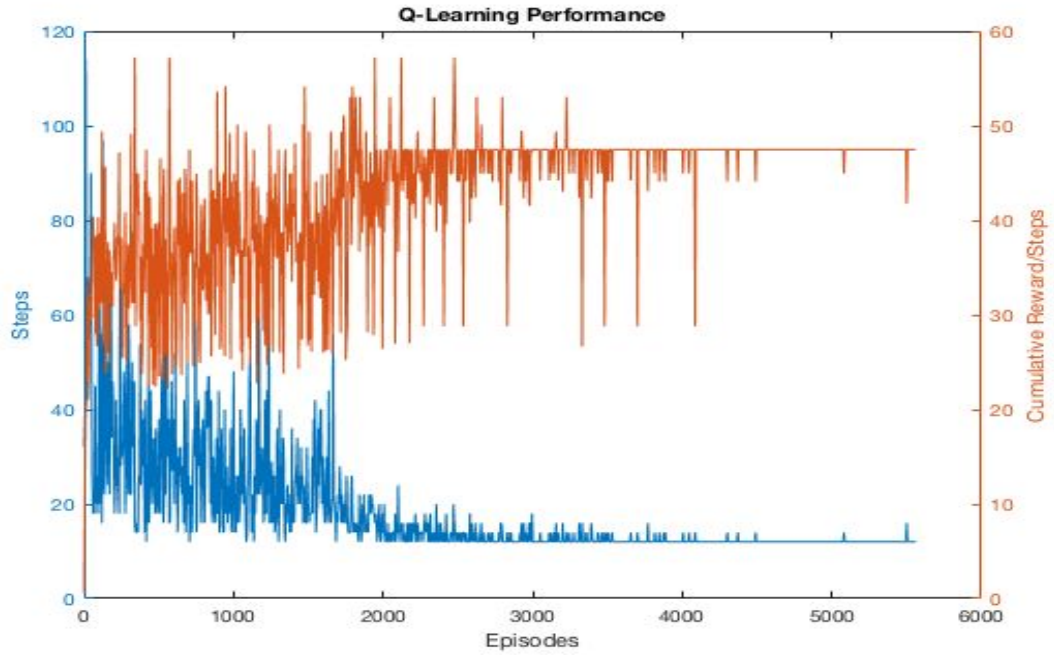


Figure 9. – Cumulative Reward/Step and Steps vs Episodes

Our agent was able to converge to the optimal 12 steps without difficulty. At the end of the learning, the Q-Matrix (Figure 7.) reveals the transition with the highest reward is from state 6⟶21, the shortcut ladder to the goal state. On the other hand, the lowest and only negative reward is from state 25⟶15, the only double-level snake. Interestingly, the reward for the snake transition from state 12⟶7 is positive. This is due to the agent learning that is the only way to the shortcut and therefore the optimal path.

It takes our agent about 2000 episodes to settle (Figures 8 and 9). Until then, there are considerable variations for both steps, cumulative reward, and cumulative reward / Steps. This is due to our agent prioritizing exploration at the beginning of the learning process. At around 2000 episodes, the agent clearly begins to exploit and to choose actions with the highest reward. That explains the plateau in Cumulative Reward / Steps as well as the sudden drop in the number of steps (Figure 9.).

2. Advanced cases

In this section, we test and report on different parameter values. For each set of tests, we keep the other parameter values to their default values (ε =0.8, α =0.5, γ =0.8).

2.1 Case 2: Altered Gamma values

<u>For Gamma = 0.1</u>
The agent only converges to a minimum of 14 steps, the Logical path. It does not understand that taking a negative reward will lead to a shortcut. Consequently, both the cumulative reward and cumulative reward / step are lower than for the default parameters. This is due to the fact it does not take the shortcut ladder (+120) and takes 2 extra steps.

There are a lot of negative values in the Q-Matrix(figure 10.) in comparison to the original parameters. In fact, every single snake is negative, even the one which leads to the shortcut. That explains why the agent is not able to learn the optimum 12 step shortcut path. However, the shortcut ladder and goal state values remain the exact same.

The overall decrease of the Q-values is due to the fact that the value of Gamma decrease the importance of the next states: the value of $\gamma max_a\, Q(s_{t+1}, a)$ decreases. Hence, the Q-value at the crossroad is higher from state 13 to state 18 than to state 12 which explain why the agent does not take the shortcut. So the weight of the Q-value from state 6 to state 21 does not affect enough the Q-value from state 13 to state 12 because the low value of the discount factor has diminished its importance.

Q Matrix for gamma = 0.1

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.02	0.00	1.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.17	0.00	0.17	0.00	0.00	0.00	16.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	1.67	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	10.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	-19.16	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	1.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	16.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	-19.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1.98	0.00	0.17	0.00	0.00	0.00	16.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.68	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-19.16	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-21.03	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-19.15	0.00	0.00	0.00	0.02	0.00	1.67	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.17	0.00	0.00	0.00	16.68	0.00
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.67	0.00	0.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.67	0.00	0.17	0.00	0.00
23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.67	0.00	0.02	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-19.16	0.00	0.00	0.00	0.17	0.00	0.00
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-41.65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00

Figure 10. - Q-Matrix for Gamma = 0.1

For Gamma = 0.5
Again, the agent converges to a minimum of 14 steps and is unable to learn the shortcut path. The cumulative reward and cumulative reward/step values are less than for the default parameters and close to the ones for gamma = 0.1.
For Gamma = 0.99
The agent finds the optimal 12 step shortcut. The high Gamma value makes the agent seek the highest future rewards possible. In turn, the cumulative reward and cumulative reward/step are the highest observed so far.

The values of the updated Q-Matrix are generally higher than with the default parameters and are all positive. However, the Shortcut Path transitions have considerably higher Q-Values than the Logical Path transitions (Figure 11.).

Here, we have again reached a level where the discount factor gives enough importance to the Q-value from state 6 to state 21 to impact the Q-value from state 12 to state 7 where we have a snake but where the Q-value becomes positive and very high. That also impacts the Q-value at the crossroad where going from state 13 to state 12 is higher than going to state 18.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.00	98.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	97.03	0.00	99.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	98.01	0.00	98.01	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	99.00	0.00	97.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	98.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	93.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	95.45	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	94.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	79.12	0.00	0.00	0.00	0.00	0.00	84.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	84.09	0.00	85.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	84.94	0.00	0.00	0.00	0.00	0.00	86.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	72.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	73.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	72.92	0.00	71.47	0.00	0.00	0.00	67.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	72.20	0.00	70.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	65.92	0.00	0.00	0.00	71.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	52.31	0.00	0.00	0.00	0.00	0.00	51.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	51.79	0.00	51.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	52.31	0.00	0.00	0.00	51.27	0.00	51.27	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	51.79	0.00	50.76	0.00	0.00	0.00	46.98	0.00
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	51.27	0.00	0.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15.89	0.00	14.61	0.00	0.00
23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15.68	0.00	31.06	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	31.39	0.00	0.00	0.00	30.61	0.00	30.76
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	31.07	0.00

Figure 11. - Q-Matrix for Gamma = 0.99

2.2 Case 3: Changing Alpha values

For Alpha = 0.1
The agent finds the optimal 12 step route. The Q-values remain the exact same as with the default parameters. There are no noticeable changes in the rewards.

For Alpha = 0.8
The agent finds the optimal 12 step route. The Q-values remain the exact same as with the default parameters. There are no noticeable changes in the rewards.

For Alpha = 1
The agent finds the optimal 12 step route. The Q-values remain the exact same as with the default parameters. There are no noticeable changes in the rewards.

2.3 Case 4: Changing ε values

Changing the ε -value will alter the action selection policy. The lowest it is set, the more the agent will exploit from the beginning. The highest it is, the more it will explore.

For ε = 0.1

The agent finds the optimal 12 step route. The final cumulative reward and cumulative reward/step values are similar to those of default parameters. However, because it favors exploiting from the start, the agent finds the Shortcut path much faster than with default parameters. Subsequently, the cumulative reward has much less variations than by default (Figure 12.).

Exploit count: 67599
Explore count: 1007

The Crossroad (State 13) is a key point in the Q-Matrix. Passed that state and into the Shortcut path (13→12), the Q-values remain the same as by default. However, going into the Logical Path (13→18), the Q-values are lower than with default parameters. This is due to the change in policy and the exploiting mindset early on. Interestingly, the downward ladder transition back to the Crossroad (18→13) is the first negative Q-value we observe for ladders from all tested parameters thus far. Additionally, the two-level snake (25→15), which was originally the only negative value, is now zero.

