

2019101050_2019101049_partB_Report

Team Name: Trees.Love

Team Number: 5

Team Members:

Pulak Malhotra, 2019101050

Ashwin Rao, 2019101049

[Question 1](#)

[Question 2](#)

[Question 3](#)

[Question 4](#)

[Question 5](#)

Question 1

We have 128 states.

Nomenclature followed through out the report is

```
S<agent row><agent col><target row><target col><call on/off>
```

Eg S11001 means agent is at (1, 1) and target is at (0,0) and call is ON

If we target is in (1, 0) and we observe o6 this means agent can be in any other state than (0, 0), (1, 1) and (1, 0). There are 5 such valid positions and for each position call can be ON or OFF. Hence total 10 such valid positions with 0.1 probability of each.

The following is the probability for initial belief state in this question

```
S00000 0
S00001 0
S00010 0
S00011 0
S00020 0
S00021 0
S00030 0
S00031 0
S00100 0
S00101 0
S00110 0
S00111 0
S00120 0
S00121 0
S00130 0
S00131 0
S01000 0
S01001 0
S01010 0
S01011 0
S01020 0
S01021 0
```

S01030 0
S01031 0
S01100 0.1
S01101 0.1
S01110 0
S01111 0
S01120 0
S01121 0
S01130 0
S01131 0
S02000 0
S02001 0
S02010 0
S02011 0
S02020 0
S02021 0
S02030 0
S02031 0
S02100 0.1
S02101 0.1
S02110 0
S02111 0
S02120 0
S02121 0
S02130 0
S02131 0
S03000 0
S03001 0
S03010 0
S03011 0
S03020 0
S03021 0
S03030 0
S03031 0
S03100 0.1
S03101 0.1
S03110 0
S03111 0
S03120 0
S03121 0
S03130 0
S03131 0
S10000 0
S10001 0
S10010 0
S10011 0
S10020 0
S10021 0
S10030 0
S10031 0
S10100 0
S10101 0
S10110 0
S10111 0
S10120 0
S10121 0
S10130 0
S10131 0
S11000 0
S11001 0
S11010 0
S11011 0
S11020 0
S11021 0
S11030 0
S11031 0
S11100 0

```

S11101 0
S11110 0
S11111 0
S11120 0
S11121 0
S11130 0
S11131 0
S12000 0
S12001 0
S12010 0
S12011 0
S12020 0
S12021 0
S12030 0
S12031 0
S12100 0.1
S12101 0.1
S12110 0
S12111 0
S12120 0
S12121 0
S12130 0
S12131 0
S13000 0
S13001 0
S13010 0
S13011 0
S13020 0
S13021 0
S13030 0
S13031 0
S13100 0.1
S13101 0.1
S13110 0
S13111 0
S13120 0
S13121 0
S13130 0
S13131 0

```

Question 2

In this case target can be in (1, 0), (0, 1), (1, 1) or (1, 2) with call OFF.

The initial belief state for this case is

```

S00000 0
S00001 0
S00010 0
S00011 0
S00020 0
S00021 0
S00030 0
S00031 0
S00100 0
S00101 0
S00110 0
S00111 0
S00120 0
S00121 0
S00130 0

```

S00131 0
S01000 0
S01001 0
S01010 0
S01011 0
S01020 0
S01021 0
S01030 0
S01031 0
S01100 0
S01101 0
S01110 0
S01111 0
S01120 0
S01121 0
S01130 0
S01131 0
S02000 0
S02001 0
S02010 0
S02011 0
S02020 0
S02021 0
S02030 0
S02031 0
S02100 0
S02101 0
S02110 0
S02111 0
S02120 0
S02121 0
S02130 0
S02131 0
S03000 0
S03001 0
S03010 0
S03011 0
S03020 0
S03021 0
S03030 0
S03031 0
S03100 0
S03101 0
S03110 0
S03111 0
S03120 0
S03121 0
S03130 0
S03131 0
S10000 0
S10001 0
S10010 0
S10011 0
S10020 0
S10021 0
S10030 0
S10031 0
S10100 0
S10101 0
S10110 0
S10111 0
S10120 0
S10121 0
S10130 0
S10131 0
S11000 0
S11001 0

```
S11010 0.25
S11011 0
S11020 0
S11021 0
S11030 0
S11031 0
S11100 0.25
S11101 0
S11110 0.25
S11111 0
S11120 0.25
S11121 0
S11130 0
S11131 0
S12000 0
S12001 0
S12010 0
S12011 0
S12020 0
S12021 0
S12030 0
S12031 0
S12100 0
S12101 0
S12110 0
S12111 0
S12120 0
S12121 0
S12130 0
S12131 0
S13000 0
S13001 0
S13010 0
S13011 0
S13020 0
S13021 0
S13030 0
S13031 0
S13100 0
S13101 0
S13110 0
S13111 0
S13120 0
S13121 0
S13130 0
S13131 0
```

Question 3

Running **pomdpeval**:

```

-----
#Simulations | Exp Total Reward
-----
10           3.20063
20           3.66471
30           3.95338
40           3.89978
50           4.12574
60           4.41018
70           4.27445
80           4.296
90           4.27271
100          4.31637
-----

```

Finishing ...

```

-----
#Simulations | Exp Total Reward | 95% Confidence Interval
-----
100          4.31637          (3.78181, 4.85093)
-----

```

For question 1

#Simulations	Exp Total Reward
10	7.73091
20	7.86528
30	7.96606
40	8.17876
50	8.29295
60	8.31069
70	8.398
80	8.36638
90	8.40314
100	8.43137

Finishing ...

#Simulations	Exp Total Reward	95% Confidence Interval
100	8.43137	(8.101, 8.76175)

Question 4

When Agent is (0, 0) and target is in (0,1), (0, 2), (1, 1) and (1, 2)

O_2 when target is at (0, 1) otherwise O_6

$$O_1 = 0$$

$$O_2 = 0.25$$

$$O_3 = 0$$

$$O_4 = 0$$

$$O_5 = 0$$

$$O_6 = 0.75$$

When Agent is (1, 3) and target is in (0,1), (0, 2), (1, 1) and (1, 2)

O_4 when target is at (1, 2) otherwise O_6

$$O1 = 0$$

$$O2 = 0$$

$$O3 = 0$$

$$O4 = 0.25$$

$$O5 = 0$$

$$O6 = 0.75$$

Taking weighted average of the probabilities with the probability of Agent being in (0, 0) with 0.4 and (1, 3) with 0.6

$$O1 = 0.1$$

$$O2 = 0$$

$$O3 = 0$$

$$O4 = 0.15$$

$$O5 = 0$$

$$O6 = 0.75$$

The most likely observation is $O6$ with a prob of 0.75

Question 5

Solving pomdp with given start state we get:

Time	#Trial	#Backup	LBound	UBound	Precision	#Alphas	#Beliefs
0.01	15	81	5.26253	5.26351	0.000973218	40	18

Here #Trial can be treated as the time horizon T

$$|A| = 5$$

$$|O| = 6$$

$$T = 15$$

$$\text{Nodes in the tree } N = \frac{|O|^T - 1}{|O| - 1} = \frac{6^{15} - 1}{5} = 94036996915$$

Approximate number of policy trees = $|A|^N = 5^{94036996915}$ is a huge number