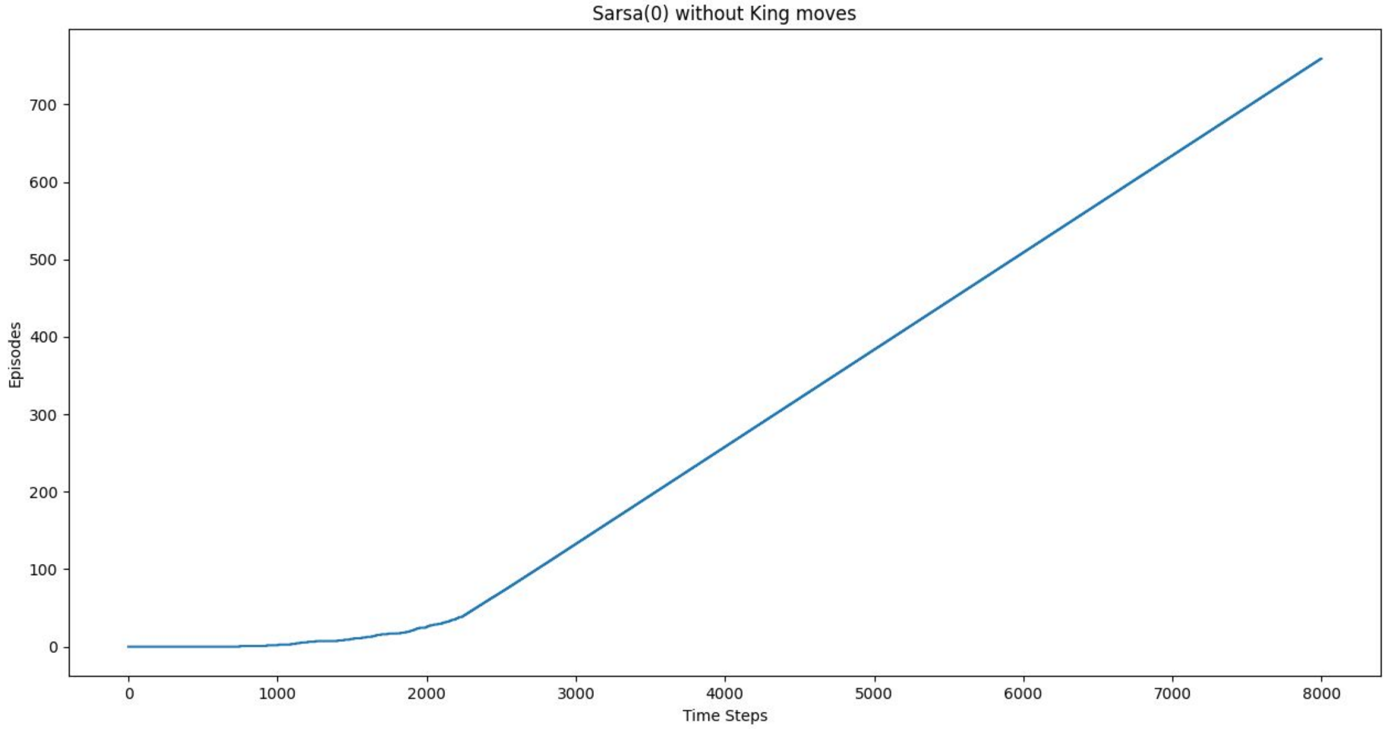
CS747 Assignment 3

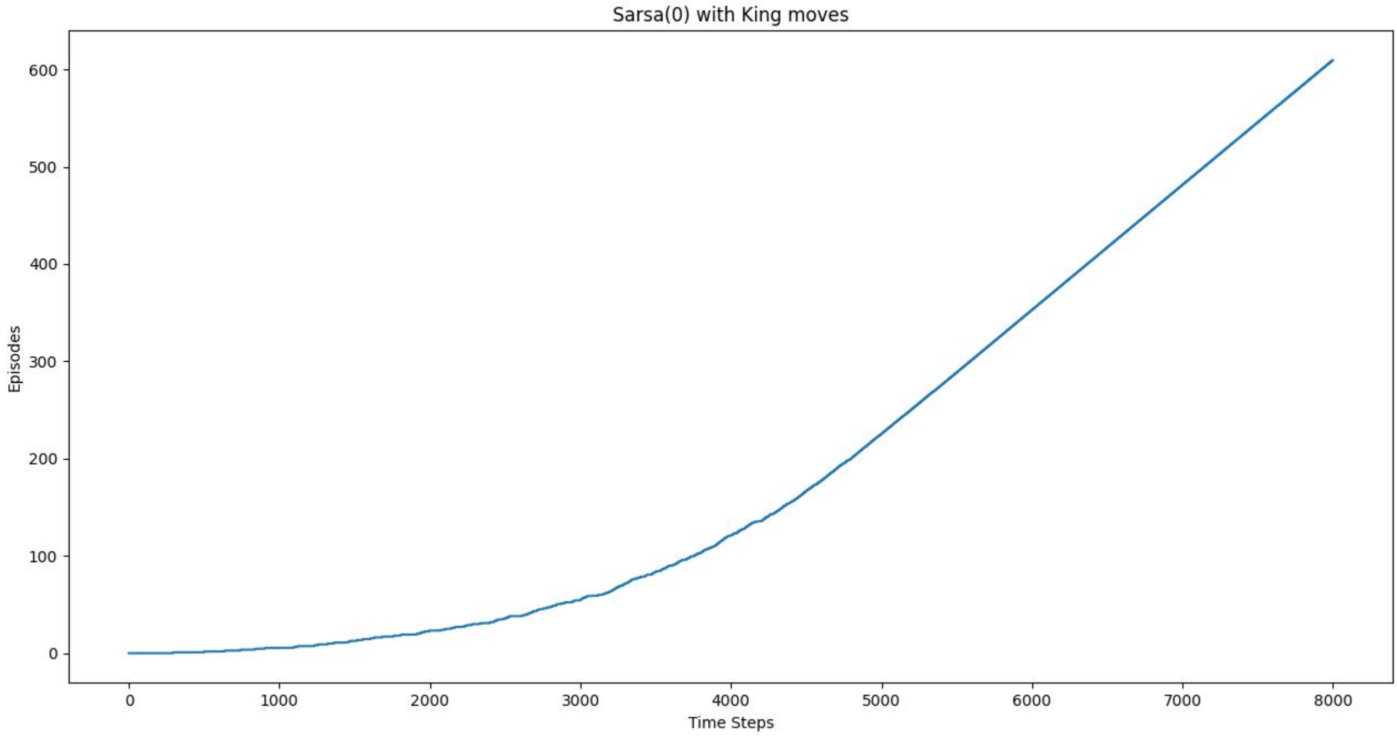
Agulla Surya Bharath **Roll.No : 17D070055**