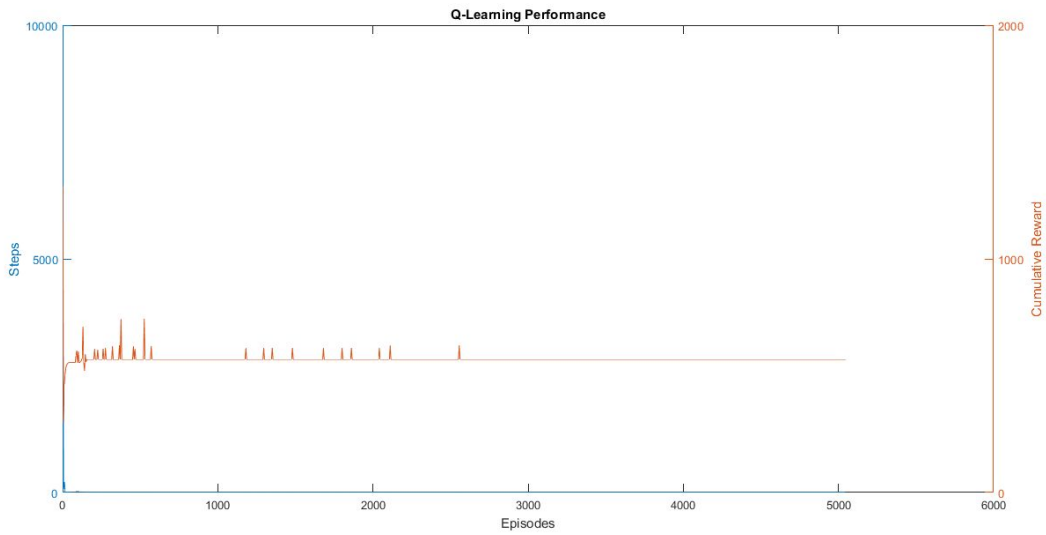


Figure 12. – Cumulative Reward and Steps vs Episodes for E = 0.1

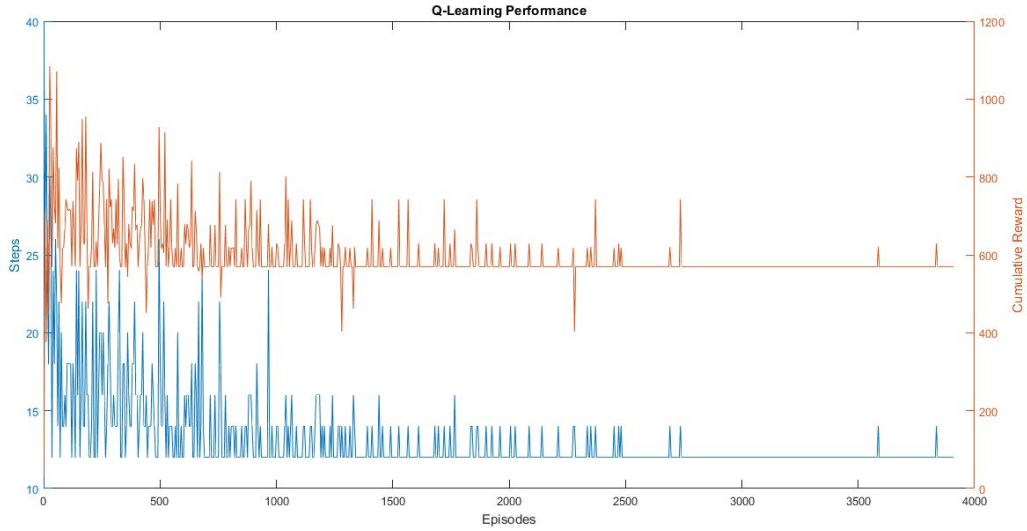


Figure 13. – Cumulative Reward and Steps vs Episodes for E = 0.5

For ε = 0.5

The agent finds the optimal 12 step path. The cumulative reward and cumulative reward/step are similar to the ones by default parameters. The agent converges to the optimal path faster than with default parameters and has very little variations in the number of steps. It only takes about 1000 episodes for the agent to stabilize. During those, it only ever reaches a maximum of 40 steps, the lowest observed thus far with all tested parameters. After those 1000 episodes, the agent rarely exceeds 15 steps for the rest of the learning period (Figure 13.).

Exploit count: 46264
Explore: 4928

Again, the Crossroad (state 13) is key in the Q-Matrix . We observe interesting variations from the previous ε = 0.1 test. Both the ladders located on levels above the Crossroads have negative values, although they are extremely small (smaller than for ε = 0.1). The Q-values for the double-level snake (25→15) change again and become the largest negative value (zero for ε = 0.1). It is however still lower than with the default parameters.

For ε = 1

The agent converges to 12 steps in about the same time as with default parameters. In this setup, we even out the exploring and exploiting mindsets. Both cumulative reward and cumulative reward/step start off with a very high variation but converges to about the same as with default parameter.

Exploit: 65259
Explore: 55227

The Q-matrix remains identical to the one with default parameters.

2.4 Case 4: Altered State Function

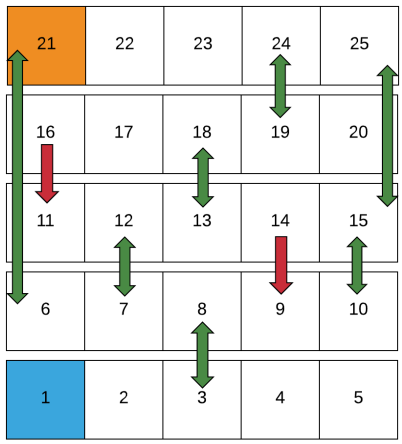


Figure 14. – Altered State Transition Function

The agent is now free to move between states 7 and 8. The two-level snake (25→15) has been changed into a ladder. The snake between 12 and 7 has also been turned into a ladder. We added a snake between 14 and 9 (Figure 14.).

As there are many more route options to the goal state, it takes the agent many more learning episodes to converge to the optimal 6 steps. The Q-values are all positive apart from the downward double-ladder transition (25→15). That same ladder has a large reward when taken upwards. Additionally, the cumulative reward/step is higher than with default parameters (figure 15.)

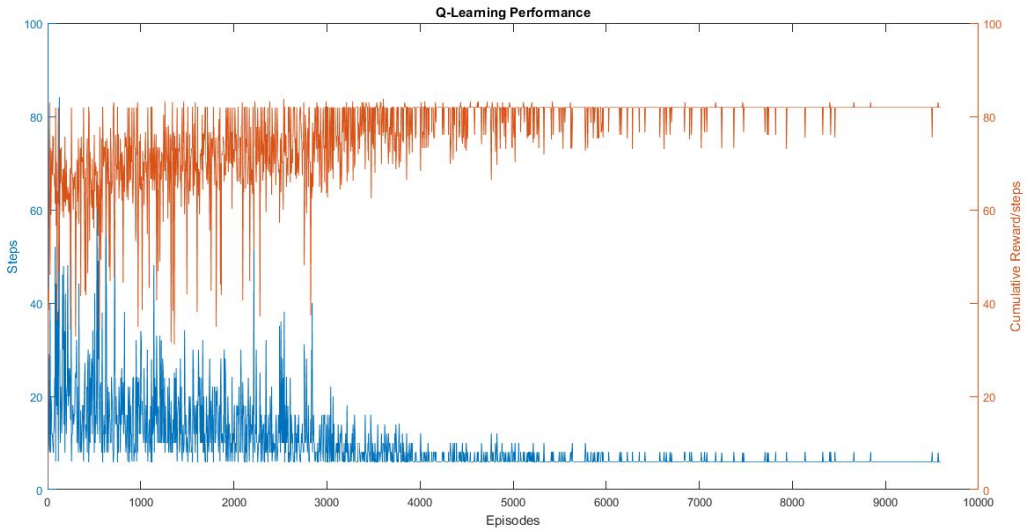


Figure 15. – Cumulative Reward/Step and Steps vs Episodes

2.5 Case 5: Changing the Reward Function

We keep the original domain but change the rewards (Figure 16.). The ladders have equally opposite rewards for going up or down. The reward for taking the shortcut ladder has been set down to +80.

