MDL Assignment 3, Part 1

Team: **Room543**

Members : **Mihir Bani, 2019113003**

Amul Agrawal, 2019101113

Roll Number used : 2019113003

$x = 1 - (3003 \% 30 + 1)/100 = 0.96$

$y = (03 \% 4) + 1 = 4$

P(action) - Probability of an action being successful or failure.

Action / Result	Success	Failure
LEFT	0.96	0.04
RIGHT	0.96	0.04

P(o|s) - Probability of an observation given current state.

	s = Red	s = Green
o = Red	0.8	0.05
o = Green	0.2	0.95

Formula used for calculating the next Belief state:

Ub' denotes Unnormalized Beliefs | **b(s)** = Current Beliefs

T(s',a,s) = Transition Probability | **O(s',a,o)** = **P(o|s)** = Observation probability

$$Ub'[s'] = O(s', a, o) \sum_{s} T(s', a, s)b(s)$$

$$sum = \sum_{s} Ub'[s] \text{ and then } b'[\,] = Ub'[\,] / sum$$

The normalized value is found by taking the sum of all entries of **Ub'[ ]** and then dividing each value of **Ub'** by *sum*.

Initial Beliefs: Initially the agent is any one of the **Red** states, and so the initial belief value of each of them is 1/3 and that of the **Green** states is 0.

S1	S2	S3	S4	S5	S6
0.33333	0	0.33333	0	0	0.33333

After Action 1 : **RIGHT** and observed **Green**

- UB'[S1] = 0.2 * [(0.04000 * 0.33333) + (0.04000 * 0.00000) + (0.00000 * 0.33333) + (0.00000 * 0.00000) + (0.00000 * 0.00000) + (0.00000 * 0.33333) +] = 0.00267
- UB'[S2] = 0.95 * [(0.96000 * 0.33333) + (0.00000 * 0.00000) + (0.04000 * 0.33333) + (0.00000 * 0.00000) + (0.00000 * 0.00000) + (0.00000 * 0.33333) +] = 0.31666
- UB'[S3] = 0.2 * [(0.00000 * 0.33333) + (0.96000 * 0.00000) + (0.00000 * 0.33333) + (0.04000 * 0.00000) + (0.00000 * 0.00000) + (0.00000 * 0.33333) +] = 0.00000
- UB'[S4] = 0.95 * [(0.00000 * 0.33333) + (0.00000 * 0.00000) + (0.96000 * 0.33333) + (0.00000 * 0.00000) + (0.04000 * 0.00000) + (0.00000 * 0.33333) +] = 0.30400
- UB'[S5] = 0.95 * [(0.00000 * 0.33333) + (0.00000 * 0.00000) + (0.00000 * 0.33333) + (0.96000 * 0.00000) + (0.00000 * 0.00000) + (0.04000 * 0.33333) +] = 0.01267
- UB'[S6] = 0.2 * [(0.00000 * 0.33333) + (0.00000 * 0.00000) + (0.00000 * 0.33333) + (0.00000 * 0.00000) + (0.96000 * 0.00000) + (0.96000 * 0.33333) +] = 0.06400

Total Sum = 0.69999

On Normalizing with Total Sum:

New Beliefs (B') calculated:

$$B'[S1] = UB'[S1]/Total\_sum$$
$$= 0.00267/0.69999 = 0.00381$$
$$B'[S2] = UB'[S2]/Total\_sum$$
$$= 0.31666/0.69999 = 0.45238$$
$$B'[S3] = UB'[S3]/Total\_sum$$
$$= 0.00000/0.69999 = 0.00000$$
$$B'[S4] = UB'[S4]/Total\_sum$$
$$= 0.30400/0.69999 = 0.43429$$
$$B'[S5] = UB'[S5]/Total\_sum$$
$$= 0.01267/0.69999 = 0.01810$$
$$B'[S6] = UB'[S6]/Total\_sum$$
$$= 0.06400/0.69999 = 0.09143$$

Updated Beliefs:

S1	S2	S3	S4	S5	S6
0.00381	0.45238	0.00000	0.43429	0.01810	0.09143

After Action 2 : **LEFT** and observed **Red**

- UB'[S1] = 0.8* [(0.96000 * 0.00381) + (0.96000 * 0.45238) + (0.00000 * 0.00000) + (0.00000 * 0.43429) + (0.00000 * 0.01810) + (0.00000 * 0.09143) +] = 0.35035
- UB'[S2] = 0.05* [(0.04000 * 0.00381) + (0.00000 * 0.45238) + (0.96000 * 0.00000) + (0.00000 * 0.43429) + (0.00000 * 0.01810) + (0.00000 * 0.09143) +] = 0.00001
- UB'[S3] = 0.8* [(0.00000 * 0.00381) + (0.04000 * 0.45238) + (0.00000 * 0.00000) + (0.96000 * 0.43429) + (0.00000 * 0.01810) + (0.00000 * 0.09143) +] = 0.34801
- UB'[S4] = 0.05* [(0.00000 * 0.00381) + (0.00000 * 0.45238) + (0.04000 * 0.00000) + (0.00000 * 0.43429) + (0.96000 * 0.01810) + (0.00000 * 0.09143) +] = 0.00087
- UB'[S5] = 0.05* [(0.00000 * 0.00381) + (0.00000 * 0.45238) + (0.00000 * 0.00000) + (0.04000 * 0.43429) + (0.00000 * 0.01810) + (0.96000 * 0.09143) +] = 0.00526
- UB'[S6] = 0.8* [(0.00000 * 0.00381) + (0.00000 * 0.45238) + (0.00000 * 0.00000) + (0.00000 * 0.43429) + (0.04000 * 0.01810) + (0.04000 * 0.09143) +] = 0.00350

Total Sum = 0.70800

On Normalizing with Total Sum:

New Beliefs (B') calculated:

B'[S1] = UB'[S1]/Total_sum
= 0.35035/0.70800 = 0.49485
B'[S2] = UB'[S2]/Total_sum
= 0.00001/0.70800 = 0.00001
B'[S3] = UB'[S3]/Total_sum
= 0.34801/0.70800 = 0.49154
B'[S4] = UB'[S4]/Total_sum
= 0.00087/0.70800 = 0.00123
B'[S5] = UB'[S5]/Total_sum
= 0.00526/0.70800 = 0.00743
B'[S6] = UB'[S6]/Total_sum
= 0.00350/0.70800 = 0.00495

Updated Beliefs:

S1	S2	S3	S4	S5	S6
0.49485	0.00001	0.49154	0.00123	0.00743	0.00495

After Action 3 : **LEFT** and observed **Green**

- UB'[S1] = 0.2* [(0.96000 * 0.49485) + (0.96000 * 0.00001) + (0.00000 * 0.49154) + (0.00000 * 0.00123) + (0.00000 * 0.00743) + (0.00000 * 0.00495) +] = 0.09501
- UB'[S2] = 0.95* [(0.04000 * 0.49485) + (0.00000 * 0.00001) + (0.96000 * 0.49154) + (0.00000 * 0.00123) + (0.00000 * 0.00743) + (0.00000 * 0.00495) +] = 0.46709
- UB'[S3] = 0.2* [(0.00000 * 0.49485) + (0.04000 * 0.00001) + (0.00000 * 0.49154) + (0.96000 * 0.00123) + (0.00000 * 0.00743) + (0.00000 * 0.00495) +] = 0.00024
- UB'[S4] = 0.95* [(0.00000 * 0.49485) + (0.00000 * 0.00001) + (0.04000 * 0.49154) + (0.00000 * 0.00123) + (0.96000 * 0.00743) + (0.00000 * 0.00495) +] = 0.02545
- UB'[S5] = 0.95* [(0.00000 * 0.49485) + (0.00000 * 0.00001) + (0.00000 * 0.49154) + (0.04000 * 0.00123) + (0.00000 * 0.00743) + (0.96000 * 0.00495) +] = 0.00456
- UB'[S6] = 0.2* [(0.00000 * 0.49485) + (0.00000 * 0.00001) + (0.00000 * 0.49154) + (0.00000 * 0.00123) + (0.04000 * 0.00743) + (0.04000 * 0.00495) +] = 0.00010

Total Sum = 0.59244

On Normalizing with Total Sum:

New Beliefs (B') calculated:

    B'[S1] = UB'[S1]/Total_sum
        = 0.09501/0.59244 = 0.16038
    B'[S2] = UB'[S2]/Total_sum
        = 0.46709/0.59244 = 0.78840
    B'[S3] = UB'[S3]/Total_sum
        = 0.00024/0.59244 = 0.00040
    B'[S4] = UB'[S4]/Total_sum
        = 0.02545/0.59244 = 0.04296
    B'[S5] = UB'[S5]/Total_sum
        = 0.00456/0.59244 = 0.00770
    B'[S6] = UB'[S6]/Total_sum
        = 0.00010/0.59244 = 0.00017

Updated Beliefs:

S1	S2	S3	S4	S5	S6
0.16038	0.78840	0.00040	0.04296	0.00770	0.00017

Beliefs at each stage after the three actions:

Initial Beliefs:

S1	S2	S3	S4	S5	S6
0.33333	0	0.33333	0	0	0.33333

After Action 1:

S1	S2	S3	S4	S5	S6
0.00381	0.45238	0.00000	0.43429	0.01810	0.09143

After Action 2:

S1	S2	S3	S4	S5	S6
0.49485	0.00001	0.49154	0.00123	0.00743	0.00495

After Action 3:

S1	S2	S3	S4	S5	S6
0.16038	0.78840	0.00040	0.04296	0.00770	0.00017