

MDL Assignment 3

(g2) The utility matrix is $\begin{bmatrix} 0, 1, -1 \end{bmatrix}$

$[0, 0, 0]$

$[0, 9, 0]$

$[0, 0, 0]$

Here we assign the penalty state with the initial value of -1 , the goal state with the value of 1 . The wall is assigned with a value of 9 .

ITERATION 1

→ Initially the utility of all other cells except the ~~near~~ goal, penalty and wall, to zero.

1) Cell $(0,0)$ - Boundary to left and above

$$\text{UP: } 0(0.7) + 0(0.15) + 1(0.15) = 0.15$$

up left right

→ Here up is $0(0.7)$ as there is a ~~no~~ boundary above the cell and thus this action results us to stay in the same cell.

→ Similarly left is $0(0.15)$ as ~~it is~~ to the left there is a boundary. We take the utility to be 0 as we stay in the cell itself and take 0.15 for going in a perpendicular direction.

→ Here right is $1(0.15)$ as the next cell to the right has utility 1 and the direction is perpendicular to UP.

$$\text{DOWN: } 0(L) + 1(R) + 0(D) = 0.15$$

$$\text{LEFT: } 0(L) + 0(R) + 0(D) = 0$$

$$\text{RIGHT: } 1(L) + 0(R) + 0(D) = 0.7$$

~~utility~~ utility $[0][0] = -0.04 + 0.95 \max(\text{UP}, \text{DOWN}, \text{LEFT}, \text{RIGHT})$

$$\begin{aligned} \text{utility}[0][0] &= -0.04 + 0.95 \max(0.15, 0.7, 0.15, 0) \\ &= -0.04 + 0.95 \times 0.7 \\ &= 0.625 \end{aligned}$$

*cells $(0,1)$ and $(0,2)$ are reward & penalty states

2) cell $(0,1)$ so we skip them *their utilities stay the same*

~~$\text{UP: } 1(L) + 0(R) + 0(D) = 0.15 / 0.85$~~

~~$\text{DOWN: } 0$~~

2) cell ~~(0,0)~~ $(1,1)$

$$\text{UP: } 0(L) + 0(R) + 0(D) = 0$$

$$\text{DOWN: } 0(L) + 0(R) + 0(D) = 0$$

$$\text{LEFT: } 0(L) + 0(R) + 0(D) = 0$$

$$\text{RIGHT: } 0(L) + 0(R) + 0(D) = 0$$

$$\text{utility}[1][1] = -0.04 + 0.95 \max(0, 0, 0, 0)$$

$$= -0.04 + 0.95(0)$$

$$= -0.04$$

3) cell $(1,2)$

$$\text{UP: } 1(L) + 0(R) + 0(D) = 0.7$$

$$\text{DOWN: } 0(0.7) + 0(0.15) + 0 \cdot (0.15) = 0$$

↓
wall below

$$\text{LEFT: } 0(0.7) + 1(0.15) + 0(0.15) = 0.15$$

↓
wall below

$$\text{RIGHT: } 0(0.7) + 1(0.15) + 0(0.15) = 0.15$$

$$\begin{aligned}\text{utility [1][1]} &= -0.04 + 0.95[\max(0, 0.7, 0.15, 0.15)] \\ &= -0.04 + 0.95(0.7) \\ &= 0.625\end{aligned}$$

3) cell (1, 2)

$$\text{UP: } -1(0.7) + 0(0.15) + 0(0.15) = -0.7$$

$$\text{DOWN: } 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\text{LEFT: } 0(0.7) + 0(0.15) + \cancel{0(0.15)} - 1(0.15) = -0.15$$

$$\text{RIGHT: } 0(0.7) + 0(0.15) - 1(0.15) = -0.15$$

$$\begin{aligned}\text{utility [1][2]} &= -0.04 + 0.95[\max(-0.7, 0, -0.15, -0.15)] \\ &= -0.04 + 0.95(0) \\ &= -0.04\end{aligned}$$

4) cell (2, 0)

↑ right is wall

$$\text{UP: } 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\text{DOWN: } 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\text{RIGHT: } 0(0.7) + 0.15(0) + 0(0.15) = 0$$

$$\text{LEFT: } 0(0.7) + 0.15(0) + 0(0.15) = 0$$

$$\begin{aligned}\text{utility}[2][0] &= -0.04 + 0.95 [\max(0, 0, 0, 0)] \\ &= -0.04 + 0 \\ &= -0.04\end{aligned}$$

cell (2,1) is a wall and thus its utility does not change

5) cell (2,2)