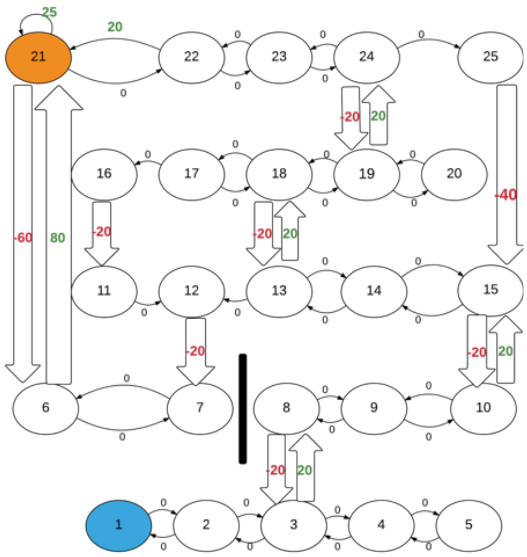


Figure 16. – Altered Reward function graph

In this setup, the agent does not find the optimal 12 step route. Nonetheless, the shortcut ladder is the highest Q-value after convergence and the snake leading up to it the second highest. Still, the agent does not converge to the Shortcut path. For the logical path, the ladder rewards are higher than with original parameters. The Crossroad ladder (13→18) also has a very high Q-value (Figure 19.) which explains why the agent decides to take that path over the shortcut. The cumulative reward/step is lower than with default parameters and varies less at the beginning of the learning process (Figure 17.).

