Reward Trajectory for Needle50 - Low 0.6 -Algorithm Mean Reward DynamicEpsilonGreedy DynamicGreedy ExploreThenExploit ThompsonSampling UCB 0.4 -250 500 750 0 1000 time