Assignment-3

P. Aditya Pavani
2021101133.

Given, Grid

	0	1	2
0	0	1	-1
1	0	0	0
2	0	wall	0
3	0	0	0

Grid values → utility of each cell.

Initial utility of each cell is zero, except ~~wall = 0~~
reward = 1
penalty = -1

we have,

$$U_{t+1}(I) = \max\left(R(I,A) + \gamma \sum_j P(j|I, A) \, U_t(j)\right)$$

$I \rightarrow$ current state

$R(I, A) \rightarrow$ cost / reward

$A \rightarrow$ Action performed

$j \rightarrow$ Next state

Given,

Step cost, $R(I,A) = -0.04$

Probability of going in the direction of action = 0.7

Probability of going in the direction perpendicular in action = 0.15.

Iteration - 1

$$U_1(0,0) = \max \begin{cases} -0.04 + 0.95(0.7 \times 1 + 0.15 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 1 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 1 + 0.15 \times 0) \end{cases}$$

$$= \max \begin{cases} 0.625 \\ -0.04 \\ -0.1625 \\ = 0.1025 \end{cases}$$

$$= 0.625.$$

$$U_1(1,0) = \max \begin{cases} -0.04 + 0.95(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \end{cases}$$

$$= \max \begin{cases} -0.04 \\ -0.04 \\ -0.04 \\ -0.04 \end{cases}$$

$$= -0.04$$

$$U_1(1,1) = \max \begin{cases} -0.04 + 0.95(0.7 \times 0 + 0.15 \times 1 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 1 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 1 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \end{cases}$$

$$= \max \begin{cases} 0.1025 \\ 0.1025 \\ 0.625 \\ -0.04 \end{cases}$$

$$= 0.625$$

$$U_1(1,2) = \max \begin{cases} -0.04 + 0.95(0.7 \times 0 + 0.15 \times (-1) + 0.15 \times 0 \\ -0.04 + 0.95(0.7 \times 0 + 0.15(-1) + 0.15 \times 0 \\ -0.04 + 0.95(0.7 \times (-1) + 0.15(0) + 0.15(0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15(0) + 0.15(0) \end{cases}$$

$$= \max \begin{cases} -0.1825 \\ -0.1825 \\ -0.705 \\ -0.04 \end{cases}$$

$$= -0.04$$

$$U_1(2,0) = \max \begin{cases} -0.04 + 0.95(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \end{cases}$$

$$= \max \begin{cases} -0.04 \\ -0.04 \\ -0.04 \\ -0.04 \end{cases}$$

$$= -0.04$$

$$V_1(2,2) = \max \begin{cases} -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \end{cases}$$

$$= \max \begin{cases} -0.04 \\ -0.04 \\ -0.04 \\ -0.04 \end{cases}$$

$$= -0.04$$

$$V_1(3,0) = \max \begin{cases} -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \end{cases}$$

$$= \max \begin{cases} -0.04 \\ -0.04 \\ -0.04 \\ -0.04 \end{cases}$$

$$= -0.04$$

$$V_1(3,1) = \max \begin{cases} -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \end{cases}$$

$$= \max \begin{cases} -0.04 \\ -0.04 \\ -0.04 \\ -0.04 \end{cases}$$

$$= -0.04$$

$$V_1(3,2) = \max \begin{cases} -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \\ -0.04 + 0.95 \,(0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0) \end{cases}$$

$$= \max \begin{cases} -0.04 \\ -0.04 \\ -0.04 \\ -0.04 \end{cases}$$

$$= -0.04$$

Grid after Iteration 1:

0.625	1	-1
-0.04	0.625	-0.04
-0.04	wall	-0.04
-0.04	-0.04	-0.04

Iteration 2:

$U_3 (0,0)$ = max $\begin{cases} -0.04 + 0.95 (0.7 \times 1 + 0.15 \times 0.625 + 0.15 (-0.04) \\ -0.04 + 0.95 (0.7 \times 0 + 0.625 \times 0.15 + 0.15 (-0.04) \\ -0.04 + 0.95 (0.7 \times 0.625 + 0.15 \times 1 + 0.15 (0.625) \\ -0.04 + 0.95 (0.7 \times (-0.04) + 0.15 \times 1 + 0.15 (0.625) \end{cases}$

= Max $\begin{cases} 0.7083625 \\ 0.4589875 \\ 0.6071875 \\ \cancel{0.0018625} \; 0.1649625 \end{cases}$

= 0.7083625.

