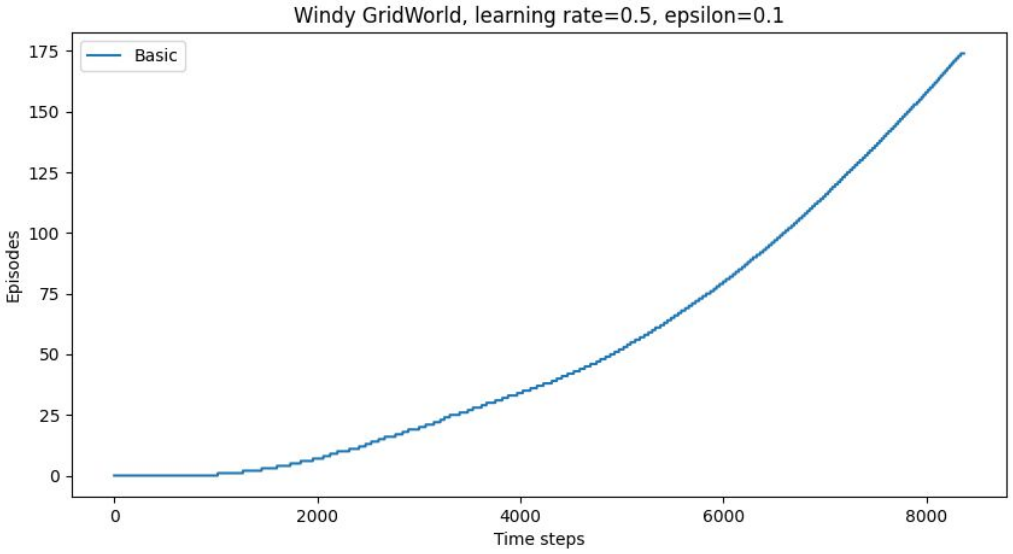
Report for Programming Assignment 3, CS747

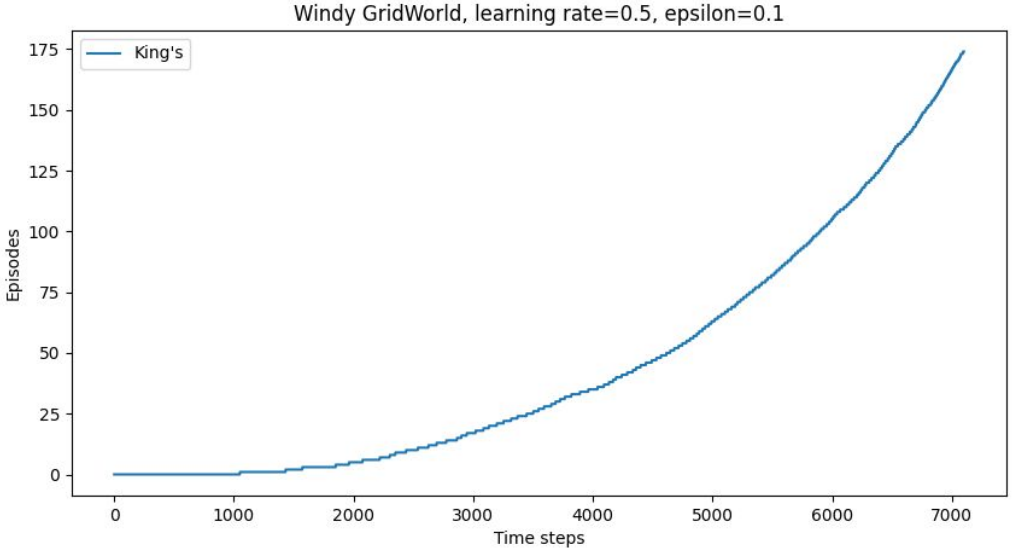
Task 2:

Sarsa(0) on the basic environment:

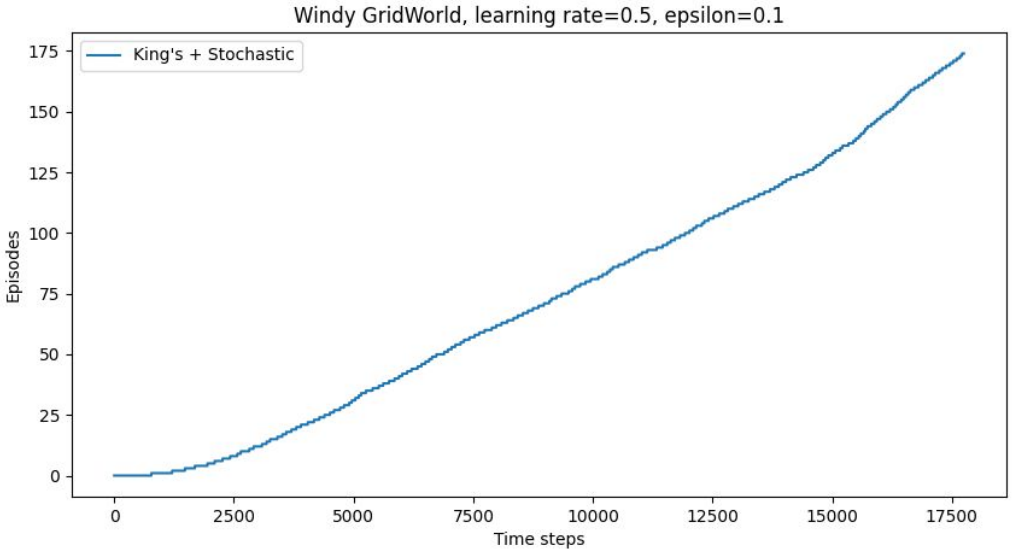


Task 3:

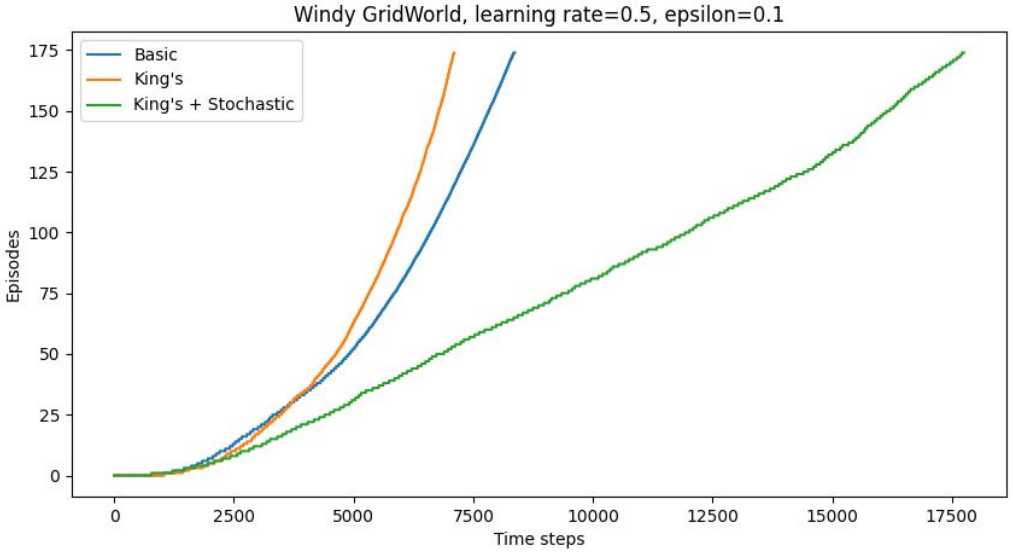
Sarsa(0) on the environment with permitted king's moves:



Task 4:
Sarsa(0) on the environment with permitted king's moves along with stochastic winds:



Comparison of environments on sarsa(0) algorithm



Observations:

1. In the comparison plot above it can be observed that the king's moves experiment has the least total timesteps and the maximum asymptotic slope compared to the basic plot.
2. Another observation is that the king's +stochastic experiment takes substantially more timesteps as compared to the base case with a much smaller slope