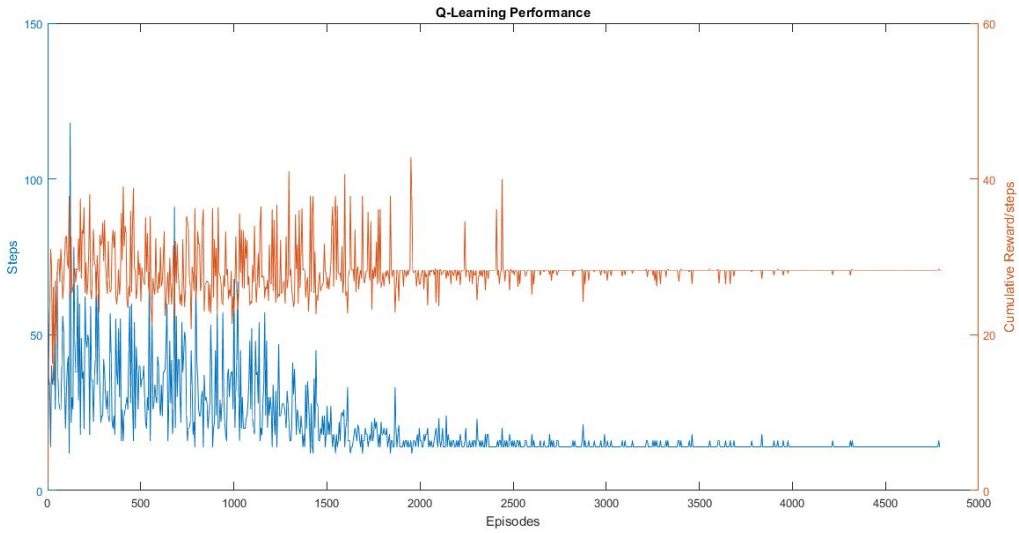


Figure 17. – Cumulative Reward/Step and Steps vs Episodes

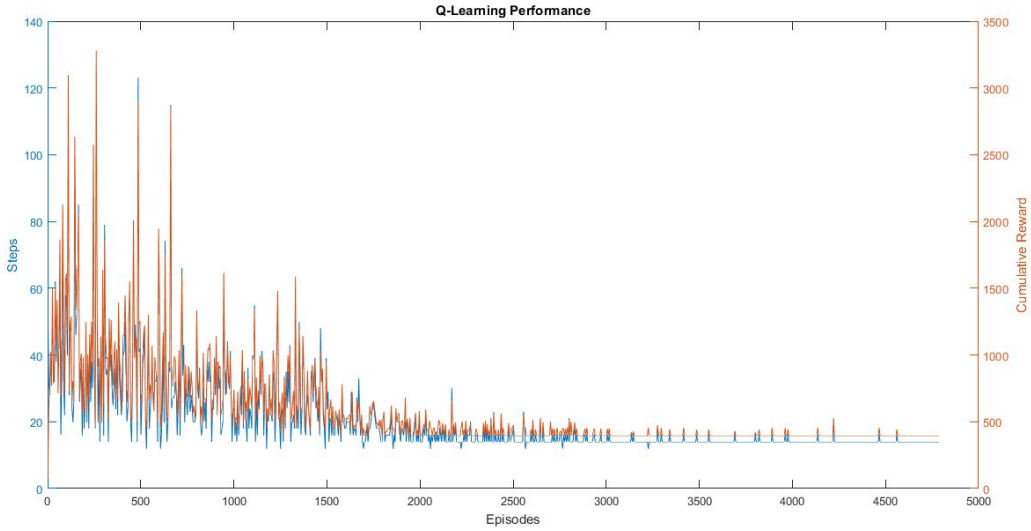


Figure 18. – Cumulative Reward and Steps vs Episodes

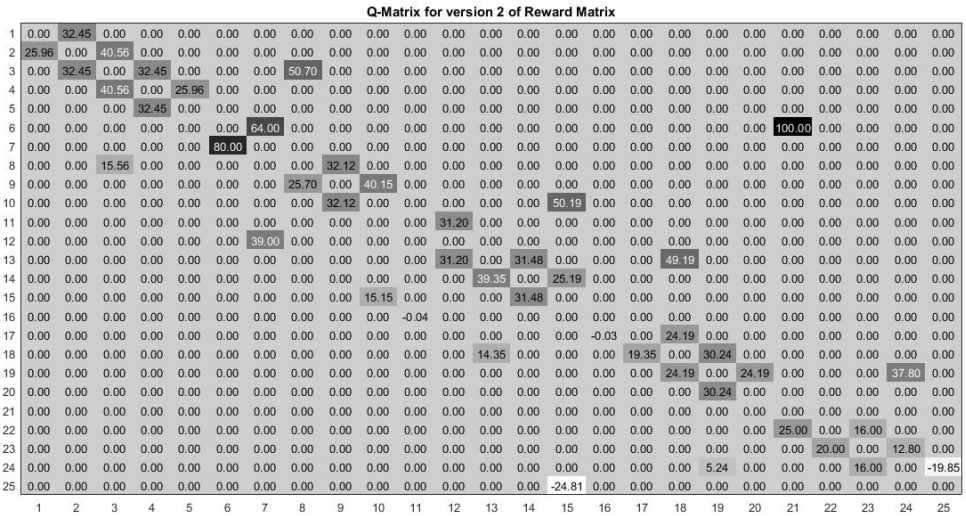


Figure 19. – Q-Matrix

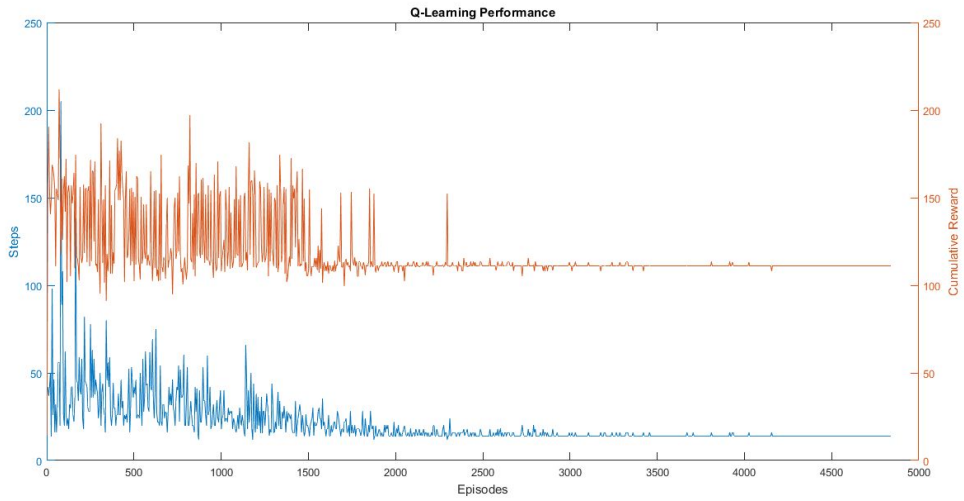
3. Appendix

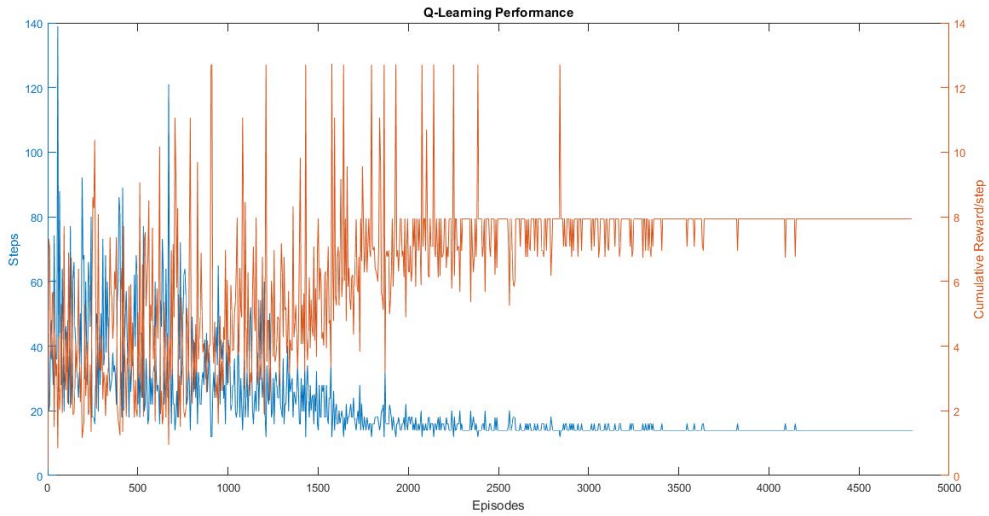
3.1 Altered Gamma Values graphs

Gamma = 0.1

Q Matrix for gamma = 0.1

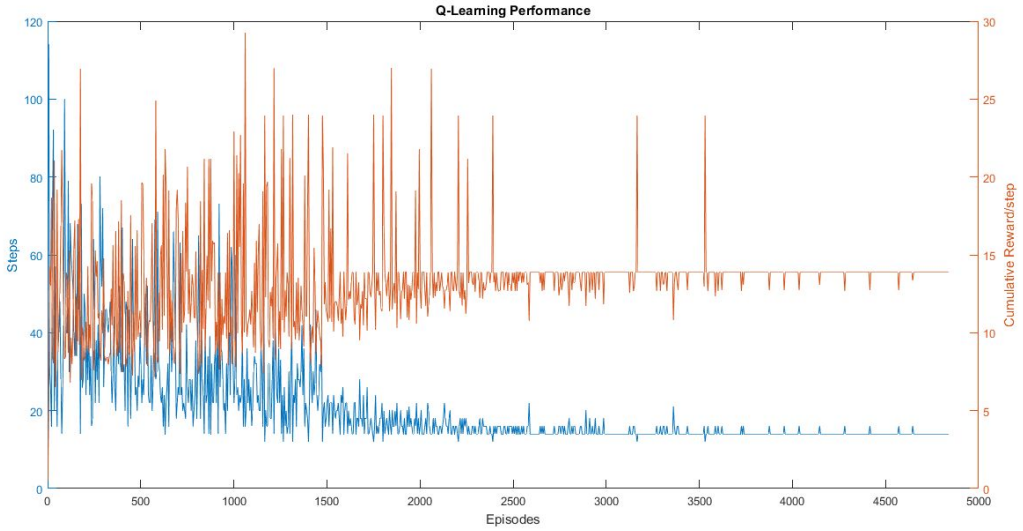
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.02	0.00	1.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.17	0.00	0.17	0.00	0.00	0.00	16.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	1.67	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	10.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	-19.16	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	1.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	16.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	-19.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1.98	0.00	0.17	0.00	0.00	0.00	16.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.68	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-19.16	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-21.03	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-19.15	0.00	0.00	0.00	0.02	0.00	1.67	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.17	0.00	0.00	0.00	16.68	0.00
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.67	0.00	0.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.67	0.00	0.17	0.00	0.00
23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.67	0.00	0.02	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-19.16	0.00	0.00	0.00	0.17	0.00	0.00
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-41.65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00

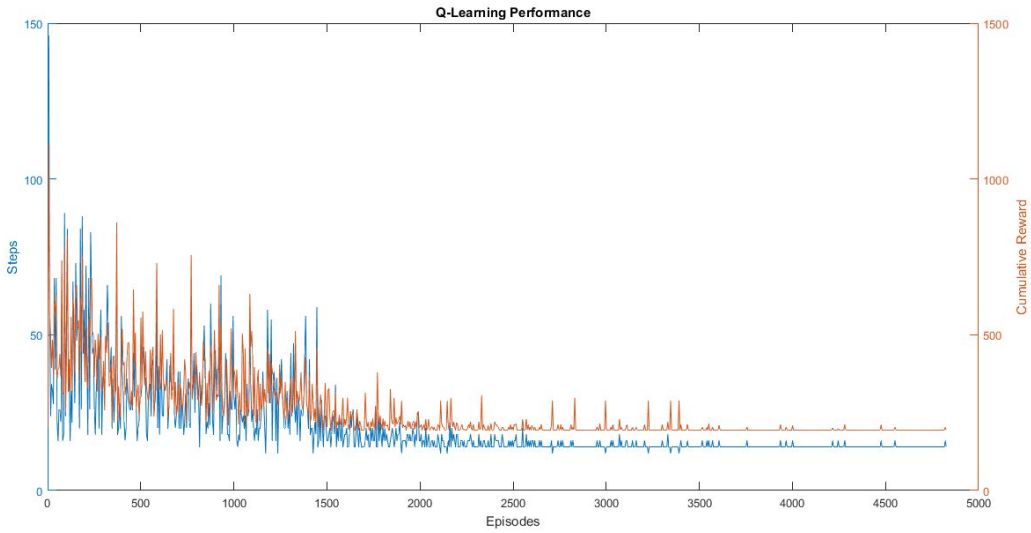




gamma = 0.5

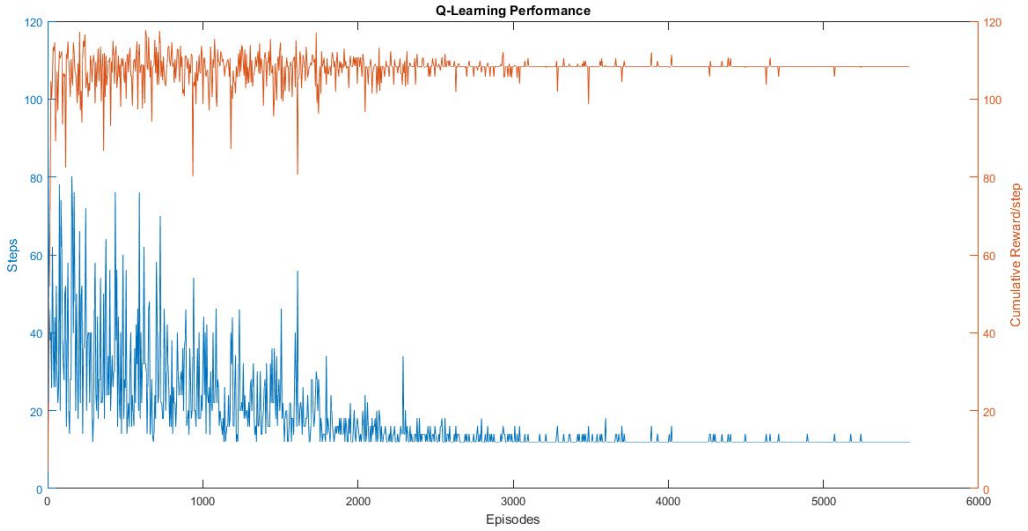
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.00	4.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	2.39	0.00	9.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	4.77	0.00	4.77	0.00	0.00	0.00	19.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	9.54	0.00	2.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	4.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	25.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	50.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	-11.29	0.00	0.00	0.00	0.00	0.00	4.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.42	0.00	9.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.83	0.00	0.00	0.00	0.00	0.00	19.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	4.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.08	0.00	5.34	0.00	0.00	0.00	21.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.68	0.00	2.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-11.17	0.00	0.00	0.00	5.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-19.79	0.00	0.00	0.00	0.00	0.00	2.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.17	0.00	4.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-10.16	0.00	0.00	0.00	2.34	0.00	9.38	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.69	0.00	4.69	0.00	0.00	0.00	18.75	0.00
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	9.38	0.00	0.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.67	0.00	4.17	0.00	0.00
23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.33	0.00	2.08	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-11.46	0.00	0.00	0.00	4.17	0.00	1.04
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-39.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.08	0.00

