Assignment 3 - Part 2

Deadline: 16th April, 11:55 PM

1 Instructions:

- You will need to install SARSOP solver for this assignment, the instructions for the same are given here. The link also explains various functionalities that are needed for the assignment.

- You can get to know about the solver more over here.

2 Problem:

In this assignment you will frame a POMDP for a tracking problem and use SARSOP solver to get the optimal policy. In the tracking problem an agent needs to find out the position of a target. The agent will be rewarded if it is in the same cell of the target when the target makes a call.

The instructions for framing the POMDP are as follows:

1. Agent and target move in a **2x4** grid. Each cell is represented as **(row number, column number)**. Each state of the POMDP is represented as a tuple **(Agent Position, Target Position, Call)**. Here agent position and target position are tuples of position in the grid.
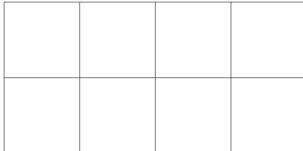
2. At each state we have five possible actions: **Stay, Up, Down, Left, Right**.

3. The target can either move Up, Down, Left, Right with equal probability of 0.1 (or) stay in the same position with a probability of 0.6. It can make a call with a probability of 0.5 and turn off the call with a probability of 0.1. If the agent reaches the target while the call is on , the agent gets its reward and the call is turned off by the target. Motion of target and the agent are independent of each other.

4. Transition probabilities for the agent are :

   - If it wants to stay at the same location then the action is executed perfectly,that is the agent's position doesn't change.

   - On the other hand, if it wants to move then it moves in the desired direction with a probability of x and with a probability of 1-x it moves in the opposite direction. Here x = 1 - (((LastFourDigitsOfRollNumber)%30 + 1) / 100)

5. The agent and target cannot move outside the grid, when they try to move outside the grid , the agent will stay back in the same state with a probability of x, and the target

will stay back in the same cell with a probability of 0.1.

6. Sensors on the agent can detect the following 6 observations with 100% accuracy:

   - o1 is observed when the target is in the same cell as the agent.

   - o2 is observed when the target is in the cell to the right of the agent's cell.

   - o3 is observed when the target is in the cell below agent's cell.

   - o4 is observed when the target is in the cell to the left of agent's cell.

   - o5 is observed when the target is in the cell above the agent's cell.

   - o6 is observed when the target is not in the 1 cell neighbourhood of the agent.

7. The rewards for the agent will be as follows:

   - -1 for each step that it takes.

   - (RollNumber%90 + 10) for reaching the target before the call is turned off.

8. To calculate x and reward, please use the roll number of any one of the team members and make sure that you **mention the roll number being used in the report.**

3 Questions:

Now you need to create the appropriate POMDP file as per the instructions and answer the following questions :

1. If you know the target is in (1,0) cell and your observation is o6 , what will be the initial belief state? Please submit the optimal policy file named **RollNumber.policy** for the POMDP taking into account the initial belief state you obtained.

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

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

4. If your agent is in (0,0) with probability 0.4 and in (1,3) with probability 0.6 and the target is in (0,1), (0,2), (1,1) and (1,2) with equal probability, which observation are you most likely to observe? Explain.

5. How many policy trees are obtained in the case of question 4, explain?

4 Submission instructions:

- You need to submit a typed or handwritten report containing the answers for the questions, please provide all the steps and explanations involved in answering a particular question. Name the report as **partB-Report.pdf** .

- Policy file mentioned in the 1st question, name it as **RollNumber1_RollNumber2.policy**.

- **Put the above two deliverables** in a folder named **RollNumber1_RollNumber2** and submit a compressed file named **RollNumber1_RollNumber2.zip**.

- The directory structure should be as follows:

```
RollNumber1_RollNumber2
├── partB-report.pdf
├── RollNumber1_RollNumber2.policy
```

5   Marking Scheme

```
Assignment
├── Question 1:  30%
│   ├── Belief state:  10%
│   ├── Policy File:  20%
├── Question 3:  30%
│   ├── Expected Utility question 1:  15%
│   ├── Expected Utility question 2:  15%
├── Question 4:  20%
│   ├── Probability for each observation:  10%
│   ├── Final Answer:  10%
├── Question 5:  10%
```