

无人系统设计第四次作业

1.1 内容要求

- 本次作业旨在让同学们使用Simulink完成强化学习训练，控制战机在不同的情形下击落靶机。
- 取值范围：角速度 $-\pi/6 \sim \pi/6$ ，线速度 $0 \sim 150$ ，偏航角 $0 \sim 2\pi$ ，俯仰角 $-\pi/6 \sim \pi/6$ （角速度单位为rad/s，系统内所有角度数值均为弧度值）。
- 远距离固定靶机中，敌机初始坐标为（-1600，0，0）本机初始坐标为（1600，0，0）。
- x、y、z坐标没有最大值的限制。

1.2 训练流程

- 训练前需要修改的地方：
 - 修改envs.m文件中observations和actions的数量以及取值范围，一般情况下范围为-1~1。
 - 设计并实现observations的优化，逻辑在get_my_obs.m文件中实现，原始的observations共有27个量，每个量对应的意义在get_my_obs.m文件中的注释里有相应的解释。
 - 设计并实现优化后observations的归一化处理，逻辑在normalize_my_obs.m文件中实现。
 - 设计并实现reward计算，逻辑在get_my_reward.m文件中实现，get_my_reward函数的输入是当前回合的原始observations和上一回合的原始observations，第一回合的时候，上一回合的数据和当前回合的数据相同。如果需要把函数的输入修改为步骤二中优化后的observations可以修改rlflight.slx文件中的相应逻辑。（注：在设计奖励函数的过程中，最好将reward的最大值控制在10的3次方数量级及以下，即最大奖励值 <10000 ，最小值没有限制）
 - 打开rlflight.slx文件，根据设计修改agent的输出action处理，“TCP发送控制量”模块需要的输入是一个四维向量，代表着四个控制量，分别是油门(throttle)、俯仰(pitch)、翻滚(roll)、偏航(yaw)，如果agent的输出仅有一个偏航角，那么需要在该输出的基础上将另外三个控制量补全，补全控制量为0。
- 如何开始训练：
 - 首次训练时，打开envs.m文件，运行该文件在MATLAB工作区创建初始agent；非首次训练时，双击运行.mat文件将agent加载进MATLAB工作区。
 - 运行tcpServer.exe，输入端口号，该端口号应该与rlflight.slx文件中的tcp发送和接收模块的端口号相同，初始端口号为8080。
 - 打开UE客户端，在UE客户端中根据需要创建一个相应类型的新房间，输入比赛轮数和每轮最大回合数，进入房间后，客户端上方显示的内容为“房间X”，X为房间Uid。
 - 打开rlflight.slx文件，修改房间Uid为步骤三中创建房间的房间Uid。
 - 打开training.m文件，修改maxepisodes的值为步骤三中输入的比赛轮数，运行该文件，开始训练。
- 注意事项：
 - 如果训练在中途异常中止，可以执行saveCommand.m中的保存指令将MATLAB工作区中的内容保存在.mat文件中，也可以在工作区右键，点击“保存”来存储工作区变量。
 - 如果在训练过程中要中止训练，点击Reinforcement Learning Episode Manager中右上角的Stop Training按钮，等当前回合结束后，即可关闭训练窗口和服务端，训练正常结束后会将工作区的变量保存在“save.mat”文件中。（注：一定要等当前回合结束后再关闭训练窗口和服务端，不然训练会异常中止，异常中止的处理方法参考上一点）

2.1 任务一（20分）

- 控制偏航角训练击落近距离固定靶机，在作业报告中写明observation和奖励函数的设计以及训练结果截图，UE下的训练结果录视频放在“结果视频”文件夹中，视频命名为“一维近距离固定靶机.mp4”。

2.2 任务二（30分）

- 控制偏航角和俯仰角训练击落近距离固定靶机，在作业报告中写明observation和奖励函数的设计以及训练结果截图，UE下的训练结果录视频放在“结果视频”文件夹中，视频命名为“二维近距离固定靶机.mp4”。

2.3 任务三（30分）

- 控制油门和偏航角训练击落远距离固定靶机，在作业报告中写明observation和奖励函数的设计以及训练结果截图，UE下的训练结果录视频放在“结果视频”文件夹中，视频命名为“二维远距离固定靶机.mp4”。

2.4 任务四（10分）

- 了解并修改envs.m文件，优化训练性能，要有修改前后的参数和训练结果对比截图，并在作业报告中写清做出的哪些修改以及原因，例：更改网络结构。
- 关于envs.m文件中定义网络结构以及训练参数的各类函数可以通过在[MATLAB官网帮助中心](#)中搜索相应函数查看函数功能和参数。

2.5 任务五（10分）

- 控制油门、俯仰角、偏航角击落远距离固定靶机，在此基础上可以考虑在控制量中加入翻滚角，也可以选择训练击落远距离固定轨迹靶机，在作业报告中写明observation和奖励函数的设计以及训练结果截图，UE下的训练结果录视频放在“结果视频”文件夹中，视频根据训练的控制量以及目标进行命名，命名规则为“X维远距离固定（固定轨迹）靶机”，X是控制量的个数。
- 注：如果实现了任务五，可以不用实现任务一、二、三，也可获得前三个任务的所有分数。

3. 提交内容

- 将作业发布时的文件夹整体打包进行提交，需要完成的内容有哪些以及需要创建的新文件，如：视频文件应该保存在哪个文件夹里在任务介绍中均已写明。
- 本次作业的相关助教为吴泽一。
- 注：本次作业为小组作业，每组仅需组长提交一份.zip文件，文件命名为“小组组长的学号+姓名.zip”。