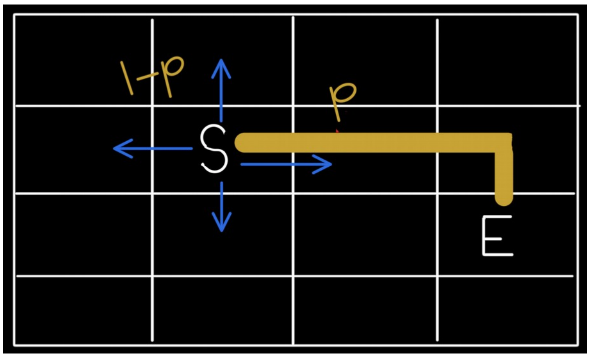
Information Theory and Statistics Project

Problem Statement

What is the expected number of steps required to reach from ‘S’ to ‘E’, when the guy is taking shortest path from ‘S’ to ‘E’ with probability p and a random decision with probability 1-p

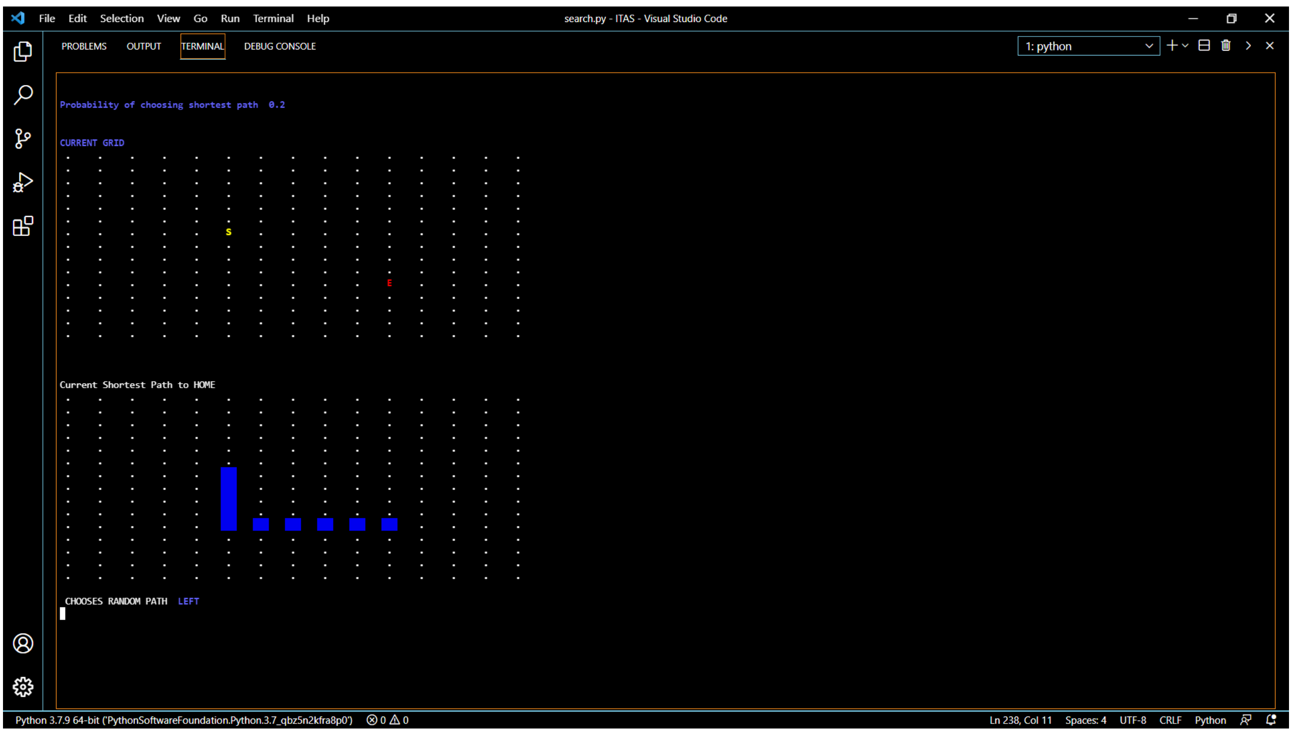
From the figure person is at ‘S’:

- Can take shortest path (represented via yellow line) with probability p
- Can take any random path (represented via blue arrows) with probability 1-p



What is the exptected number of steps required to reach destination i.e. ‘E’