Parameters used in solving experiment :
Gamma = 1 , alpha = 0.5 , epsilon = 0.1 .

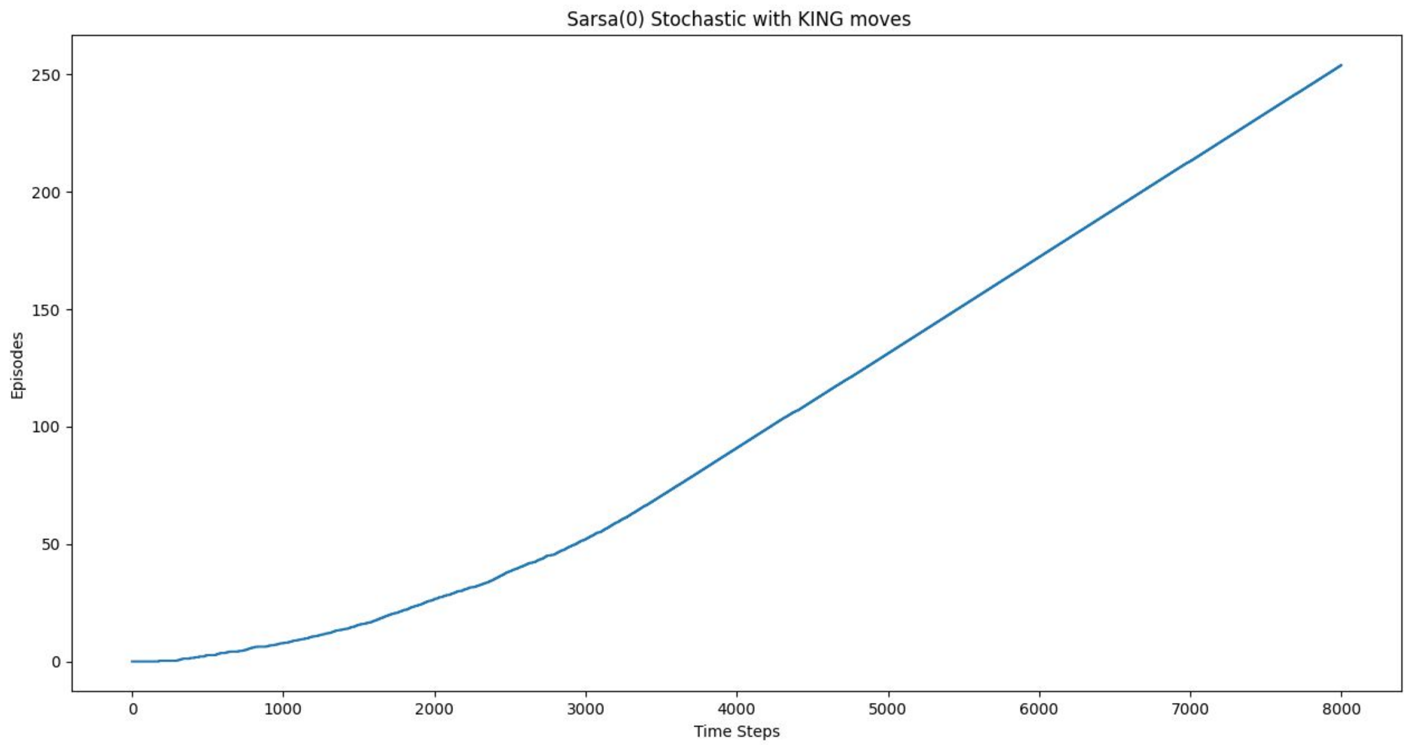
Task 2 :



