**MDL** by Alapan Sau Apr 16, 2021

# MDL Assignment 3 | Part 1

The roll number used for the calculations is 2019101033.
The values of $x$ and $y$ are as follows:

$$x = 1 - (((1033\%30) + 1)/100) = 1 - 0.14 = 0.86$$

$$y = (33)\%4 + 1 = 2$$

The probability values based on x and y are calculated as follows:

$$P(success) = 1 - x = 0.86$$

$$P(failure) = 1 - 0.86 = 0.14$$

In the case of failure, the agent moves in direction opposite to the original intended direction

From the value of $y$, the observation probabilities are as follows:

$$P(Observation = Red \mid State = Red) = 0.9$$

$$P(Observation = Green \mid State = Green) = 0.85$$

From these values, the follwoing can be calculated:

$$P(Observation = Red \mid State = Green) = 1 - 0.85 = 0.15$$

$$P(Observation = Green \mid State = Red) = 1 - 0.9 = 0.1$$

We know that the agent is initially present in one of the red states ($S1$, $S3$ or $S6$). The inital beliefs are as follows:

$$b(S1) = 0.3333$$

$$b(S2) = 0.0000$$

$$b(S3) = 0.3333$$

$$b(S4) = 0.0000$$

$$b(S5) = 0.0000$$

$$b(S6) = 0.3333$$

## Step 1:

***Agent took the action Right and observed Green.***

We calculate the new belief states as follows: $b'(s') = \alpha * Pr(o|s') * \sum Pr(s'|s, action) * b(s)$

$$\begin{aligned} b'(S_1) &= \alpha * Pr(O_G|S_1) * \sum Pr(S_1|s, right) * b(s) \\ &= \alpha * 0.1 * (0.14 * 0.33333 + 0.14 * 0 + 0 * 0.3333 + 0 * 0 + 0 * 0 + 0 * 0.3333) = 0.04666667\alpha \end{aligned}$$

$$\begin{aligned} b'(S_2) &= \alpha * Pr(O_G|S_2) * \sum Pr(S_2|s, right) * b(s) \\ &= \alpha * 0.85 * (0.86 * 0.33333 + 0 * 0 + 0.14 * 0.3333 + 0 * 0 + 0 * 0 + 0 * 0.3333) = 0.28333333\alpha \end{aligned}$$

$$\begin{aligned} b'(S_3) &= \alpha * Pr(O_G|S_3) * \sum Pr(S_3|s, right) * b(s) \\ &= \alpha * 0.1 * (0 * 0.33333 + 0.86 * 0 + 0 * 0.3333 + 0.14 * 0 + 0 * 0 + 0 * 0.3333) = 0 \end{aligned}$$

$$\begin{aligned} b'(S_4) &= \alpha * Pr(O_G|S_4) * \sum Pr(S_4|s, right) * b(s) \\ &= \alpha * 0.85 * (0 * 0.33333 + 0 * 0 + 0.86 * 0.3333 + 0 * 0 + 0.14 * 0 + 0 * 0.3333) = 0.243666667\alpha \end{aligned}$$

$$\begin{aligned} b'(S_5) &= \alpha * Pr(O_G|S_5) * \sum Pr(S_5|s, right) * b(s) \\ &= \alpha * 0.85 * (0 * 0.33333 + 0 * 0 + 0 * 0.3333 + 0.86 * 0 + 0 * 0 + 0.14 * 0.3333) = 0.0396666667\alpha \end{aligned}$$

$$\begin{aligned} b'(S_6) &= \alpha * Pr(O_G|S_6) * \sum Pr(S_6|s, right) * b(s) \\ &= \alpha * 0.1 * (0 * 0.33333 + 0 * 0 + 0 * 0.3333 + 0 * 0 + 0.86 * 0 + 0.86 * 0.3333) = 0.0286666667\alpha \end{aligned}$$

The value of $\alpha$ can be calculated as follows:
$b'(S1) + b'(S2) + b'(S3) + b'(S3) + b'(S4) + b'(S5) + b'(S6) = 1$
$0.04666667\alpha + 0.28333333\alpha + 0 + 0.243666667\alpha + 0.0396666667\alpha + 0.0286666667\alpha = 1$
$\alpha = 1/0.5999999999999999 = 1.666666666666667$

