Reward Trajectory for Needle50 - Low 1 0.6 -Algorithm Mean Reward DynamicEpsilonGreedy DynamicGreedy ExploreThenExploit ThompsonSampling UCB 0.4 -0.2 -5000 2000 3000 0 1000 4000 time