Reward Trajectory for Needle50 – High 3 0.8 -Algorithm Mean Reward DynamicEpsilonGreedy DynamicGreedy ExploreThenExploit ThompsonSampling UCB 0.4 -1000 0 2000 5000 3000 4000 time