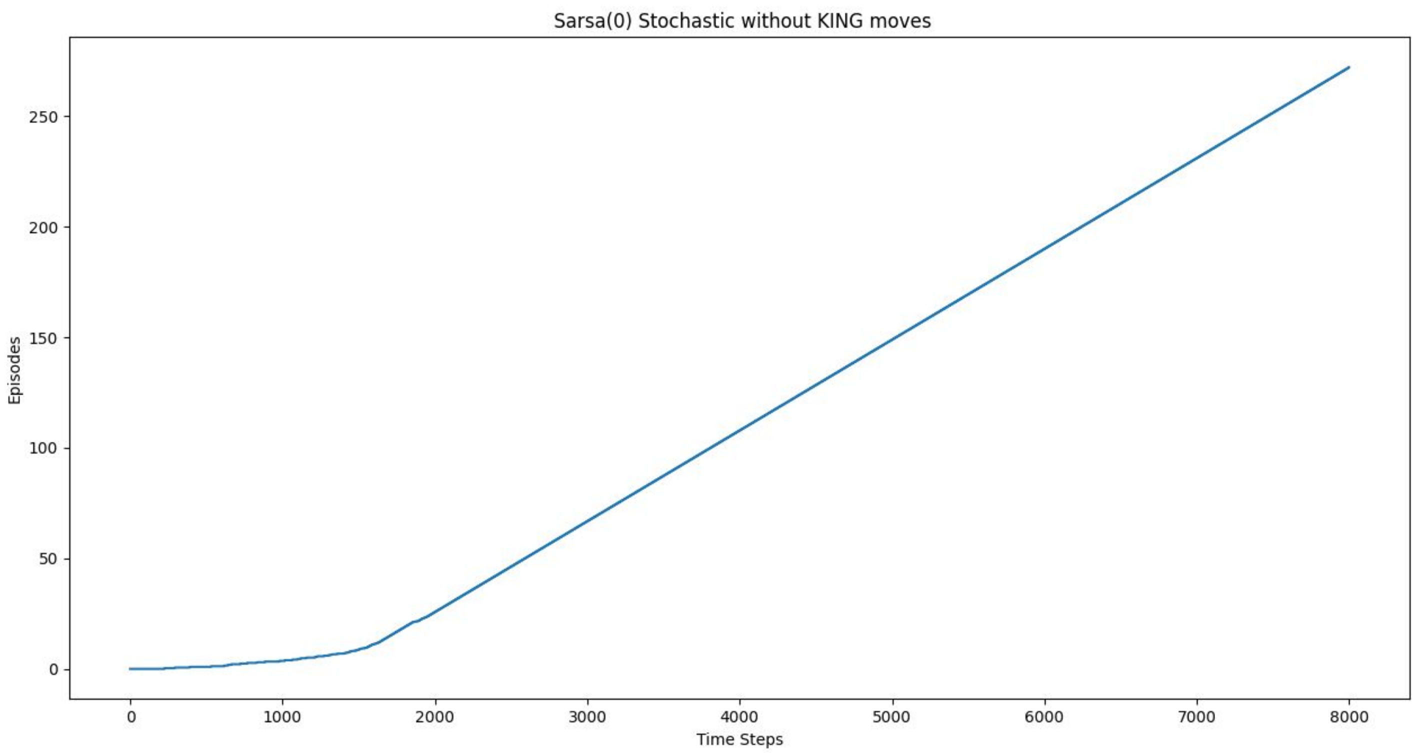
Task 3 :



Task 4 :

