Veeral agarwal - 2019114009
P.Sahithi Reddy - 2020121011

Roll number being used = 2019101017
x = 1 - (((4009 % 30) + 1) / 100)
  = 1 - ((19 + 1) / 100)
  = 1 - 0.2 = 0.8
y = (09 % 4) + 1 = 1 + 1 = 2
States, S = [S1, S2, S3, S4, S5, S6]
Actions, A = [Right, Left]

states	Correct observation	Incorrect observation
S1	0.9	0.1
S2	0.85	0.15
S3	0.9	0.1
S4	0.85	0.15
S5	0.85	0.15
S6	0.9	0.1

Initial Belief State, b = [0.3333, 0, 0.3333, 0, 0,0.3333]

For calculating beliefs of every state after every action we are calculating numerator and denominator (formulae are mentioned in calculation below)

Belief after action 1 , b = [0.0111,0.4722,0.0000,0.3777,0.0944,0.0444]
Belief after action 2 , b = [0.4589,0.0004,0.4707,0.0149, 0.0218,0.0328]
Belief after action 3 , b = [0.0659,0.7146,0.0021,0.1704,0.0447,0.0019]

MDL - Assignment - 3 Part - 1

VEERAL AHARWAL (2019114009)
P. SAHITHI REDDY (2020121011)

we are using roll no. 1 (2019114009) for calculation of x and y

$x = 1 - (((\text{last 4 digits of roll no.}) \% 30 + 1)/100)$

$= 1 - (((4009) \% 30 + 1)/100)$

$\boxed{x = 0.89}$

$Y = (\text{last two digits of roll no.}) \% 4 + 1$

$= (09) \% 4 + 1$

$\boxed{Y = 2}$

Table : Y i.e. Table : 2

$P(\text{observation} = \text{Red} \mid \text{State} = \text{red}) = 0.9$

$P(\text{observation} = \text{green} \mid \text{state} = \text{green}) = 0.85$

Initially the agent knows he is in one of the red states i.e. S_1, S_3, S_6. The agent has taken 3 actions consecutively

1. Agent took the right action and observed green.
2. Agent took the left action and observed Red.
3. Agent took the left action and observed green.

Initial belief = [0.3333, 0, 0.3333, 0, 0, 0.3333]

Action 1 : Agent took action Right & observed green

$O(s_1', a, \text{green}) = 0.1$, $a = \text{Right}$

$$b'(s_1) = O(s_1', a, o) * \sum_{s \in S} T(s, a, s_1')\, b(s)$$

$$= 0.1 \times \left((1-x)\, b(s_1) + (1-x)\, b(s_2) \right)$$

$$= 0.1 * \left(0.2 \times \frac{1}{3} + 0.2 \times 0 \right)$$

$$= 0.1 \times \frac{0.2}{3} = \frac{0.02}{3}$$

$$b'(s_1) = 0.0066$$

$$b'(s_2) = O(s_2', a, o) * \sum_{s \in S} T(s, a, s_2')\, b(s)$$

$$= 0.85 \times \left(x \times \underset{b(s_1)}{\frac{1}{3}} + (1-x) \times \underset{b(s_3)}{\frac{1}{3}} \right)$$

