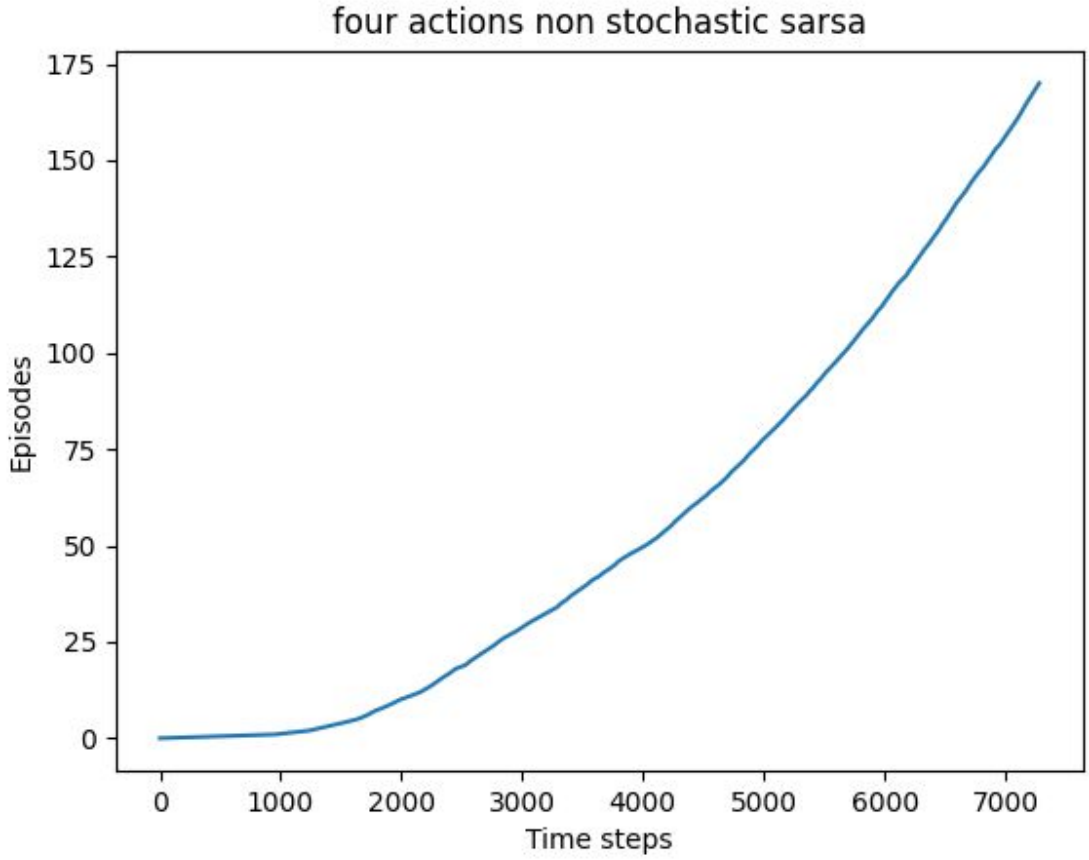
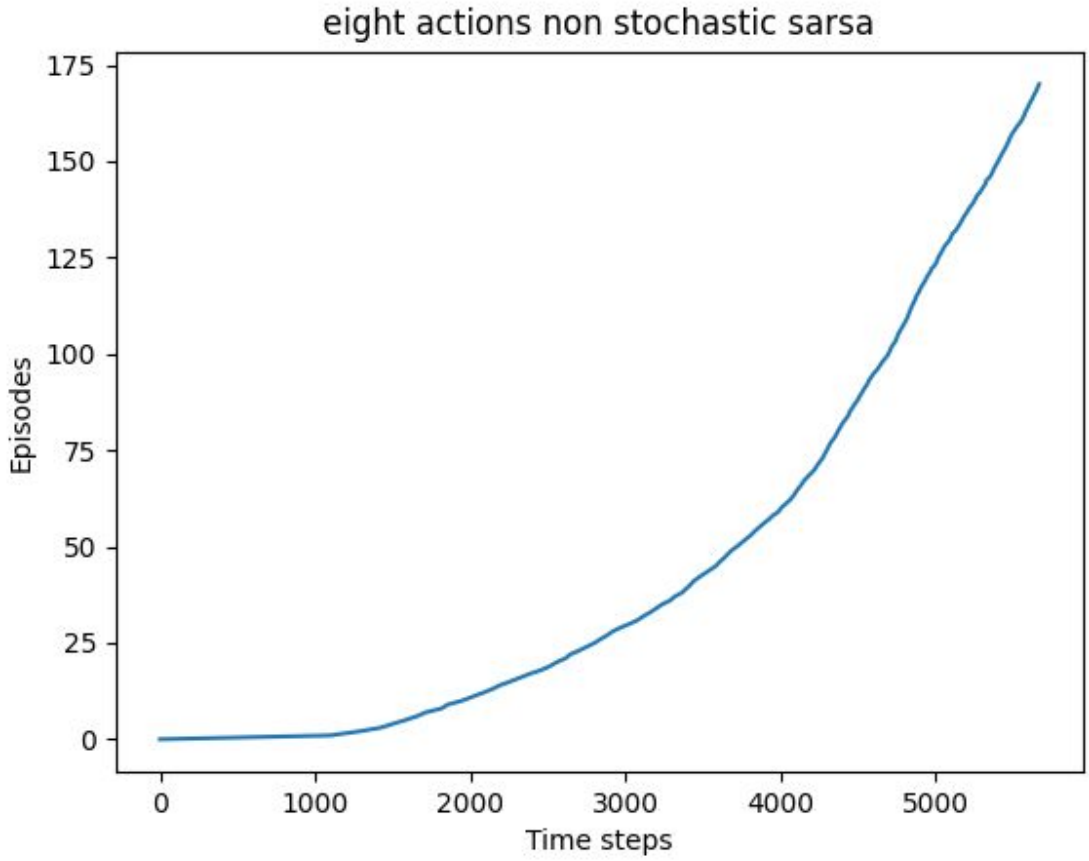
Assignment-3 170050067

