MDL Assignment 5

Kalp Shah : 2018113003

Part 1

Algorithm

Definitions

The action update is the update that occurs when an action takes place. For example, if the agent wants to take left, then the action is left. Another concept is observation, that is the observed result after an action takes place. For example, if the agent reports that it sees Red than the observation is color being red.

These are the two things that can vhange the belief state of a POMDP, and it is very clear why.

Formulation

POMDP algorithm is used, which is a MDP with belief (probability) states b, which is a continuous distribution over the states S of the MDP.

The updates are as follows :

- Action Update of belief system:

$$b'(s) = \Sigma p(s|a, s')p(s')$$

  but b(s) = p(s), hence :

$$b'(s) = \Sigma p(s|a, s')b(s')$$

- Observation Update of belief system:

$$b'(s) = p(s|o, b)\ b'(s) = \frac{p(o|s)b(s)}{\Sigma p(o,s')b(s')}$$

These equations can be then combined to form :

$$b(s') = p(s|o, a, b)$$
$$\implies b(s') = \frac{p(s|a,b)p(o|s,a,b)}{p(o|b)}$$

The final equation is the one that will be used to solve the problem.

Solution

Global Values

Roll No : 2018113003

$\implies$ x = 1 - (3%40 + 1)/100
= .96

$\implies$ y = (3) % 3
= 0

$\therefore$ Table 0 is to be taken

p(O=Red—S=Red)	0.9
p(O=Green—S=Green)	0.85

Initialization

Initially, the agent can be at any one of the three Red states, hence the initial beliefs are :

b(S_1)	1/3
b(S_2)	1/3
b(S_3)	0
b(S_4)	0
b(S_2)	1/3

Actions and Observations

Action = Right & Observation = Red

First calculating belief after **Action = Right**
Formulae : $b(s') = \Sigma b(s')p(s|a, s')$

$b(S_1)$ = 0.333*0.04 + 0.333*0.04
= 0.027

$b(S_2)$ = 0.333*0.96 + 0.000*0.04
= 0.320

$b(S_3)$ = 0.333*0.96 + 0.000*0.04
= 0.320

$b(S_4)$ = 0.000*0.96 + 0.333*0.04
= 0.013

$b(S_5) = 0.000*0.96 + 0.333*0.96$
$= 0.320$

Now, belief after **observation = Red**
Formulae = $b(s') = \frac{p(o|s)*b(s)}{\Sigma b(s')p(o|s')}$
$= \frac{p(o|s)*b(s)}{0.650}$

$b'(S_1) = \frac{0.027*0.900}{0.650} = 0.037$

$b'(S_2) = \frac{0.320*0.900}{0.650} = 0.443$

$b'(S_3) = \frac{0.320*0.150}{0.650} = 0.074$

$b'(S_4) = \frac{0.013*0.150}{0.650} = 0.003$

$b'(S_5) = \frac{0.320*0.900}{0.650} = 0.443$

Action = Left & Observation = Green

First calculating belief after **Action = Left**
Formulae : $b(s') = \Sigma b(s')p(s|a, s')$

$b(S_1) = 0.037*0.96 + 0.443*0.96$
$= 0.461$

$b(S_2) = 0.037*0.04 + 0.074*0.96$
$= 0.072$

$b(S_3) = 0.443*0.04 + 0.003*0.96$
$= 0.021$

$b(S_4) = 0.074*0.04 + 0.443*0.96$
$= 0.428$

$b(S_5) = 0.003*0.04 + 0.443*0.04$
$= 0.018$

Now, belief after **observation = Green**
Formulae = $b(s') = \frac{p(o|s)*b(s)}{\Sigma b(s')p(o|s')}$
$= \frac{p(o|s)*b(s)}{0.437}$

$b'(S_1) = \frac{0.461*0.100}{0.437} = 0.106$

$b'(S_2) = \frac{0.072*0.100}{0.437} = 0.017$

$b'(S_3) = \frac{0.021*0.850}{0.437} = 0.040$

$b'(S_4) = \frac{0.428*0.850}{0.437} = 0.834$

$b'(S_5) = \frac{0.018*0.100}{0.437} = 0.004$

Action = Left & Observation = Green

First calculating belief after **Action = Left**
Formulae : $b(s') = \Sigma b(s')p(s|a, s')$

$b(S_1) = 0.106*0.96 + 0.017*0.96$
$= 0.117$

$b(S_2) = 0.106*0.04 + 0.040*0.96$
$= 0.043$

$b(S_3) = 0.017*0.04 + 0.834*0.96$
$= 0.801$

$b(S_4) = 0.040*0.04 + 0.004*0.96$
$= 0.006$

$b(S_5) = 0.834*0.04 + 0.004*0.04$
$= 0.034$

Now, belief after **observation = Green**
Formulae = $b(s') = \frac{p(o|s)*b(s)}{\Sigma b(s')p(o|s')}$
$= \frac{p(o|s)*b(s)}{0.705}$

$b'(S_1) = \frac{0.117*0.100}{0.705} = 0.017$

$b'(S_2) = \frac{0.043*0.100}{0.705} = 0.006$

$b'(S_3) = \frac{0.801*0.850}{0.705} = 0.966$

$b'(S_4) = \frac{0.006*0.850}{0.705} = 0.007$

$b'(S_5) = \frac{0.034*0.100}{0.705} = 0.005$

Result

Hence, after these calculations, the final result is :

Belief	Update 1	Update 2	Update 3
b(S_1)	0.037	0.106	0.017
b(S_2)	0.443	0.017	0.006
b(S_3)	0.074	0.040	0.966
b(S_4)	0.003	0.834	0.004
b(S_2)	0.443	0.004	0.005