$$= 0.85 \times \left(0.8 \times \frac{1}{3} + 0.2 \times \frac{1}{3} \right)$$

$$= 0.85 \times (0.2666 + 0.666)$$

$$b'(s_2) = 0.2833$$

$$b'(s_3) = O(s_3, a, o) * \sum_{s \in S} T(s, a, s_3)\, b(s)$$

$$= 0.1 * (x \times b(s_2) + (1-x) \times b(s_4))$$

$$= 0.1 * (0 + 0)$$

$$= 0$$

$$b'(s_3) = 0.0$$

$$b'(s_4) = O(s_4, a, o) * \sum_{s \in S} T(s, a, s_4)\, b(s)$$

$$= 0.85 * (x \times b(s_3) + (1-x) \times b(s_5))$$

$$= 0.85 * \left(0.8 \times \frac{1}{3} + 0.2 \times 0\right)$$

$$= 0.85 \times 0.8 \times \frac{1}{3}$$

$$b'(s_4) = 0.2266$$

$$b'(s_5) = O(s_5, a, o) * \sum_{s \in S} T(s, a, s_5)\, b(s)$$

$$= 0.85 * [x \times b(s_4) + (1-x) \times b(s_6)]$$

$$= 0.85 * \left[0.8 \times 0 + 0.2 \times \frac{1}{3}\right]$$

$$= 0.85 \times 0.2 \times \frac{1}{3}$$

$$b'(s_5) = 0.0566$$

$$b'(s_6) = O(s_6, a, o) * \sum_{s \in S} T(s, a, s_6)\, b(s)$$

$$= 0.1 * [\alpha \times b(s_6)]$$

$$= 0.1 \times 0.8 \times \frac{1}{3}$$

∴ from s_6 if it take right then return to s_6.

$$b'(s_6) = 0.0266$$

$$\Rightarrow \text{Denominator} = \sum_{s' \in S} b'(s')$$

$$= \cancel{0.0067} + b'(s_1) + b(s_2) + b'(s_3) + b'(s_4) + b'(s_5) + b'(s_6)$$

$$= 0.0066 + 0.2833 + 0 + 0.2266 + 0.0566 + 0.0266$$

$$\text{Denominator} = \cancel{0} 0.5997$$

$b(s_1) = b'(s_1) / 0.5997$

$= 0.0111$

$b(s_2) = b'(s_2) / 0.5997$

$= 0.4722$

$b(s_3) = b'(s_3) / 0.5997$

$= 0.0$

$b(s_4) = b'(s_4) / 0.5997$

$= 0.3777$

$b(s_5) = b'(s_5) / 0.5997$

$= 0.0944$

$b(s_6) = b'(s_6) / 0.5997$

$= 0.0444$

Action 2

Agent took the action left and observed Red.

