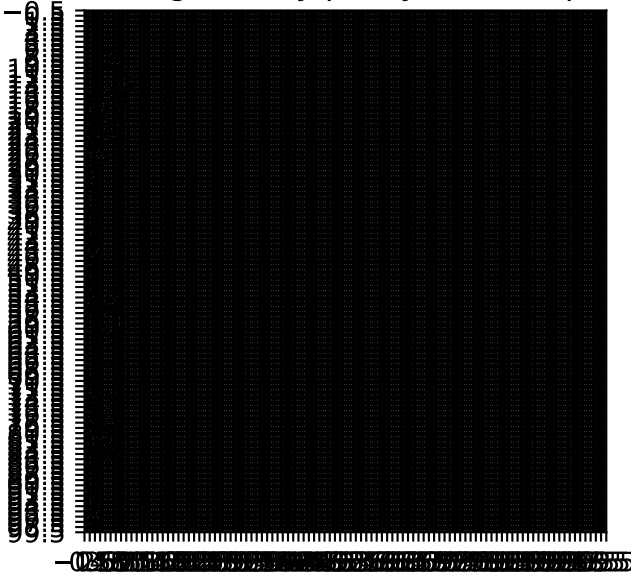
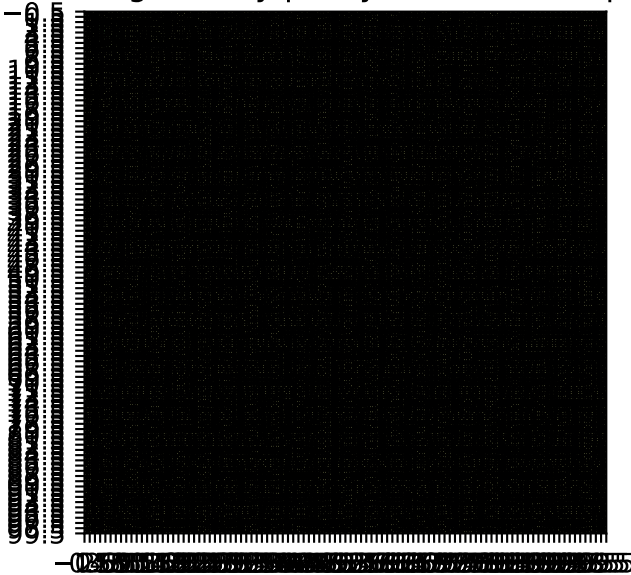


Q-learning Greedy policy at 0 th episode



Q-learning Greedy policy at 10000 th episode



Q-learning Greedy policy at 20000 th episode