Simulation results

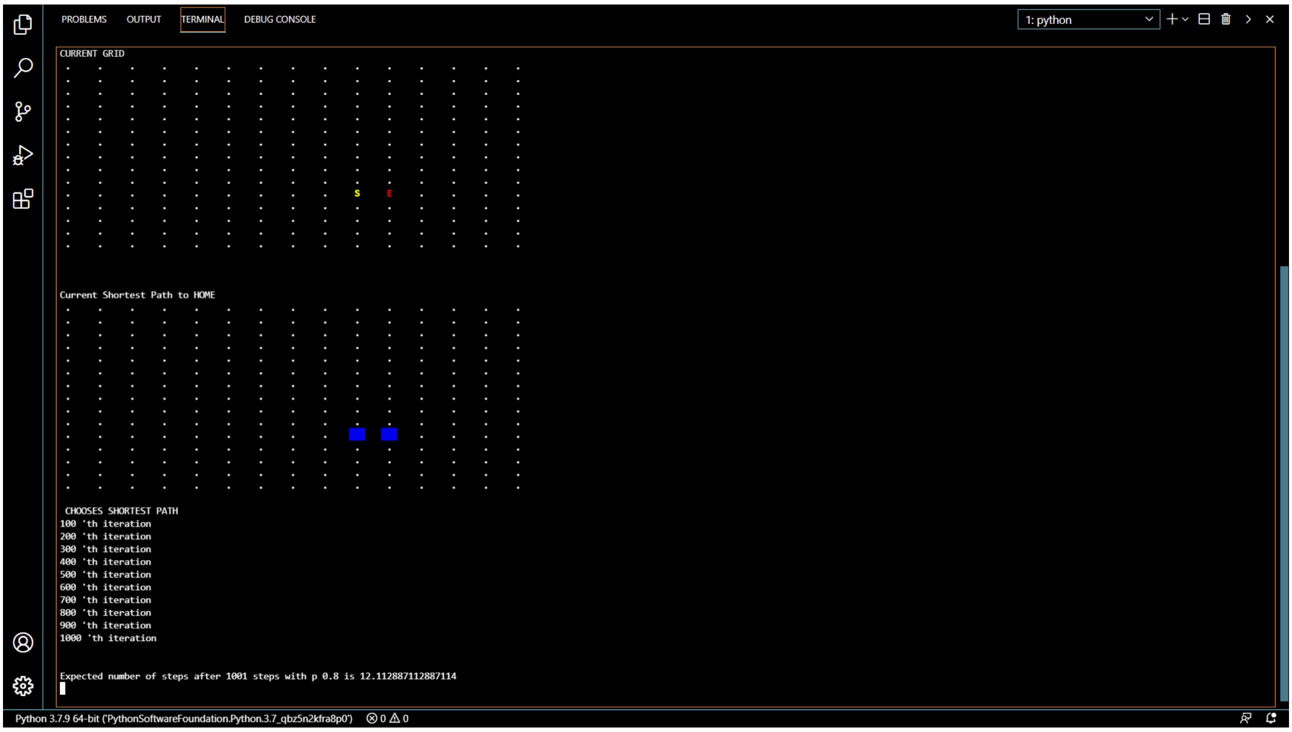


Notations:

- ‘S’ denotes the current position of the person and ‘E’ denotes the destination.
- Blue lines denotes the current shortest path from ‘S’ to ‘E’.
- At every step person either chooses the path using probability p.

Simulation:

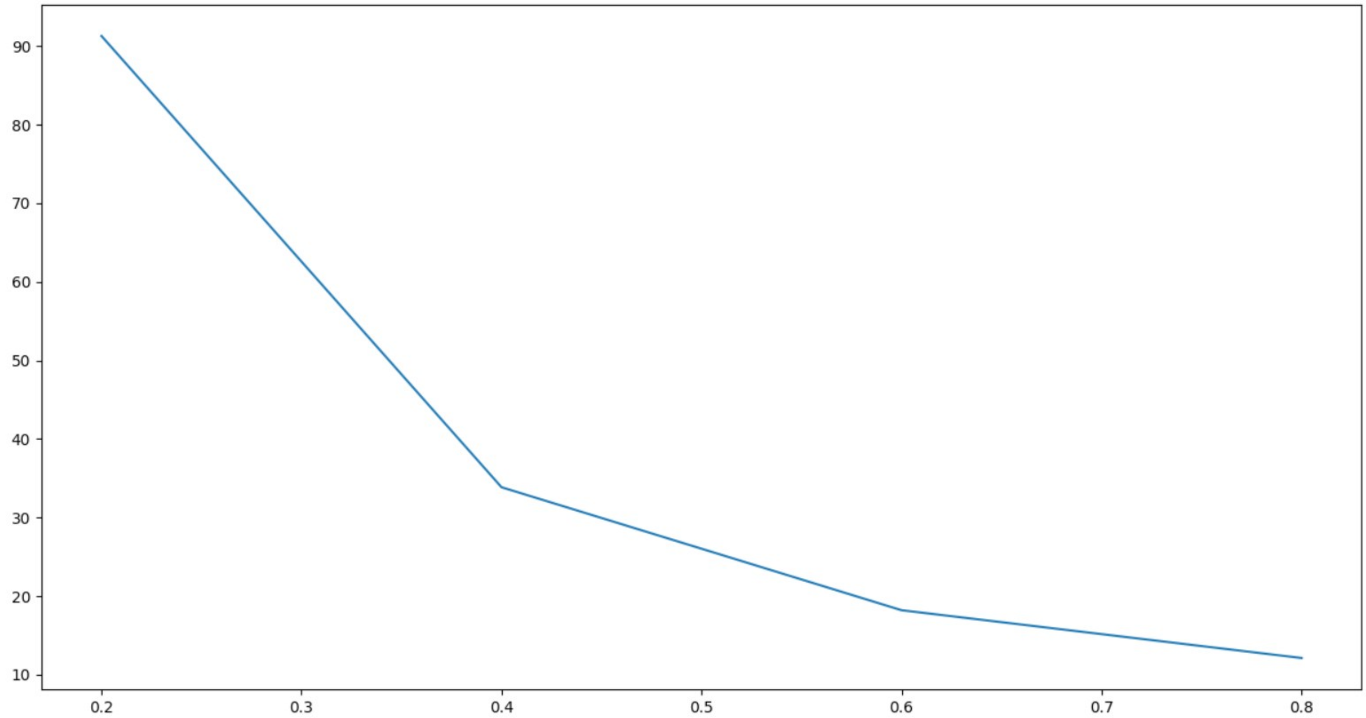
1. At every step we are doing Breadth First Search (bfs) to find the shortest path from ‘S’ to ‘E’
2. Then with probability p person choose to go with this path or a random path with 1-p probability.
3. Once person reaches the destination 1 iteration/simulation ends.
4. We are printing the full simulation to reach ‘E’ from ‘S’ only for 1 iterations. There are 1001 iterations done to find the expected value.
5. After all the iterations we find the expected number of steps required to reach ‘E’ with probability p.
6. p value changes 0.2 , 0.4, 0.6 and 0.8
7. We don’t choose 0.0 because in that case player always choose a random direction and it will take a lot of steps (time too) for the simulation to complete.
8. Similarly for p=1.0, person always choose the shortest path hence no of steps are always same (shortest distance between ‘S’ and ‘E’)



Inference:

#Exptected number of steps with:
**#	p = 0.2 -> 94.025**
**#	p = 0.4 -> 33.933**
**#	p = 0.6 -> 18.437**
**#	p = 0.8 -> 12.091**

With p=1.0 it should be 9 for all the simulations (i.e. shortest path)



Obvious statement: As the p increaseas the expected number of steps required to reach destination decreases because the probability to take shortest path is increasing.

Not to obvious statement: With the same increase of probability p = 0.2 to 0.4 and p = 0.4 to 0.6, we see that expected number of steps decreases exponentially. It goes from 94 to 34 to 18.5. Hence there is small difference in expected number of steps from p=0.6 to 0.8

Conclusion:

- We get insights of how a random and not so random walk can be simulated in real life scenario.
- We assumed that all the steps are independent of each other and the randomness doesn't depend on the location of the person. For example in real life p should be higher if person is near the destination (like near his/her home etc)

Cons:

- We could not simulate a theoretical model to prove the accuracy of our model. (We tried on a straight line but not on a matrix with taking independence and probability into consideration).
- We could not get the expected number of steps for p = 0.0 (i.e. for complete random walk) because the number of steps are too large to simulate (can be done but will take too much time).

Improvements:

- Variance of number of steps coming should also be analysed for more insights and better understanding of the results.
- Number of iterations could be increased with better cpu to get better expectation value.
- Could be done on bigger graph and more dynamic graphs.