

on-off AWAC_GAE, intrinsic reward 0.01
on-off AWAC_Q_lambda_Peng, intrinsic reward 0.01
on-off AWAC_Q_lambda_TB, intrinsic reward 0.01

PPO