19/04/24

MDL ASSIGNMENT 3.2

Question 3.2

→ Given probability = 0.7, perpendicular_prob = 0.15 $\left(\frac{1-0.7}{2}\right)$

→ Intial Grid =

$$\begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Grid[0][1] = Penalty

Grid[0][2] = Goal

Grid[2][1] = Wall

$U_t((i,j)) = \max_A [Cost((i,j),A) + \gamma \sum_J P(J|A,(i,j)) U_{t-1}(J)]$

Grid(i)(j) = $U_t(i,j)$

Step cost = -0.04

(i,j) ≡ state

γ = discount Rate ≡ 0.95

A ≡ Action

So for Iteration 1

Right up down left

Grid[0][0] = $-0.04 + 0.95 \times \max$ [$0.7 \times (-1) + 0.15 \times 0 + 0.15 \times 0$; $0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0$; $0.7 \times 0 + 0.15 \times -1 + 0.15 \times 0$; $0.7 \times 0 + 0.15 \times 0 + 0.15 \times 0$]

$= -0.04 + 0.95 \times 0 = -0.04$

Grid[10][0] = $-0.04 + 0.95 \times \max$ [$0.7 \times 0 + 0.15(0+0)$; $0.7 \times 0 + 0.15(0+0)$; $0.7 \times 0 + 0.15(0+0)$; $0.7 \times 0 + 0.15(0+0$]

$= -0.04$

Grid[1][1] = $-0.04 + 0.95 \times \max$ [$0.7 \times 0 \times 0.15(0+0)$; $0.7 \times 0 + 0.15 \times (0+0)$; $0.7 \times 0 + 0.15(0+0)$; $0.7 \times 0 + 0.15(0+0)$]

$= -0.04$

Grid[1][2] = $-0.04 + 0.95 \times \max$ [$0.7 \times 0 + 0.15(1+0)$; $0.7 \times 1 + 0.15(0+0)$; $0.7 \times 0 \times 0.15(0+0)$; $0.7 \times 0 + 0.15(1+0)$]

$= -0.04 + 0665$

$= 0.625$

Grid[2][0] = $-0.04 + 0.95 \times \max$ [$0.7 \times 0 + 0.15(0+0)$; $0.7 \times 0 + 0.15(0+0)$; $0.7 \times 0 + 0.15(0+0)$; $0.7 \times 0 + 0.15(0+0)$]

$= -0.04$

__/__/____

Grid[2][2] = -0.04 + 0.95 × max [0.7×0 + 0.15(0+0), 0.7×0 + 0.15(0+0), 0.7×0 + 0.15(0+0), 0.7×0 + 0.15(0+0)]

= -0.04

Grid[3][0] = -0.04 + 0.95 × max [.7×0 + .15×(0+0), .7×0 + .15×(0+0), .7×0 + .15×(0+0), .7×0 + .15(0+0)]

= -0.04

Grid[3][1] = -0.04 + .95 × max [.7×0 + .15(0+0), .7×0 + .15(0+0), .7×0 + .15(0+0), .7×0 + .15(0+0)]

= -0.04

Grid[3][2] = -0.04 + 0.95 × max [.7×0 + .15(0+0), .7×0 + .15(0+0), .7×0 + .15(0+0), .7×0 + .15(0+0)]

= -0.04

$$Grid = \begin{bmatrix} -0.04 & -1 & 1 \\ -0.04 & -0.04 & 0.625 \\ -0.04 & 0 & -0.04 \\ -0.04 & -0.04 & -0.04 \end{bmatrix}$$

```
Iteration 1:
Utility Values:
[[-0.04  -1.     1.   ]
 [-0.04  -0.04   0.625]
 [-0.04   0.    -0.04 ]
 [-0.04  -0.04  -0.04 ]]
------------
```

The values generated by code

Iteration 2

Grid[0][0] = -0.04 + 0.95 × max

= -0.04 + (-0.038)

