

整体代码框架

整体运行框架

mpx/Project/DRL\_Quadruped\_Robot\_Isaac\_Sim

/legged\_gym

/scripts

plat\_flat.py 平坦环境下测试，模型运行时间、失败次数  
play\_terrain.py 复杂环境下测试，模型运行时间、失败次数  
play.py 综合测试  
train.py 平坦环境、复杂环境的训练  
创建训练环境、加载环境参数 make\_env  
创建PPO算法、加载训练参数 make\_alg\_runner  
算法与环境交互 learn

/tests

test.py 机器人导入测试  
test\_terrain.py 地形测试  
test\_env.py 训练前测试

/utils

helpers.py 初始参数设定、辅助函数设置 任务初始参数设置 get\_args  
logger.py 运行结果打印和记录  
math.py 常用数学计算  
task\_registry.py 任务主体函数  
环境、交互、算法参数设置 register  
创建环境 make\_env  
创建算法 make\_alg\_runner  
terrain.py 地形环境文件  
随机环境 randomized\_terrain  
训练环境 curriculum  
测试环境 curriculum2

/logs

/flat 平坦环境下，历史较优版本  
/flat\_QuadrupedArm 平坦环境下，当前最优版本  
/rough 复杂环境下，历史较优版本  
/Rough\_QuadrupedArm 复杂环境下，当前最优版本

/resources

/robots/quadruped\_arm  
/meshes 各关节STL文件  
/urdf 模型URDF文件

/envs

/base  
base\_config.py 外部类初始化  
base\_task.py 强化学习交互任务基类  
执行 step  
重置 reset  
可视化 render  
legged\_robot.py 强化学习交互任务派生类  
奖励计算 compute\_reward  
状态计算 computr\_observation  
创建环境 create\_sim  
重置检测 check\_termination  
legged\_robot\_config.py 强化学习环境、算法基类  
地形 terrain  
命令 commands  
初始状态 init\_state  
控制 control  
模型加载 asset  
随机扰动 domain\_rand  
奖励 rewards  
状态、动作规范化 normalization  
状态空间噪声 noise  
摄像机参数 viewer  
仿真环境 sim  
强化学习算法 LeggedRobotCfgPPO

/quadruped\_arm

quadruped\_arm\_flat\_config.py 平坦环境下，强化学习环境、算法派生类  
初始位置、关节状态设置 init\_state  
地形设置 terrain  
控制模式选择 control  
模型文件获取 asset  
奖励函数设置 reward  
命令空间 commands  
外部干扰设置 domain\_rand  
状态空间噪声设置 noise  
强化学习算法设置 LeggedRobotCfgPPO  
网络设置 policy  
PPO算法超参设置 algorithm  
算法运行超参设置 runner  
quadruped\_arm\_rough\_config.py 复杂环境下，强化学习环境、算法派生类  
初始位置、关节状态设置 init\_state  
地形设置 terrain  
控制模式选择 control  
模型文件获取 asset  
奖励函数设置 reward  
强化学习算法设置 LeggedRobotCfgPPO  
网络设置 policy  
PPO算法超参设置 algorithm  
算法运行超参设置 runner

PPO算法框架

mpx/rsi\_rl

/algorithms

ppo.py 强化学习过程  
动作执行 act  
损失函数计算与网络更新 update  
策略网络误差计算 surrgate\_loss  
价值网络误差计算 value\_loss

/env

vec\_env.py 矢量化环境基类

/modules

actor\_critic.py 策略网络值函数网络的搭建  
分布更新 update\_distribution  
从分布中采样得到动作 act  
actor\_critic\_recurrent.py 加入RNN的策略网络值函数网络的搭建

/runners

on\_policy\_runner.py  
学习阶段 learn  
结果记录 log

/storage

rollout\_storage.py  
回报计算 compute\_returns  
采样数据批处理 mini\_batch\_generator

/utils

utils.py 辅助函数