CS747:Programming Assignment 3 Report

Debabrata Biswal
Roll no : 203050024

November 13, 2020

1 Introduction

The report contains some of the implementation details of the algorithms that I have coded for this assignment including the assumptions,some observations.

2 Formulating The MDP (Task 1)

I formulated the the windy grid to an episodic MDP as mentioned in the RL book by Sutton and Barto(same reward , same starting and ending state etc.).So I am not reiterating them.I have kept alpha = 0.5,epsilon = 0.1 and the number of episodes = 200.

3 Task 2

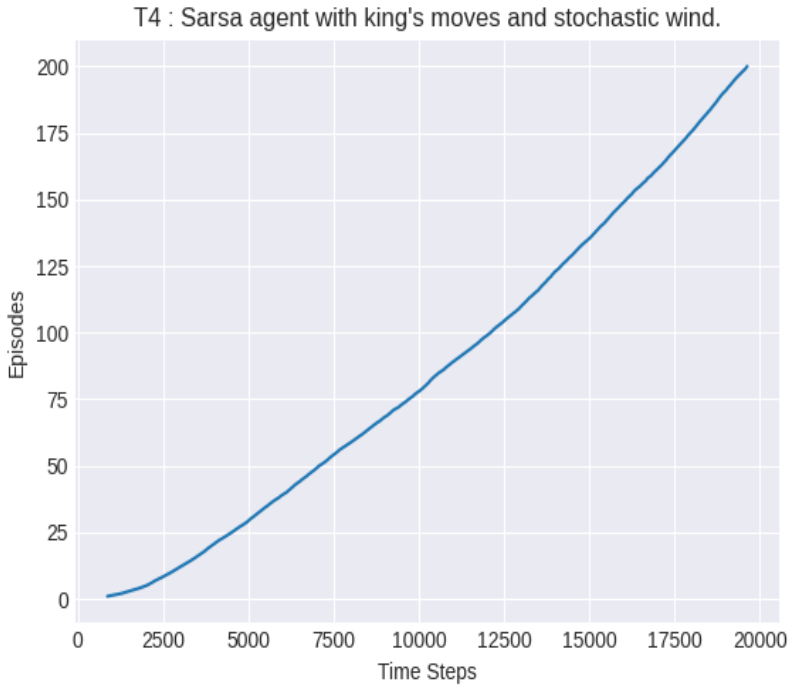
I implemented the sarsa agent and plotted the graph using the averaged data over 50 random seeds.Alpha is set to 0.5 and epsilon is set to 0.1 .

4 Task 3

With king's move it took less timesteps for same number of episodes.And that is expected because of increased freedom of movement inside the grid.

5 Task 4

After adding stochastic wind in the grid,even with king's move Sarsa took much longer timesteps.This might be because of difficulty in choosing optimal policy in an random environment.



6 Task 5

In this task I wrote another two agents that use Q learning and Expected sarsa and compared all the three algorithms with basic moves and no added stochasticity . And Q learning came out on top followed closely by Expected sarsa which was followed by sarsa.I noticed that for small values of epsilon 0.01 the performance of q learning and expected sarsa was almost same. you can generate this plot by doing

```
./plot_for.sh 0.01 0.5 200 0 0 sarsa q_learning e_sarsa
```


T5 : Comparison of Q learning, sarsa, expected sarsa agents with basic moves
Episodes
Time Steps
0
25
50
75
100
125
150
175
200
1000
2000
3000
4000
5000
6000
7000
8000
9000