

Name,
Oluwatomilayo Adegbite,
Nikolas Maier,

Student Number
500569283
500461990

CPS841 Assignment 3 Report

Question 1:

First Visit = 10 as $G_{10} = 10$ when $\gamma = 1$.

Every Visit Average = 5.5

$$G_n = Q_n + \frac{1}{n} (R_t + Q_n)$$

$$G_0 = 0$$

$$G_1 = 10 + \frac{1}{1} * (10 - 0)$$

$$G_2 = 9.5 + \frac{1}{2} * (9 - 10)$$

$$G_3 = 9 + \frac{1}{3} * (8 - 9.5)$$

$$G_4 = 8.5 + \frac{1}{4} * (7 - 9)$$

$$G_5 = 8 + \frac{1}{5} * (6 - 8.5)$$

$$G_6 = 7.5 + \frac{1}{6} * (5 - 8)$$

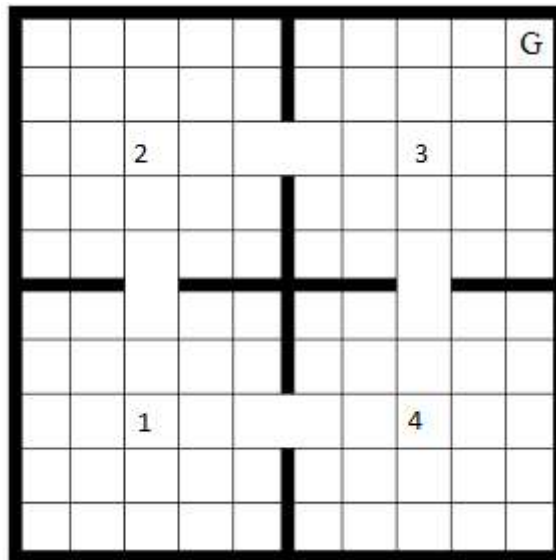
$$G_7 = 7 + \frac{1}{7} * (4 - 7.5)$$

$$G_8 = 6.5 + \frac{1}{8} * (3 - 7)$$

$$G_9 = 6 + \frac{1}{9} * (2 - 6.5)$$

$$G_{10} = 5.5$$

Question 2:



Grid Order as pictured

We found that Monte Carlo would, once it found a correct path, have its values converge quickly and find a path that worked best for it. There were some scenarios that would cause infinite loops, such as when it got stuck in a corner and due to a maximum number for steps would have its episode terminated. Now that whole corner's values would be incredibly negative

