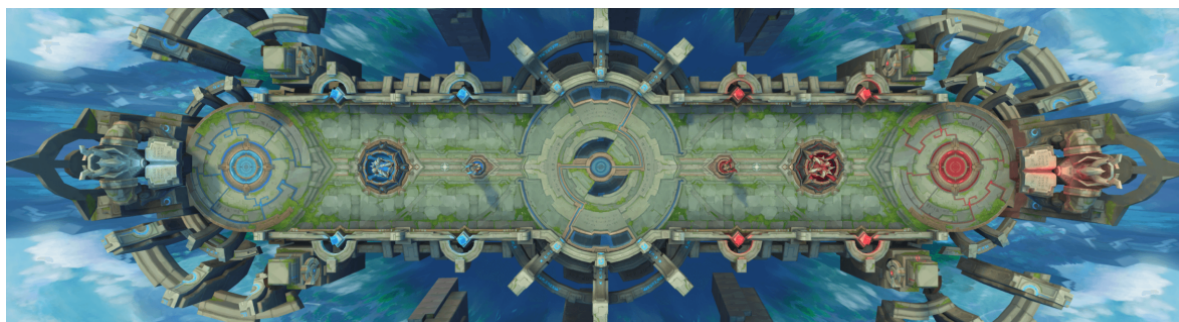


# 实践作业三：王者荣耀开悟平台——1v1实验

## 概述

通过算法训练一个英雄（智能体），在地图获取经济和经验来提高英雄的攻击力，率先摧毁对方的水晶以获得胜利。若己方水晶被摧毁，则对方获得胜利。



有关该环境的详情请移步开悟平台详见[开发指南](#)

## 开发指引

本次实验意在让同学们尝试调优PPO算法。

### • PPO算法的改进

- 本实验在代码库中，给出了tencent在2020年训练王者荣耀AI的dual-clip ppo算法。具体算法内容可对照下述论文。（但本次环境只需要训练单一英雄）

[Towards playing full moba games with deep reinforcement learning](#)

- 在此基础上同学们可以尝试一些PPO算法上的训练trick以稳定训练或提升性能。推荐同学们阅读2022年iclr上的这个PPO的37个实现细节，结合实际环境进行实验。

[The 37 Implementation Details of Proximal Policy Optimization](#)

### • 奖励上的改进

- 开发指南中已经给出了一些奖励权重的指引，大家可以自行观看。

下面附一段之前实验室同学的参赛心得。

### 奖励体系

奖励我们认为是复赛内容中很重要的一块，元老队伍之所以那么强，我们认为他们不仅在网络部分有所积累，也应该有保留过很多很有效果的奖励设置，这是赛事中一个重大提升点，一个好的奖励能让智能体更好的处理相关的场景情况，在本次赛事中，以下是我们使用的相关奖励设置。

- 时序奖励的退火机制，退火机制经过实测确实会带来训练速度的极大提升，但是在训练后期，我们发现智能体在处理后期情况的时候相对来说会较难以处理，分析原因主要有为：退火程度过高，后期奖励与前期的误差较大，导致智能体在训练过程对各个情况没有一个很好的处理方式。退火机制下，我们的智能体前期作战能力很强，然后后期会出现对相关情况没有一个很好处理的现象。

- 死亡击杀奖励机制：在训练过程中，我们发现当智能体完成击杀的时候，会出现原地迷茫的情况，针对于这种特定的情况，我们放大了智能体己方死亡与对方死亡时间段内的推塔和金币奖励，以及死亡惩罚，让他能够在完成击杀的时候，尽快完成兵线的清理与推塔进度的推进。
- 实践证明，这个奖励能有效的提高智能体在对面死亡过程中情况的处理，我们的智能体能够完成残血推塔以及卡在对方上线前回家清兵。
- 技能命中奖励，在训练过程中，智能体技能的准头并没有非常好，因此我们提高了针对于这部分内容，给予了技能使用惩罚与命中奖励，这部敌方英雄塔下决斗内容有效的提高了智能体的命中效率（感谢中级赛道）
- 另外，我们认为，1v1场景下的小项奖励系数，baseline的奖励设置就已经很好很合理，可以不进行相关调整。
- 建筑惩罚递增奖励：事实上，在进行对战的过程中，游戏的核心本质还是推塔游戏，因此守护防御塔应该为一个核心任务，然后在实际中，单纯调高塔的守护奖励并不能有效的告诉智能体如何处理情况。根据游戏机制，英雄的推塔速度相比于小兵要快很多，因此这部分奖励主要分为三个内容：
  - 建筑摧毁递增：当建筑不断的摧毁的时候，塔的奖励值不断地提高，从而激发智能体守塔和推塔的欲望。
  - 敌方英雄塔下决斗：当敌方英雄处于塔下的时候，判断此时为危险环境，正常训练下，智能体会优先解决好打的小兵，但这是错误的，因此在这里需要调高对敌方英雄击杀的奖励。
  - （未实现）小兵击杀判断：当敌方小兵血量较少的时候，可以优先击杀塔内的小兵，从而能够让敌方英雄对塔伤害大幅降低，这是一种特殊情况，可以考虑实现

## • 奖励上的改进

- 网络上的改进
- 超参数改进
- .....

落到实处上的代码开发过程，请参考开发指南中**建议修改的代码**结合上述方向与你自己的想法进行开发。

## 任务

和作业二一样本次分数的占比分为测评分（30%）与报告分（70%）两部分组成

### 报告部分（40%）

- 根据目前提交的部分作业二的报告情况，我认为同学们的实验报告都相当完备，也很有创新思路。因此报告部分请同学们自由发挥。  
(心得式、论坛式或者technical report类的报告模式也非常不错，希望同学们在写的时候能带着闭包的性质，能让读者对你的改进有从0到1的见解，你们的探索 and 技巧也会成为明年、后年学弟学妹的宝贵财富)
- 本次实验分组进行（五人一组），希望同学们在报告中体现组内的分工，或者组员在讨论或者不同分支路径中的尝试。  
(尽管一个技巧可能没有生效，但是也提供了避免二次踩坑的独特价值)

### 模型测评部分（60%）

将根据开悟平台上的教学测评进行排名赋分。同学们需要在规定时间前完成开发生命周期，并提交你最强的模型至开悟平台的教学测评之中，后续助教将会通知开放提交的时间。

告报告文件以 `学号_作业2.pdf` 命名 发送到邮箱[13868629768@163.com](mailto:13868629768@163.com)

提交截至时间：**2025-6-25**

## 问题反馈

---

有任何疑问可以发邮件至助教邮箱: [kekq@mail2.sysu.edu.cn](mailto:kekq@mail2.sysu.edu.cn)，每天固定时间检查并回复邮件。