# Assignment5 - Part1

**By 2018101075**

## Dataset

x = 1 - (075 % 40 + 1) / 100 = 0.64
y = 75 % 3 = 0

| | |
|---|---|
| P(Observation = Red \| State = Red) | 0.9 |
| P(Observation = Green \| State = Green) | 0.85 |

There are 5 states and two actions (Left, Right).
States S1, S2, S5 are Red and S3, S5 are green.
When an agent takes action Left , it goes to the left state with a probability x and goes to the right state with a probability 1- x . Similarly when the action Right is chosen, the agent goes to the right state with a probability x and goes to the left state with a probability 1- x. When the agent is in one of the extreme states (S1 or S5) and decides to go in the direction of the wall, it remains in the same state.

## Calculating Belief State

$b'(s') = Pr(s' \mid o, a, b)$
$= ( O(s', a, o) * \sum T(s, a, s') b(s) ) / Pr(o \mid a, b)$
Where s = old state,

b = old belief state, and b(s) probability of s given belief state b,
a = action,
b’ = new belief state,
b’(s’) = probability of s’ given b’,
o = observation.

We first calculate Ub’[s’] = O(s’, a, o) * ∑ T(s, a, s’) b(s)
Then b’ = Ub’ / ∑ Ub’[s’]

**Step 1**

Initially, the agent knows that it is in one of the red states i.e. S1, S2 or S5.
Therefore belief b = [1 / 3, 1 / 3, 0, 0, 1 / 3]

Agent took the action Right and observed Red

Ub’(s1) = Pr(Red | s1, Right) * [Pr(s1 | s1, Right) * b(s1) + Pr(s1 | s2, Right) * b(s2) + Pr(s1 | s3, Right) * b(s3) + Pr(s1 | s4, Right) * b(s4) + Pr(s1 | s5, Right) * b(s5)]
= 0.9 * [(1 - x) * 0.33 + (1 - x) * 0.33 + 0 + 0 + 0]
= 0.216

Ub’(s2) = Pr(Red | s2, Right) * [Pr(s2 | s1, Right) * b(s1) + Pr(s2 | s2, Right) * b(s2) + Pr(s2 | s3, Right) * b(s3) + Pr(s2 | s4, Right) * b(s4) + Pr(s2 | s5, Right) * b(s5)]
= 0.9 * [x * 0.33 + 0 + 0 + 0 + 0]
= 0.192

Ub’(s3) = Pr(Red | s3, Right) * [Pr(s3 | s1, Right) * b(s1) + Pr(s3 | s2, Right) * b(s2) + Pr(s3 | s3, Right) * b(s3) + Pr(s3 | s4, Right) * b(s4) + Pr(s3 | s5, Right) * b(s5)]
= 0.1 * [0 + x * 0.33 + 0 + 0 + 0]
= 0.02133

Ub’(s4) = Pr(Red | s4, Right) * [Pr(s4 | s1, Right) * b(s1) + Pr(s4 | s2, Right) * b(s2) + Pr(s4 | s3, Right) * b(s3) + Pr(s4 | s4, Right) * b(s4) + Pr(s4 | s5, Right) * b(s5)]
= 0.1 * [0 + 0 + 0 + 0 + (1 - x) * 0.33]
= 0.012

Ub’(s5) = Pr(Red | s5, Right) * [Pr(s5 | s1, Right) * b(s1) + Pr(s5 | s2, Right) * b(s2) + Pr(s5 | s3, Right) * b(s3) + Pr(s5 | s4, Right) * b(s4) + Pr(s5 | s5, Right) * b(s5)]
= 0.9 * [0 + 0 + 0 + 0 + x * 0.33]
= 0.192

Let p = sum(Ub’) = 0.63333
b’ = Ub’ / p
Therefore b’ = [0.34105, 0.30315, 0.03368, 0.01894, 0.30315]

**Step 2**

Now belief b = [0.34105, 0.30315, 0.03368, 0.01894, 0.30315]

Agent took the action Left and observed Green

Ub’(s1) = Pr(Green | s1, Left) * [Pr(s1 | s1, Left) * b(s1) + Pr(s1 | s2, Left) * b(s2) + Pr(s1 | s3, Left) * b(s3) + Pr(s1 | s4, Left) * b(s4) + Pr(s1 | s5, Left) * b(s5)]

= 0.15 * [x * b[s1] + x * b[s2] + 0 + 0 + 0]

= 0.06184

Ub’(s2) = Pr(Green | s2, Left) * [Pr(s2 | s1, Left) * b(s1) + Pr(s2 | s2, Left) * b(s2) + Pr(s2 | s3, Left) * b(s3) + Pr(s2 | s4, Left) * b(s4) + Pr(s2 | s5, Left) * b(s5)]

