<u>CS7.301 (Machine Data and Learning)</u>

Vikrant Dewangan
Roll Number - 2018111024

28th April,2020
Assignment - 5, part A

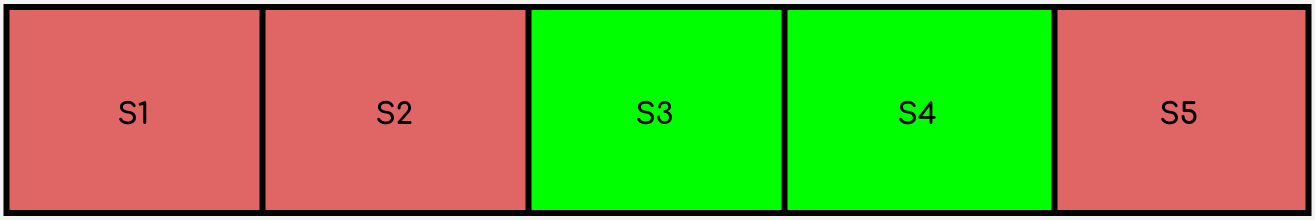
Parameters Involved

- x = 1 - (((2018111024) % 40 + 1) / 100)
  = **0.75**
  Thus the agent moves to the state in the direction of the intended action with probability **0.75** and to the state in opposite to the action with probability **0.25** (or to the same state if at extreme states).
- y = (24) % 3
  = **0**
  Thus the following observation table is chosen as reference -

P (observation = Red | State = Red)	0.9
P (observation = Green | State = Red)	0.1
P (observation = Green | State = Green)	0.85
P (observation = Red | State =Green)	0.15

Table 0

Formulating the POMDP problem



We would need to formally define the problem as a POMDP problem. Accordingly, each PoMDP can be defined as -

POMDP **: < S, A, P, R, Ω, O >**

- **S (set of states) :** { S1, S2, S3, S4, S5 }
- **A (set of actions)** : { Left, Right }
- **P (transition probabilities table) :** Given below is the transition table. Each cell (i,j) under action a represents transition probability P(Sj|Si,a) -

Left	S1	S2	S3	S4	S5
S1	0.75	0.25	0.0	0.0	0.0
S2	0.75	0.0	0.25	0.0	0.0
S3	0.0	0.75	0.0	0.25	0.0
S4	0.0	0.0	0.75	0.0	0.25
S5	0.0	0.0	0.0	0.75	0.25

Right	S1	S2	S3	S4	S5
S1	0.25	0.75	0.0	0.0	0.0
S2	0.25	0.0	0.75	0.0	0.0
S3	0.0	0.25	0.0	0.75	0.0
S4	0.0	0.0	0.25	0.0	0.75
S5	0.0	0.0	0.0	0.25	0.75

- **R (set of rewards) :** Information not provided here.
- **Ω (set of observations) :** { red, green }
- **O (table of observation probabilities) :** Table 0 is the observation table. Each cell (i,o) under represents observation probability O(o|Si)

1. Action 1

Initial belief state **B** = { 0.33333, 0.33333, 0.0, 0.0, 0.33333 }. Agent took the action **Right** and observed **Red.**

We have the formula,

$$b(s') = \alpha P(e \mid s') \sum P(s'|s,a)\, b(s)$$

α is the normalizing constant and e is the evidence we have. We shall compute the numerator for each one of the states -

1. S1

b(S1) = α P(Red | S1) (0.333 x P(S1|S1,right) + 0.333 x P(S1|S2,right) + 0.0 x P(S1|S3,right) + 0.0x P(S1|S4,right) + 0.333 x P(S1|S5,right))
= α (0.9) (0.16665)
= α x 0.14999

2. **S2**
b(S2) = α P(Red | S2) (0.333 x P(S2|S1,right) + 0.333 x P(S2|S2,right) + 0.0 x P(S2|S3,right) + 0.0x P(S2|S4,right) + 0.333 x P(S2|S5,right))
= α (0.9) (0.24999)
= α x 0.22499
3. **S3**
b(S3) = α P(Red | S3) (0.333 x P(S3|S1,right) + 0.333 x P(S3|S2,right) + 0.0 x P(S3|S3,right) + 0.0x P(S3|S4,right) + 0.333 x P(S3|S5,right))
= α (0.15) (0.24999)
= α x 0.03749
4. **S4**
b(S4) = α P(Red | S4) (0.333 x P(S4|S1,right) + 0.333 x P(S4|S2,right) + 0.0 x P(S4|S3,right) + 0.0x P(S4|S4,right) + 0.333 x P(S4|S5,right))
= α (0.15) (0.08333)
= α x 0.01249
5. **S5**
b(S5) = α P(Red | S5) (0.333 x P(S5|S1,right) + 0.333 x P(S5|S2,right) + 0.0 x P(S5|S3,right) + 0.0x P(S5|S4,right) + 0.333 x P(S5|S5,right))
= α (0.9) (0.24999)
= α x 0.22499