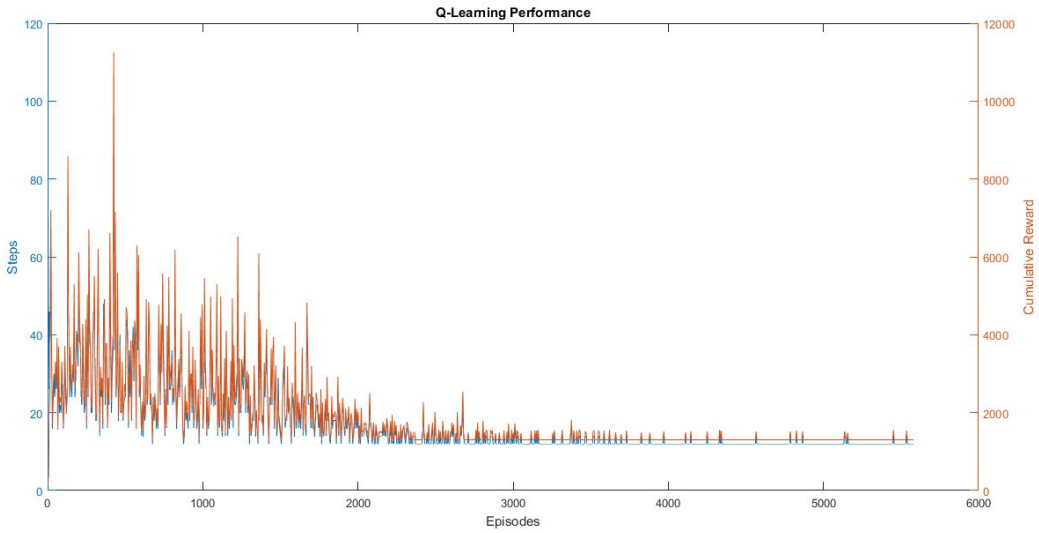




gamma = 0.99

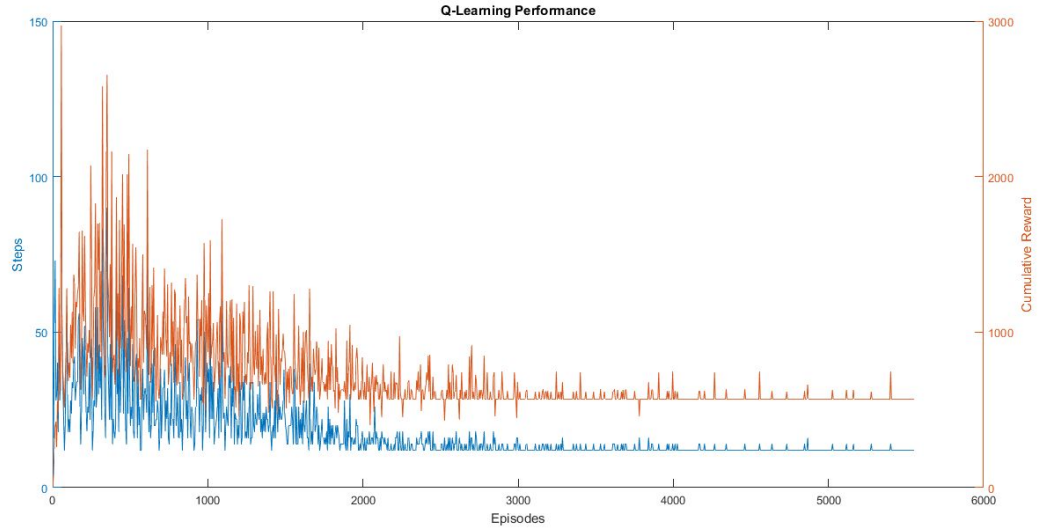
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.00	98.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	97.03	0.00	99.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	98.01	0.00	98.01	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	99.00	0.00	97.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	98.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	93.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	95.45	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	94.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	79.12	0.00	0.00	0.00	0.00	0.00	84.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	84.09	0.00	85.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	84.94	0.00	0.00	0.00	0.00	0.00	86.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	72.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	73.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	72.92	0.00	71.47	0.00	0.00	0.00	67.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	72.20	0.00	70.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	65.92	0.00	0.00	0.00	71.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	52.31	0.00	0.00	0.00	0.00	0.00	51.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	51.79	0.00	51.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	52.31	0.00	0.00	0.00	51.27	0.00	51.27	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	51.79	0.00	50.76	0.00	0.00	0.00	46.98	0.00
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	51.27	0.00	0.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15.89	0.00	14.61	0.00	0.00
23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15.68	0.00	31.06	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	31.39	0.00	0.00	0.00	30.61	0.00	30.76
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	31.07	0.00