→ wall to the left

$$\underline{\text{UP}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{DOWN}}: 0(0.7) + \cancel{0(0.15)} + 0(0.15) = 0$$

$$\underline{\text{LEFT}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{RIGHT}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\begin{aligned}\text{utility}[2][2] &= -0.04 + 0.95 [\max(0, 0, 0, 0)] \\ &= -0.04 + 0 \\ &= -0.04\end{aligned}$$

6) cell (3,0)

$$\underline{\text{UP}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{LEFT}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{RIGHT}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{DOWN}}: 0(0.7) + 0.15(0) + 0(0.15) = 0$$

$$\begin{aligned}\text{utility}[3][0] &= -0.04 + 0.95 [\max(0, 0, 0, 0)] \\ &= -0.04 + 0 \\ &= -0.04\end{aligned}$$

7) cell (3,1) \rightarrow wall above

$$\underline{\text{UP}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{DOWN}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{LEFT}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{RIGHT}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$8) \text{utility}[3][1] = -0.04 + 0.95 [\max(0, 0, 0; 0)]$$

$$= -0.04 + 0$$

$$= -0.04$$

8) cell (3,2)

$$\underline{\text{UP}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{DOWN}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{LEFT}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\underline{\text{RIGHT}}: 0(0.7) + 0(0.15) + 0(0.15) = 0$$

$$\text{utility}[3][2] = -0.04 + 0.95 [\max(0, 0, 0, 0)]$$

$$= -0.04 + 0$$

$$= -0.04$$

8. The utility matrix after 1 iteration =

$$\begin{bmatrix} [0.625 & 1 & -1] \\ [-0.04 & 0.625 & -0.04] \\ [-0.04 & 9 & -0.04] \\ [-0.04 & -0.04 & -0.04] \end{bmatrix}$$

ITERATION 2

In this iteration we use the utility values generated in iteration 1.

1) $U(0,0)$

~~$UP: 0(0.625) + 0(0.15) + 1(0.6) + (0.15) = 0.775$~~

~~$DOWN: 0.7(-0.04) + 0.15(0.625) + 0.15(1) =$~~

~~$= 0.028 + 0.09375 + 0.15 = 0.21575$~~

~~$UP: 0.7(0.625) + 0.625(0.15) + 1(0.15) = 0.625$~~

~~$LEFT: 0.7(0.625) + 0.15(0.625) + 0.15(-0.04) = 0.52525$~~

~~$DOWN: 0.7(-0.04) + 0(0.625)(0.15) + 1(0.15) = 0.21575$~~

~~$RIGHT: 0.7(1) + 0.15(0.625) + 0.15(-0.04) = 0.78775$~~

$$\text{utility}[0][0] = -0.04 + 0.95 [\max(0.21575, 0.68125, 0.52525, 0.78775)]$$

$$= -0.04 + 0.95(0.78775)$$

$$= 0.70836$$

② $U(0,1)$ and $U(0,2)$ are reward & penalty states
so their utility remains the same

2) $U(1,0)$

~~$UP: 0.7(0.625) + 0.15(-0.04) + 0.15(0.625) = 0.52525$~~

~~$DOWN: 0.7(-0.04) + 0.15(0.625) + 0.15(-0.04) = 0.05975$~~

~~$LEFT: 0.7(-0.04) + 0.15(0.625) + 0.15(0.625) = 0.05975$~~

~~$RIGHT: 0.7(0.625) + 0.15(0.625) + 0.15(-0.04) = 0.52525$~~

$$\text{utility}[1][0] = -0.04 + 0.95 [\max(\text{UP}, \text{DOWN}, \text{LEFT}, \text{RIGHT})]$$

$$= -0.04 + 0.95 \times 0.52525 = 0.52525$$

$$= 0.4589$$

3) cell (1, 1)

