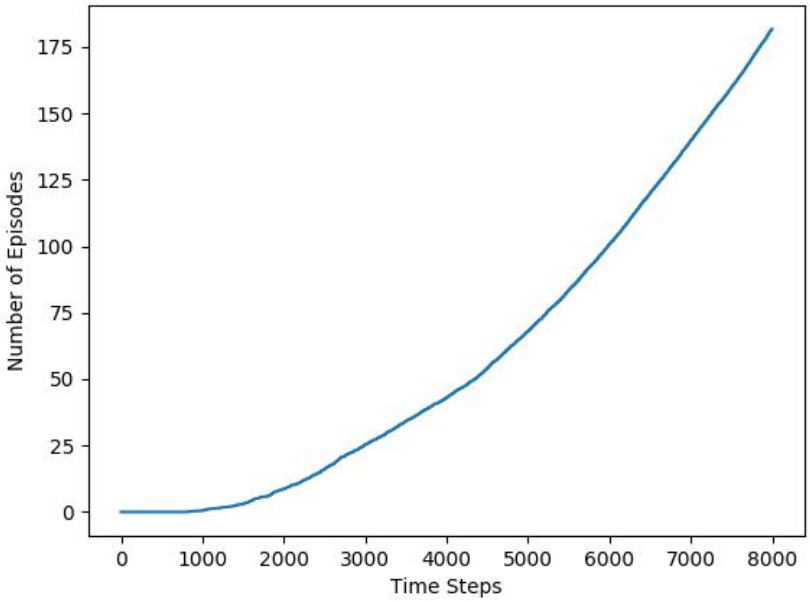
CS 747 - Foundations of Intelligent and Learning Agents

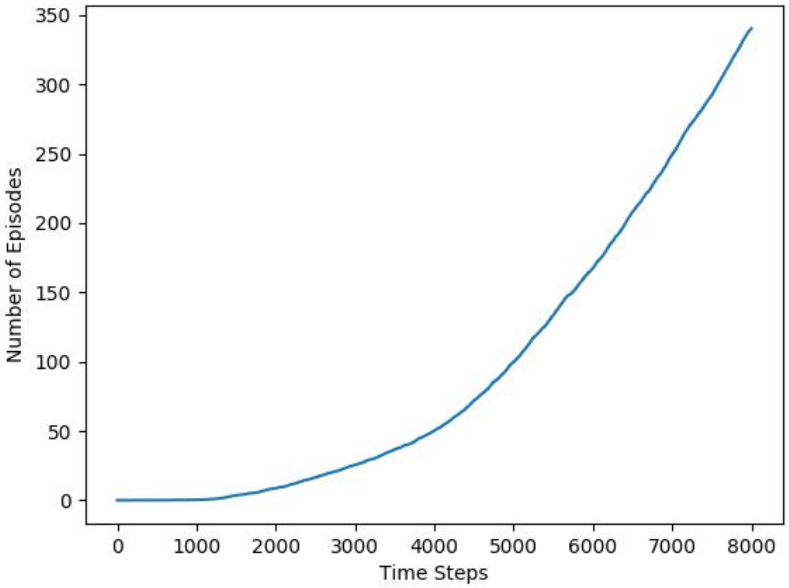
Assignment 4

Kumar Ashutosh, 16D070043

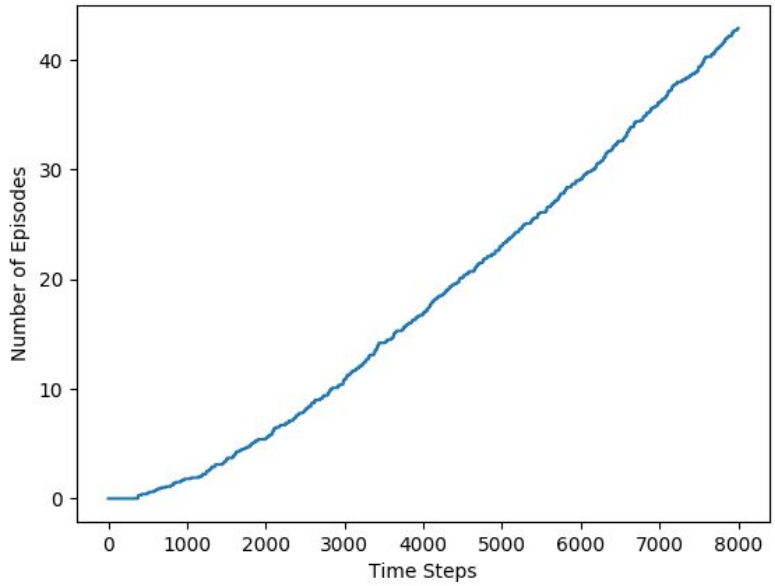
Part A



Part B



Part C



From the above graphs, we can conclude the following -

1. We see that as time step increases, the agent is able to reach the end of episode more easily. Hence this graph is evident of the fact the the agent is learning to move optimally. This is a result of the Q update that we are doing.
2. Increasing the number of allowed moves (in part B) resulted in more number of episodes, this is also expected since the number of allowed moves increases and it would be easier to reach the end state in less number of steps. But this comes at a cost of more learning time. Since the number of movements is more, it will result in more time to learn and longer episode run initially.
3. Adding stochasticity degrades the performance tremendously. As seen in part C, the number of episodes went to do around 40 from 170. And also the time to learn also increases.

Some Other observations in general-

1. Increasing the time step from 8000 to a higher number makes the graph more and more linear, since after a while, the agent has learnt enough and it takes a constant number of steps for each iteration.
2. Decreasing alpha to a lesser value decreases the number of iterations, for a value of alpha=1e-5, the number of episodes decreased to less than 10 in part C.
3. Increasing epsilon means more exploration, which means the number of episode decreases significantly. For part B, and for epsilon 0.9, the number of episodes decreases to less than 10, while for epsilon 0.1 it is approximately 350