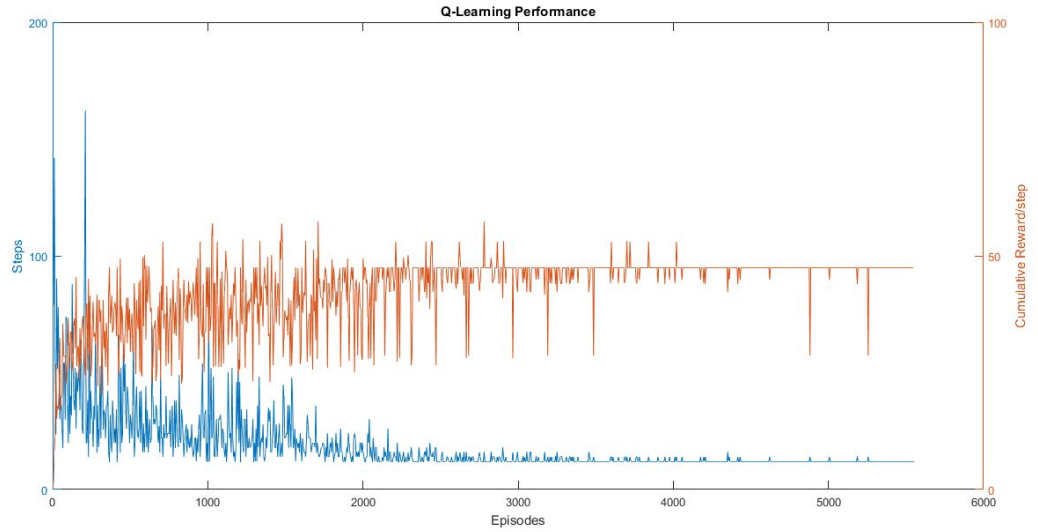




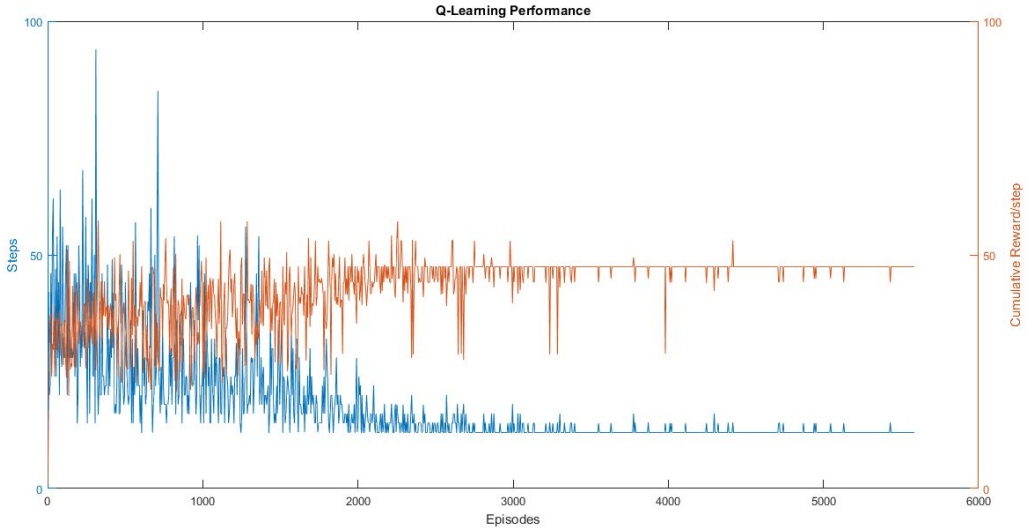
3.2 Altered Alpha values graphs

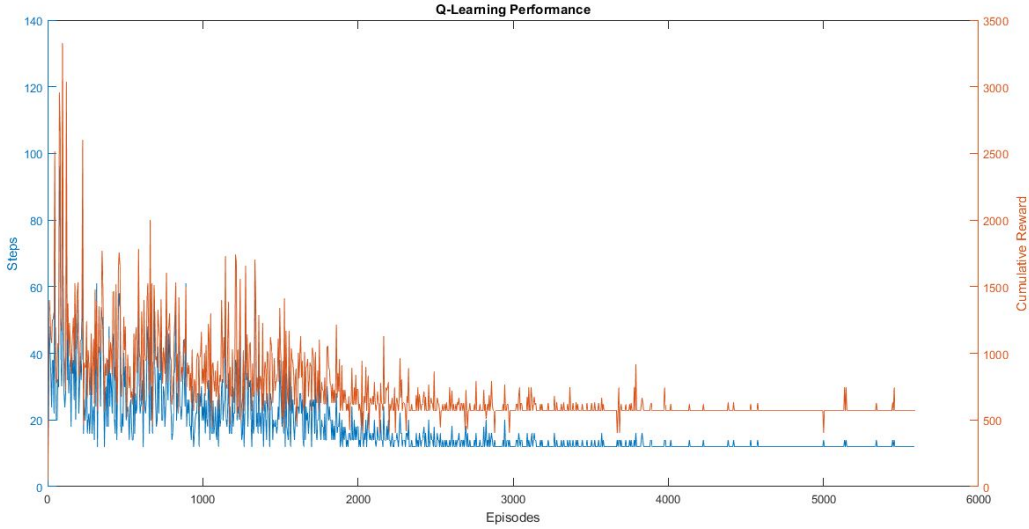
Alpha = 0.1



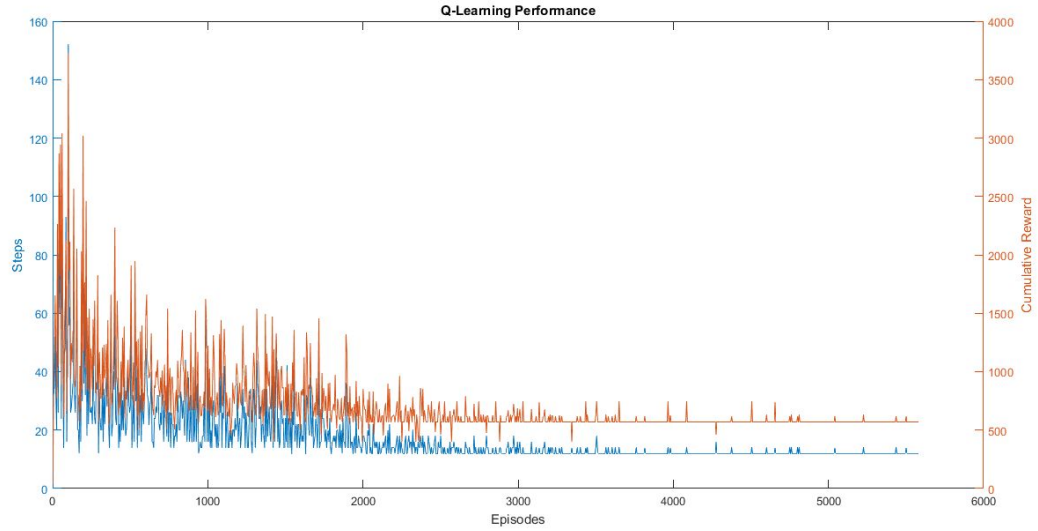


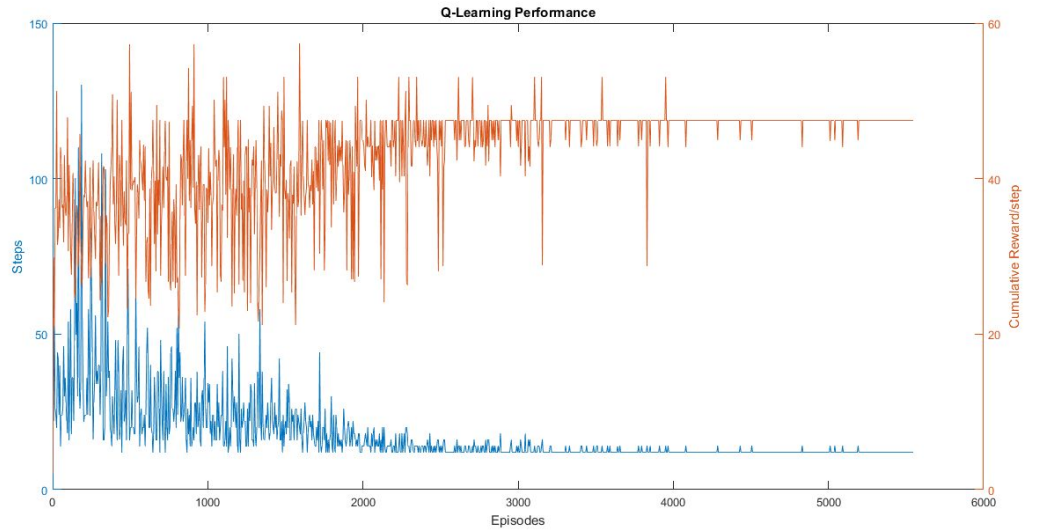
Alpha = 0.8


Q-Learning Performance
Steps
Cumulative Reward/step
Episodes
0
1000
2000
3000
4000
5000
6000
0
50
100


Q-Learning Performance
Steps
Cumulative Reward
Episodes
0
1000
2000
3000
4000
5000
6000
0
20
40
60
80
100
120
140
0
500
1000
1500
2000
2500
3000
3500

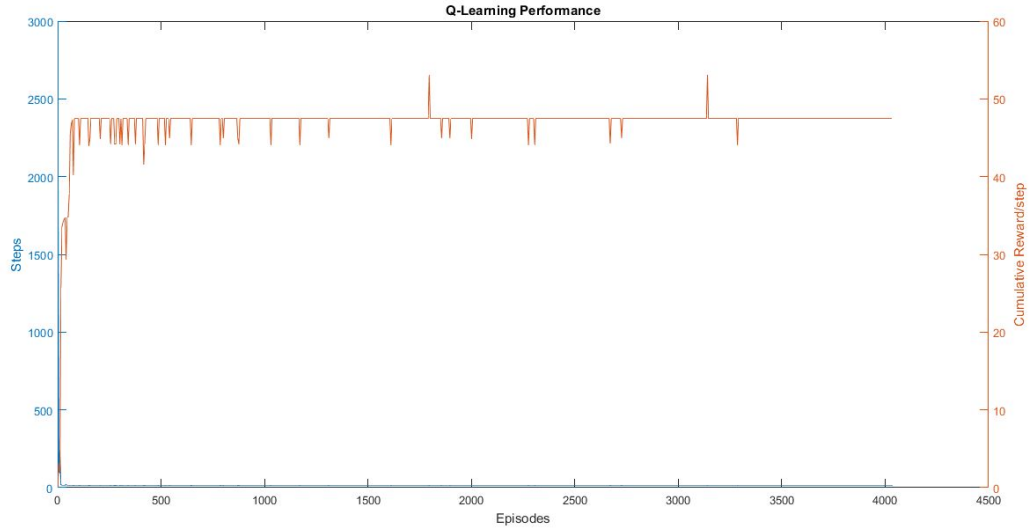
Alpha = 1

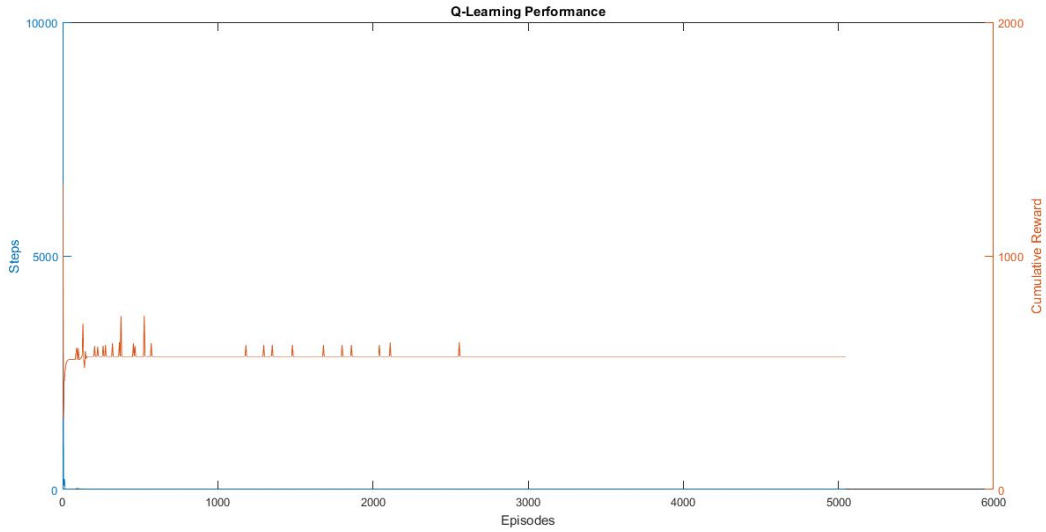

Q-Learning Performance
Steps
Cumulative Reward
Episodes
0
20
40
60
80
100
120
140
160
0
500
1000
1500
2000
2500
3000
3500
4000
0
1000
2000
3000
4000
5000
6000


