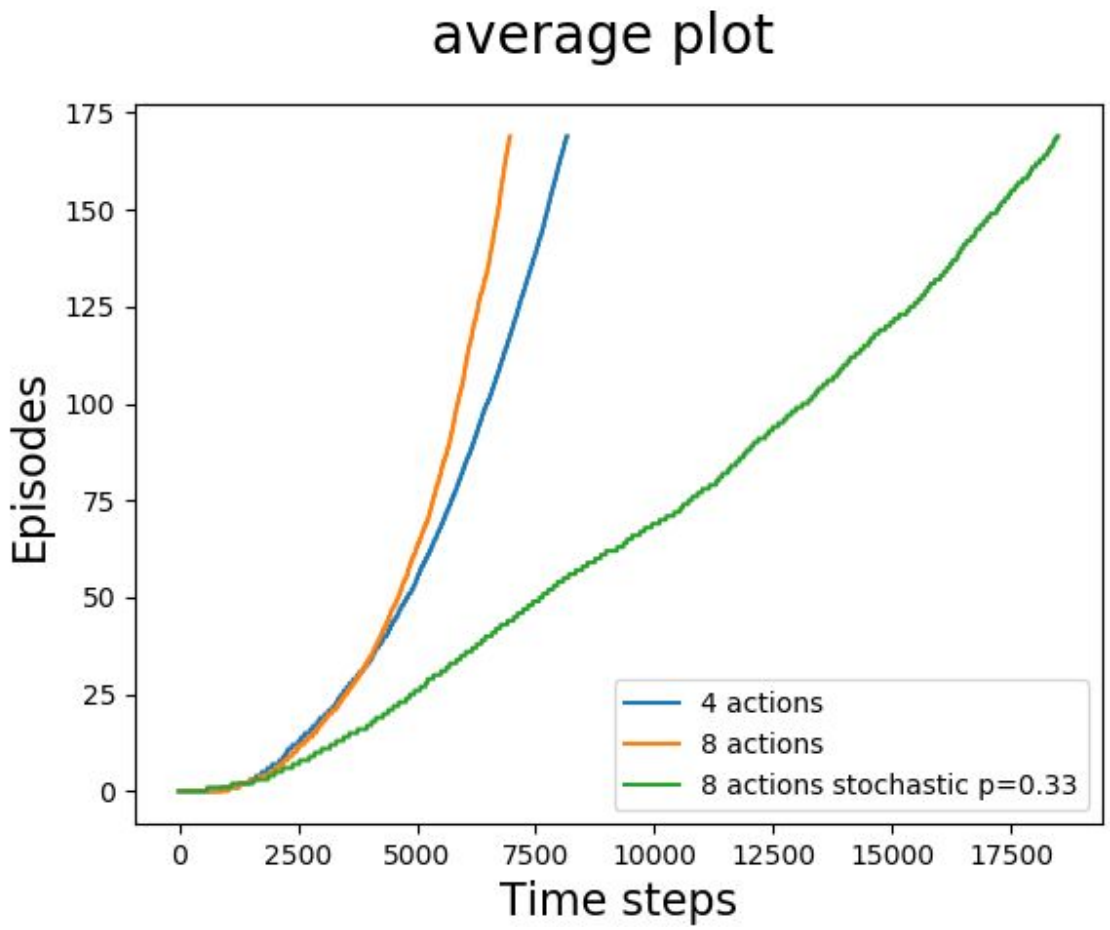
CS 747

ASSIGNMENT 4



NOTE: 4 actions - basic up,down,left,right movements, 8 actions- including king moves,

