Name: Omar Youssef

ID:43

Name: Islam Mostafa Gaber

ID:13

Name: Omar Emam

ID:42

Assignment 3

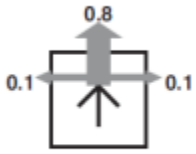
Problem Statement:

Consider the 3x3 world shown in the following figure:

r	-1	+10
-1	-1	-1
-1	-1	-1

The agent has four actions Up, Down, Right and Left.

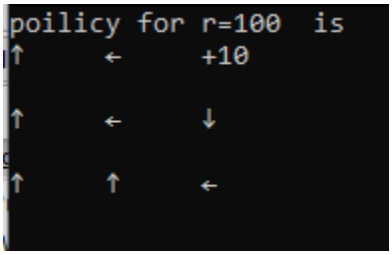
The transition model is: 80% of the time the agent goes in the direction it selects; the rest of the time it moves at right angles to the intended direction. A collision with a wall result in no movement.



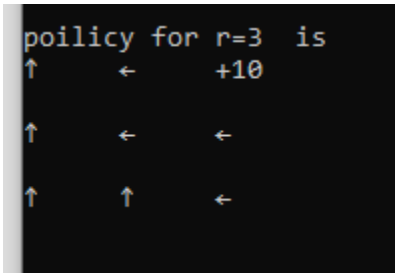
Requirements:

Gamma=0.99:

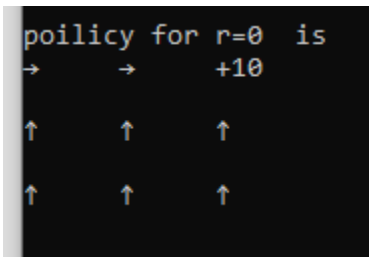
r=100:



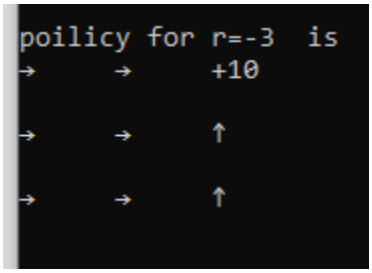
r=3:



r=0:



r=-3:



Explain intuitively why the value of r leads to each policy:

For r=100

We do the above steps as it is the max reward I can have as the other reward is 10 so we try to get the maximum one.

For r=3

It goes in this direction as it goes the number of rewards increase so try to collect the max.

For r=0

The num of rewards is small so it tries to go and take the 10,

For r=-3

It tries to finish the game it takes the 10 fastest.