$U_2(1,0)$ = max $\begin{cases} -0.04 + 0.95(0.7(0.625) + 0.15(0.625) + 0.15(-0.04) \\ -0.04 + 0.95 (0.7(-0.04) + 0.15(0.625) + 0.15(0.04) \\ -0.04 + 0.95 (0.7(0.625) + 0.15(-0.04) + 0.15(0.625) \\ -0.04 + 0.95 (0.7(0.04) + 0.15(-0.04) + 0.15(0.625) \end{cases}$

= max $\begin{cases} 0.4589875 \\ 0.0167625 \\ 0.4589875 \\ 0.0167625 \end{cases}$

= 0.4589875

$U_2(1,1)$ = max $\begin{cases} -0.04 + 0.95 (0.7(-0.04) + 0.15 \times 1 + 0.15 (0.625) \\ -0.04 + 0.95 (0.7 (-0.04) + 0.15 \times 1 + 0.15 (0.625) \\ -0.04 + 0.95 (0.7(1) + 0.15 (-0.04) + 0.15 (-0.04) \\ -0.04 + 0.95 (0.7 (0.625) + 0.15 (-0.04) + 0.15 (0.04) \end{cases}$

= max $\begin{cases} 0.1649625 \\ 0.1649625 \\ 0.6136 \\ 0.364225 \end{cases}$

= 0.6136.

$$U_2(1,2) = \max \begin{cases} -0.04 + 0.95(0.7(-0.04) + 0.15\cdot(-1) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(0.625) + 0.15(-1) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-1) + 0.15(0.625) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(0.625) + 0.15(-0.04) \end{cases}$$

$$= \max \begin{cases} -0.2148 \\ 0.227425 \\ -0.6216375 \\ 0.0167625 \end{cases}$$

$$= 0.227425.$$

$$U_2(2,0) = \max \begin{cases} -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \end{cases}$$

$$= \max \begin{cases} -0.078 \\ -0.078 \\ -0.078 \\ -0.078 \end{cases}$$

$$= -0.078$$

$$U_2(2,2) = \max \begin{cases} -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \end{cases}$$

$$= \max \begin{cases} -0.078 \\ -0.078 \\ -0.078 \\ -0.078 \end{cases}$$

$$= -0.078$$

$$U_2(3,0) = \max \begin{cases} -0.04 + 0.95((-0.04)0.7 + 0.15(-0.04) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) \end{cases}$$

$$= \max \begin{cases} -0.078 \\ -0.078 \\ -0.078 \\ -0.078 \end{cases}$$

$$= -0.078.$$

$$U_2(3,1) = \max \begin{cases} -0.04 + 0.95(0.7(-0.04) + 0.15(+0.04) + 0.15(-0.04)) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04)) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04)) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04)) \end{cases}$$

$$= \max \begin{cases} -0.078 \\ -0.078 \\ -0.078 \\ -0.078 \end{cases}$$

$$= -0.078.$$

$$U_2(3,2) = \max \begin{cases} -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04)) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04)) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04)) \\ -0.04 + 0.95(0.7(-0.04) + 0.15(+0.04) + 0.15(-0.04)) \end{cases}$$

$$= \max \begin{cases} -0.078 \\ -0.078 \\ -0.078 \\ -0.078 \end{cases}$$

$$= -0.078.$$

Grid after Iteration 2

0.7083625	1	-1
0.4589875	0.6136	0.227425
-0.078	Wall	-0.078
-0.078	+0.078	-0.078