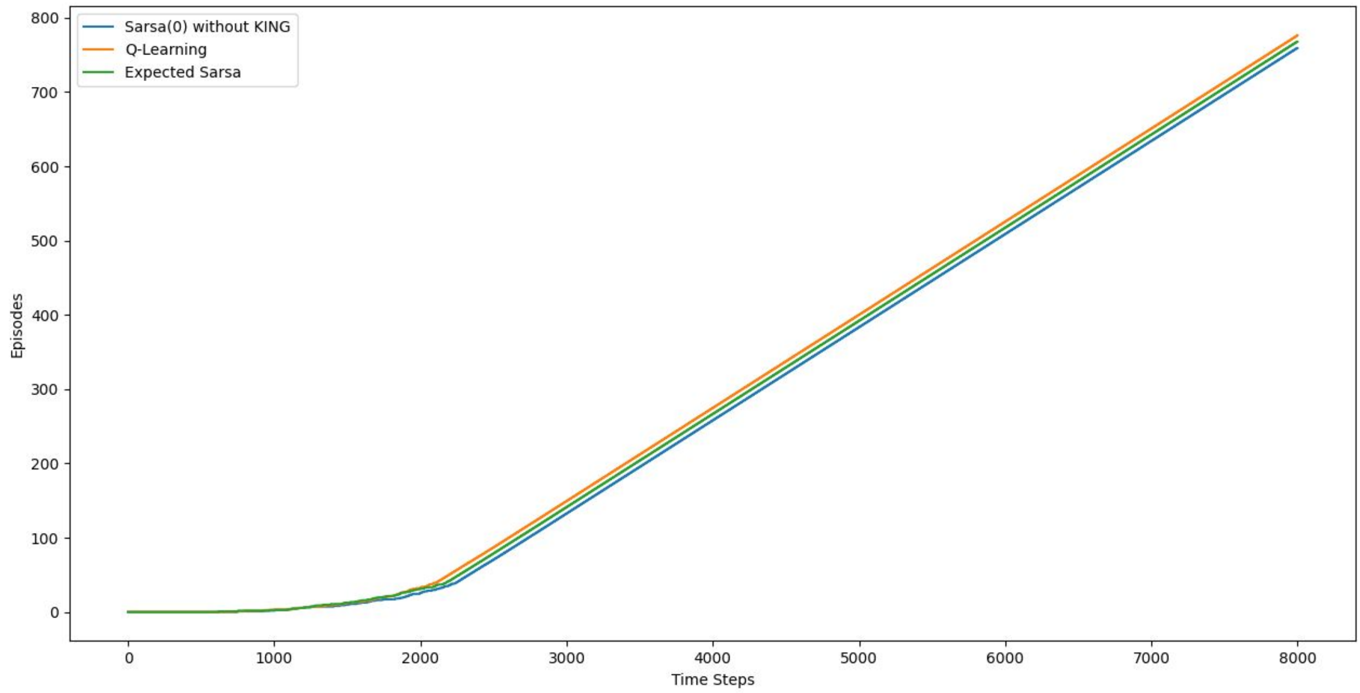


Sarsa(0) Stochastic with KING moves
Episodes
Time Steps
0
50
100
150
200
250
0
1000
2000
3000
4000
5000
6000
7000
8000



