Cartpole swingup - greedy policy evaluation Difficulty k = 1