TASK-1



Observation:
From result of timesteps of last 5 episodes are
7125. 7201. 7223. 7243. 7262. 7282.
avg steps per episode
((7201-7125)+(7223-7201)+(7243-7223)+(7262-7243)+(7282-7262)/5)=21
This is much less compared to initial avg time steps of the episodes
This is because initially agent doest know the optimal policy and its takes more time for each episode.After some time episodes the policy gets better and steps for next episode would become constant in text book mentioned at 17 steps for episode.

Task-2

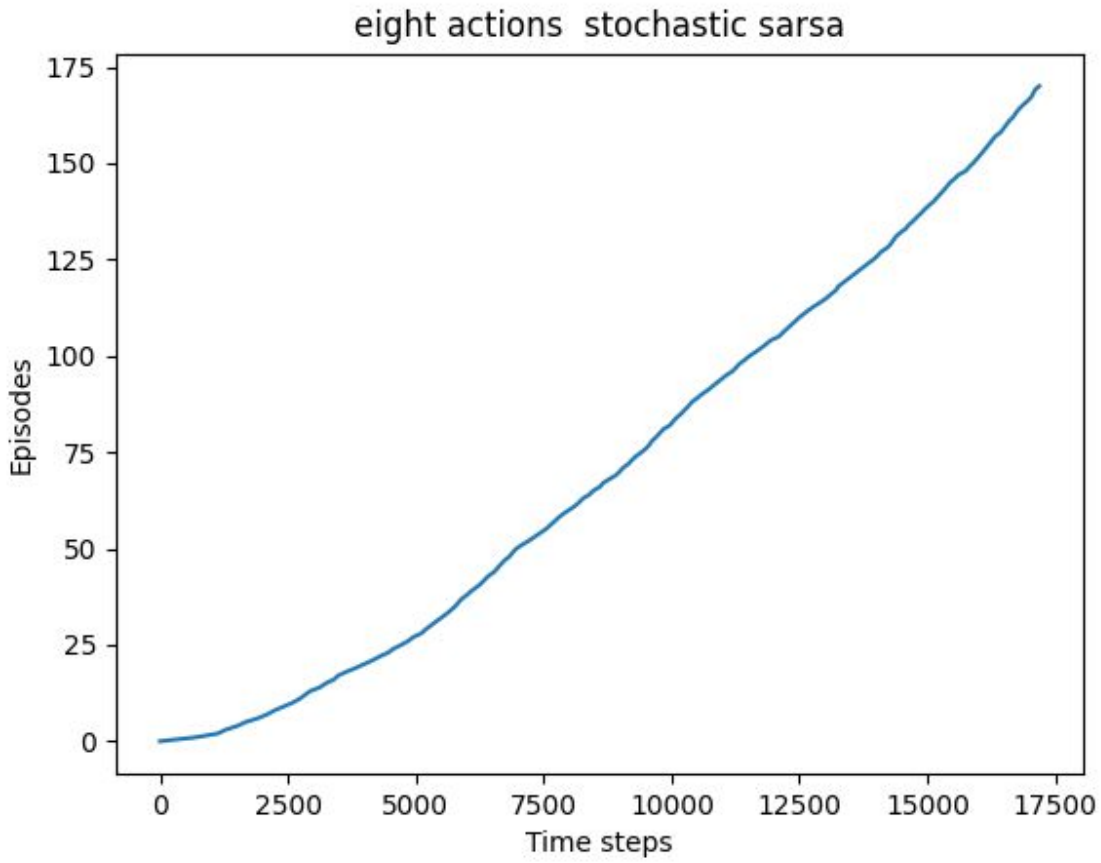


Observation:
