Reward Trajectory for Needle In Haystack 1 Medium 1 High inst 0.80 -0.75 -Instantaneous Mean Reward Algorithm DynamicEpsilonGreedy DynamicGreedy NonBayesianEpsilonGreedy 0.70 -ThompsonSampling UCB1WithConstantOne 0.65 -0.60 -5000 0 2500 7500 10000 time