MDL

Assignment-3 Part-1 REPORT

Team members:

Name: Gokul Vamsi Thota (Roll no: 2019111009)

Name: T.H. Arjun (Roll no: 2019111012)

We know that, POMDP models depict are generalizations of MDP models, with a sensor-based approach.

A POMDP can be represented by 7-tuple $(S, A, T, R, \Omega, O, \gamma)$. In our case, the entries of the tuple are explained as follows:

→ Set of states, $S = \{S_1, S_2, S_3, S_4, S_5, S_6\}$

→ Set of actions, $A = \{\text{Left}, \text{Right}\}$

→ Transition probabilities, 'T', can be given by the transition table below:

** Note that, for our case,

Roll number considered = 2019111009

Last 4 digits = 1009, Last 2 digits : 09

$\boxed{x} = 1 - \left(\frac{19+1}{100}\right) = 1 - \left(\frac{20}{100}\right) = \boxed{0.8}$

$\boxed{y} = (09)\% 4 + 1 = 1+1 = \boxed{2}$

Scanned with CamScanner

Current state	Action	Next State	Transition probability
S_1	Left	S_1	0.8
S_1	Right	S_2	0.8
S_1	Right	S_1	0.2
S_2	Left	S_1	0.8
S_2	Left	S_3	0.2
S_2	Right	S_3	0.8
S_2	Right	S_1	0.2
S_3	Left	S_2	0.8
S_3	Left	S_4	0.2
S_3	Right	S_4	0.8
S_3	Right	S_2	0.2
S_4	Left	S_3	0.8
S_4	Left	S_5	0.2
S_4	Right	S_5	0.8
S_4	Right	S_3	0.2
S_5	Left	S_4	0.8
S_5	Left	S_6	0.2
S_5	Right	S_6	0.8
S_5	Right	S_4	0.2
S_6	Left	S_5	0.8
S_6	Left	S_6	0.2
S_6	Right	S_6	0.8
S_6	Right	S_5	0.2
S_1	Left	S_2	0.2

→ Set of observations, $\Omega = \{Green, Red\}$

→ Set of conditional observation probabilities 'O' are given by the following table

(Table: 2 of pdf as we have $y = 2$)

State	Colour	Red is observed	Green is observed
S_1	Red	0.9	0.1
S_2	Green	0.15	0.85
S_3	Red	0.9	0.1
S_4	Green	0.15	0.85
S_5	Green	0.15	0.85
S_6	Red	0.9	0.1

→ We ignore 'R' and 'r' as they are not relevant for computation of belief state.

→ We know that, belief state at each time step is updated by the formulation:

$$b'(s') = \frac{P(o' \mid s', a_m)\left(\sum_{s \in S} P(s' \mid a_m, s)\, b(s)\right)}{P(o' \mid a_m, b)}$$

where 'b', is current belief state, a_m is action taken, $s, s' \in S$

→ Initial belief state:

Initially, agent knows that it is in one of red states : S_1, S_3, S_6. Hence, initial belief state of the POMDP would be given

$$\boxed{\begin{aligned} b(S_1) &= \tfrac{1}{3} \\ b(S_2) &= 0 \\ b(S_3) &= \tfrac{1}{3} \\ b(S_4) &= 0 \\ b(S_5) &= 0 \\ b(S_6) &= \tfrac{1}{3} \end{aligned}}$$

→ Thus, based on all above information, we can start with the computations :

Scanned with CamScanner

* For a state, lets define,

$$\boxed{b''(s) = \sum_{s' \in S} P(s|a,s')\, b(s')\, P(o|s)}$$

Here, 'o' is the observation observed on taking action 'a' from state s' to reach 's'.

$P(s|a,s')$ → Transition probability of reaching state 's' from s' on taking action 'a'.

$P(o|s)$ → probability that 'o' is observed on reaching 's'.

$b(s')$ → probability that agent was initially in state s'.

Lets define normalization factor. $N = \left(\dfrac{1}{\sum_{s \in S} b''(s)} \right)$

Hence, current belief state for a state s', can now be given as

$$\boxed{b'(s) = b''(s) \cdot N}$$

Lets use this approach in all computation below.