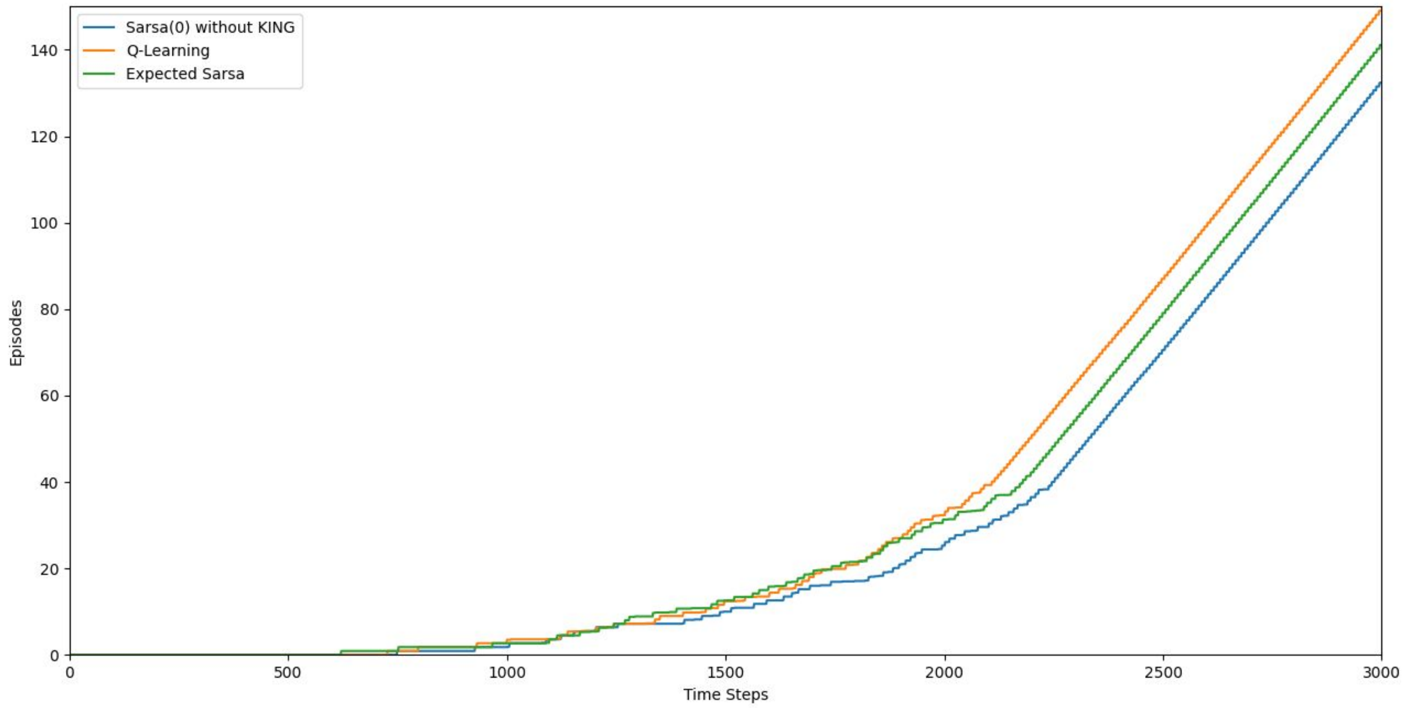
Sarsa(0) Stochastic without KING moves
Episodes
Time Steps
0
50
100
150
200
250
0
1000
2000
3000
4000
5000
6000
7000
8000

Task 5 :

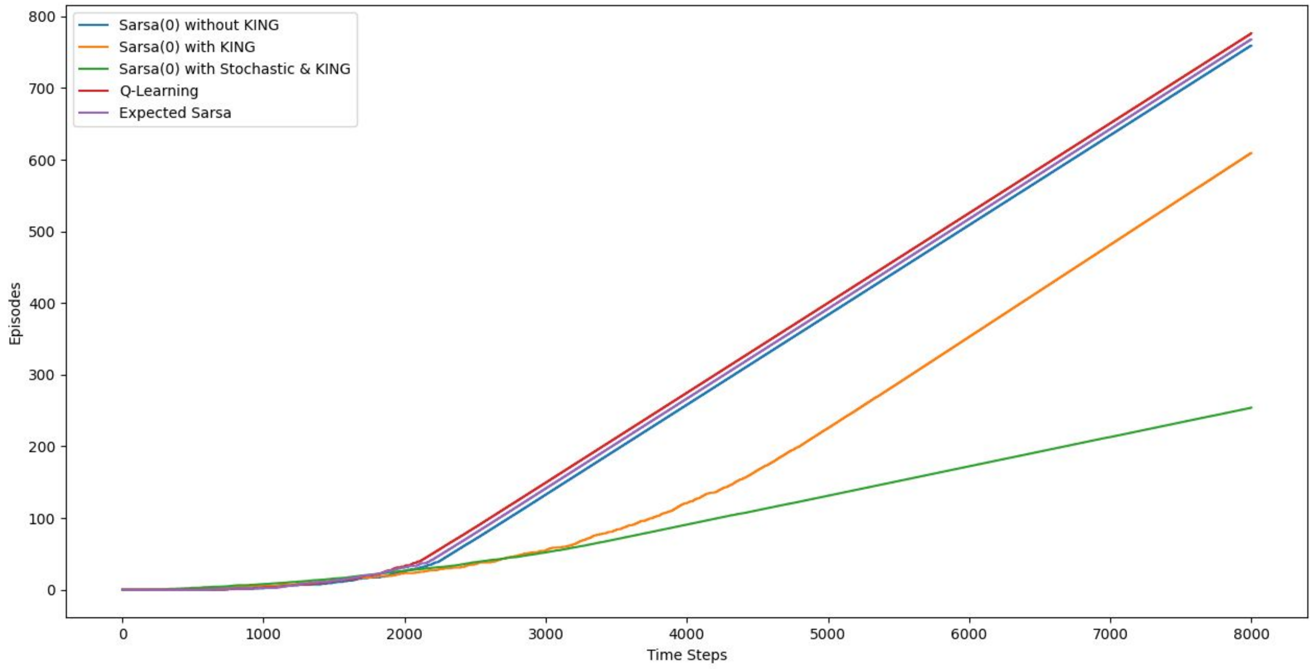


Performance :

Q-Learning > Expected Sarsa > Sarsa(0) without KING moves



Gamma = 1 , alpha = 0.5 , epsilon = 0.1 .