From this we get
α = 1 / (0.64935)
= 1.53857
Thus we get beliefs states as
b(s) = { 0.23077,0.34616,0.05768,0.01923,0.34616 }

2. Action 2

Agent took the action left and observed Green.

1. **S1**
b(S1) = α P(Green | S1) (0.23076 x P(S1|S1,left) + 0.34615 x P(S1|S2,left) + 0.0577 x P(S1|S3,left) + 0.01923 x P(S1|S4,left) + 0.34615 x P(S1|S5,left))
= α (0.1) (0.43269)
= α x 0.04327

2. **S2**
b(S2) = α P(Green | S2) (0.23076 x P(S2|S1,left) + 0.34615 x P(S2|S2,left) + 0.0577 x P(S2|S3,left) + 0.01923 x P(S2|S4,left) + 0.34615 x P(S2|S5,left))

= α (0.1) (0.10095)
= α x 0.01009

3. **S3**

b(S3) = α P(Green | S2) (0.23076 x P(32|S1,left) + 0.34615 x P(S3|S2,left) + 0.0577 x P(S3|S3,left) + 0.01923 x P(S3|S4,left) + 0.34615 x P(S3|S5,left))
= α (0.85) (0.10096)
= α x 0.08581

4. **S4**

b(S4) = α P(Green | S2) (0.23076 x P(S4|S1,left) + 0.34615 x P(S4|S2,left) + 0.0577 x P(S4|S3,left) + 0.01923 x P(S4|S4,left) + 0.34615 x P(S4|S5,left))
= α (0.85) (0.27404)
= α x 0.23293

5. **S5**

b(S5) = α P(Green | S2) (0.23076 x P(S5|S1,left) + 0.34615 x P(S5|S2,left) + 0.0577 x P(S5|S3,left) + 0.01923 x P(S5|S4,left) + 0.34615 x P(S5|S5,left))
= α (0.1) (0.091347)
= α x 0.00913

From this we get

α = 1 / (0.38123)
= 2.62309

The new belief states as

b(s) = {0.11350 , 0.02646, 0.22508, 0.61099, 0.02394}

3. Action 3

Agent took the action Left and observed Green.

1. **S1**

b(S1) = α P(Green | S1) (0.11350 x P(S1|S1,left) + 0.02646 x P(S1|S2,left) + 0.22508 x P(S1|S3,left) + 0.61099 x P(S1|S4,left) + 0.02394 x P(S1|S5,left))
= α (0.1) (0.10497)
= α x 0.01049

2. **S2**

b(S2) = α P(Green | S1) (0.11350 x P(S2|S1,left) + 0.02646 x P(S2|S2,left) + 0.22508 x P(S2|S3,left) + 0.61099 x P(S2|S4,left) + 0.02394 x P(S2|S5,left))
= α (0.1) (0.19718)
= α x 0.01971

3. **S3**

b(S3) = α P(Green | S1) (0.11350 x P(S3|S1,left) + 0.02646 x P(S3|S2,left) + 0.22508 x P(S3|S3,left) + 0.61099 x P(S3|S4,left) + 0.02394 x P(S3|S5,left))
= α (0.85) (0.46485)
= α x 0.39512

4. **S4**

b(S4) = α P(Green | S1) (0.11350 x P(S4|S1,left) + 0.02646 x P(S4|S2,left) + 0.22508 x P(S4|S3,left) + 0.61099 x P(S4|S4,left) + 0.02394 x P(S4|S5,left))

= α (0.85) (0.07422)

= α x 0.06309

5. **S5**

b(S5) = α P(Green | S1) (0.11350 x P(S5|S1,left) + 0.02646 x P(S5|S2,left) + 0.22508 x P(S5|S3,left) + 0.61099 x P(S5|S4,left) + 0.02394 x P(S5|S5,left))

= α (0.1) (0.15873)

= α x 0.01587

Thus, we get

α = 1 / (0.50428)

= 1.98302

Thus the new belief state is

b(s) = { 0.020801, 0.03908, 0.78353, 0.12510, 0.03147 }