MDL ASSIGNMENT-2 REPORT (PART - 2)

Team Number: 44
Team Name: LearningMachinesFromData
Team Members:
T.H.Arjun (Roll number: 2019111012),
Gokul Vamsi Thota (Roll number: 2019111009)

PART 2:

Task 1:

Analysis -

- ***WEST*** - When IJ is in WEST square, he prefers to take the action ***'RIGHT'*** *if MM is in* ***Dormant State,*** *otherwise he would prefer to* ***'STAY'*** *there itself*. This is because he wants to get to the CENTRE square as quickly as possible, just so that he can attack MM better from a closer distance provided MM is in **Dormant State** and wouldn't attack him**.** Otherwise, IJ prefers to ***'STAY'*** away so that he is not attacked. He also chooses to ***'SHOOT'*** from a far off distance when there is not much difference in utility between ***'SHOOT'*** and ***'STAY'***, and attacking MM with arrows via shooting would definitely not have a negative impact, and might be more optimal when sufficient arrows are present.

- ***NORTH*** - When IJ is in NORTH square, he prefers to move ***'DOWN'*** if he **doesn't possess any material** and if MM is in **Dormant State**, or prefers to ***'STAY'*** till he becomes **Dormant** so that he can move ***'DOWN'*** to attack in a better way from a closer distance**,** because in this case it is optimal to get closer to MM for attacking him as he cannot make any arrows here. But, if IJ is in possession of sufficient material, he prefers to ***'CRAFT'*** and makes use of this

material to obtain arrows for attacking MM in the next time step when he moves ***'DOWN'.***

- ***EAST*** - When IJ is in EAST square. He keeps on trying to ***'HIT'*** MM from this square, or tries to ***'SHOOT'*** him (if there are sufficient arrows) so that he can finish the battle off quickly and get the reward as fast as possible. He uses this greedy strategy as there is depreciation of utility with time and he wants to acquire the reward as soon as possible, (the action chosen here also depends on tradeoff between ***'SHOOT'*** & ***'HIT',*** because loss of health of MM is higher with ***'HIT'*** but accuracy is better with ***'SHOOT'***). But in most cases, the action ***'HIT'*** is preferred as there is **no loss of any property** of IJ in choosing this action (whereas he would lose arrows on choosing the action *'SHOOT'*).

- ***SOUTH*** - When IJ is in SOUTH square, he attempts to move ***'UP'*** to reach CENTRE square, so that he can then further attempt to ***'HIT'*** or ***'SHOOT'*** MM from that position. He also chooses to **'GATHER'** if he is **low on materials** and MM is in **Ready State** so as to stay away from his attack. He chooses to ***'GATHER'*** even if has full capacity since it is more optimal compared to ***'STAY'*** as there is no chance of failure and he still retains a safe position with respect to MM's attack (Note that gathering wouldn't increase his materials in possession to a value higher than 2). He also decides to ***'STAY'*** sometimes when MM is in ***READY state,*** just so he is safe and away from MM's attacks, which is highly likely if the latter is in **READY** state.

- ***CENTRE*** - When IJ is in CENTRE square, he mostly attempts to take the ***'RIGHT'*** when MM is in ***Dormant State*** and thus reach the EAST square, so that he has better probabilities of success / accuracy of actions ***'SHOOT'*** and ***'HIT'.*** If MM is in **Ready State** he moves **'DOWN'** to get away from MM (to stay safe) if he is **low on material,** because in this case he would reach the SOUTH square where he can ***'GATHER'*** material or ***'STAY'*** to remain safe (gathering is more optimal)**,** and prefers to move **'UP'** to NORTH if he is **high on materials,** so that he would reach NORTH square where he can ***'CRAFT'*** arrows and thus gain enough ammunition to attack MM or choose ***'STAY'*** to stay safe**.** Sometimes IJ prefers

to ***'STAY'*** even if ***'CRAFT'*** or ***'GATHER'*** clearly appear more optimal, this is because, if IJ is in possession of sufficient resources and he chooses to ***'STAY'***, then he would reach EAST square on failure of action, and he can now attempt to attack MM more comfortably (provided MM is in **Dormant state**) as he has sufficient ammunition. Also, if the health of MM is low, then he tries to ***'SHOOT'*** if sufficient arrows are present because shooting has better accuracy and extent of damage to MM is not important here (Health of MM cannot go below 0 anyway).

Simulation -

1) **Start state = (W, 0, 0, D, 100):**
   Output of Simulation:

Step Number: 1 Current State: (W,0,0,D,100) Taking Action: RIGHT
Step Number: 2 Current State: (C,0,0,R,100) Taking Action: DOWN
Step Number: 3 Current State: (S,0,0,R,100) Taking Action: STAY
Step Number: 4 Current State: (S,0,0,R,100) Taking Action: STAY
Step Number: 5 Current State: (S,0,0,D,100) Taking Action: UP
Step Number: 6 Current State: (E,0,0,D,100) Taking Action: HIT
Step Number: 7 Current State: (E,0,0,D,100) Taking Action: HIT
Step Number: 8 Current State: (E,0,0,D,100) Taking Action: HIT
Step Number: 9 Current State: (E,0,0,D,100) Taking Action: HIT
Step Number: 10 Current State: (E,0,0,D,100) Taking Action: HIT
Step Number: 11 Current State: (E,0,0,R,50) Taking Action: HIT
Step Number: 12 Current State: (E,0,0,R,50) Taking Action: HIT
Step Number: 13 Current State: (E,0,0,R,50) Taking Action: HIT
Step Number: 14 Current State: (E,0,0,D,75) Taking Action: HIT
Step Number: 15 Current State: (E,0,0,D,75) Taking Action: HIT
Step Number: 16 Current State: (E,0,0,D,75) Taking Action: HIT
Step Number: 17 Current State: (E,0,0,D,25) Taking Action: HIT
Step Number: 18 Current State: (E,0,0,R,25) Taking Action: HIT
Step Number: 19 Current State: (E,0,0,R,25) Taking Action: HIT
Step Number: 20 Current State: (E,0,0,R,25) Taking Action: HIT
Step Number: 21 Current State: (E,0,0,R,0) Taking Action: NONE

