**Question 5**

Brainstorming Strategies for Agent 8

Before approaching how we wanted to design Agent 8, we noticed that there were situations for both Agent 6 and 7 where the following would happen: the agent would examine a cell and have a negative result, the entire belief state is updated, and then it was determined that the new cell with the highest probability of containing the target (or the cell where it was most likely to find the target for Agent 7) was all the way on the other side of the board! More or less. So, we would see the trajectory length skyrocket for some trials, leading to a fairly high cost of operations that the agent must execute.

Our first thought for Agent 8 was something the professor had also expressed: “if you determine that there is a cell with a high probability that is far away from you that you should examine, as you are moving towards it, you may find cells of lower probability (but still good). Should you examine them prior to getting to the maximum probability cell?” Essentially, why don’t we make some examinations along the way to potentially save a costly trip across the entire grid-world?

Naturally extended from this line of thinking, we asked: what if we were able to choose a new cell to travel towards from the very start (of planning a new path), considering not only cells with high probability of successfully locating the target but also the cost for the agent to get there and examine? And from there, the mantra of Agent 8 was **cost-effective destination planning**.

To formalize this approach, we had to come up with an effective way of representing this idea of the most “cost-effective” destination for our agent’s next planning phase. Of course, the probability still factors into this determination, but we also want to consider the cost (i.e. the trajectory length to travel to the destination cell as well as the examination once we arrive). The higher the probability, the more the agent wants to go there, but the higher the cost, the less the agent will want to go there. So, we created a relationship for a cell’s “c” value being the following:

Probability of finding the target in this cell / ((planned path length from current position to there) + 1)

The +1 is for the examination that occurs at the end of the path which has a cost of 1 based on the metrics provided by the project description. With this relationship, we could determine the cell with the best probability to cost ratio, and this would be the next cell that we would want to travel to.

Algorithm

A new destination cell is determined anytime there is some update to the belief system. This is triggered by either an examination, finding a cell to be blocked, or finding that a cell is unreachable via A\*. (It’s worth mentioning that given the Agent 7 behavior, updating the probability of finding the cell in a previously unvisited cell also would factor into this. However, in our current structure, if the agent finds that the cell it’s currently in actually just became the most optimal place to find the target, the agent simply sets the new destination to be where it is at (very cost-effective planning), and it will examine anyway, which is covered by the examination trigger listed above).

The “event cell” (i.e. the cell that is examined, the cell that is blocked, or the cell that is found to be unreachable) has its probability updated, and then every other cell is updated as well. While we’re updating all the probabilities of the cells in the grid-world, we determine the cell with the best metric to be our next destination. For Agent 6, it was P(target in ij). For Agent 7, it was P(finding target in ij). And now for Agent 8, it will be determined via the “c” value described above. Consider the following algorithm:

* Some event triggers an update of the knowledge base
* Event cell’s probability is updated
* Event cell’s c-value = event cell’s updated probability / 1
  + *// If the event cell is the current cell the agent occupies, that means the triggering event was an examination, and so the denominator, which represents the cost of the next plan to travel and examine, is essentially 0 steps to get from current position to event cell + 1 examination*
  + *// If the event cell is a blocked cell or if it is a cell that is unreachable, then the updated probability will be 0, and the corresponding c-value will be 0.*
* Current destination = event cell
* Current max c-value = event cell’s c-value
  + *// Based on our belief system update so far, this is the best place to go and examine, but except in rare circumstances, it will almost surely change*
* For all remaining cells, “temp”, to be updated that aren’t the event cell
  + Temp’s probability is updated
  + Temp’s c-value = temp’s updated probability / ((Manhattan distance from current cell occupied by agent to temp) + 1)
    - *// The original idea was to attempt planning using A\* so we know the actual path length to get there, but this turned out to be way too costly to the runtime (with potentially upwards of 200 A\* runs after every update to the belief state, even after filtering out unpromising cells). This just wasn’t feasible. In the same way that it’s used as a heuristic anyway, the Manhattan distance is used to save having to plan a path for many cells in the grid-world which would be quite costly. This is still an appropriate and helpful metric to gauge cost-effectiveness because if the Manhattan distance is a lower bound for the path length and a cell’s c-value is already lower than the “current max c-value”, then the actual path length can only be equal to or longer, and we know that this cell is definitely not going to be our next destination.*
  + If temp’s c-value > current max c-value
    - Current destination = temp
    - Current max c-value = temp’s c-value
* Using the destination determined above, plan a new path to get there
* Normal probability agent behavior, etc.

So, by the end of the update to the belief system, agent 8 will have determined the most cost-effective destination in the grid-world to shoot for next, factoring in not only the probability of finding the target there but also the cost of travelling there and examining it. This will cut down on the overall cost endured by the agent while searching for the target.