# MDL Assignment 3

POMDP

Part 1


- Dayitva Goel: 2019101005
- Prajneya Kumar: 2019114011


## 1 Overview of POMDP

A partially observable Markov decision process (POMDP) is a generalization of a Markov decision process (MDP). A POMDP models an agent decision process in which it is assumed that the system dynamics are determined by an MDP, but the agent cannot directly observe the underlying state. Instead, it must maintain a probability distribution over the set of possible states, based on a set of observations and observation probabilities, and the underlying MDP.

If b(s) was the previous belief state, and the agent does action a and then perceives evidence e, then the new belief state is given by

$$b'(s') = \alpha P(e|s') \sum_{s} P(s'|s,a)b(s),$$

where $\alpha$ is a normalizing constant that makes the belief state sum to 1.

## 2 Calculation

The roll number we have used is 2019114011.

$$x = 1 - ((21+1)/100)) = 1 - \frac{22}{100} = 1 - 0.22 = 0.78$$
$$y = 11\%4 + 1 = 3 + 1 = 4$$

| **State** | **Color** | **Red Observed** | **Green Observed** |
|---|---|---|---|
| S1 | Red | 0.8 | 0.2 |
| S2 | Green | 0.05 | 0.95 |
| S3 | Red | 0.8 | 0.2 |
| S4 | Green | 0.05 | 0.95 |
| S5 | Green | 0.05 | 0.95 |
| S6 | Red | 0.8 | 0.2 |

Initial Belief States:

| **State** | **Probability** |
|---|---|
| S1 | 1/3 |
| S2 | 0 |
| S3 | 1/3 |
| S4 | 0 |
| S5 | 0 |
| S6 | 1/3 |

# 3 Belief States for Action-Observation Pairs

These are the steps we have used for each action-observation pair.

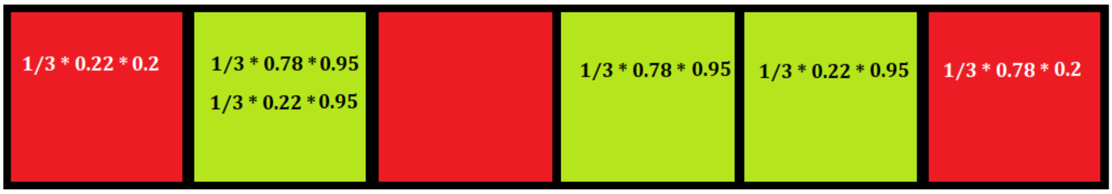
Raw probabilities are calculating using:

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

We then normalize the probabilities by multiplying with the normalizing constant which is the sum of all raw probabilities:

$$\alpha(\text{Normalizing constant}) = 1 / \sum_{s \epsilon S} b(s),$$

## Agent took the action Right and observed Green




| State | Probability |
|---|---|
| b($S_1$) | 11/750 |
| b($S_2$) | 19/60 |
| b($S_3$) | 0/1 |
| b($S_4$) | 247/1000 |
| b($S_5$) | 209/3000 |
| b($S_6$) | 13/250 |

### Probabilities after Normalization

b($S_1$) $= \frac{11}{750} * \frac{1}{0.7} = 0.0209$

b($S_2$) $= \frac{19}{60} * \frac{1}{0.7} = 0.4523$

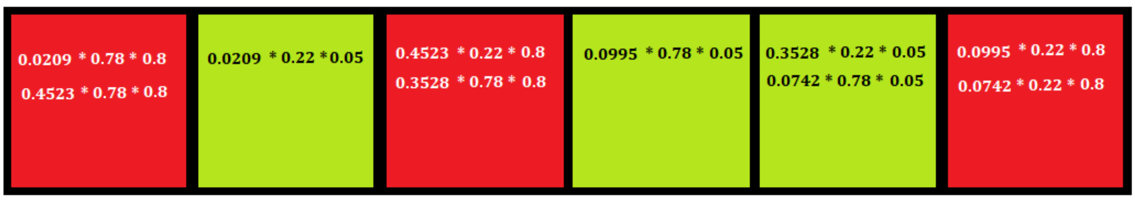
b($S_3$) $= \frac{0}{1} * \frac{1}{0.7} = 0.0$

b($S_4$) $= \frac{247}{1000} * \frac{1}{0.7} = 0.3528$

b($S_5$) $= \frac{209}{3000} * \frac{1}{0.7} = 0.0995$

b($S_6$) $= \frac{13}{250} * \frac{1}{0.7} = 0.0742$

## Agent took the action Left and observed Red



| State | Probability |
|---|---|
| b($S_1$) | 923/3125 |
| b($S_2$) | 121/525000 |
| b($S_3$) | 39349/1312505 |
| b($S_4$) | 2717/700000 |
| b($S_5$) | 949/140000 |
| b($S_6$) | 803/26250 |

**Probabilities after Normalization**

$$b(S_1) = \frac{923}{3125} * \frac{1}{0.6366} = 0.4639$$

$$b(S_2) = \frac{121}{525000} * \frac{1}{0.6366} = 0.0003$$

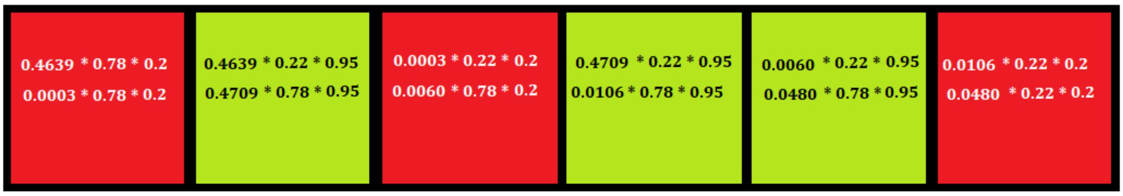
$$b(S_3) = \frac{39349}{131250} * \frac{1}{0.6366} = 0.4709$$

$$b(S_4) = \frac{2717}{700000} * \frac{1}{0.6366} = 0.0060$$

$$b(S_5) = \frac{949}{140000} * \frac{1}{0.6366} = 0.0106$$

$$b(S_6) = \frac{803}{26250} * \frac{1}{0.6366} = 0.0480$$

## Agent took the action Left and observed Green




| State | Probability |
|---|---|
| b($S_1$) | 403481/5570625 |
| b($S_2$) | 4139967/9284375 |
| b($S_3$) | 323213/334237500 |
| b($S_4$) | 142131191/1336950000 |
| b($S_5$) | 16435133/445650000 |
| b($S_6$) | 34529/13369500 |

**Probabilities after Normalization**

$$b(S_1) = \frac{403481}{5570625} * \frac{1}{0.6650} = 0.1089$$

$$b(S_2) = \frac{4139967}{9284375} * \frac{1}{0.6650} = 0.6704$$

$$b(S_3) = \frac{323213}{334237500} * \frac{1}{0.6650} = 0.0014$$

$$b(S_4) = \frac{142131191}{1336950000} * \frac{1}{0.6650} = 0.1598$$

$$b(S_5) = \frac{16435133}{445650000} * \frac{1}{0.6650} = 0.0554$$

$$b(S_6) = \frac{34529}{13369500} * \frac{1}{0.6650} = 0.0038$$