Analysis: We followed the policy, since there are zero arrows and zero material at WEST starting state, it's better to take the RIGHT action here and move to CENTRE square from where he can choose the action HIT, but since MM is in Ready State he moves to SOUTH and waits till he reaches Dormant state. When MM reaches Dormant state, IJ Moves UP. But the action failed and he reached EAST, but since MM is in Dormant State, it is optimal to keep on choosing the HIT action to defeat MM. He keeps on hitting until he is defeated (since MM is in Dormant State and IJ has nothing to lose). When MM eventually changes to Ready State, he is already on the verge of defeat, so IJ chooses to take risk and continues to hit and get done with the fight, so that he can acquire the reward before the depreciation takes place.

2) **Start state = (C, 2, 0, R, 100):**

Output of Simulation:

Step Number: 1 Current State: (C,2,0,R,100) Taking Action: UP
Step Number: 2 Current State: (C,2,0,D,100) Taking Action: RIGHT
Step Number: 3 Current State: (E,2,0,D,100) Taking Action: HIT
Step Number: 4 Current State: (E,2,0,D,100) Taking Action: HIT
Step Number: 5 Current State: (E,2,0,D,100) Taking Action: HIT
Step Number: 6 Current State: (E,2,0,D,100) Taking Action: HIT
Step Number: 7 Current State: (E,2,0,D,100) Taking Action: HIT
Step Number: 8 Current State: (E,2,0,D,100) Taking Action: HIT
Step Number: 9 Current State: (E,2,0,D,100) Taking Action: HIT
Step Number: 10 Current State: (E,2,0,D,50) Taking Action: HIT
Step Number: 11 Current State: (E,2,0,D,0) Taking Action: NONE

Analysis:

Initially IJ decided to move up but MM attacks and his action fails. Now, MM is in Dormant State. So IJ moves east using RIGHT action to HIT MM while he is in Dormant State to defeat him. He keeps on hitting until he is defeated since MM is in Dormant State and by choosing the action HIT in EAST square, IJ has nothing to lose.

Task 2:

Case 1:

In this case, when IJ chooses the action ‘LEFT’ at EAST square, he would end up at WEST square if successful. This changes the policy in an interesting way. When IJ is in EAST square and MM has high health and he is in READY STATE, IJ prefers to take ‘LEFT’ action more often, as he can get away directly from MM by moving to WEST Square and wait till he goes back to Dormant State. Otherwise the policy remains the same. That is ‘LEFT’ action has more utility now since it takes to WEST directly where he can stay safe now, whereas previously he still moved to CENTRE square which was vulnerable to attack.

Case 2:

In this case, the step cost when ‘IJ’ chooses to ‘STAY’ in the same square is 0. Here, the policy would change. Evidently, IJ Prefers 0 step cost over negative step costs since the goal of IJ is to maximize expected reward, and hence would rather prefer to ‘STAY’ in the same square and not gain any reward, in comparison with taking an action which changes his state but makes him lose reward. He also relies on failure of STAY to go to other states. He is now more selective in picking his actions, and would filter out the possible actions from current state such that he would choose only those actions which are worthy of the negative step cost (actions which would assist in gaining higher reward by attacking MM). Such actions would typically be chosen if the health of MM is low and IJ has a decent chance to reduce his health significantly. Otherwise ‘ IJ’ would choose to ‘STAY’ (especially when MM attacks), preferring the 0 reward, over negative costs and a depreciated reward for destroying MM. But if he somehow reaches a

square close to MM (if ‘STAY’ fails), he tries to kill MM. IJ also moves closer to MM if he has a lot of arrows and health of MM is low.

Case 3:

In this case, Gamma is reduced to 0.25 from the high value of 0.999. The algorithm now converges faster (in lesser iterations) because of the decreasing impact of the values of next states along iterations. Subtle changes in policy are observed here. We start to observe ‘SHOOT’ as one of the optimal actions at WEST even if it has low accuracy. This typically occurs when health of MM is low and IJ has sufficient arrows (as he is more interested in the current gain of the positive reward) or if he is READY State, shooting from here is optimal as he can inflict damage without getting affected by MM. North is approached in a similar way as the initial policy. If he is in EAST square, IJ chooses to move ‘LEFT’ if health of MM is high and he is READY state, otherwise ‘HIT’ as ’IJ wants to deal as much damage as possible quickly, otherwise he chooses to ‘SHOOT’ as it has better accuracy. IJ further prefers to ‘GATHER’ at SOUTH, if he is low on Material, so that he can ‘CRAFT’ at NORTH. That is NORTH and SOUTH the policy remains almost the same, logically. We also notice that IJ has become more risk taking as his threshold for moving back is now higher and he takes more risks as he wants the reward as fast as possible as the depreciation is now high.