Assignment -5
Part-2
Report

Mehul Goyal
2018101018

STATES

We consider the tuple (Agent Position, Target Position, Call) as a state and we have 3x3 matrix so we gave 9 possible positions for the and for each there are 9 possible positions of target and it can be calling or not calling which further makes 2 more positions. So total 9x9x2 possible states i.e 162 states.

Possible positions of agent (0,0) - > 0, (1,0) - > 1 , (2,0) - > 2, (0,1) -> 3, (1.1) -> 4, (2,1) -> 5, (0,2) -> 6, (1,2) -> 7, (2,2) -> 8

Possible positions for Target (0,0) - > 0, (1,0) - > 1 , (2,0) - > 2, (0,1) -> 3, (1.1) -> 4, (2,1) -> 5, (0,2) -> 6, (1,2) -> 7, (2,2) -> 8

Cell no.(Encoding)

For call two possibilities 0 and 1 .

A state is presented in following format:

- abc

a -> represents position of agent
b -> represents position of target
c -> call on or off

I assumed the Grid This way.

(0,2)	(1,2)	(2,2)
(0,1)	(1,1)	(2,1)
(0,0)	(1,0)	(2,0)

All possible states(in order):

States : [(0,0,0) , (0,0,1) , (0,1,0) , (0,1,1) , (0,2,0) , (0,2,1) , (0,3,0) , (0,3,1) , (0,4,0) , (0,4,1) , (0,5,0) , (0,5,1) , (0,6,0) , (0,6,1) , (0,7,0) , (0,7,1) , (0,8,0) , (0,8,1) , (1,0,0) , (1,0,1) , (1,1,0) , (1,1,1) , (1,2,0) , (1,2,1) , (1,3,0) , (1,3,1) , (1,4,0) , (1,4,1) , (1,5,0) , (1,5,1) , (1,6,0) , (1,6,1) , (1,7,0) , (1,7,1) , (1,8,0) , (1,8,1) , (2,0,0) , (2,0,1) , (2,1,0) , (2,1,1) , (2,2,0) , (2,2,1) , (2,3,0) , (2,3,1) , (2,4,0) , (2,4,1) , (2,5,0) , (2,5,1) , (2,6,0) , (2,6,1) , (2,7,0) , (2,7,1) , (2,8,0) , (2,8,1) , (3,0,0) , (3,0,1) , (3,1,0) , (3,1,1) , (3,2,0) , (3,2,1) , (3,3,0) , (3,3,1) , (3,4,0) , (3,4,1) , (3,5,0) , (3,5,1) , (3,6,0) , (3,6,1) , (3,7,0) , (3,7,1) , (3,8,0) , (3,8,1) , (4,0,0) , (4,0,1) , (4,1,0) , (4,1,1) , (4,2,0) , (4,2,1) , (4,3,0) , (4,3,1) , (4,4,0) , (4,4,1) , (4,5,0) , (4,5,1) , (4,6,0) , (4,6,1) , (4,7,0) , (4,7,1) , (4,8,0) , (4,8,1) , (5,0,0) , (5,0,1) , (5,1,0) , (5,1,1) , (5,2,0) , (5,2,1) , (5,3,0) , (5,3,1) , (5,4,0) , (5,4,1) , (5,5,0) , (5,5,1) , (5,6,0) , (5,6,1) , (5,7,0) , (5,7,1) , (5,8,0) , (5,8,1) , (6,0,0) , (6,0,1) , (6,1,0) , (6,1,1) , (6,2,0) , (6,2,1) , (6,3,0) , (6,3,1) , (6,4,0) , (6,4,1) , (6,5,0) , (6,5,1) , (6,6,0) , (6,6,1) , (6,7,0) , (6,7,1) , (6,8,0) , (6,8,1) , (7,0,0) , (7,0,1) , (7,1,0) , (7,1,1) , (7,2,0) , (7,2,1) , (7,3,0) , (7,3,1) , (7,4,0) , (7,4,1) , (7,5,0) , (7,5,1) , (7,6,0) , (7,6,1) , (7,7,0) , (7,7,1) , (7,8,0) , (7,8,1) , (8,0,0) , (8,0,1) , (8,1,0) , (8,1,1) , (8,2,0) , (8,2,1) , (8,3,0) , (8,3,1) , (8,4,0) , (8,4,1) , (8,5,0) , (8,5,1) , (8,6,0) , (8,6,1) , (8,7,0) , (8,7,1) , (8,8,0) , (8,8,1)]