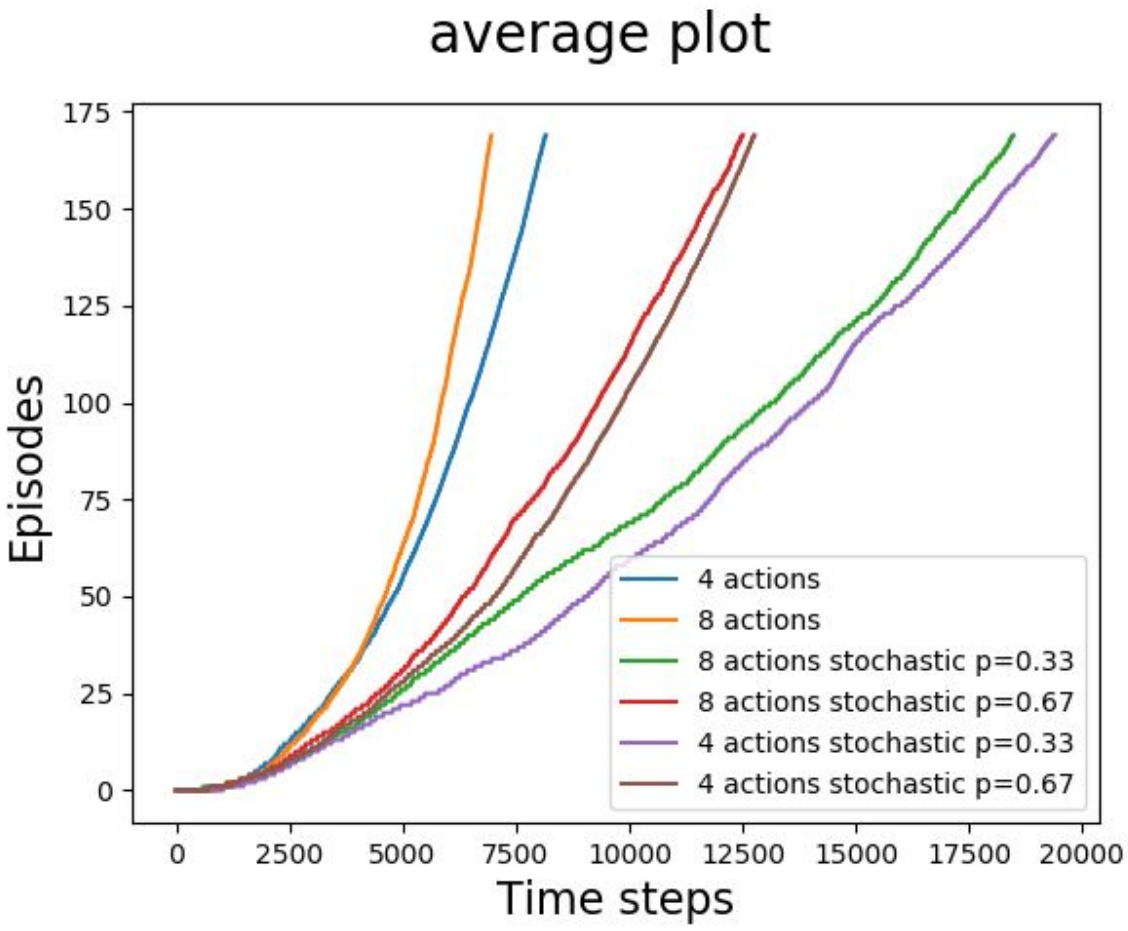
- For a fixed number of episodes(170 here), 4 actions takes more time steps than 8 actions. This is expected because, to make a diagonal movement in it takes 2 steps (without wind) in 4 actions, and it takes 1 step in 8 actions. Consider there is a wind of strength 1 upwards, to move right from current position, it takes 2 steps in 4 actions and 1 step in 8 actions.
- The slope increases over time in all cases, which shows that goal is reached more quickly. This is expected because the policy is improved over time.

- The stochastic case takes more number of steps to reach the goal because of the randomness.

