# README

## 参数设置

星机多智能体默认配置为5个卫星，2个WRJ，场景时间设置为3600s，任务奖励（1-10）、任务所需观测时间（3-6）均在一定范围内随机生成，每个场景默认训练1000批次，仿真软件场景存放在scenario/RLSTAR.sc中；在默认配置下，共设计有四个实验， 分别为lab\_50，lab\_100，lab\_200，lab\_400，每批次分别传入50、100、200、400个观测任务。

## 运行实验

运行no\_similate\_train.py，读取场景下的指定实验文件，记录训练过程中不同批次的单个智能体奖励函数收益和总奖励函数收益，以CSV格式存放到data/reward/中，并将训练过程中奖励函数收益最高的模型存储到models/中；

输入：data/lab/lab\_50.csv，data/lab/lab\_100.csv，data/lab/lab\_200.csv，data/lab/lab\_400.csv

输出：data/reward/lab\_50\_rewards.csv，data/reward/lab\_100\_rewards.csv，data/reward/lab\_200\_rewards.csv，data/reward/lab\_400\_rewards.csv

运行plot\_training\_rewards.py，读取data/reward/中的收益文件并绘制模型奖励函数曲线。

输入：data/reward/lab\_50\_rewards.csv，data/reward/lab\_100\_rewards.csv，data/reward/lab\_200\_rewards.csv，data/reward/lab\_400\_rewards.csv

输出：pic/lab/lab\_50\_Agent1(1-7) Reward\_reward\_plot.png，pic/lab/lab\_50\_total\_reward\_plot.png，

pic/lab/lab\_100\_Agent1(1-7) Reward\_reward\_plot.png，pic/lab/lab\_100\_total\_reward\_plot.png，

pic/lab/lab\_200\_Agent1(1-7) Reward\_reward\_plot.png，pic/lab/lab\_200\_total\_reward\_plot.png，

pic/lab/lab\_400\_Agent1(1-7) Reward\_reward\_plot.png，pic/lab/lab\_400\_total\_reward\_plot.png