# MDL Assignment 5 Part 1

Tanvi Karandikar
2018101059

x=1-(59%40+1)/100=1-0.2=0.8
y=59%3=2

P(action is taken successfully)

| Action | Succes | Failure |
|---|---|---|
| Left | 0.8 | 0.2 |
| Right | 0.8 | 0.2 |

P(Observation=o | State=s)

| | o=Red | o=Green |
|---|---|---|
| s=Red | 0.85 | 0.15 |
| s=Green | 0.1 | 0.9 |

Formula Used:

$$\frac{O(s', a, o) \sum T(s, a, s') b(s)}{Pr(o \mid a, b)}$$

Note: in the calculations first the numerator is calculated then it is normalized

Initial Beliefs:
He is in a red state, so all of them must have equal initial belief 1/3

| S1 | S2 | S3 | S4 | S5 |
|---|---|---|---|---|
| ⅓=0.333333 | ⅓=0.333333 | 0 | 0 | ⅓=0.333333 |

**On taking action 1:** Agent took the action **Right** and observed **Red**

Ub'[S1]=0.85[ (0.2 * 0.333333) + (0.2 * 0.333333) + (0 * 0) + (0 * 0) + (0 * 0.333333) ]=0.11333322

Ub'[S2]=0.85[ (0.8 * 0.333333) + (0 * 0.333333) + (0.2 * 0) + (0 * 0) + (0 * 0.333333) ]=0.22666644

Ub'[S3]=0.1[ (0 * 0.333333) + (0.8 * 0.333333) + (0 * 0) + (0.2 * 0) + (0 * 0.333333) ]=0.02666664

Ub'[S4]=0.1[ (0 * 0.333333) + (0 * 0.333333) + (0.8 * 0) + (0 * 0) + (0.2 * 0.333333) ]=0.00666666

Ub'[S5]=0.85[ (0 * 0.333333) + (0 * 0.333333) + (0 * 0) + (0.8 * 0) + (0.8 * 0.333333) ]=0.22666644

Normalizing:
sum of all Ub' =0.5999994

So new belief b'[i]=Ub'[i]/sum of all Ub'
Thus,
b'[S1]=0.188889
b'[S2]=0.377778
b'[S3]=0.044444
b'[S4]=0.011111
b'[S5]=0.377778

Belief States after action 1:

| S1 | S2 | S3 | S4 | S5 |
|---|---|---|---|---|
| 0.188889 | 0.377778 | 0.04444444 | 0.011111111 | 0.377778 |

**On taking action 2:** Agent took the action **Left** and observed **Green**

Ub'[S1]=0.15[ (0.8 * 0.188889) + (0.8 * 0.377778) + (0 * 0.044444) + (0 * 0.011111) + (0 * 0.377778) ]=0.068

Ub'[S2]=0.15[ (0.2 * 0.188889) + (0 * 0.377778) + (0.8 * 0.044444) + (0 * 0.011111) + (0 * 0.377778) ]=0.011

Ub'[S3]=0.9[ (0 * 0.188889) + (0.2 * 0.377778) + (0 * 0.044444) + (0.8 * 0.011111) + (0 * 0.377778) ]=0.076

Ub'[S4]=0.9[ (0 * 0.188889) + (0 * 0.377778) + (0.2 * 0.044444) + (0 * 0.011111) + (0.8 * 0.377778) ]=0.28

Ub'[S5]=0.15[ (0 * 0.188889) + (0 * 0.377778) + (0 * 0.044444) + (0.2 * 0.011111) + (0.2 * 0.377778) ]=0.0116666

Normalizing:
sum of all Ub' =0.4466666

So new belief b'[i]=Ub'[i]/sum of all Ub'

Thus,
b[S1]=0.152239
b[S2]=0.024627
b[S3]=0.170149
b[S4]=0.626866
b[S5]=0.026119

Belief States after action 2:

| S1 | S2 | S3 | S4 | S5 |
|---|---|---|---|---|
| 0.152239 | 0.024627 | 0.170149 | 0.626866 | 0.026119 |

**On taking action 3:** Agent took the action **Left** and observed **Green**

Ub'[S1]=0.15[ (0.8 * 0.152239) + (0.8 * 0.024627) + (0 * 0.170149) + (0 * 0.626866) + (0 * 0.026119) ]=0.021224

Ub'[S2]=0.15[ (0.2 * 0.152239) + (0 * 0.024627) + (0.8 * 0.170149) + (0 * 0.626866) + (0 * 0.026119) ]=0.024985

Ub'[S3]=0.9[ (0 * 0.152239) + (0.2 * 0.024627) + (0 * 0.170149) + (0.8 * 0.626866) + (0 * 0.026119) ]=0.455776

Ub'[S4]=0.9[ (0 * 0.152239) + (0 * 0.024627) + (0.2 * 0.170149) + (0 * 0.626866) + (0.8 * 0.026119) ]=0.049433

Ub'[S5]=0.15[ (0 * 0.152239) + (0 * 0.024627) + (0 * 0.170149) + (0.2 * 0.626866) + (0.2 * 0.026119) ]=0.019589

Normalizing:
sum of all Ub' = 0.571007

So new belief b'[i]=Ub'[i]/sum of all Ub'

Thus,
b[S1]=0.037169
b[S2]=0.043756
b[S3]=0.798196
b[S4=0.086571
b[S5]=0.034307

Thus beliefs at each stage are:

Initial:

| S1 | S2 | S3 | S4 | S5 |
|---|---|---|---|---|
| ⅓=0.333333 | ⅓=0.333333 | 0 | 0 | ⅓=0.333333 |

After action 1:

| S1 | S2 | S3 | S4 | S5 |
|---|---|---|---|---|
| 0.188889 | 0.377778 | 0.04444444 | 0.011111111 | 0.377778 |

After action 2:

| S1 | S2 | S3 | S4 | S5 |
|---|---|---|---|---|
| 0.152239 | 0.024627 | 0.170149 | 0.626866 | 0.026119 |

After action 3:

| S1 | S2 | S3 | S4 | S5 |
|---|---|---|---|---|
| 0.037169 | 0.043756 | 0.798196 | 0.086571 | 0.034307 |