Substituting the value of $\alpha$, the values of the beliefs are as follows:

$b'(S1) = 0.04666667\alpha = 0.04666667 * 1.666666666666667 = 0.0078$
$b'(S2) = 0.28333333\alpha = 0.28333333 * 1.666666666666667 = 0.4722$
$b'(S3) = 0$
$b'(S4) = 0.243666667\alpha = 0.243666667 * 1.666666666666667 = 0.4061$
$b'(S5) = 0.0396666667\alpha = 0.0396666667 * 1.666666666666667 = 0.0661$
$b'(S6) = 0.02866666666\alpha = 0.02866666666 * 1.666666666666667 = 0.0478$

## Step 2:

***Agent took the action Left and observed Red.***

We calculate the new belief states as follows: $b'(s') = \alpha * Pr(o|s') * \sum Pr(s'|s, action) * b(s)$

$$\begin{aligned} b'(S_1) &= \alpha * Pr(O_R|S_1) * \sum Pr(S_1|s, left) * b(s) \\ &= \alpha * 0.9 * (0.86 * 0.0078 + 0.86 * 0.4722 + 0 * 0 + 0 * 0.4061 + 0 * 0 + 0 * 0.0478) = 0.3715200\alpha \end{aligned}$$

$$\begin{aligned} b'(S_2) &= \alpha * Pr(O_R|S_2) * \sum Pr(S_2|s, left) * b(s) \\ &= \alpha * 0.15 * (0.14 * 0.0078 + 0 * 0.4722 + 0.86 * 0 + 0 * 0.4061 + 0 * 0 + 0 * 0.0478) = 0.00016333333\alpha \end{aligned}$$

$$\begin{aligned} b'(S_3) &= \alpha * Pr(O_R|S_3) * \sum Pr(S_3|s, left) * b(s) \\ &= \alpha * 0.9 * (0 * 0.0078 + 0.14 * 0.4722 + 0 * 0 + 0.86 * 0 + 0 * 0.4061 + 0 * 0.0478) = 0.37383 \end{aligned}$$

$$\begin{aligned} b'(S_4) &= \alpha * Pr(O_R|S_4) * \sum Pr(S_4|s, left) * b(s) \\ &= \alpha * 0.15 * (0 * 0.0078 + 0 * 0.4722 + 0.14 * 0 + 0 * 0 + 0.86 * 0.4061 + 0 * 0.0478) = 0.008528333\alpha \end{aligned}$$

$$\begin{aligned} b'(S_5) &= \alpha * Pr(O_R|S_5) * \sum Pr(S_5|s, left) * b(s) \\ &= \alpha * 0.15 * (0 * 0.0078 + 0 * 0.4722 + 0 * 0 + 0.14 * 0 + 0 * 0.4061 + 0.86 * 0.0478) = 0.01435000000\alpha \end{aligned}$$

$$\begin{aligned} b'(S_6) &= \alpha * Pr(O_R|S_6) * \sum Pr(S_6|s, left) * b(s) \\ &= \alpha * 0.9 * (0 * 0.0078 + 0 * 0.4722 + 0 * 0 + 0 * 0 + 0.14 * 0.4061 + 0.14 * 0.0478) = 0.0143500000\alpha \end{aligned}$$

The value of $\alpha$ can be calculated as follows:
$b'(S1) + b'(S2) + b'(S3) + b'(S3) + b'(S4) + b'(S5) + b'(S6) = 1$
$0.3715200\alpha + 0.00016334\alpha + 0.37383\alpha + 0.0085283334\alpha + 0.014691666\alpha + 0.01435000\alpha = 1$
$\alpha = 1/0.7830833333 = 1.27700330$

Substituting the value of $\alpha$, the values of the beliefs are as follows:

