Algoritmo di guida End-to-End Critic Actor Rumore $\nabla_a Q$ **Actor** Critic a_t **TORCS Target Target** a_{t+1} **Actor** Critic (s_t,s_{t+1}) a_t s_{t+1} s_{t+1} **Experience** Reward Replay $(s_t, a_t, s_{t+1}, r_{t+1})$