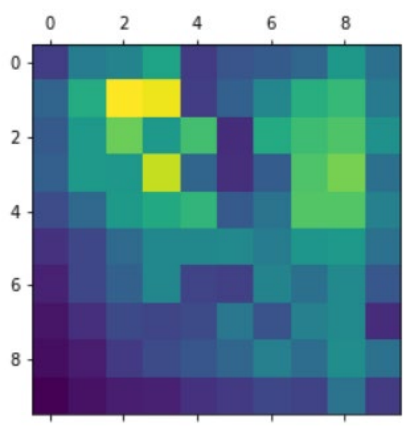
Report
Project 3 - Probability in Grid world

Parvathi Mahesh Hedathri: pm850
Kavya Kavuri: kk1069
Harish Udhayakumar: hu33
November 14 2021

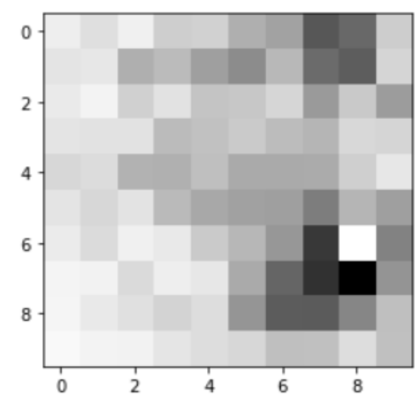
Code:

The grid world is represented as a numpy matrix. The values assigned to the grid world are randomly assigned based on the probability of blocked and unblocked cells with randomly distributed terrain type.

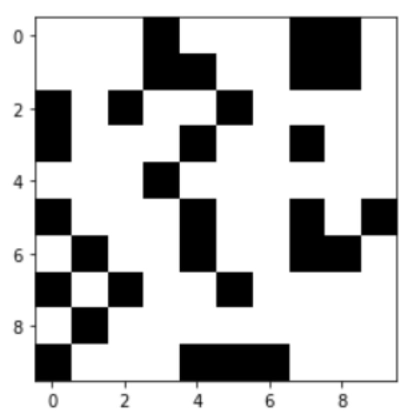
There are 5 types of matrices that represent the probabilistic gridworld are:



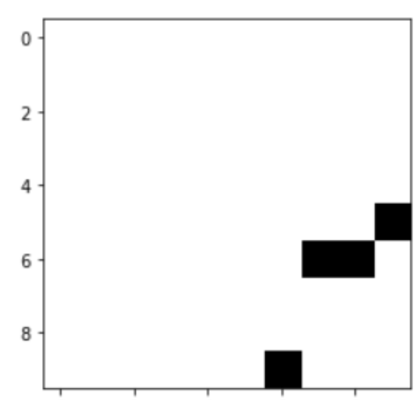
Confidence matrix



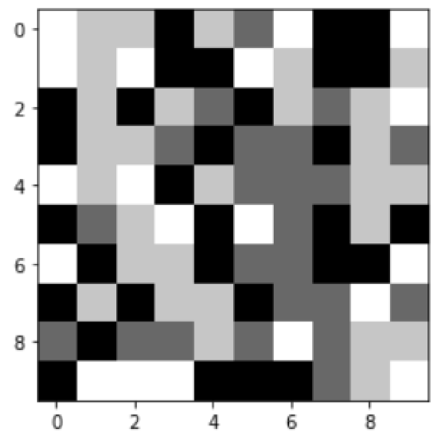
Belief matrix



Full Grid world



Agent Grid world



Terrain matrix

The Belief matrix shows the probability of how confident the agent is in finding the target in a particular cell.

The Confidence matrix shows the probability of the target being found in a particular cell. The shade of belief and confidence determines the probability of the target in the cell.

The Full gridworld matrix is the actual gridworld, showing blocked and free cells for the agent to traverse through the maze.

The Agent gridworld is the subset of fullgrid matrix, that the agent finds in the process of finding the target

The Terrain matrix holds the information about the terrain type of individual cells in the maze. The color range from lightest to darkest corresponds to terrain type from flat to hilly to forest.

Question 1: Prior to any interaction with the environment, what is the probability of the target being in a given cell?

- The target is equally probable to be in any of the cell in the grid world initially. Hence, it is given by: $1 \text{Dimension} * \text{Dimension}$

Question 2: Let $P_{(i,\ j)}(t)$ be the probability that cell (i, j) contains the target, given the observations collected up to time t. At time t+1, suppose you learn new information about cell (x; y). Depending on what information you learn, the probability for each cell needs to be updated. What should the new $P_{(i,\ j)}(t + 1)$ be for each cell (I, j) under the following circumstances:

- **At time t + 1 you attempt to enter (x, y) and find it is blocked?**

- $P_{x,y}(t)$ is initial probability of (x,y) containing target.
  Since (x, y) is blocked, $P_{x,y}(t+1) = 0$
  $P_{i,j}$(t+1) = Pi,j (t)1-Pi,j (t)