Q-Learning Performance
Steps
Cumulative Reward/step
Episodes
0
50
100
150
0
20
40
60
0
1000
2000
3000
4000
5000
6000

3.3 Altered Epsilon values graphs

Epsilon = 0.1

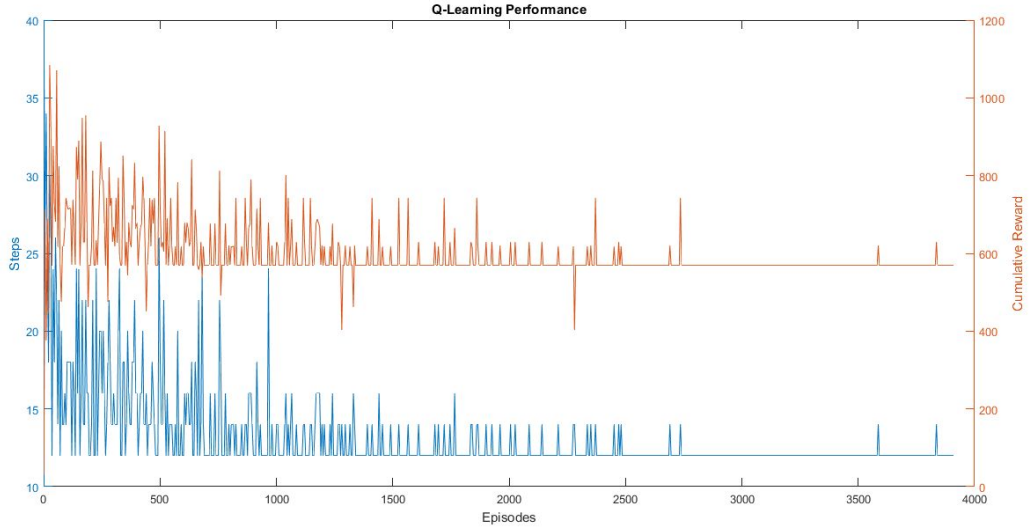


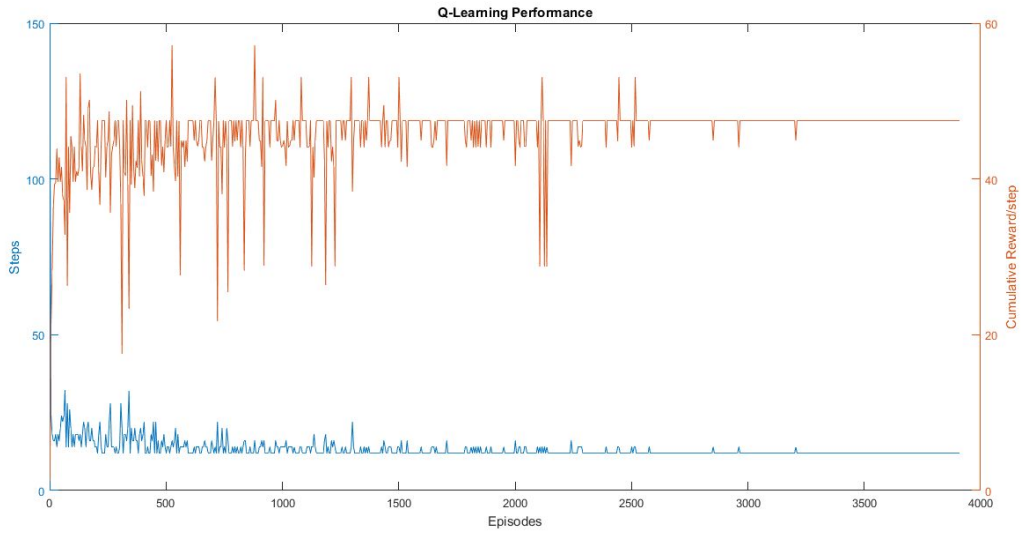


Epsilon = 0.5

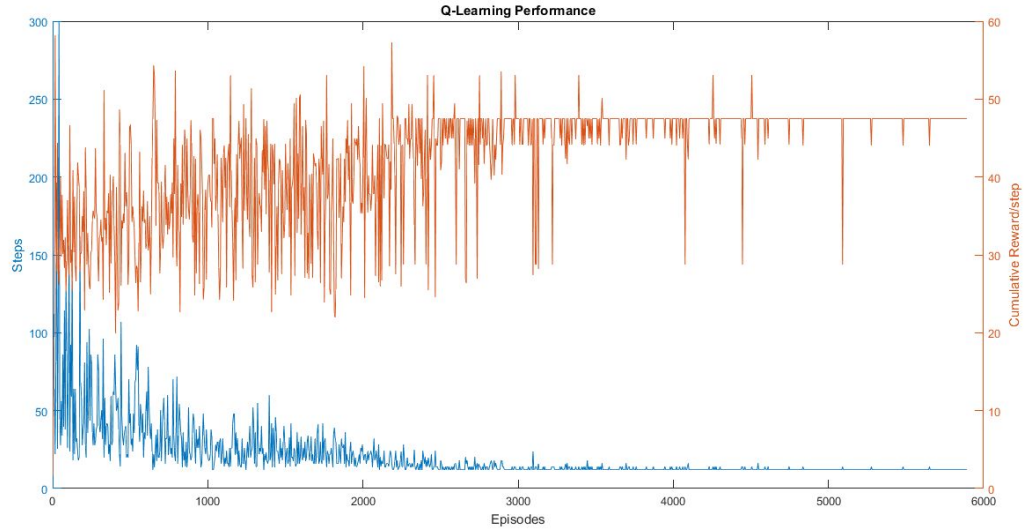
Q-matrix for epsilon = 0.5

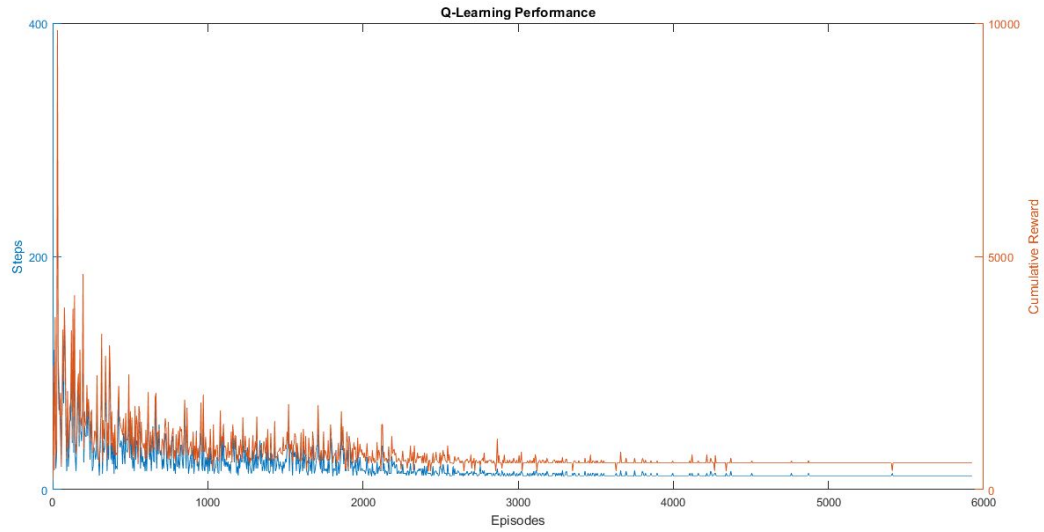
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.00	21.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	17.54	0.00	27.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	21.92	0.00	21.92	0.00	0.00	0.00	34.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	27.40	0.00	17.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	21.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	64.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	80.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	6.57	0.00	0.00	0.00	0.00	0.00	21.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	17.59	0.00	27.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	21.98	0.00	0.00	0.00	0.00	0.00	0.00	34.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	34.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	43.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	34.53	0.00	22.10	0.00	0.00	0.00	32.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	27.63	0.00	17.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.64	0.00	0.00	0.00	22.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.79	0.00	0.00	0.00	0.00	0.00	4.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.43	0.00	3.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.90	0.00	0.00	0.00	4.35	0.00	20.16	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15.65	0.00	6.72	0.00	0.00	0.00	25.20	0.00
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.66	0.00	0.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.67	0.00	10.52	0.00	0.00
23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	13.33	0.00	8.50	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1.55	0.00	0.00	0.00	10.67	0.00	0.85
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-17.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.13	0.00





