Pendulum-v0

Twin Delay DDPG (TD3) 玩Pendulum-v0游戏。TD3是改进版的DDPG：（1）增大了两个critic网络，克服了DDPG对行为Q值评估过高的缺点； （2） 延时更新actor网络，使网络寻优action过程变得更稳定； （3）critic更新Q值对行为加入噪声，target policy smoothing regularization，在一个区域中的行为，使得critic网络更稳健。