- **At time t + 1 you attempt to enter (x; y), find it unblocked, and learn its terrain type?**
- By Bayes' theorem, P(A|B) = P(A)*P(B| A) P(B)

  Initial assumption in agent's grid world is that all the cells are unblocked. So if we enter a cell (x,y) and find it unblocked, it will not affect other cells because this doesn't give any information on the position of the target. So **$P_{i,j}$(t+1) = $P_{i,j}$(t) for all (i,j)**

- **At time t + 1 you examine a cell (x; y) of terrain type flat, and fail to find the target?**
- Computing the probability
  When (i,j) ≠ (x,y) :
  A: Target in (i,j)
  B: Failed to find target at (x,y) , here (x,y) ≠ (i,j)
  Let terrain type of (x,y) be T, where T = "flat", "hilly", "forest"

  P(A) = Pi,j(t) -> We already know the initial probabilities - (1)
  P(B|A) = P(Failed to find target at (x ,y) | (i, j) contains target) = 1 - (2)
  P(B) = P(Failed to find target at (x, y))

  = Σ(i,j)=All (i,j) , (x,y) P(Failed to find target at (x,y) and target is in (i,j))
  = Σ(i,j)=All (i,j) P(Target in (i,j))*P(Failed to find target in (x,y)| Target in (i,j)) + P(Failed to find target in (x,y)| Target in (x,y))

  P(B)= 1=1dimensionj=1dimensionPi,j) + (Px,y * False_negative_rate(x,y) - (3)

  Using (1), (2), (3) we can calculate P(A|B) for all (i,j) where (i,j) ≠ (x,y)
  $P_{i,j}$(t+1) = Pi, j (t) (1=1dimensionj=1dimensionPi,j) + (Px,y * False_negative_rate(x,y)))

  When (i, j) = (x, y)
  P(A) = $P_{x,y}(t)$
  P(B|A) = False_negative_rate(x,y)
  P(B) = 1=1dimensionj=1dimensionPi,j) + (Px,y * False_negative_rate(x,y)

  So, we can calculate P(A|B) for (x, y) using above equations.

  $P_{x,y}$(t+1) = Px, y (t) * False_negative_rate(x,y) (1=1dimensionj=1dimensionPi,j) + (Px,y * False_negative_rate(x,y))

  Therefore, for a cell of terrain flat,

$P_{x,y}$(t+1) = Px, y (t) * 0.2 (1=1dimensionj=1dimensionPi,j) + (Px,y * 0.2)

- **At time t + 1 you examine cell (x; y) of terrain type hilly, and fail to find the target?**
- **$P_{x,y}$(t+1)** = Px, y (t) * 0.5 (1=1dimensionj=1dimensionPi,j) + (Px,y * 0.5)

- **At time t + 1 you examine cell (x; y) of terrain type forest, and fail to find the target?**
- **$P_{x,y}$(t+1)** = Px, y (t) * 0.5 (1=1dimensionj=1dimensionPi,j) + (Px,y * 0.5)

- **At time t + 1 you examine cell (x; y) and find the target?**
- If we find target at (x,y) the $P_{x,y}$(t+1) = 1 all other $P_{i,j}$(t+1) = 0

Question 3: At time t, with probability Pi, j(t) of cell (I, j) containing the target, what is the probability of finding the target in cell (x; y):

- $P_{x,y}$(finding the target in cell (x,y)) = $P_{x,y}$(t)*(1- False_negative_rate(x,y))

- **If (x; y) is hilly?**
- Px,y(t)*(1-0.5) = Px,y(t)*(0.5)

- **If (x, y) is flat?**
- $P_{x,y}$(t)*(1-0.2) = $P_{x,y}$(t)*(0.8)

- **If (x, y) is forest?**
- $P_{x,y}$(t)*(1-0.8) = $P_{x,y(}$t)*(0.2)

- **If (x, y) has never been visited?**
- P(finding target in cell (x,y)) = Σ for all t P(finding target in cell (x,y) and terrain = t)

$= \Sigma$ all t P(terrain = t)P(finding target in cell (x,y) | terrain = t)

= P(terrain = flat)P(finding target in (x,y) | (x,y) is flat) + P(terrain = hilly)P(finding target in (x,y) | (x,y) is hilly) + P(terrain = forest)P(finding target in (x,y) | (x,y) is forest)

$= 0.7*1/3*P_{x,y}(t)*(1-0.2) + 0.7*1/3*P_{x,y}(t)*(1-0.5) + 0.7*1/3*P_{x,y}(t)*(1-0.8)$
$= 0.35*P_{x,y}(t)$

Question 4: Implement Agent 6 and 7. For both agents, repeatedly run each agent on a variety of randomly generated boards (at constant dimension) to estimate the number of actions (movement + examinations) each agent needs on average to find the target. You will need to collect enough data to determine which of these agents is superior. Do you notice anything about the movement/examinations distribution for each agent? Note, boards where the target is unreachable from the initial agent position should be discarded.

