

AI Assignment-2

Team-10

Pragun Saxena

Sankalp Agarwal

Input

4 4
1 0 0 10
0 0 0 0
0 0 0 0
0 0 -2 0
3 2
0 0
0 3
3 2
0 1
2 2
3 0
-1

1a) Gamma = 0.99

Step cost = -1

1.000	0.000	8.480	10.000
3.947	5.755	7.276	8.480
2.941	4.270	0.000	7.127
1.727	2.352	-2.000	4.935

Number of Iterations : 16

-	-	E	-
E	E	E	N
E	N	-	N
N	N	-	N

Since Value of end state (0,3) is very high is compensates the step cost of each and every state and reaches there.

1b) Gamma = 0.1

Step cost = -1

1.000	0.000	-0.212	10.000
-0.940	-1.097	-1.029	-0.212
-1.097	-1.109	0.000	-1.038
-1.109	-1.110	-2.000	-1.109

Number of Iterations : 3

-	-	E	-
N	W	E	N
N	N	-	N
N	W	-	E

Owing to very low discount factor utility does not propagate very far. Therefore despite having the ability to reach a higher goal state, it chooses to reach one that is close (0,0).

2a) Gamma = 0.99

Step cost = 10

1.000	0.000	508.639	10.000
508.639	508.639	508.639	508.628
508.639	508.639	0.000	508.627
508.639	508.639	-2.000	508.615

Number of Iterations : 70

-	-	W	-
S	E	W	S
E	N	-	N
N	W	-	E

Since the step cost is positive, the policy prefers remaining on the board , i.e. actively avoids end states. In all these examples it has picked the option where there is no chance of it making to the exit stage, since value keeps adding to the utility.

2b) Gamma = 0.99

Step cost = -2

1.000	0.000	7.098	10.000
-0.687	1.902	4.799	7.098
-2.927	-0.874	0.000	4.515
-5.184	-3.412	-2.000	1.529

Number of Iterations : 13

-	-	E	-
E	E	E	N
N	N	-	N
N	N	-	N

Since here Step cost is -2 and gamma 0.99 , its behaviour is almost similar to **1a** but it will try to reach end state sooner because step cost is high and moving too much may be harmful for the policy

2c) Gamma = 0.99

Step cost = -2.5

1.000	0.000	6.407	10.000
-1.897	-0.008	3.561	6.407
-4.806	-3.310	0.000	3.210
-7.537	-4.896	-2.000	-0.173

Number of Iterations : 17

-	-	E	-
N	E	E	N
N	N	-	N
N	E	-	N

Tries at actively leave the board very fast as the penalty is high, so it doesn't care taking even the negative goal state if it means being able to leave the board faster. If better alternative state near enough to compensate for step cost, it prefers that state(eg 3,3)

2d) gamma = 0.99

Step cost = -10

1.000	0.000	-3.951	10.000
-------	-------	--------	--------

-13.004	-25.517	-14.992	-3.951
-25.517	-27.800	0.000	-16.350
-27.800	-15.890	-2.000	-14.647

Number of Iterations : 7

-	-	E	-
N	W	E	N
N	S	-	N
E	E	-	W

Since the Step Cost is very high It prefers to exit the board as soon as possible.Hence Taking the nearest Exit with minimum number of steps as even the best state cannot compensate for the step cost if more than one step is needed.