This plot is for 8000 timesteps , from this plot of different agents we can immediately interpret that as the number of timesteps increases the plot becomes more linear and which in turn means that the number of time steps required per episode is getting stabilized to constant . So the optimal path is found by our Learning agent as expected.

From this 8000 time steps analysis we can roughly say that average speed of convergence or average performance of agents is as follows:

Q-learning > Expected Sarsa > Sarsa (without king moves) > Sarsa (with king moves) > Sarsa (stochastic)

By performance I mean , the number of time steps required per episode converges to optimal value (for optimal path) in less number of total time steps from start of the experiment by the agent.

Interpretations from Task2 & Task3 :

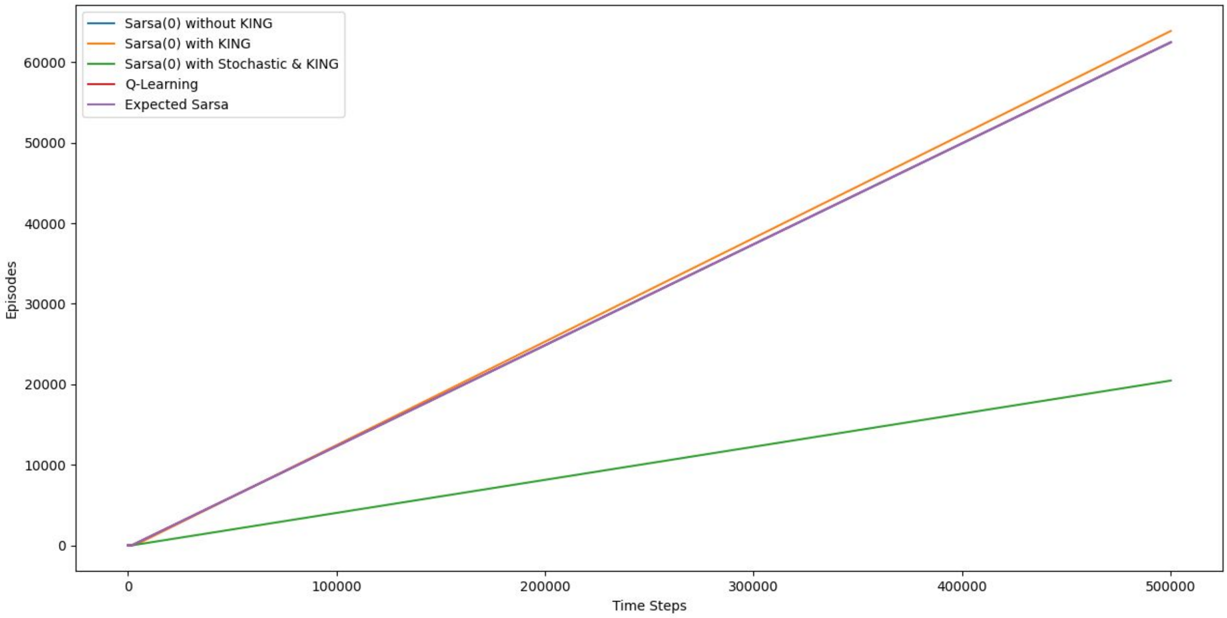
From the plots of Sarsa(0) without KING moves and Sarsa(0) with KING moves one thing we can see is that even though both of them get closer to the optimal episode length in the long run , adding KING moves makes the learning slower. This is expected because as we increase more possibilities of paths the agent experiences long episodes in the beginning of learning and it takes comparatively longer time to figure out optimal path .

Interpretations from Task3 & Task4 :

