PART B

Input:

X/10	0	0	X
0	0	0	0
0	0	0	0
0	0	-X/5	0

Green color => Goal states with positive reward.
Red color => Goal states with negative reward.
Brown color => wall states.

1. Discount factor:
   a) r = 0.1

   Utility values:

2.2	0	-0.44	22
-2.024	-2.2	-2.2	-0.44
-2.2	-2.2	0	-2.2
-2.2	-2.2	-4.4	-2.2

   Policy:

_	_	E	_
N	W	E	N
N	E	_	N
E	W	_	E

b) r = 0.99

Utility values:

2.2	0	18.654	22
8.685	12.659	16.004	18.654
6.481	9.398	0	15.678
3.824	5.185	-4.4	10.856

Policy:

_ _ E _
E E E N
E N _ N
N N _ N

If discount factor is less

- delta is more => number of iterations are less.
- It considers less past into account and gives utility.

If discount factor is more

- It does less iterations and has more consideration of the past iteration impact but number of iterations are more.

2. Step cost :

a) 22

Utility values:-

2.2	**0**	**2199.853**	**22**
2199.853	**2199.853**	**2199.853**	**2199.853**
2199.853	**2199.853**	**0**	**2199.853**
2199.853	**2199.853**	**-4.4**	**2199.853**

Policy:-

_ _ W _
S S S S
S S _ S
S W _ E

b) -4.4

Utility values:-

2.2	**0**	**15.616**	**22**
-1.499	**4.189**	**10.56**	**15.616**
-6.411	**-1.906**	**0**	**9.934**
-11.34	**-7.468**	**-4.4**	**3.366**

Policy:-

_ _ E _
E E E N
N N _ N
N N _ N

c) -5.5

Utility values:

2.2	**0**	**14.094**	**22**
-4.172	**-0.019**	**7.832**	**14.094**
-10.572	**-7.282**	**0**	**7.06**
-16.581	**-10.772**	**-4.4**	**-0.382**

Policy:

_ _ E _
N E E N
N N _ N
N E _ N

d) -22

Utility values:

2.2	0	-4.576	22
-20.258	-22	-22	-4.576
-22	-22	0	-22
-22	-22	-4.4	-22

Policy:

_ _ E _
N W E N
N E _ N
E E _ W

We can see how the policy changes , as step cost value increases.

- If step cost is so high => In a position we don’t want to reach nearest goal position fastly as that may decrease the utility values compared to other routes where more steps are involved.

- If stepcost is less, we wanted to maximise utility and so reaches the nearest goal state in less number of steps.

Step cost =>22	Step cost = -22
policy(1,3)=S	policy(1,3)=N

Here each case follows the **MEU Principle**.