Q2 Ans.

Date

(1.) There are four observable parameters. So, with 50 bins per dimension, the size of the Q-table becomes $50^4 \times 2$ and in order to get good results using this ~~obs~~ Q-table, the no. of training episodes should be large. 1000 episodes ~~do~~ are not sufficient to update such a large Q-table and most of the final values remain as ~~at~~ their initial ones ~~not~~ while for a much smaller Q-table like $5^4 \times 2$, each cell in the Q-table can be visited more than once and the values in most of the cells reach their converging point. That is why, coarse discretization gives better results for 1000 training episodes.

(2.) For 10,000 episodes, the larger Q-table gets ~~mu~~ updated in a much better way as ~~it~~ compared to case of 1000 episodes and now the reesults obtained ~~fo~~rom this Q-table are extremely good as compared to smaller Q-table case as for this Q-table, the ~~success~~ reward obtained becomes stagnant around 60 after a certain no. of ~~$~~episodes and there is no more scope of improvement in this table.

(3.) The issue with discretization can be observed at the boundary points of sub-intervals. Discretization maps a whole interval of continuous states to one discrete state and for all the states mapped to a specific discrete state, the agent is forced to take same action. Near boundary points, two almost same states can fall in different sub-intervals and actions corresponding to the two can be different. These edge effects can induce error in our policy and this is a drawback of discritization.

4.) Although, increasing the number of bins makes our discretization move closer to reality, the size of Q- table becomes extremely-extremely large and visiting each cell of this Q- table multiple times in order to make them converge is not practically feasible as it would take very long time to ~~learn~~ make the model learn. Also, the boundary issues/edge issues increase along with the number of bins.