$b'(S1) = 0.3715200\alpha = 0.3715200 * 1.27700330 = 0.4744$
$b'(S2) = 0.00016334\alpha = 0.00016334 * 1.27700330 = 0.0002$
$b'(S3) = 0.37383\alpha = 0.37383 * 1.27700330 = 0.4774$

$b'(S4) = 0.0085283334\alpha = 0.0085283334 * 1.27700330 = 0.0109$
$b'(S5) = 0.014691666\alpha = 0.014691666 * 1.27700330 = 0.0188$
$b'(S6) = 0.01435000\alpha = 0.01435000 * 1.27700330 = 0.0183$

## Step 3:

***Agent took the action Left and observed Green.***

We calculate the new belief states as follows: $b'(s') = \alpha * Pr(o|s') * \sum Pr(s'|s, action) * b(s)$

$$b'(S_1) = \alpha * Pr(O_G|S_1) * \sum Pr(S_1|s, left) * b(s)$$
$$= \alpha * 0.1 * (0.86 * 0.4744 + 0.86 * 0.0002 + 0 * 0.4774 + 0 * 0.0109 + 0 * 0.0188 + 0 * 0.0183) = 0.0408191125\alpha$$

$$b'(S_2) = \alpha * Pr(O_G|S_2) * \sum Pr(S_2|s, left) * b(s)$$
$$= \alpha * 0.85 * (0.14 * 0.4744 + 0 * 0.0002 + 0.86 * 0.4774 + 0 * 0.0109 + 0 * 0.0188 + 0 * 0.0183) = 0.405423786\alpha$$

$$b'(S_3) = \alpha * Pr(O_G|S_3) * \sum Pr(S_3|s, left) * b(s)$$
$$= \alpha * 0.1 * (0 * 0.4744 + 0.14 * 0.0002 + 0 * 0.4774 + 0.86 * 0.0109 + 0 * 0.0188 + 0 * 0.0183) = 0.000939521$$

$$b'(S_4) = \alpha * Pr(O_G|S_4) * \sum Pr(S_4|s, left) * b(s)$$
$$= \alpha * 0.85 * (0 * 0.4744 + 0 * 0.0002 + 0.14 * 0.4774 + 0 * 0.0109 + 0.86 * 0.0188 + 0 * 0.0183) = 0.0705229903\alpha$$

$$b'(S_5) = \alpha * Pr(O_G|S_5) * \sum Pr(S_5|s, left) * b(s)$$
$$= \alpha * 0.85 * (0 * 0.4744 + 0 * 0.0002 + 0 * 0.4774 + 0.14 * 0.0109 + 0 * 0.0188 + 0.86 * 0.0183) = 0.0146915675\alpha$$

$$b'(S_6) = \alpha * Pr(O_G|S_6) * \sum Pr(S_6|s, left) * b(s)$$
$$= \alpha * 0.1 * (0 * 0.0078 + 0 * 0.0002 + 0 * 0.4774 + 0 * 0.0109 + 0.14 * 0.0188 + 0.14 * 0.0183) = 0.000519208\alpha$$

The value of $\alpha$ can be calculated as follows:
$b'(S1) + b'(S2) + b'(S3) + b'(S3) + b'(S4) + b'(S5) + b'(S6) = 1$
$0.040819112\alpha + 0.40542378\alpha + 0.0009395\alpha + 0.0705229\alpha + 0.0146915\alpha + 0.00051920\alpha = 1$
$\alpha = 1/0.532916186 = 1.876467681$

Substituting the value of $\alpha$, the values of the beliefs are as follows:

$b'(S1) = 0.040819112\alpha = 0.040819112 * 1.876467681 = 0.0766$
$b'(S2) = 0.40542378\alpha = 0.40542378 * 1.876467681 = 0.7608$
$b'(S3) = 0.0009395\alpha = 0.0009395 * 1.876467681 = 0.0018$
$b'(S4) = 0.0705229\alpha = 0.0705229 * 1.876467681 = 0.1323$
$b'(S5) = 0.0146915\alpha = 0.0146915 * 1.876467681 = 0.0276$
$b'(S6) = 0.00051920\alpha = 0.00051920 * 1.876467681 = 0.0010$