Similar to above observation in terms of average timesteps for each episodes decrease and becomes constant ,because of policy improvements.

Due to increase of some action it reaches some states with fewer steps compared to part-1.Hence reach end state sooner compare to 4 actions of part-1
Calculating the avg steps per episode (5607. 5622. 5633. 5647. 5658. 5669.)=12.4

This decreases further and becomes constant.

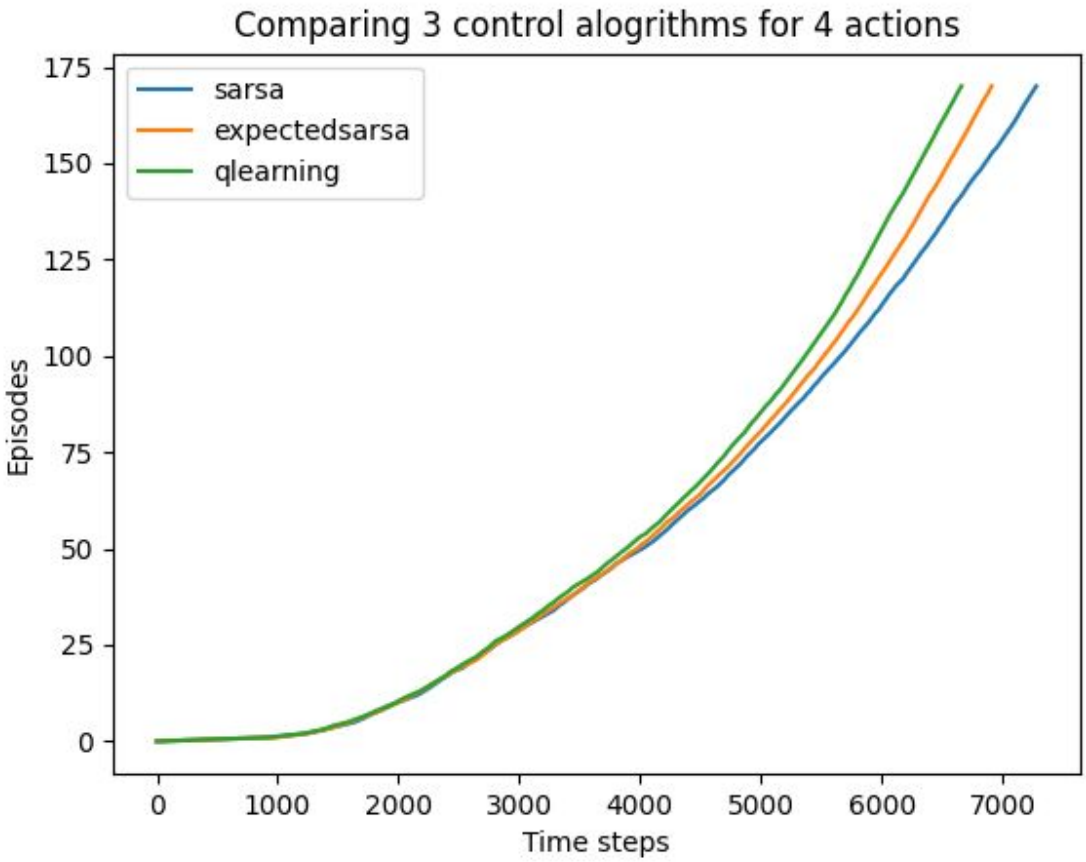
Task-3



Observation:
Similar to above observation in terms of average timesteps for each episodes decrease and becomes constant ,because of policy improvements.

