2025.5.4

1. 优先级采样增加n\_steps √

2025.5.6

1. 在维护multistep的memory中会修改length:

dic\_tmp["length"] = self.n\_steps

导致在append中添加的类型为torchtensor的length变成了int导致出错 √

2.agent.load()失败没有报错

2025.5.7

1.画reward曲线图时出错（self.history用于绘制会少一个，因为当前reward没有保存）

2025.5.9

1. 修改RLInnerModel的类继承关系，不是所有的类都有探索算法(不一定是输出state的actionq值分布)

2025.5.10

1.向量环境SyncVectorEnv的支持

2025.5.11

1. 输出向量action的DQN有问题argmaxaQ(s,a)现在是每个action分开计算的
2. InnerModel不同的action类型输出，支持不同的探索策略（离散/连续）
3. InnerModel依据gym中Action的不同类型来进行判断使用什么更底层的模型如DQN/NAF √

2025.5.12

1. 给分层强化学习可视化
2. InnerModel增加记录加载最优解
3. InnerModelCallback整合到RLBases中√
4. OnPolicy和OffPolicy的抽象

2025.6.12

1. BNAF内部网络实现有问题
2. 调整项目文件结构
3. DDPG、PPO等算法的实现
4. 增加seed
5. 输出环境当前步数/总步数

2025 6.14

1. ExplorationStrategy序列化问题√
2. ExplorationStrategy和调度器之间的关系√
3. 修改离散动作的选择
4. Tensorboard路径修正

2025.6.16

1. 给RLAgent增加Steps
2. 可配置化&超参数管理
3. 早停机制，超过若干次后没有提升结束并且保存最优模型
4. 优先采样访问的current\_value修改