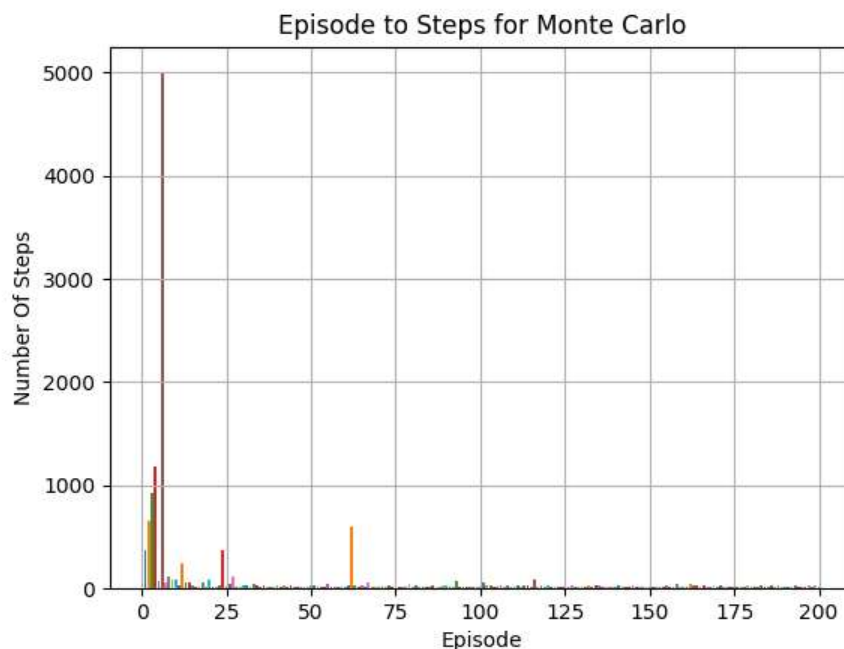
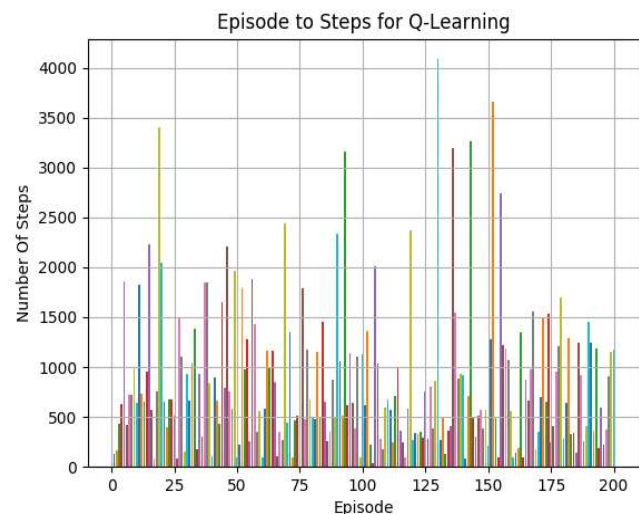
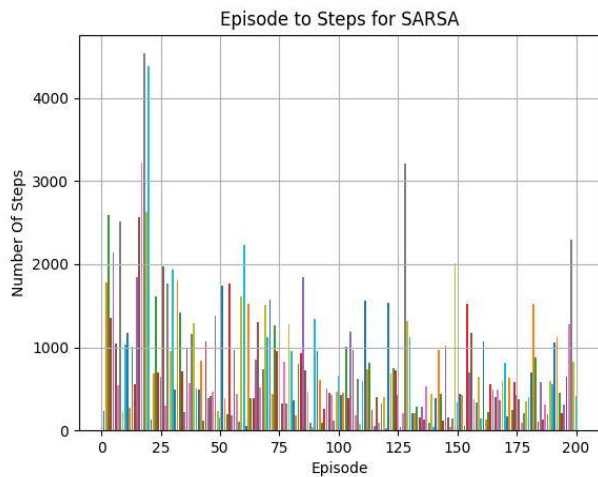
and any future episodes that involved that area could be forced into actions that prevented from leaving.

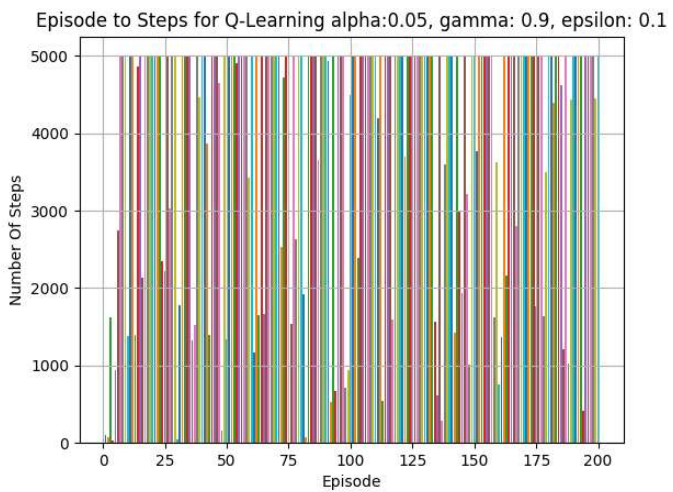
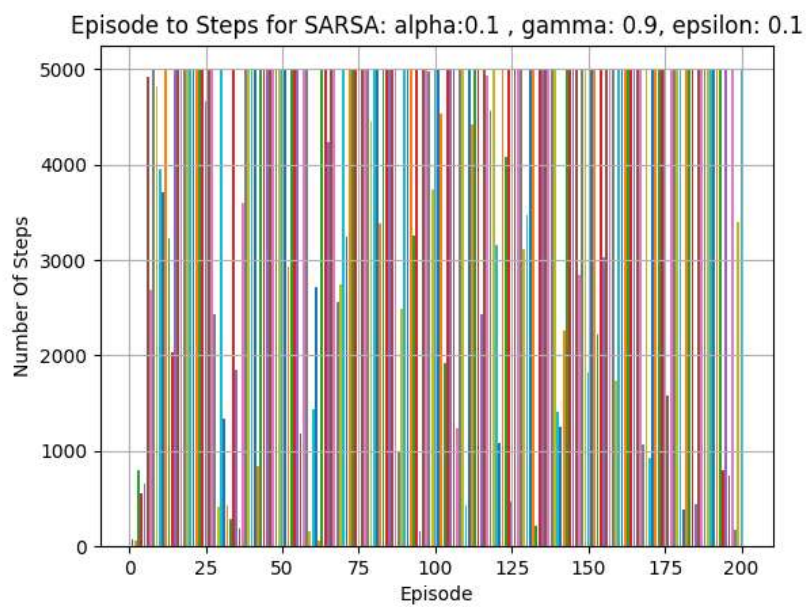
SARSA and Q-learning are both very good at getting to the goal state and given enough steps (3000 - 4000) they have almost a 100% success rate of getting to the Goal State. The issue is that these algorithms have issues creating an optimal policy that lowers the number of steps. Instead they seem to always take larger number of steps, and changing the alpha and epsilon can help increase the likelihood of reaching the Goal State.