= 0.15 * [(1 - x) * b[s1] + 0 + x * b[s3] + 0 + 0]

= 0.02165

Ub’(s3) = Pr(Green | s3, Left) * [Pr(s3 | s1, Left) * b(s1) + Pr(s3 | s2, Left) * b(s2) + Pr(s3 | s3, Left) * b(s3) + Pr(s3 | s4, Left) * b(s4) + Pr(s3 | s5, Left) * b(s5)]

= 0.85 * [0 + (1 - x) * b[s2] + 0 + x * b[s4] + 0]

= 0.10307

Ub’(s4) = Pr(Green | s4, Left) * [Pr(s4 | s1, Left) * b(s1) + Pr(s4 | s2, Left) * b(s2) + Pr(s4 | s3, Left) * b(s3) + Pr(s4 | s4, Left) * b(s4) + Pr(s4 | s5, Left) * b(s5)]

= 0.85 * [0 + 0+ (1 - x) * b[s3] + 0 + x * b[s5]]

= 0.17522

Ub’(s5) = Pr(Green | s5, Left) * [Pr(s5 | s1, Left) * b(s1) + Pr(s5 | s2, Left) * b(s2) + Pr(s5 | s3, Left) * b(s3) + Pr(s5 | s4, Left) * b(s4) + Pr(s5 | s5, Left) * b(s5)]

= 0.15 * [0 + 0 + 0 + (1 - x) * b[s4] + (1 - x) * b[s5]]

= 0.01739

Let p = sum(Ub') = 0.37918
b' = Ub' / p
Therefore b' = [0.16309, 0.05709, 0.27182, 0.46210, 0.04587]

**Step 3**

Now belief b = [0.16309, 0.05709, 0.27182, 0.46210, 0.04587]

Agent took the action Left and observed Green

Ub'(s1) = Pr(Green | s1, Left) * [Pr(s1 | s1, Left) * b(s1) + Pr(s1 | s2, Left) * b(s2) + Pr(s1 | s3, Left) * b(s3) + Pr(s1 | s4, Left) * b(s4) + Pr(s1 | s5, Left) * b(s5)]
= 0.15 * [x * b[s1] + x * b[s2] + 0 + 0 + 0]
= 0.02113

Ub'(s2) = Pr(Green | s2, Left) * [Pr(s2 | s1, Left) * b(s1) + Pr(s2 | s2, Left) * b(s2) + Pr(s2 | s3, Left) * b(s3) + Pr(s2 | s4, Left) * b(s4) + Pr(s2 | s5, Left) * b(s5)]
= 0.15 * [(1 - x) * b[s1] + 0 + x * b[s3] + 0 + 0]
= 0.03490

Ub'(s3) = Pr(Green | s3, Left) * [Pr(s3 | s1, Left) * b(s1) + Pr(s3 | s2, Left) * b(s2) + Pr(s3 | s3, Left) * b(s3) + Pr(s3 | s4, Left) * b(s4) + Pr(s3 | s5, Left) * b(s5)]
= 0.85 * [0 + (1 - x) * b[s2] + 0 + x * b[s4] + 0]

= 0.26885

Ub’(s4) = Pr(Green | s4, Left) * [Pr(s4 | s1, Left) * b(s1) + Pr(s4 | s2, Left) * b(s2) + Pr(s4 | s3, Left) * b(s3) + Pr(s4 | s4, Left) * b(s4) + Pr(s4 | s5, Left) * b(s5)]
= 0.85 * [0 + 0+ (1 - x) * b[s3] + 0 + x * b[s5]]
= 0.10813

Ub’(s5) = Pr(Green | s5, Left) * [Pr(s5 | s1, Left) * b(s1) + Pr(s5 | s2, Left) * b(s2) + Pr(s5 | s3, Left) * b(s3) + Pr(s5 | s4, Left) * b(s4) + Pr(s5 | s5, Left) * b(s5)]
= 0.15 * [0 + 0 + 0 + (1 - x) * b[s4] + (1 - x) * b[s5]]
= 0.02743

Let p = sum(Ub’) = 0.46046
b’ = Ub’ / p
Therefore b’ = [0.04590, 0.07579, 0.58388, 0.23483, 0.05957]

**Output**

2018101075
0.64 0
0.34105 0.30315 0.03368 0.01894 0.30315
0.16309 0.05709 0.27182 0.46210 0.04587
0.04590 0.07579 0.58388 0.23483 0.05957