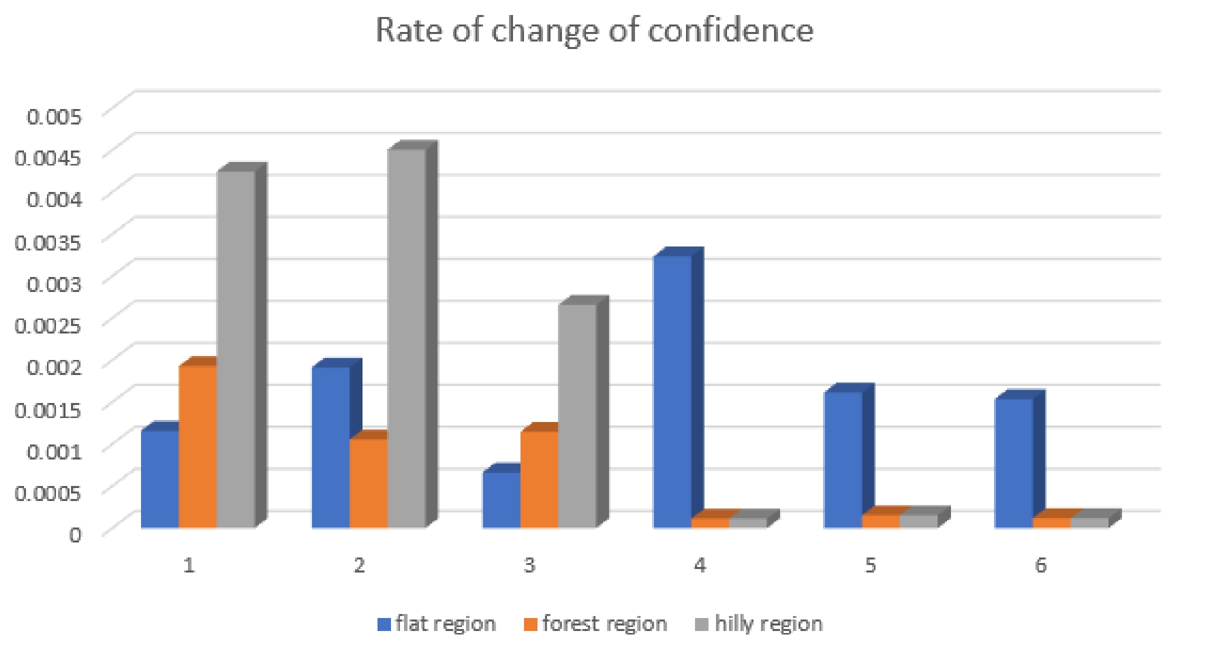
The agents make decisions based on two factors - the current belief state and Manhattan distance to the highest probability cell at any point in time. In case of agent 6, the decision to search the cell with the highest probability containing the target is made based on belief state after every iteration. Whereas, in the case of agent 7, the decision to examine a cell with highest probability is based on the confidence state at every iteration.

Note: The confidence state is updated only after the belief state is updated after every iteration.

On an average agent 7, requires lesser searches when the target is present in different terrains other than blocked cells and hence performs better. This behavior is because agent 7, has more information about the grid, as it utilizes the belief state of a particular cell along with false negative rate before making the decision.

A sample observation made on rate of change in confidence in accordance with the false negative rate of individual cell based on it's terrain type is as shown below.

Confidence[i][j] = (1 - false_negative_rate (Terrain_type) * Belief[i][j])



We see the change in confidence is a lot drastic for a forest terrain compared to a flat and hilly terrain.

Question 5: (20 points) Describe your algorithm, be explicit as to what decisions it is making, how, and why. How does the belief state ($P_{i,j}(t)$) enter into the decision making? Do you need to calculate anything new that you didn't already have available?

Question 6: (25 points) Implement Agent 8, run it sufficiently many times to give a valid comparison to Agents 6 and 7, and verify that Agent 8 is superior.

Question 7: (5 points) How could you improve Agent 8 even further? Be explicit as to what you could do, how, and what you would need.

Moving target

In this case, the target moves one step at a time to one of it's random unblocked neighboring cells. The information that the agent receives is the terrain type from which the target was previously in, and terrain type to which the target is currently in.

Initially belief is divided equally among all the cells in the gridworld.

Updating Probabilities:
The probabilities are updated until the agent is able to find the target.

Step 1: The agent checks if the cell is opened at time t + 1 and has the target. If the target is present then the agent moves to the termination state. If the target is not present, then the agent moves to step 2.

Step 2: At any time t, the belief of the corresponding terrain at time t are interchanged and equally distributed among it's neighbors of the other terrain type.

Example, if the information obtained is Type1 and Type2, the belief in terrain Type1 is distributed equally and added to it's neighbors of type2 terrain.
Similarly, the belief of Type2 terrain is distributed equally and added to beliefs of Type1 terrain that are neighbors of Type2.

The confidence state of the cell for moving the target is calculated the same way as for the stationary target case.

If the target is found, terminate.