Programming Assignment 4

Mohit Madan
15D070028

All the Plots have been generated after averaging the agents time steps over 10 instances. number of episodes for each run is kept at 170.
Epsilon = 0.01
Alpha = 0.5 constant throughout the episode.
Value of epsilon is chosen as 0.01 after running for different values it gives best results for this value among the chosen values.
Alpha = 0.5 implies equal weightage to current and previous value.

1. Windy Gridworld with 4 moves (up, down, right, left)

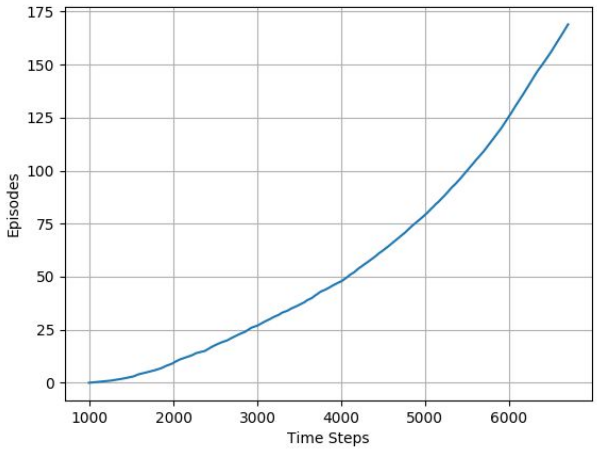

Fig 1. Windy GridWorld

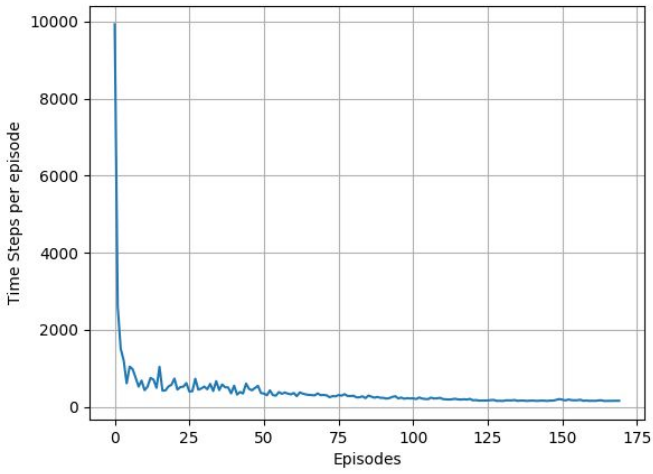

Fig 2. Time Steps per episode

2. **Kings Moves** - Windy Gridworld with 8 moves(up, down, right, left, top-right, top-left, bottom-right and bottom-left). This achieves the results in the least number of time steps. Here we can see in figure 4 that results converge very quickly.

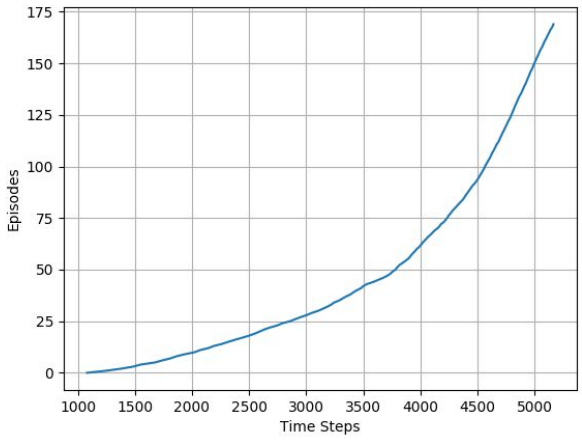

Fig 3. With Kings Moves

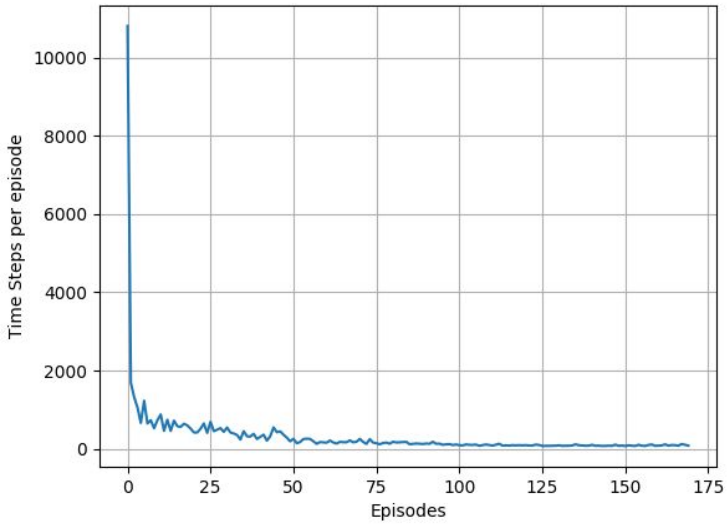

Fig 4. Time steps per episode

3. **Stochastic** - Windy Gridworld with 4 moves and stochastic wind (+1,0,-1) . From figure 6 we can observe that the number of steps doesn't converge even after 170 episodes due to the stochastic nature of the winds.

