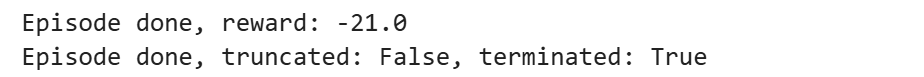
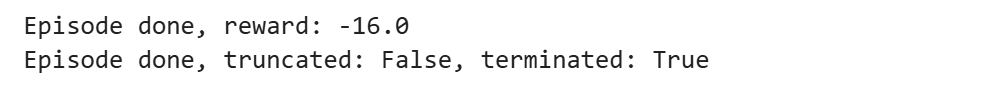


A2C:





PPO:



总的来说 PPO的准确率比A2C高一些，训练所需的时间也稍微短一点点。此外看到群里有说明Pong游戏的reward是负数的情况是正常的