Due to stochastic nature of the wind,it has to estimate various different policies and takes more steps per each episode to reach the final state.
Avg for last 3 blocks is almost 80 per episode.
[16863. 16947. 17024. 17067. 17101. 17187]

Task-4

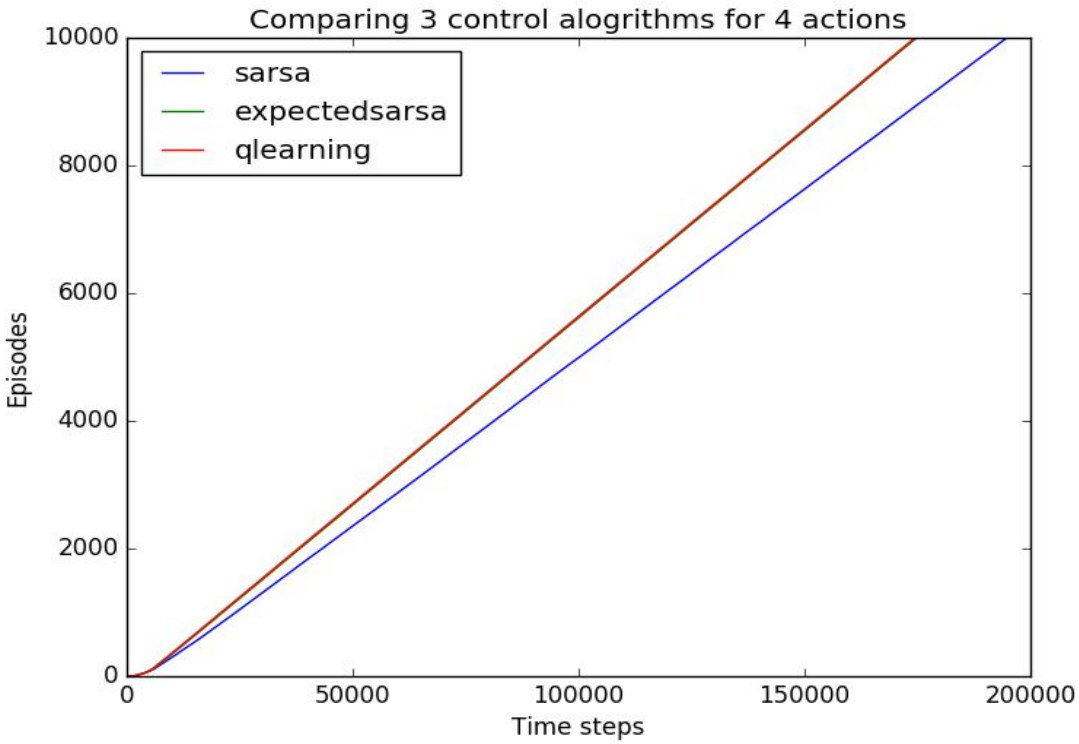


Observation

Steps to reach final state for an episode is less for Q-learning.

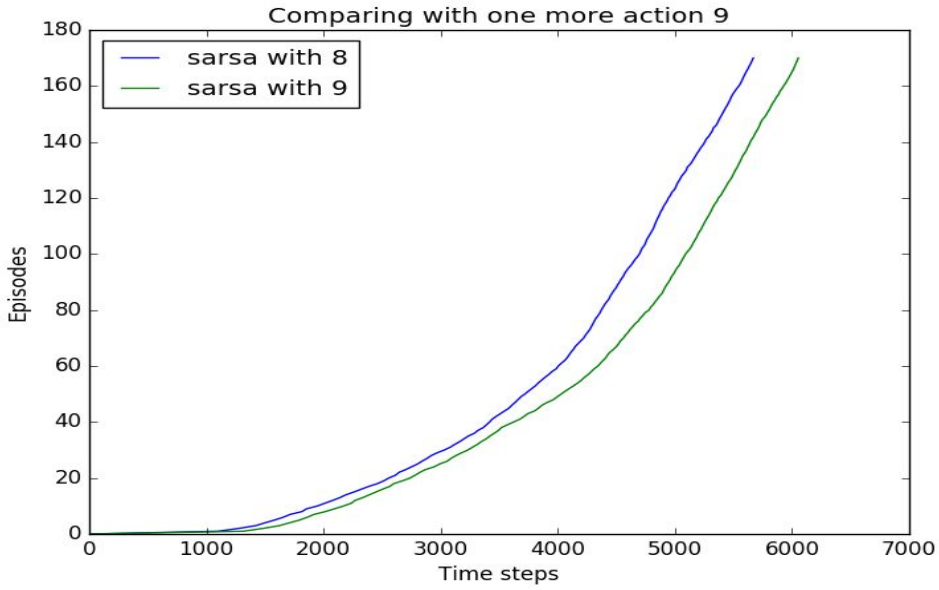
This is because this would be minimum when it attains Q* as t reaches infinity,Here only Q learning is only off policy that could converge to Q*.Remain to are on-line where update depends on the action performed.Hence Q learning would reach faster and converge.

