

Average Reward vs.  
Number of Training Iterations

