Homework 3

Zhengyang Ling(zl351), Yalin Liu(yl1255), Yujia Liu(yl1261)

April 10, 2021

1

(1)

$P(TargetinCell_i|Observations_t \wedge FailureinCell_j) = P(TargetinCell_i \wedge Observations_t \wedge FailureinCell_j)/P(Observations_t \wedge FailureinCell_j)$.

(2)

$P(TargetfoundinCell_i|Observations_t) = (1 - P(environment)) * P(TargetinCell_i|Observations_t)$

(3)

We generate 10 maps, and play through each map (with random target location and initial agent location eachtime) 10 times with each agent. We have the result: on average the cost(total distance traveled + number of searches) of agent1 is 39296.7 , agent2 is 36496.1. So agent2 works better on average

(4)

Improved agent is our Smart Agent(SA) is combined by two parts. We aim to choose the next step. Thus, we need to get the utility for choosing random points. For this part, my reward is combined by two parts. one is belief, another is distance. Thus, for a cell, if it around has more high belief cells, then it has a higher utility. Then, using this value to update value iteration (Bellman equation) to calculate the final utility. Then, we can write our second part. We want to start at one point in less step to get the higher utility. Thus, this reward is combined by the distance from this point and the utility that we calculate in the last part. Then, using the value iteration (Bellman equation) to get the value of each action. Then, we can get the action that has the highest value. Because of the limit of the computer, we will only calculate the 10 by 10 around the current cell. If we have unlimited compute resources, SA will calculate the utility for the whole graph. SA is much faster than our agent1 and 2. We generate 10 maps, and play through each map 10 times with each agent. We have the result:on average the cost(total distance traveled + number of searches) of SA is 7825.8.