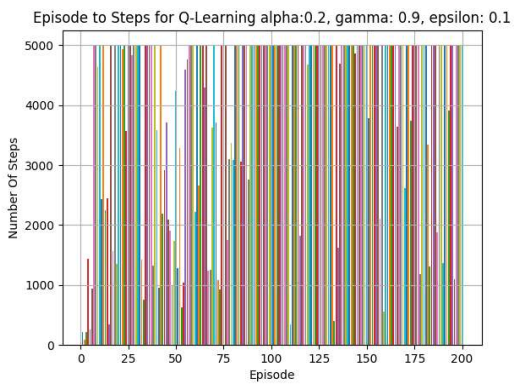
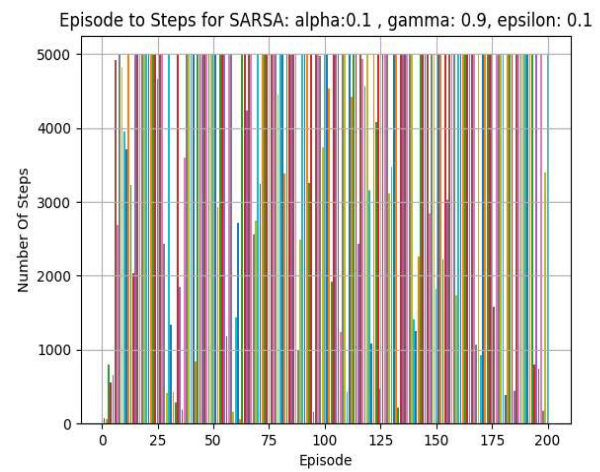
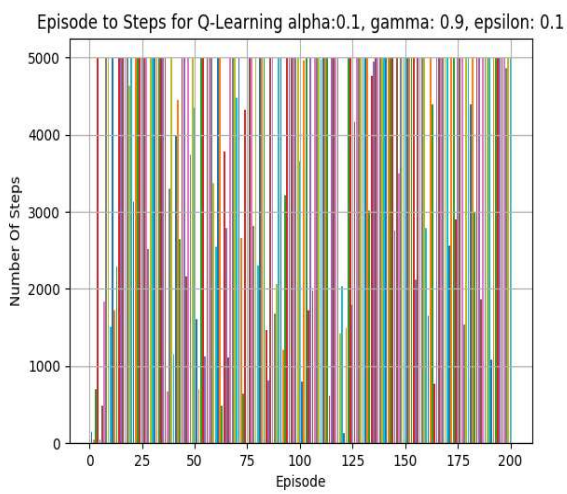
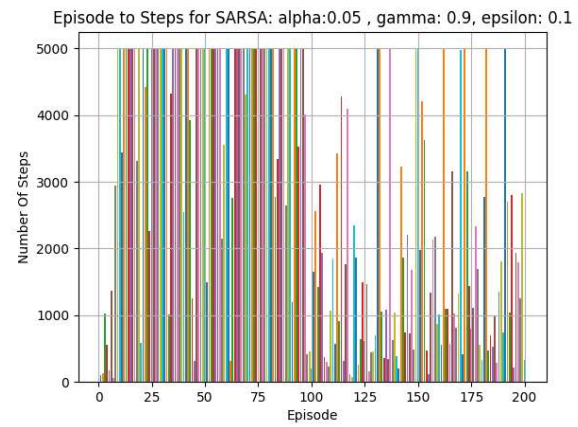
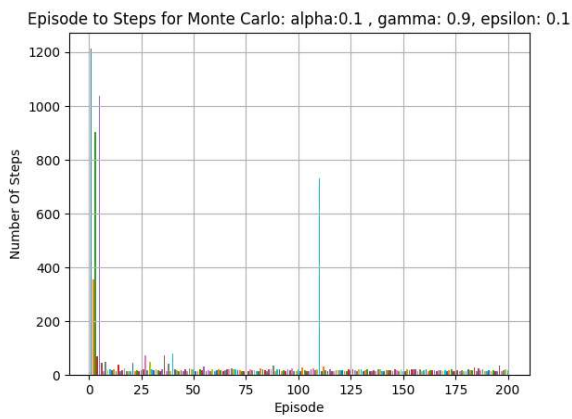
Please find below a large amount of graphs we have plotted over many runs of our algorithms that help agree with the above observations.

