partB-Report

Team Name: Trees.Love

Team Number: 5

Team Members:

Pulak Malhotra, 2019101050

Ashwin Rao, 2019101049

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

```
----------------------------------
 #Simulations  | Exp Total Reward
----------------------------------
 10            7.37056
 20            6.49078
 30            6.6604
 40            6.94866
 50            7.06029
 60            7.08135
 70            7.32471
 80            7.35423
 90            7.54491
 100           7.4488
----------------------------------

Finishing ...

-------------------------------------------------------------
 #Simulations  | Exp Total Reward | 95% Confidence Interval
-------------------------------------------------------------
 100           7.4488             (6.75078, 8.14682)
-------------------------------------------------------------
```

For Question 1

```
----------------------------------
 #Simulations  | Exp Total Reward
----------------------------------
 10             10.7154
 20             11.3089
 30             11.5038
 40             11.4369
 50             10.909
 60             10.9451
 70             10.9848
 80             10.9504
 90             10.9878
 100            11.0693
----------------------------------

Finishing ...

-------------------------------------------------------------
 #Simulations  | Exp Total Reward | 95% Confidence Interval
-------------------------------------------------------------
 100            11.0693           (10.6013, 11.5372)
-------------------------------------------------------------
```

For Question 2

Question 4

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

$O2$ when target is at (0, 1) otherwise $O6$

$$O1 = 0$$

$$O2 = 0.25$$

$$O3 = 0$$

$$O4 = 0$$

$$O5 = 0$$

$O6 = 0.75$

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

$O4$ when target is at (1, 2) otherwise $O6$

$O1 = 0$

$O2 = 0$

$O3 = 0$

$O4 = 0.25$

$O5 = 0$

$O6 = 0.75$

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
-------------------------------------------------------------------------------
 0.02    21      131      8.23261    8.23356    0.000945528 53       28
-------------------------------------------------------------------------------
```

Here #Trial can be treated as the time horizon T

$|A|$ = 5

$|O|$ = 6

T = 21

Nodes in the tree N = $\frac{|O|^T - 1}{|O| - 1}$ = $\frac{6^{21} - 1}{5}$ = $4.387390128e15$

Number of policy trees = $|A|^N$ = $5^{4.387390128e15}$ is a very huge number