= -0.078

0.7 × -1 + 0.15 (-0.04 + (-0.04

0.7 × (-0.04) + 0.15 [(-0.04) + (-0.04)]

0.7 × (0.04) + 0.15 [(-0.04) + (-1)]

0.7 × (-0.04) + 0.15 [(-0.04) + (0.04)]

Grid[0][0] = -0.04 + 0.95 × max

= -0.04 + (-0.038)

= -0.078

0.7 × (-0.04) + 0.15 [(-0.04) + (0.04)]

0.7 × (-0.04) + 0.15 [(0.04) + (-0.04)]

0.7 × (-0.04) + 0.15 [(-0.04) + (0.04)]

0.7 × (-0.04) + 0.15 [(-0.04) + (0.04)]

Grid[0][1] = -0.04 + 0.95 × max

= -0.04 + 0.267425

= 0.227425

0.7 × 0.625 + 0.15 [-1 + 0]

0.7 × -1 + 0.15 [0.04 - 0.04]

0.7 × 0 + 0.15 (-0.04 - 0.04]

0.7 × (-0.04) + 0.15 [-1 + 0]

Grid[1][2] = -0.04 + 0.95 × max

= -0.04 + 0.7483625

= 0.7083625

0.7 × 0.625 + 0.15 [1 - 0.04)

0.7 × 1 + 0.15 (0.625 - 0.04)

0.7 × (-0.04) + 0.15 [0.625 - 0.04)

0.7 × (-0.04) + 0.15 [1 - 0.04

Grid[2][0] = -0.04 + 0.95 × max

= -0.04 + (-0.038)

= -0.78

0.7 × (-0.04) + 0.15 (-0.04 - 0.04)

0.7 × (-0.04) + 0.15 (-0.04 - 0.04)

0.7 × (-0.04) + 0.15 (-0.04 - 0.04)

0.7 × (-0.04) + 0.15 (-0.04 - 0.04)

Grid[2][2] = -0.04 × 0.95 × max

= -0.04 + 0.46425

= 0.36425

0.7 × (-0.04) + 0.15 (.625 - 0.04)

0.7 × (0.625) + 0.15 (-0.04 - 0.04)

0.7 × (-0.04) + 0.15 (-0.04 - 0.04)

0.7 × (-0.04) + 0.15 (-0.04 + .625)

__/__/__

Grid[3][0] = -0.04 + 0.95 × max

= -0.04 + (0.038)

= -0.078

[0.7 × (-0.04) + .15 (-0.04 - 0.04)
0.7 × (-0.04) + .15 (-0.04 - 0.04)
0.76 (-0.04) + .15 (-0.04 - 0.04)
0.7 × (-0.04) + .15 (-0.04 - 0.04)]

Grid[3][1] = (-0.04) + 0.95 × max

= -0.044 (-0.038)

= -0.078

[.7 (-0.04) + .15 (-0.04 - 0.04)
.7 (-0.04) + .15 (-0.04 - 0.04)
.7 (-0.04) + .15 (-0.04 - 0.04)
.7 (-0.04) + .15 (-0.04 - 0.04)]

Grid[3][2] = -0.04 + 0.95 max

= -0.04 + (0.038)

= -0.078

[.7(-0.04) + .15 (-0.04 - 0.04)
.7(-0.04) + .15 (-0.04 - 0.04)
.7 (-0.04) + .15 (-0.04 - 0.04)
.7 (-0.04) + .15 (-0.04 - 0.04)]

Grid = [-0078 -1 1
-0.078 0.227425 0.7083625
-0.078 0 0.364225
-0.078 -0.078 -0.078]

```
Iteration 2:
Utility Values:
[[-0.078     -1.          1.       ]
 [-0.078      0.227425    0.7083625]
 [-0.078      0.          0.364225 ]
 [-0.078     -0.078      -0.078    ]]
-------------
```

The values generated by code