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