Average Reward vs. Number of Training Iterations 4.0 3.5 discount=1 discount=0.999 1.0 discount=0.99 10 20 30 40 Iteration