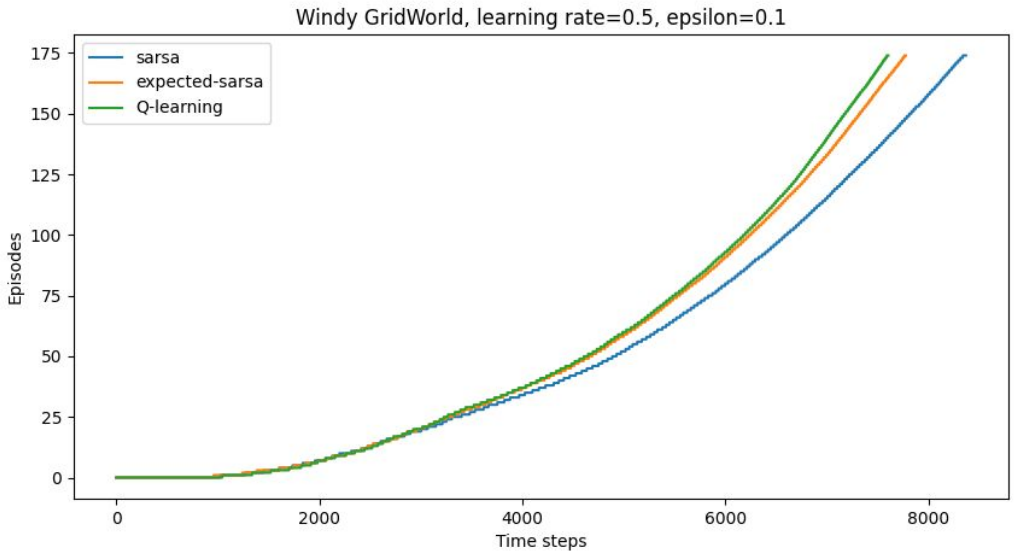
‘

Interpretations:

1. Observation 1 implies that the optimal steps are smaller in kings case than the base case which is what is expected
2. The observation of lesser slope in point 2 implies that adding stochasticity significantly increases the optimal steps

Task 5:

Comparison of training algorithms on basic environment



Observations:

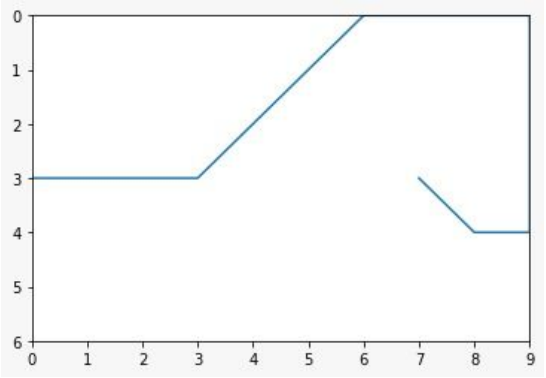
1. It can be observed that q learning which is an off-policy algorithm performs better than expected sarsa and sarsa which are on policy algorithms
2. Between expected sarsa and sarsa, the former performs better

Handling of corner cases:

After computing the next position taking into the action and effect wind, if one or more coordinates go out of the grid, they are set to the edge value.

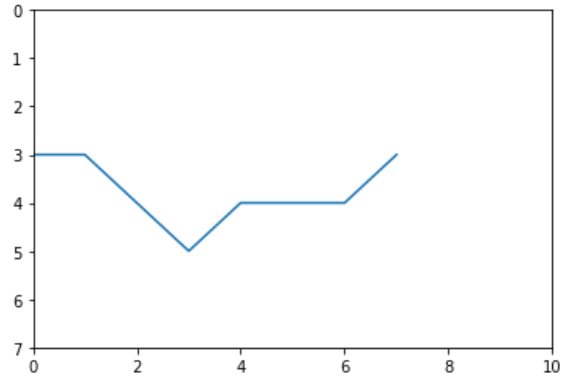
Optimal paths:

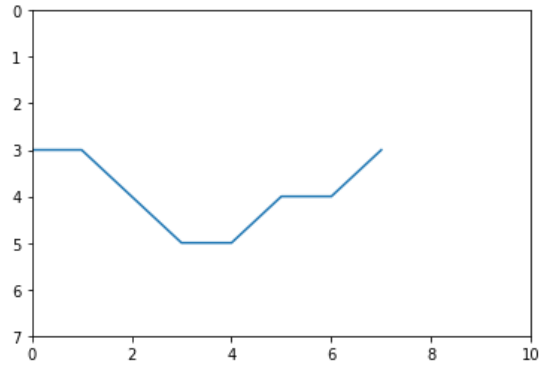
For the basic environment:



There is a unique optimal path with optimal length = 15

For Kings moves:





According to my observations, there is no unique optimal path and the length of the optimal path is 7.