Ones with “Drop” in the title are ones where the epsilon was lowered to 0 over the run

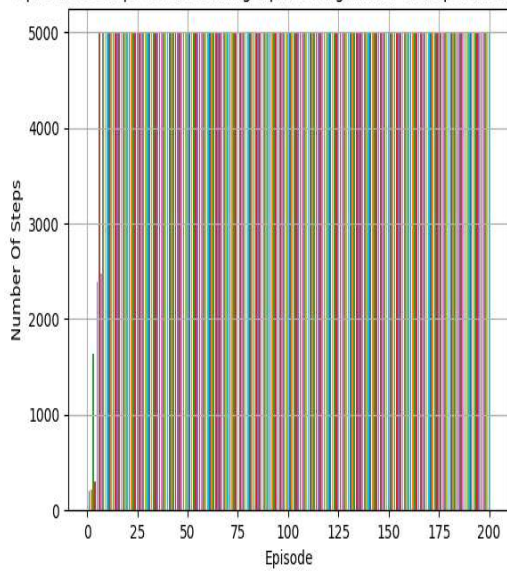
Agent starting {X:2,Y:2,Grid:1}



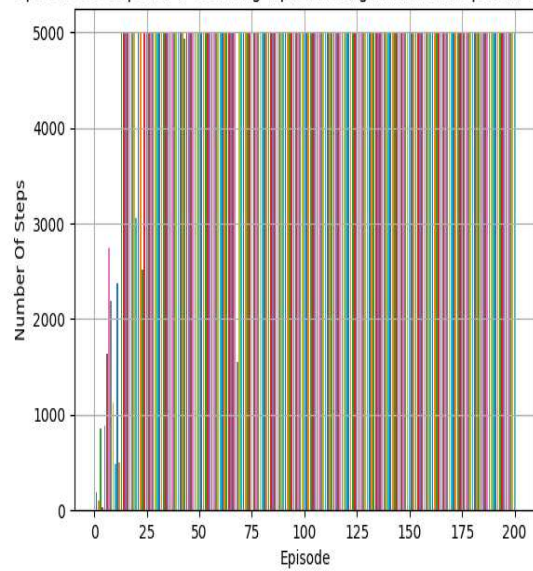




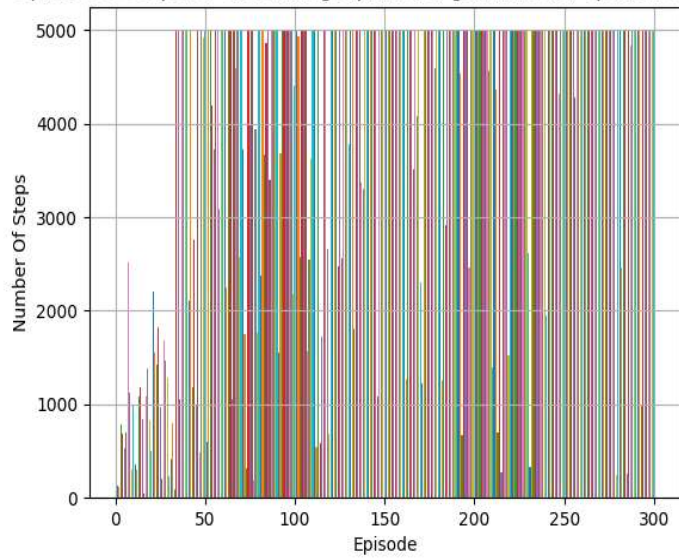
Episode to Steps for Q-Learning alpha:0.2, gamma: 0.9, epsilon: 0.1 Drop



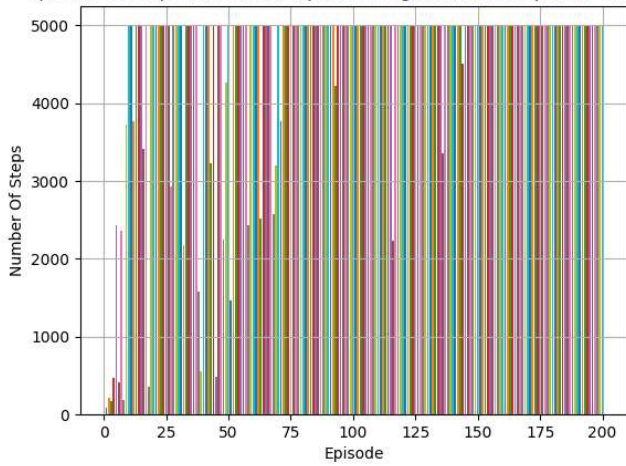
Episode to Steps for Q-Learning alpha:0.05, gamma: 0.9, epsilon: 0.1 Drop



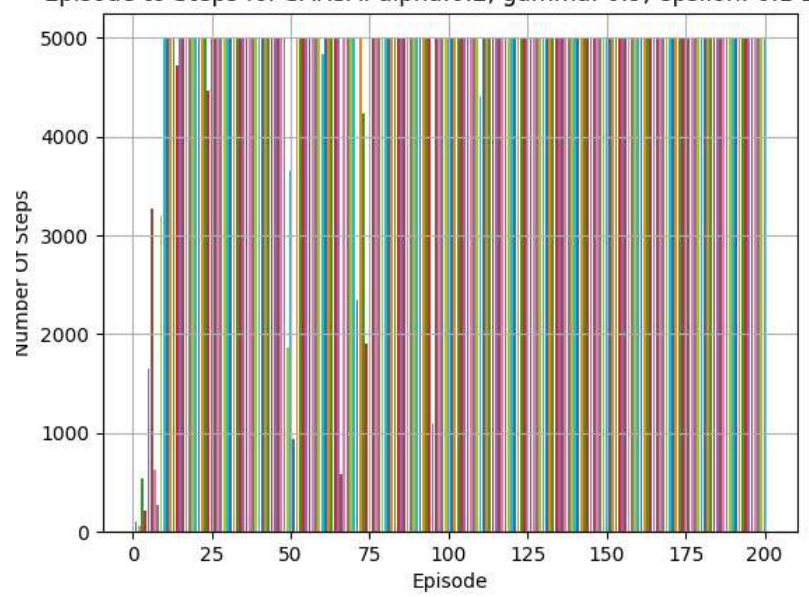
Episode to Steps for Q-Learning alpha:0.2 , gamma: 0.9, epsilon: 0.2 Drop



Episode to Steps for SARSA: alpha:0.05, gamma: 0.9, epsilon: 0.1 Drop



Episode to Steps for SARSA: alpha:0.2, gamma: 0.9, epsilon: 0.1 Drop



Episode to Steps for Q-Learning alpha:0.1, gamma: 0.9, epsilon: 0.1 Drop

