

# CS747-pa4

Avinash Reddy,183070010

November 2019

Every task is implemented to run for 170 episodes.  
 $\alpha=0.5$   $\epsilon=0.1$  Implemented algorithm = *sarsa*(0) with  $\epsilon$  - *greedy*

1. Normal moves (up,down,left,right). After training the agent took a minimum of 15 steps to reach the terminal state. With above actions, the agent took almost 8000 steps to complete 170 episodes.

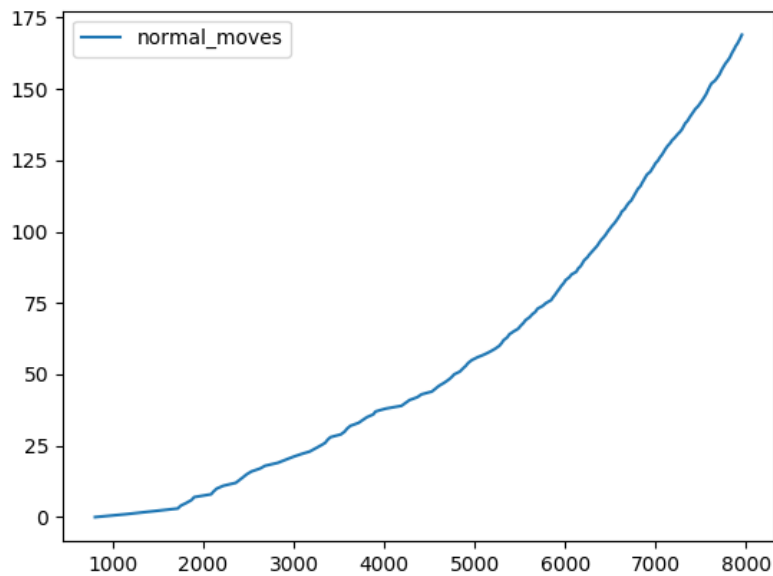


Figure 1: Normal Moves without stochastic

2. King moves (up,down,left,right,top-left,top-right,down-left,down-right). After training the agent took a minimum of 7 steps to reach the terminal state. With above actions, the agent took almost 7000 steps to complete 170 episodes.

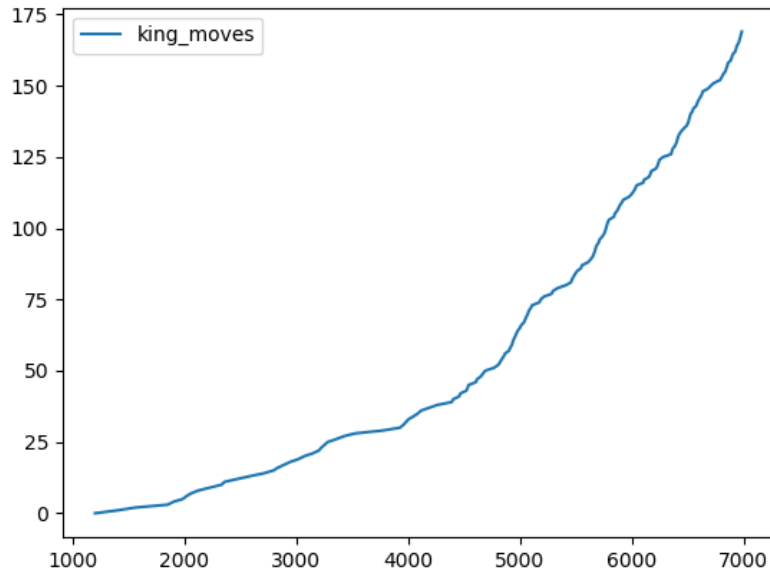


Figure 2: King Moves without stochastic

3. King moves (up,down,left,right,top-left,top-right,down-left,down-right,no-move or stable).With above actions, the agent took almost 8000 steps to complete 170 episodes.The reason of larger no. steps is the agent tries to explore the no-move action also. But eventually it has learnt, no-move is not optimal. Finally the agent took a minimum of 7 steps to reach the terminal state.

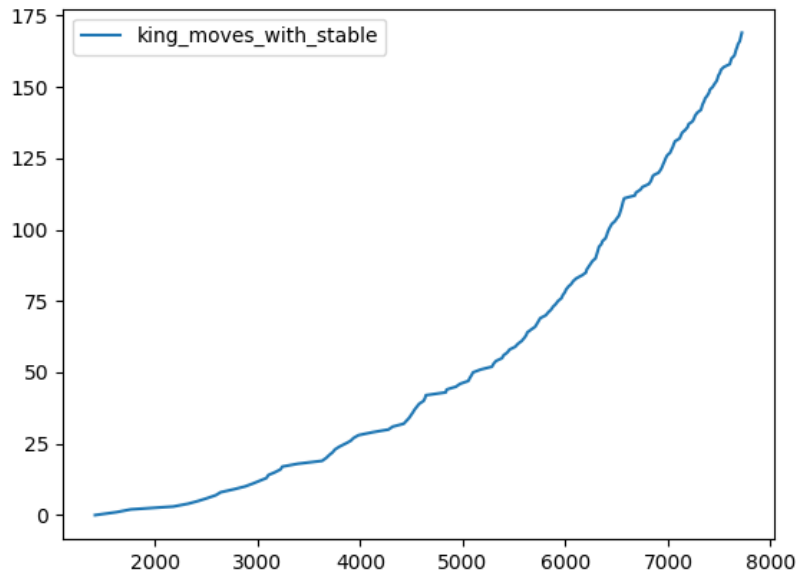


Figure 3: King Moves with no-move action and without stochastic

4. King moves (up,down,left,right,top-left,top-right,down-left,down-right). The agent took more than 20000 steps to complete 170 episodes. This lag is caused by the stochasticity.

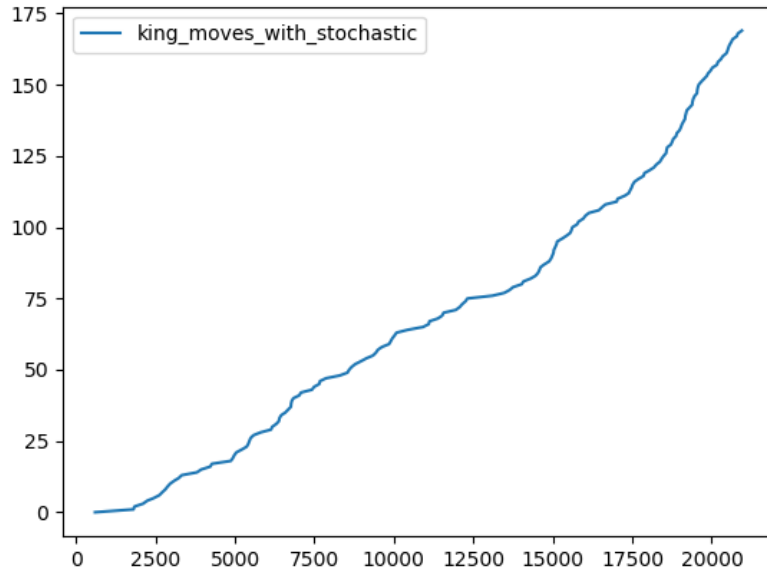


Figure 4: King Moves with stochastic