NOTE: conventions followed:

I. Let w be given wind strength at a column. In the stochastic case, wind strength is w with probability p; w+1 with probability (1-p)/2, w-1 with probability (1-p)/2. In the above plot, with p=1/3, w,w-1,w+1 are all equiprobable

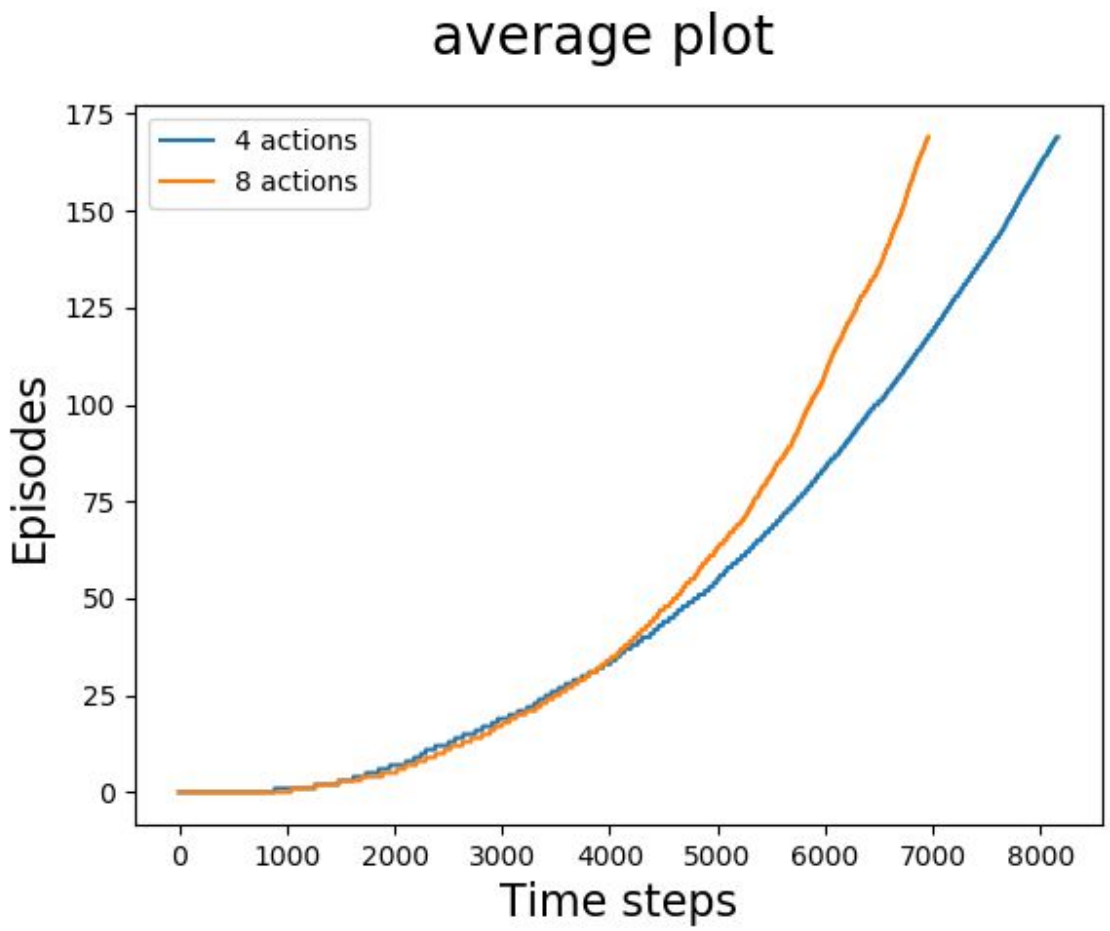
II. Movement across boundary is not allowed, let the current state be (x,y) and a diagonal move is taken, the next state should be (x+1,y+1), but if x+1 is out of bounds, then the next position will be (x,y+1)



- For p=0.33 , 8 actions took less time than than 4 actions. For p=0.67, these 2 are very close.

Average data:

Experiment	Number of steps
4 actions deterministic	8165
8 actions deterministic	6962
8 actions p=1/3	18487
8 actions p=2/3	12516
4 actions p=1/3	19404
4 actions p=2/3	12771
8 actions p=1/2	16072
4 actions p=1/2	16768



- In the initial stages, 4 actions is quicker, but later on 8 actions is quicker to reach the goal. This is explainable because, in 8 actions the number of actions is more, so it takes a bit more time initially to explore them and learn.