① Agent took 'Right' and observed 'green':

$$b''(s) = \sum_{s' \in S} P(s|a,s')\, b(s')\, P(o|s)$$

$$= \sum_{s' \in S} b(s')\, P(s|a,s')\, P(o|s)$$

$$\therefore b''(S_1) = P(S_1|\text{Right}, S_1)\, b(S_1)\, P(\text{Green}|S_1) + P(S_1|\text{Right}, S_2)\, b(S_2)\, P(\text{Green}|S_1)$$

$$= (0.2)\left(\tfrac{1}{3}\right)(0.1) + (0.8)(0) = \frac{1}{150}$$

$$b''(S_2) = P(S_2|\text{Right}, S_1)\, b(S_1)\, P(\text{Green}|S_2) + P(S_2|\text{Right}, S_3)\, b(S_3)\, P(\text{Green}|S_2)$$

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

$$= \frac{85}{300}$$

$$b''(S_3) = P(S_3|\text{Right}, S_2)\, b(S_2)\, P(\text{Green}|S_3) + P(S_3|\text{Right}, S_4)\, b(S_4)\, P(\text{Green}|S_3)$$

$$= 0$$

$$b''(s_4) = P(s_4 \mid Right, s_3)\, b(s_3)\, P(Green \mid s_4) + P(s_4 \mid Right, s_5)\, b(s_5)\, P(Green \mid s_4)$$

$$= (0.8)\left(\frac{1}{3}\right)(0.85) + 0 = \frac{68}{300}$$

$$b''(s_5) = P(s_5 \mid Right, s_4)\, b(s_4)\, P(Green \mid s_5) + P(s_5 \mid Right, s_6)\, b(s_6)\, P(Green \mid s_5)$$

$$= 0 + (0.2)\left(\frac{1}{3}\right)(0.85) = \frac{17}{300}$$

$$b''(s_6) = P(s_6 \mid Right, s_5)\, b(s_5)\, P(Green \mid s_6) + P(s_6 \mid Right, s_6)\, b(s_6)\, P(Green \mid s_6)$$

$$= (0.8)(0) + (0.8)\left(\frac{1}{3}\right)(0.1) = \frac{8}{300}$$

$$\therefore \sum_{s \in S} b''(s) = \frac{2}{300} + \frac{85}{300} + 0 + \frac{68}{300} + \frac{17}{300} + \frac{8}{300}$$

$$= \frac{180}{300} = 0.6$$

$$\therefore N = \frac{1}{\sum_{s \in S} b''(s)} = \frac{300}{180}$$

$$\therefore b'(s_1) = \left(\frac{1}{150}\right)\left(\frac{300}{180}\right) = \frac{1}{90}$$

$$b'(s_2) = \left(\frac{85}{300}\right)\left(\frac{300}{180}\right) = \frac{85}{180} = \frac{17}{36}$$

$b'(s_3) = 0$

$b'(s_4) = \left(\frac{68}{300}\right)\left(\frac{300}{180}\right) = \frac{17}{45}$

$b'(s_5) = \left(\frac{17}{300}\right)\left(\frac{300}{180}\right) = \frac{17}{180}$

$b'(s_6) = \left(\frac{8}{300}\right)\left(\frac{300}{180}\right) = \frac{2}{45}$

$\therefore$ Belief states after the given action are

$$\boxed{\begin{aligned} b(S_1) &\approx 0.011111 &= \tfrac{1}{90} \\ b(S_2) &\approx 0.472222 &= \tfrac{17}{36} \\ b(S_3) &= 0 &= 0 \\ b(S_4) &\approx 0.377778 &= \tfrac{17}{45} \\ b(S_5) &\approx 0.094444 &= \tfrac{17}{180} \\ b(S_6) &\approx 0.044444 &= \tfrac{2}{45} \end{aligned}}$$

(2) Agent took action 'Left' and observed 'Red':

$b''(s_1) = P(S_1 | Left, S_1)\, b(S_1)\, P(Red | S_1) +$

$P(S_1 | Left, S_2)\, b(S_2)\, P(Red | S_1)$

$= (0.8)\left(\frac{1}{90}\right)(0.9) + (0.8)\left(\frac{17}{36}\right)(0.9)$

Scanned with CamScanner

$$\Rightarrow b''(S_1) = \frac{8}{1000} + \frac{340}{1000} = \frac{348}{1000}$$

$$b''(S_2) = P(S_2 \mid \text{left}, S_3)\, b(S_3)\, P(\text{Red} \mid S_2) + P(S_2 \mid \text{left}, S_1)\, b(S_1)\, P(\text{Red} \mid S_2)$$

$$= (0.8)(0) + (0.2)\left(\frac{1}{90}\right)(0.15) = \frac{1}{3000}$$

$$b''(S_3) = P(S_3 \mid \text{left}, S_4)\, b(S_4)\, P(\text{Red} \mid S_3) + P(S_3 \mid \text{left}, S_2)\, b(S_2)\, P(\text{Red} \mid S_3)$$