$O(obs) = red, \quad a = left$

$b'(s_1) = O(s_1, a, o) * \sum_{s \in S} T(s, a, s_1) b(s)$

$= 0.9 * (x \times b(s_2) + x \times b(s_1))$

$= 0.9 * [0.8 \times 0.4722 + 0.8 \times 0.0111]$

$\Rightarrow b'(s_1) = 0.3480$

$$b'(s_2) = O(s_2, a, o) * \sum_{s \in S} T(s, a, s_2) b(s)$$

$$= (1 - 0.85) * ((1 - x) b(s_1) + x\, b(s_3))$$

$$= 0.15 \times 0.2 \times 0.0111$$

$$= 0.0003$$

$$b'(s_3) = O(s_3, a, o) * \sum_{s \in S} T(s, a, s_3) b(s)$$

$$= 0.9 * [x \times b(s_4) + (1 - x) \times b(s_2)]$$

$$= 0.9 * [0.8 \times 0.3777 + 0.2 \times 0.4722]$$

$$= 0.3570$$

$$b'(s_4) = O(s_4, a, o) * \sum_{s \in S} T(s, a, s_4) b(s)$$

$$= 0.15 * [(1 - x) b(s_3) + x\, b(s_5)]$$

$$= 0.15 * [0 + 0.8 \times 0.0944$$

$$= 0.0113$$

$$b'(s_5) = O(s_5, a, o) * \sum_{s \in S} T(s, a, s_5) b(s)$$

$$= 0.15 * [(1 - x) * b(s_4) + x\, b(s_6)]$$

$$= 0.15 * [0.2 \times 0.3777 + 0.8 \times 0.0444]$$

$$= 0.0166$$

$$b'(s_6) = o(s_6, a, o) * \sum_{s \in S} T(s, a, s_6)\, b(s)$$

$$= 0.9 * [(1-x)\, b(s_5) + (1-x)\, b(s_6)]$$

$$= 0.1 * [0.2 \times 0.9 \; 0.0944 + 0.2 \times 0.0444]$$

$$= 0.0249$$

$$\text{Denominator} = \sum_{s' \in S} b'(s')$$

$$= 0.3480 + 0.0003 + 0.3570 + 0.0113 + 0.0166 + 0.0249$$

$$= 0.7581$$

$$b(s_1) = b'(s_1) / \text{denominator} = 0.4589$$

$$b(s_2) = b'(s_2) / \text{denom.} = 0.0004$$

$$b(s_3) = b'(s_3) / \text{den.} = 0.4707$$

$$b(s_4) = b'(s_4) / \text{den.} = 0.0149$$

$$b(s_5) = b'(s_5) / \text{den.} = 0.0218$$

$$b(s_6) = b'(s_6) / \text{den} = 0.0328$$

Action 3

Agent took the action left and observed green.

o(observe) = green, a = left.

$$b'(s_1) = O(s_1, a, o) * \sum_{s \in S} T(s, a, s_1)\, b(s)$$

$$= 0.1 * [x \times b(s_1) + x \times b(s_2)]$$

$$= 0.1 * [0.8 \times 0.4589 + 0.8 \times 0.0004]$$

$$= 0.0367$$

$$b'(s_2) = O(s_2, a, o) * \sum_{s \in S} T(s, a, s_2)\, b(s)$$

$$= 0.85 * [(1-x)\, b(s_1) + x\, b(s_3)]$$

$$= 0.85 * [0.2 \times 0.4589 + 0.8 \times 0.4707]$$

$$= 0.3981$$

$$b'(s_3) = O(s_3, a, o) * \sum_{s \in S} T(s, a, s_3)\, b(s)$$

$$= 0.1 * [(1-x)\, b(s_2) + x\, b(s_4)]$$

$$= 0.1 * [0.2 \times 0.0004 + 0.8 \times 0.0149]$$

$$= 0.0012$$

$$b'(s_4) = O(s_4, a, o) * \sum_{s \in S} T(s, a, s_4)\, b(s)$$

$$= 0.85 * [(1-x)\, b(s_3) + x\, b(s_5)]$$

$$= 0.85 * [0.2 \times 0.4707 + 0.8 \times 0.0218]$$

$$= 0.0949$$

$$b'(s_5) = O(s_5, a, o) * \sum_{s \in S} T(s, a, s_5) b(s)$$

$$= 0.85 * [(1-x) b(s_4) + x\, b(s_6)]$$

$$= 0.85 * [0.2 \times 0.0149 + 0.8 \times 0.0328]$$

$$= 0.0249$$

$$b'(s_6) = O(s_6, a, o) * \sum_{s \in S} T(s, a, s_6) b(s)$$

$$= 0.1 * [(1-x) b(s_5) + (1-x) b(s_6)]$$

$$= 0.1 * [0.2 \times 0.0218 + 0.2 \times 0.0328)$$

$$= 0.00109$$

$$\text{Denominator} = \sum_{s' \in S} b'(s)$$

$$= 0.00109 + 0.0249 + 0.0949 + 0.0012 + 0.3981 + 0.0367$$

$$= 0.55689$$

$b(s_1) = b'(s_1)/den = 0.0659$

$b(s_2) = b'(s_2)/den = 0.7146$

$b(s_3) = b'(s_3)/den = 0.0021$

$b(s_4) = b'(s_4)/den = 0.1704$

$b(s_5) = b'(s_5)/den = 0.0447$

$b(s_6) = b'(s_6)/den = 0.0019$

beliefs after action 1

[0.0111, 0.4722, 0.0, 0.3777,
0.0944, 0.0444]

beliefs after action 2

[0.4589, 0.0004, 0.4707, 0.0149,
0.0218, 0.0328]

beliefs after action 3

[0.0659, 0.7146, 0.0021, 0.1704,
0.0447, 0.0019]