A1. Target is in (1,1) cell and observation is o6

o6 is observed when the target is not in the 1 cell neighbourhood of the agent.
Hence the agent cannot be in (0,1),(1,0),(1,2),(2,1) so the agent can be in (0,0),(0,2),(2,0),(2,2) with equal probability. Also call can be on or off . So eight states with equal probabilities.

Agent		Agent
	Target	
Agent		Agent

Initial belief state b : (agent , target , call)
States : [(0,4,0) , (0,4,1) , (6,4,0) , (6,4,1) , (2,4,0) , (2,4,1) , (8,4,0) , (8,4,1)]

These states have equal Probability = ⅛ = 0.125 and all other zero.

b : [0 0 0 0 0 0 0 0 0.125 0.125 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.125 0.125 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.125 0.125 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.125 0.125 0 0 0 0 0 0 0 0]

A2. Agent in (0,1) and Target is in one neighborhood and is not making a call.

Target		
Agent \ Target	Target	
Target		

Target can be in (0,0) , (0,1) , (0,2) , (1,1)
c is 0
States : [(3,0,0) , (3,3,0) , (3,6,0) , (3,4,0)]

These states have equal Probability = $\frac{1}{4}$ = 0.25 and all other zero.

b : [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0.25 0 0 0 0 0 0.25 0 0.25 0 0 0 0.25 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]

A3. Expected utility for initial belief states in 1 and 2

1.

```
------------------------------------------------------------
#Simulations   | Exp Total Reward | 95% Confidence Interval
------------------------------------------------------------
1000             3.70183            (3.51431, 3.88935)
------------------------------------------------------------
```

2.

```
------------------------------------------------------------
#Simulations   | Exp Total Reward | 95% Confidence Interval
------------------------------------------------------------
1000             8.54886            (8.35421, 8.74351)
------------------------------------------------------------
```

A4. Agent is in (0,1) with probability 0.6 and in (2,1) with probability 0.4 and the target is in the 4 corner cells with equal probability.

Target		Target
Agent(0.6)		Agent(0.4)
Target		Target

This gives the observations o3 and o5 with probability with probability $\frac{1}{4}$ and o6 with probability $\frac{1}{2}$ when the agent is in (0, 1).
This gives the observations o3 and o5 with probability with probability $\frac{1}{4}$ and o6 with probability $\frac{1}{2}$ when the agent is in (2, 1).

P(o1)= 0
P(o2)=0
P(o3)=0.6*0.25+0.4*25=0.25
P(o4)=0
P(o5)=0.6*0.25+0.4*25=0.25
P(o6)=0.6*0.5+0.4*0.5=0.5

Hence O6 is most likely to be observed

A5. Total no. of Policy trees :

The formula for calculating the number of trees in POMDP is $|A|^N$
where N = Number of nodes in the tree and A = All possible actions
Formula to calculate N:

N = $\frac{|O|^{T-1}-1}{|O|-1}$

where T = Levels in the tree and O = Number of observations
Number of trees obtained is dependent on the horizon T . The value of N increases as the value of T increases.In the SARSOP Solver, the horizon is not a fixed value, but depends on the precision value obtained. The program terminates when a target precision is obtained for a certain converging calculated value. Hence, the number of policy trees obtained is not a fixed value and can not be calculated without running the program.