$$= (0.8)\left(\frac{17}{45}\right)(0.9) + (0.2)\left(\frac{17}{36}\right)(0.9)$$

$$= \frac{272}{1000} + \frac{85}{1000} = \frac{357}{1000}$$

$$b''(S_4) = P(S_4 \mid \text{left}, S_5)\, b(S_5)\, P(\text{Red} \mid S_4) + P(S_4 \mid \text{left}, S_3)\, b(S_3)\, P(\text{Red} \mid S_4)$$

$$= (0.8)\left(\frac{17}{180}\right)(0.15) + (0.2)(0)$$

$$= \frac{17}{1500}$$

$$b''(S_5) = P(S_5 \mid \text{left}, S_6)\, b(S_6)\, P(\text{Red} \mid S_5) + P(S_5 \mid \text{left}, S_4)\, b(S_4)\, P(\text{Red} \mid S_5)$$

$$= (0.8)\left(\frac{2}{45}\right)(0.15) + (0.2)\left(\frac{17}{45}\right)(0.15)$$

Scanned with CamScanner

$$b''(s_5) = \frac{8}{1500} + \frac{17}{1500} = \frac{25}{1500}$$

$$b''(s_6) = P(s_6 \mid left, s_6)\, b(s_6)\, P(Red \mid s_6) + P(s_6 \mid left, s_5)\, b(s_5)\, P(Red \mid s_6)$$

$$= (0.2)\left(\frac{2}{45}\right)(0.9) + (0.2)\left(\frac{17}{180}\right)(0.9)$$

$$= \frac{25}{1000}$$

$$\therefore \sum_{s \in S} b''(s) = \frac{348}{1000} + \frac{1}{3000} + \frac{357}{1000} + \frac{17}{1500} + \frac{25}{1500} + \frac{25}{1000}$$

$$= \frac{2275}{3000} \simeq 0.758333$$

$$\therefore N = \frac{1}{\left(\frac{2275}{3000}\right)} = \frac{3000}{2275}$$

$$b'(s_1) = \left(\frac{348}{1000}\right)\left(\frac{3000}{2275}\right) = \frac{1044}{2275}$$

$$b'(s_2) = \left(\frac{1}{3000}\right)\left(\frac{3000}{2275}\right) = \frac{1}{2275}$$

$$b'(s_3) = \left(\frac{357}{1000}\right)\left(\frac{3000}{2275}\right) = \frac{1071}{2275}$$

$$b'(s_4) = \left(\frac{17}{1500}\right)\left(\frac{3000}{2275}\right) = \frac{34}{2275}$$

$$b'(s_5) = \left(\frac{25}{1500}\right)\left(\frac{3000}{2275}\right) = \frac{50}{2275}$$

$$b'(s_6) = \left(\frac{25}{1000}\right)\left(\frac{3000}{2275}\right) = \frac{75}{2275}$$

∴ Belief states after the given action are:

$$\boxed{\begin{aligned}
b(s_1) &= \frac{1044}{2275} \approx 0.458901 \\
b(s_2) &= \frac{1}{2275} \approx 0.000439 \\
b(s_3) &= \frac{1071}{2275} \approx 0.470769 \\
b(s_4) &= \frac{34}{2275} \approx 0.014945 \\
b(s_5) &= \frac{50}{2275} \approx 0.021978 \\
b(s_6) &= \frac{75}{2275} \approx 0.032967
\end{aligned}}$$

③ Agent took action 'left' and observed 'green':

$$b''(s_1) = P(s_1 \mid \text{left}, s_1)\, b(s_1)\, P(\text{Green} \mid s_1) + P(s_1 \mid \text{left}, s_2)\, b(s_2)\, P(\text{Green} \mid s_1)$$

$$\cancel{= (0.8)\left(\frac{1}{90}\right)(0.1) + (0.8)\left(\frac{17}{36}\right)(0.1)}$$

$$\cancel{= \frac{8}{9000} + \frac{340}{9000} = \frac{348}{9000}}$$

$$\cancel{b''(s_2}$$

$$b''(s_1) = (0.8)\left(\frac{1044}{2275}\right)(0.1) + (0.8)\left(\frac{1}{2275}\right)(0.1)$$

$$= \frac{8352}{227500} + \frac{8}{227500} = \frac{8360}{227500}$$

$$b''(S_2) = P(S_2 \mid left, S_3)\, b(S_3)\, P(Green \mid S_2) + P(S_2 \mid left, S_1)\, b(S_1)\, P(Green \mid S_2)$$

