Roll No : 2018113003, 2018114001

MDL Assignment 2

TASK 1

The policy adopted by the protagonist is quite obvious as protecting health and having arrows. This affinity of having more resources and attacking only when the opposing party is vulnerable has its roots in the greedy way one should approach these situations.

There are no sharp differences in the penalty function. Hence, the strategy that attack when stamina is high and health of dragon is low, recharge when stamina is low and dodge if low on arrows applies.

TASK 2

Part 1

The convergence of this task is much faster than previously, hence the lowered step cost and favorable action of shooting allows the model to learn by attacking and only in the worst case retreat. This emphasis on attacking allows the model to lean towards attacking as it has more incentive, thus improving the score

Part 2

With no incentive to kill the dragon, the familiar action is to dodge and retreat, and the huge discount factor does that. Hence, due to this dragon will be attacked when his health is low and every other action will be equally preferred. Due to the strategy of retreating is more suitable, the policy is fairly long and the number of steps it takes to converge is larger due to every other action being equally prevalent.

Part 3

As the motivation was already removed, the change in delta will make the agent take more number of iterations to get to the desired policy which will not be so much dissimilar to the policy described in part 2.