Some other graphs

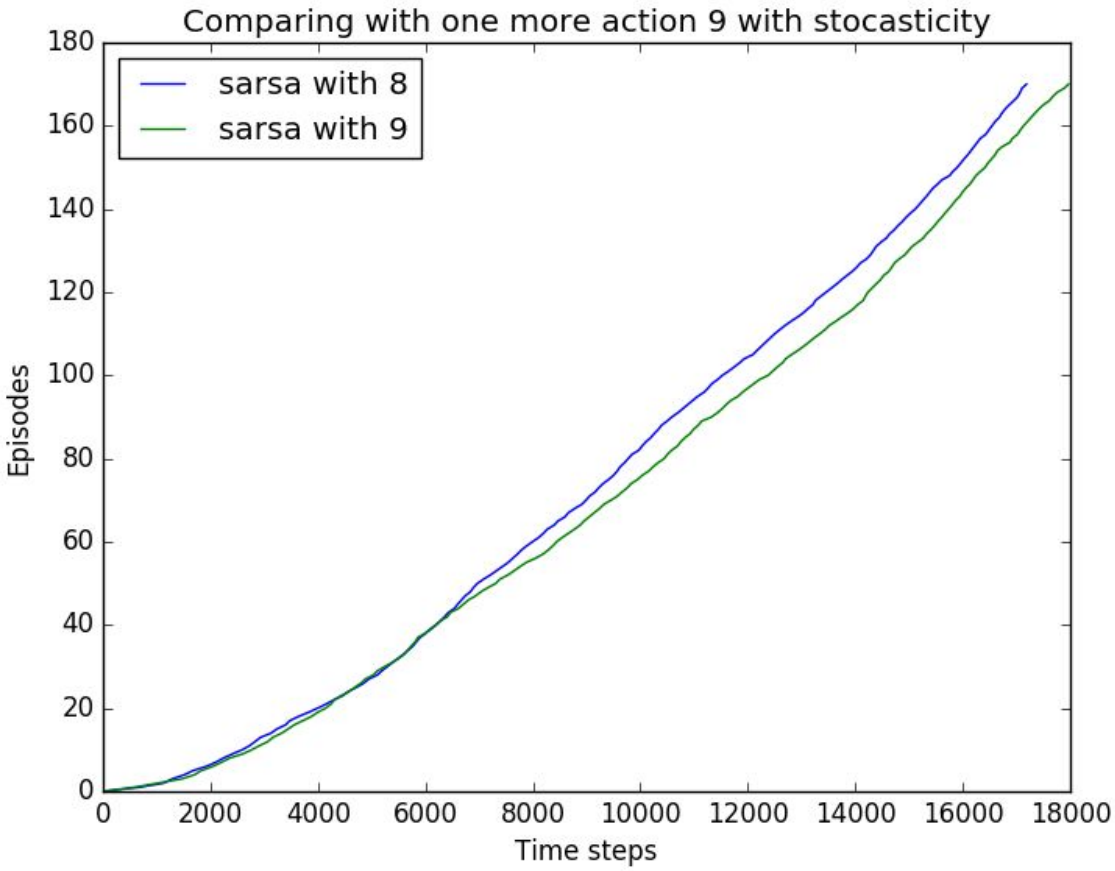


There graphs are straight line because after evalution of many episodes,the avg time for episode is constant.

Including one more action 9,no movement.



The above is 9 actions without stochasticity



This shows there is action 9 doesnt have any improvement compare to 8.