MDL Assignment 5

Kalp Shah : 2018113003

Part 2

Question 1

If you know the target is in (1,1) cell and your observation is o6 , what will be the initial belief state? Please submit the optimal policy file named (Roll-Number).policy for the POMDP taking into account the initial belief state you obtained.

Answer 1

Answer is in the policy file

Question 2

If you are in (0,1) and you know the target is in your one neighborhood and is not making a call what is your initial belief state?

Answer 2

Given :

Call : Off
Agent Coordinates : (0,1)
Possible Target Coordinates : (0,1), (0,0), (1,1) or (2,0)

The only quantitiy that changes is the Target Coordinates.

$\therefore$ For the system with Agent at (0,1) and Call : Off, the probabilities are as follows :

- Target (0,1)
  P(E) = 0.25
- Target (0,0)
  P(E) = 0.25

- Target (1,1)
  P(E) = 0.25
- Target (2,0)
  P(E) = 0.25

Where E ≡ Given Event

Question 3

What is the expected utility for initial belief states in questions 1 and 2?

Answer 3

The expected utility is :

- Question 1
  Utility = 1.045
- Question 2
  Utility = 3.267

Question 4

If your 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, which observation are you most likely to observe? Explain.

Answer 4

Given :
Agent : P(S = (0,1)) = 0.6 , P(S = (2,1)) = 0.4
Target : (0,0), (0,2), (2,0) or (2,2)

All probabilities are initially 0, hence :

$$P(O_i) = 0, \forall i \in [0, 6]$$

∴ For Agent at (0,1), the probabilities are :

- Target (0,0)
  P(O_4) += 0.6 * 0.25
  += 0.15
- Target (0,2)
  P(O_2) += 0.6 * 0.25
  += 0.15

- Target (2,0)
  P(O_6) += 0.6 * 0.25
  += 0.15

- Target (2,2)
  P(O_6) += 0.6 * 0.25
  += 0.15

$\therefore$ For Agent at (2,1), the probabilities are :

- Target (0,0)
  P(O_6) += 0.4 * 0.25
  += 0.10

- Target (0,2)
  P(O_6) += 0.4 * 0.25
  += 0.10

- Target (2,0)
  P(O_4) += 0.4 * 0.25
  += 0.10

- Target (2,2)
  P(O_2) += 0.4 * 0.25
  += 0.10

The final probabilities, thus are :
P(O_2) = 0.25
P(O_4) = 0.25
P(O_6) = 0.50

Hence, $max(P(O_i)) \equiv P(O_6) = 0.50$, which is believable since for each agent coordinate, the probabilites remain same, but the observation corrosponding to O_6 is twice. (Due to there being 2 target states giving O_6)

Question 5

How many policy trees are obtained in this case, explain?

Answer 5

No of policy trees is given by this formulae :

$$N \equiv \sum_{i=0}^{T-1} |O|^i = \frac{|O|^T - 1}{|O| - 1}$$

$$N_{Tree} = A^N$$

In the computations done above, the values required in the equations are :

T = 1000
$|O| = 6$
A = 5

$$\implies N = \frac{6^{1000} - 1}{6 - 1}$$
$$\equiv \frac{6^{1000}}{5}$$
$$\rightarrow \infty$$

$$\therefore N_{Tree} = A^{\infty}$$
$$\rightarrow \infty$$

There are problems with this computation, as not all actions are available for each observation and thus the actual value would be a little less than the computed one.