

22-23 学年春季强化学习第二次实验作业

（环境/算法集成）

授课教师：赵冬斌，朱圆恒，张启超；

教师助教：陈亚冉

2023 年 4 月

作业任务—MindSpore 接入强化学习的新环境/新算法

强化学习的目的是智能体通过与环境的交互过程，学习优化策略来使得累计回报最大。环境是强化学习中的重要要素。强化学习使用的环境种类繁多：Gym、MuJoCo、MPE、Atari、PySC2、SMAC、TORCS、Isaac 等，目前 MindSpore 架构上已接入了 Gym、SMAC 两个环境。MindSpore 已有的强化学习算法有 DQN、SAC 等。要求：

- 1) 选择除上述所列的 Gym、SMAC 之外的一种环境，或其他的典型环境，接入 MindSpore 架构，利用课堂上讲述的强化学习方法，解决环境中的典型任务。
- 2) 在 MindSpore 框架下，编写除 [MindSpore 架构中包含的强化学习代码](#) 外的强化学习算法，实现 Gym 或 SMAC 环境下任务的控制。

要求完成上述一种任务即可，平台接口基于 Python 语言，设计的强化学习系统使用昇思 MindSpore 架构。

将整个过程整理成完整的报告，要求有任务的介绍，方法描述，定量的实验结果，如训练过程曲线、最终得分，以及结论等。

同时要求每位同学提交一份关于自己在整个小组研究过程中的分工与贡献。

在规定的时间内提交报告和源代码，完成实验作业。

注意：严禁抄袭代码和报告，若发现，本部分为零分！

平台搭建

使用昇思 MindSpore 架构搭建仿真环境和强化学习算法。

参考资料：https://gitee.com/mindspore/reinforcement/blob/master/README_CN.md,

MindSpore 华为工程师的技术支持，<https://gitee.com/mindspore/mindspore/issues> 提问。