Training results for several runs run 0 200 run 1 run 2 run 3 run 4 150 mean of means Episode reward +STD -STD 100 50 0 0 25 50 75 100 125 150 175 200 Training episode