$$= (0.8)\left(\frac{1071}{2275}\right)(0.85) + (0.2)\left(\frac{1044}{2275}\right)(0.85)$$

$$= \frac{72828}{227500} + \frac{17748}{227500} = \frac{90576}{227500}$$

$$b''(S_3) = P(S_3 \mid left, S_4)\, b(S_4)\, P(Green \mid S_3) + P(S_3 \mid left, S_2)\, b(S_2)\, P(Green \mid S_3)$$

$$= (0.8)\left(\frac{34}{2275}\right)(0.1) + (0.2)\left(\frac{1}{2275}\right)(0.1)$$

$$= \frac{272}{227500} + \frac{2}{227500} = \frac{274}{227500}$$

$$b''(S_4) = P(S_4 \mid left, S_5)\, b(S_5)\, P(Green \mid S_4) + P(S_4 \mid left, S_3)\, b(S_3)\, P(Green \mid S_4)$$

$$= (0.8)\left(\frac{50}{2275}\right)(0.85) + (0.2)\left(\frac{1071}{2275}\right)(0.85)$$

$$= \frac{3400}{227500} + \frac{18207}{227500} = \frac{21607}{227500}$$

$$b''(S_5) = P(S_5 \mid left, S_6)\, b(S_6)\, P(Green \mid S_5) + P(S_5 \mid left, S_4)\, b(S_4)\, P(Green \mid S_5)$$

$$= (0.8)\left(\frac{75}{2275}\right)(0.85) + (0.2)\left(\frac{34}{2275}\right)(0.85)$$

$$= \frac{5100}{227500} + \frac{578}{227500} = \frac{5678}{227500}$$

$$b''(S_6) = P(S_6 \mid \text{left}, S_6)\, b(S_6)\, P(\text{Green} \mid S_6) + P(S_6 \mid \text{left}, S_5)\, b(S_5)\, P(\text{Green} \mid S_6)$$

$$= (0.2)\left(\frac{75}{2275}\right)(0.1) + (0.2)\left(\frac{50}{2275}\right)(0.1)$$

$$= \frac{150}{227500} + \frac{100}{227500} = \frac{250}{227500}$$

$$\therefore \sum_{s \in S} b''(s) = \frac{8360}{227500} + \frac{90576}{227500} + \frac{274}{227500} + \frac{21607}{227500} + \frac{5678}{227500} + \frac{250}{227500}$$

$$= \frac{126745}{227500} \simeq 0.557120$$

$$\therefore \eta = \frac{1}{\sum_{s \in S} b''(s)} = \frac{227500}{126745}$$

$$\therefore b'(s_1) = \left(\frac{8360}{227500}\right)\left(\frac{227500}{126745}\right) = \frac{8360}{126745}$$

$$b'(s_2) = \left(\frac{90576}{227500}\right)\left(\frac{227500}{126745}\right) = \frac{90576}{126745}$$

$$b'(s_3) = \left(\frac{274}{227500}\right)\left(\frac{227500}{126745}\right) = \frac{274}{126745}$$

$$b'(s_4) = \left(\frac{21607}{227500}\right)\left(\frac{227500}{126745}\right) = \frac{21607}{126745}$$

$$b'(s_5) = \left(\frac{5678}{227500}\right)\left(\frac{227500}{126745}\right) = \frac{5678}{126745}$$

$$b'(s_6) = \left(\frac{250}{227500}\right)\left(\frac{227500}{126745}\right) = \frac{250}{126745}$$

∴ Belief states after the given action are:

$b(s_1) = \frac{8360}{126745} \simeq 0.065959$

$b(s_2) = \frac{90576}{126745} \simeq 0.714631$

$b(s_3) = \frac{274}{126745} \simeq 0.002161$

$b(s_4) = \frac{21607}{126745} \simeq 0.170476$

$b(s_5) = \frac{5678}{126745} \simeq 0.044798$

$b(s_6) = \frac{250}{126745} \simeq 0.001972$

Hence, based on above computations, the belief states after each action can be represented in the following table (each probability approximated to 4 decimal places)

After Action	$b(s_1)$	$b(s_2)$	$b(s_3)$	$b(s_4)$	$b(s_5)$	$b(s_6)$
1	0.0111	0.4722	0	0.3778	0.0944	0.0444
2	0.4589	0.0004	0.4708	0.0149	0.0220	0.0330
3	0.0660	0.7146	0.0022	0.1705	0.0448	0.0020