UP: ~~(0.7)(0.625)~~

$$\underline{\text{UP}}: 1(0.7) + 0.15(-0.04) + 0.15(-0.04) =$$

$$\underline{\text{UP}}: 1(0.7) + 0.15(-0.04) + 0.15(-0.04) = 0.628$$

$$\underline{\text{DOWN}}: 0(0.625) + 0.7(0.625) + 0.15(-0.04) + 0.15(-0.04)$$

$$= 0.4255$$

$$\underline{\text{LEFT}}: 0.7(-0.04) + 0.15(1) + 0.15(0.625) = 0.21575$$

$$\underline{\text{RIGHT}}: 0.7(-0.04) + 0.15(1) + 0.15(0.625) = 0.21575$$

$$\text{utility}[1][1] = -0.04 + 0.95 [\max(\text{UP}, \text{DOWN}, \text{LEFT}, \text{RIGHT})]$$

$$= -0.04 + 0.95 (0.628) (0.688)$$

$$= -0.04 + 0.6536$$

$$= 0.6136$$

4) cell (1, 2)

$$\underline{\text{UP}}: 0.7(-1) + 0.15(-0.04) + 0.15(0.625) = -0.61225$$

$$\underline{\text{DOWN}}: 0.7(-0.04) + 0.15(0.625) + 0.15(-0.04) = 0.05975$$

$$\underline{\text{LEFT}}: 0.7(0.625) + 0.15(-1) + 0.15(-0.04) = 0.2815$$

$$\underline{\text{RIGHT}}: 0.7(-0.04) + 0.15(0.04) + 0.15(-1) = -0.184$$

$$\text{utility}[1][2] = -0.04 + 0.95 [\max(\text{UP}, \text{DOWN}, \text{LEFT}, \text{RIGHT})]$$

$$= -0.04 + 0.95 (0.6136) (0.2815)$$

$$= 0.227425$$

5) cell (2,0)

$$\underline{\text{UP}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{DOWN}}: 0.7(-0.04) + 0.15(0.04) + 0.15(-0.04) = 0.04$$

$$\underline{\text{RIGHT}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{LEFT}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\begin{aligned}\text{utility}[2][0] &= -0.04 + 0.95 [\max(\text{UP}, \text{DOWN}, \text{LEFT}, \text{RIGHT})] \\ &= -0.04 + 0.95(-0.04) \\ &= -0.078\end{aligned}$$

(6) cell (2,1) is a wall

6) cell (2,2)

$$\underline{\text{UP}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{DOWN}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{RIGHT}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{LEFT}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\begin{aligned}\text{utility}[2][2] &= -0.04 + 0.95 [\max(\text{UP}, \text{DOWN}, \text{LEFT}, \text{RIGHT})] \\ &= -0.04 + 0.95(-0.04) \\ &= -0.078\end{aligned}$$

7) cell (3, 0)

$$\underline{\text{UP}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{DOWN}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{LEFT}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{RIGHT}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\begin{aligned}\text{utility}[3][0] &= -0.04 + 0.95 \cdot \max(\text{UP}, \text{DOWN}, \text{LEFT}, \text{RIGHT}) \\ &= -0.04 + 0.95(-0.04) \\ &= -0.078\end{aligned}$$

8) cell (3, 1)

$$\underline{\text{UP}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{DOWN}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{LEFT}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{RIGHT}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\begin{aligned}\text{utility}[2][1] &= -0.04 + 0.95 \cdot \max(\text{UP}, \text{DOWN}, \text{LEFT}, \text{RIGHT}) \\ &= -0.04 + 0.95(-0.04) \\ &= -0.078\end{aligned}$$

9) cell (3, 2)

$$\underline{\text{UP}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{DOWN}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{LEFT}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\underline{\text{RIGHT}}: 0.7(-0.04) + 0.15(-0.04) + 0.15(-0.04) = -0.04$$

$$\begin{aligned}\text{utility}[3][2] &= -0.04 + 0.95 * [\max(\text{UP}, \text{DOWN}, \text{LEFT}, \text{RIGHT})] \\ &= -0.04 + 0.95 * (-0.04) \\ &= -0.078\end{aligned}$$

∴ The utility matrix after 2 iterations =

$$\begin{bmatrix} 0.70836 & -0.0051 & -1 \\ 0.4589 & 0.6136 & 0.227425 \\ -0.078 & 9 & -0.078 \\ -0.078 & -0.078 & -0.078 \end{bmatrix}$$

The values we calculated are extremely close to the ones obtained from the code.