Reward Trajectory for Needle In Haystack Medium inst 0.54 -Instantaneous Mean Reward Algorithm DynamicGreedy NonBayesianEpsilonGreedy, 0.05 ThompsonSampling UCB1WithConstantOne DynamicEpsilonGreedy, 0.05 0.52 -

4000

time

6000

0.51 -

0

2000