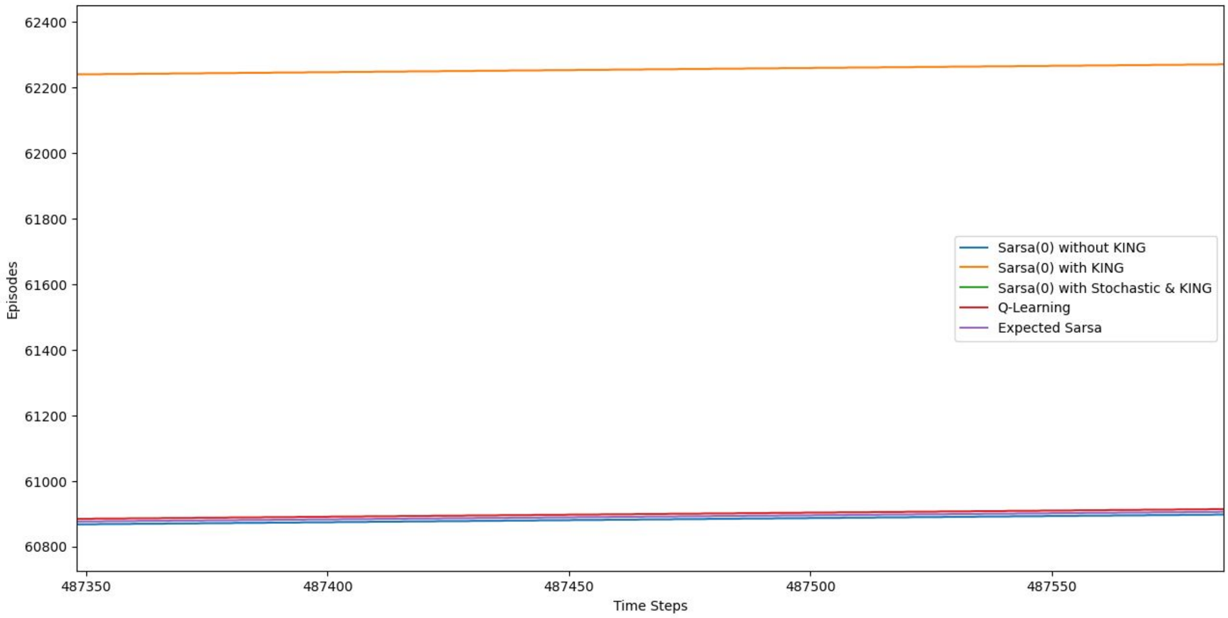
From the plots of Sarsa(0) with KING moves and Sarsa(0) with KING moves & stochasticity one thing we can clearly notice is that adding stochasticity in the moves has extremely affected the learning rate . Stochasticity reduces the learning rate due to more unwanted exploration in the moving steps . This is also expected .

Once again we can witness the effect of KING moves from the two plots in the Task4 section . This is inline with what we have expected from the interpretation of Task2 & Task3.

But wait, don't conclude that KING moves are useless. They are indeed useful if we are interested in getting a small episode length and not concerned much about the number of time steps it takes .. See below plots , they show that adding KING moves gives the largest slope (i.e small episode length) at larger time steps.



Zoomed version of above plot :



Interpretations from Task5:

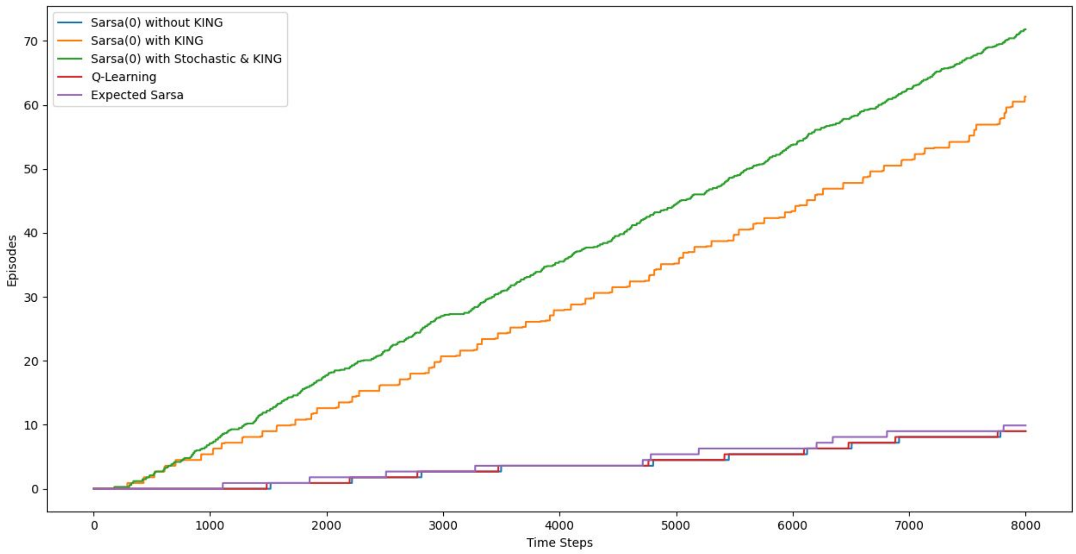
From the plots we can say that convergence of Q-learning is fastest among three and this is expected as we discussed in class Q-learning is an off-policy update and converges to the optimal Q matrix always .