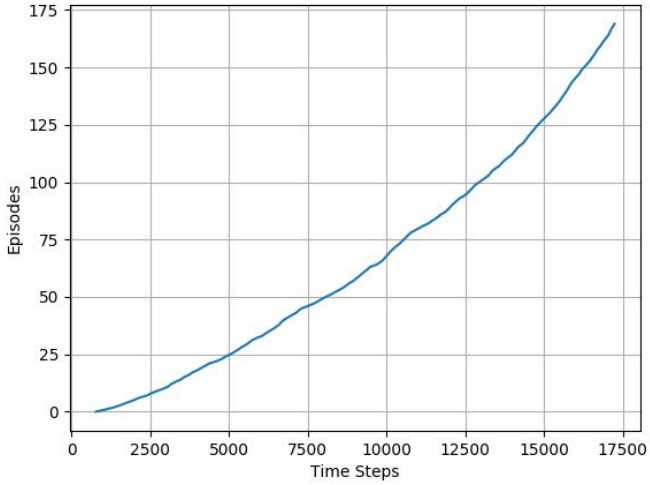

Fig 5. Stochastic

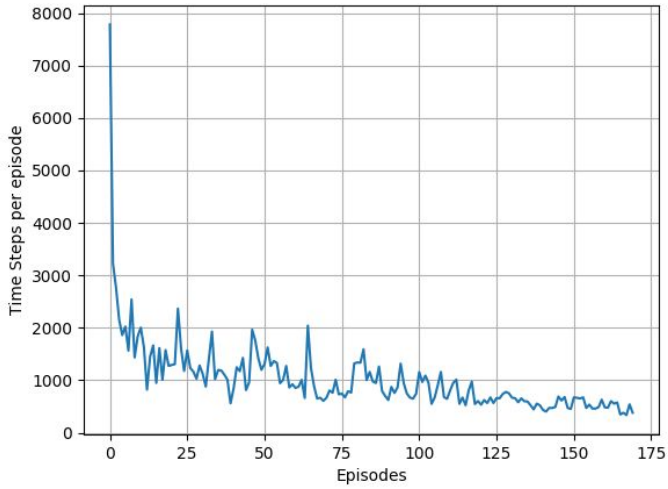

Fig 6. Time steps per episode

4. **Stochastic with Kings Moves :** Here we can see that kings is always better than the normal moves since it has the capability to move diagonally. Here with kings moves also, it does not converge but the results are still better than the one without kings moves.

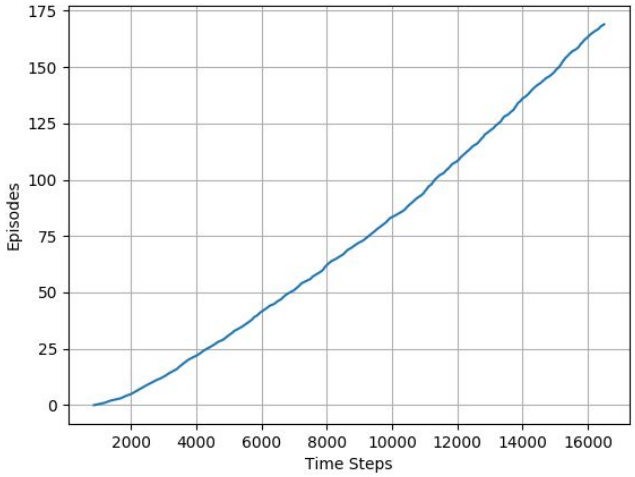

Fig 7. Stochastic With Kings Moves

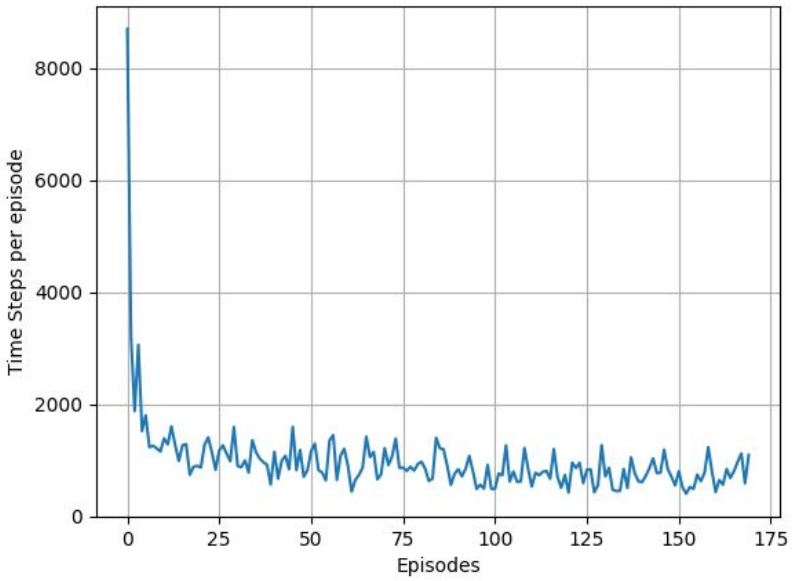

Fig 8. Time steps per episode