Epsilon = 1

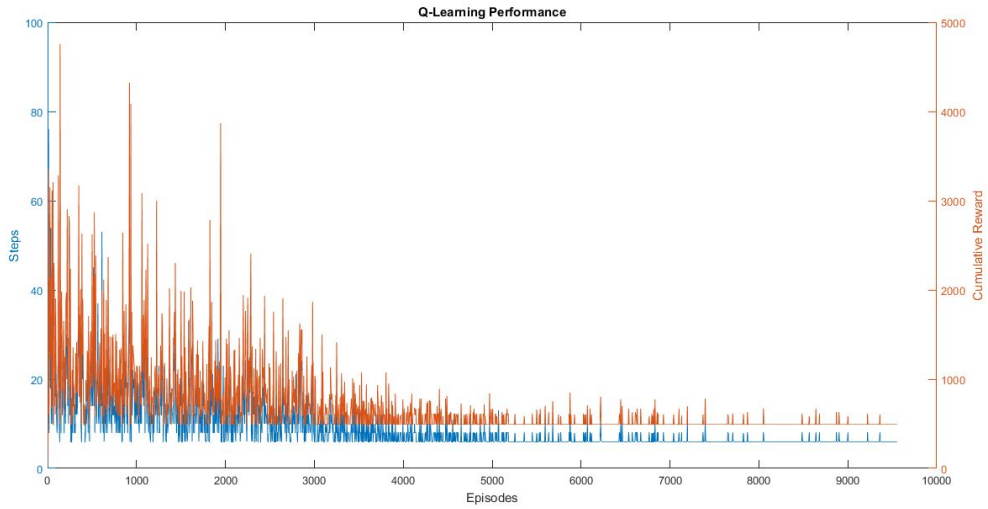

Q-Learning Performance
Steps
Cumulative Reward/step
Episodes
0
50
100
150
200
250
300
0
10
20
30
40
50
60
0
1000
2000
3000
4000
5000
6000

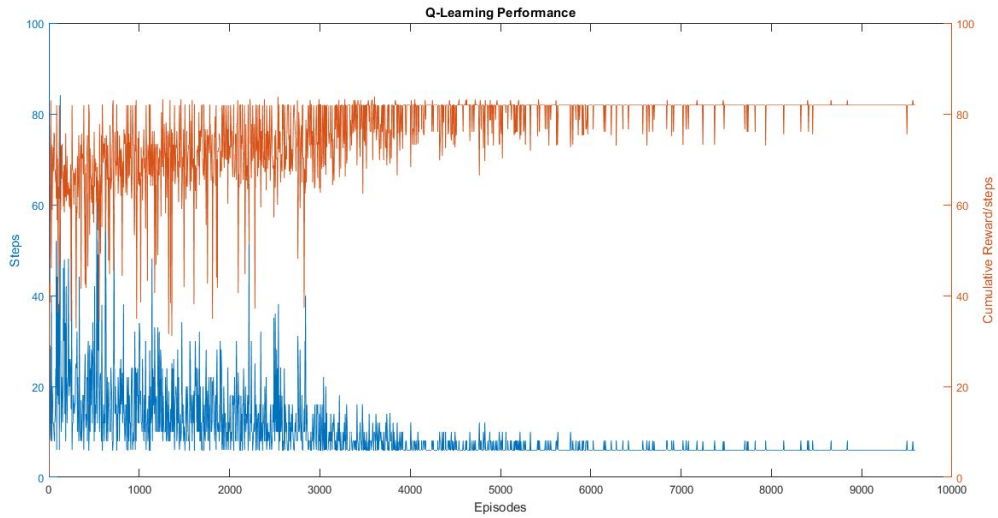

Q-Learning Performance
Steps
Cumulative Reward
Episodes
0
200
400
0
5000
10000
0
1000
2000
3000
4000
5000
6000

3.4 Altered State Transition function graphs

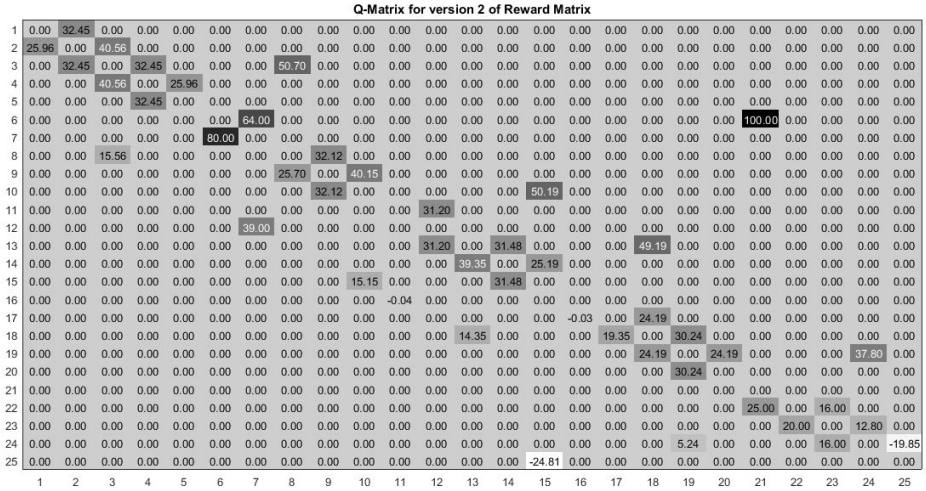
Q-Matrix new states

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.00	43.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	34.75	0.00	54.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	43.43	0.00	43.43	0.00	0.00	0.00	67.87	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	54.29	0.00	34.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	43.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	64.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	80.00	0.00	51.20	0.00	0.00	0.00	51.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	33.46	0.00	0.00	0.00	64.00	0.00	40.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	51.20	0.00	39.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	40.96	0.00	0.00	0.00	0.00	0.00	48.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	34.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	43.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	34.26	0.00	0.00	0.00	0.00	0.00	24.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	20.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	18.20	0.00	0.00	0.00	16.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	40.16
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.77	0.00	14.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.92	0.00	0.00	11.15	0.00	20.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.13	0.00	16.13	0.00	0.00	0.00	25.20	0.00	
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	20.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.67	0.00	10.67	0.00	0.00
23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	13.33	0.00	8.53	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.67	0.00	0.00	0.00	10.67	0.00	6.83
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-9.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.53	0.00





3.5 Altered Reward function graphs



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.00	32.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	25.96	0.00	40.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	32.45	0.00	32.45	0.00	0.00	0.00	50.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	40.56	0.00	25.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	32.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	64.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	80.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	15.56	0.00	0.00	0.00	0.00	0.00	32.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	25.70	0.00	40.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	32.12	0.00	0.00	0.00	0.00	0.00	50.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	31.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	39.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	31.20	0.00	31.48	0.00	0.00	0.00	49.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	39.35	0.00	25.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15.15	0.00	0.00	0.00	31.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.03	0.00	24.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	14.35	0.00	0.00	0.00	19.35	0.00	30.24	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	24.19	0.00	24.19	0.00	0.00	0.00	37.80	0.00
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.24	0.00	0.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	25.00	0.00	16.00	0.00	0.00
23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	20.00	0.00	12.80	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.24	0.00	0.00	0.00	16.00	0.00	-19.85
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-24.81	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00