Also between sarsa and Expected sarsa plots ,Expected sarsa performed well . This is also expected because as we discussed in class the expected sarsa works deterministically as sarasa in expectation . So variance incurred by choosing actions by epsilon greedy (which might not be optimal action) tends to decrease by using Expected sarsa as it gives some weight to all actions in update mapping and don't just merely take argmax.

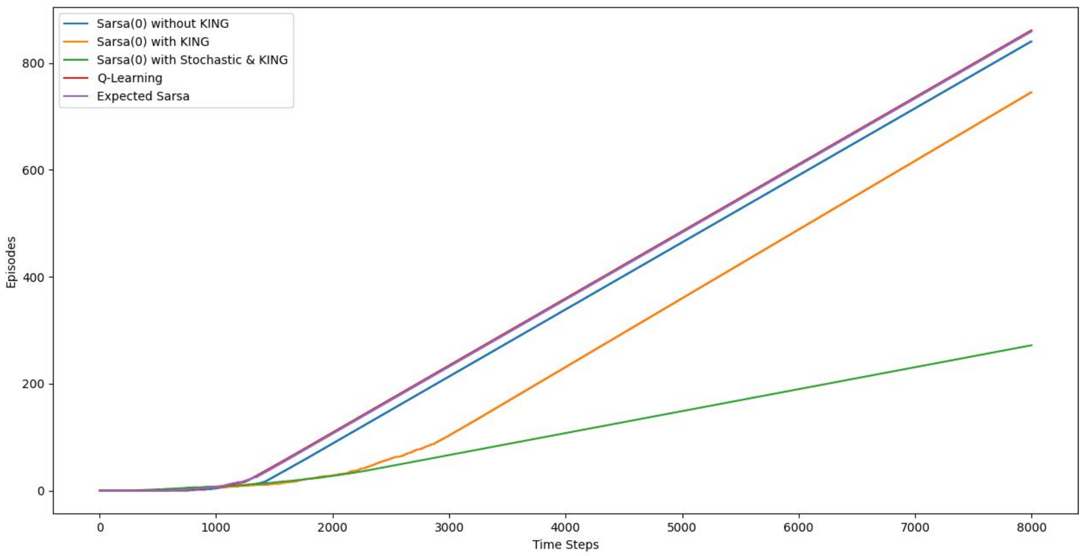
Effect of epsilon and alpha :

By decreasing alpha learning becomes slow and episode length becomes larger and takes long time steps to converge to optimal .

For Gamma = 1 , alpha = 0.001 , epsilon = 0.1

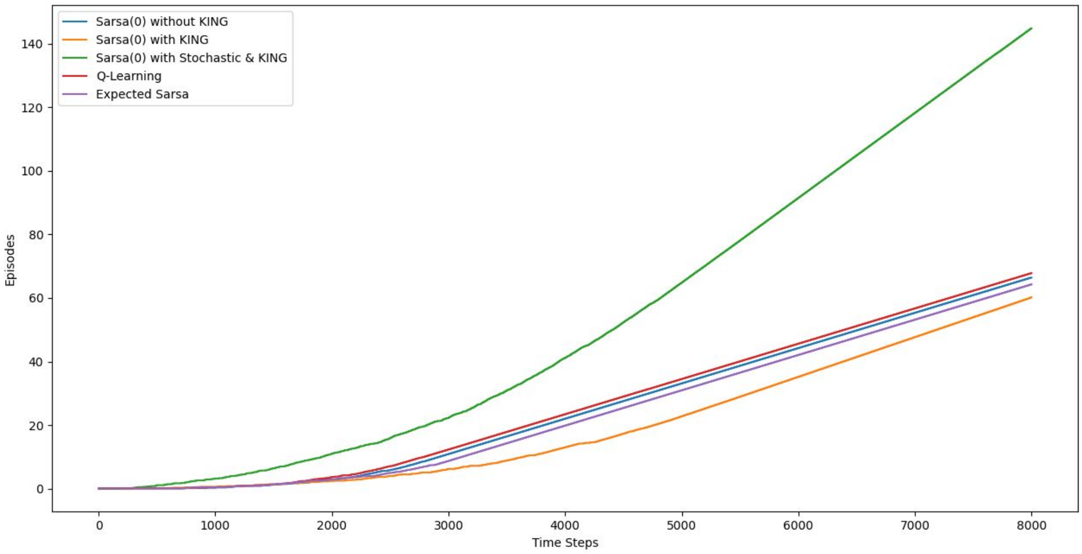


For Gamma = 1 , alpha = 0.9 , epsilon = 0.1



Increasing epsilon enables more exploration and which inturn allows more steps to explore before reaching an optimal path . This is also evident from the plot below.

For Gamma = 1 , alpha = 0.5 , epsilon = 0.9



NOTE : All these plots can be found in the "plots" folder in the submission folder.