

22-23 学年春季强化学习第二次实验作业

(机器人方向)

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-强化学习机器人导航

利用课堂讲述的强化学习算法实现机器人导航任务, 环境提供机器人的位置和朝向信息 (x, y, α) 以及目标的位置 (gx, gy) 信息, 设计强化学习策略控制机器人速度、导航并避障到达目标点, 且朝向目标点。

我们将提供 MindSpore 架构下的 DQN 算法的简单实现, 有一定学习效果但不能稳定学习到有效的策略, 可能由于以下原因: 环境奖励稀疏, 仅在到达目标位置时有奖励; DQN 不适合解决连续动作空间的问题, 等等。

针对以上问题, 自己定义合适的辅助密集奖励, 或使用更适合连续动作空间的强化学习算法 (MindSpore 提供的 RL 算法接口, 如 DDPG) 来实现有效的导航策略学习 (300 个 episode 内), 相比于给定的 DQN 有提升, 完成机器人导航任务。

要求使用昇思 MindSpore 架构设计强化学习导航系统。将整个过程整理成完整的报告, 包括 MindSpore 框架的结构分析以及导航结果分析, 整个小组研究过程中的分工与贡献等, 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

DQN 导航示例代码使用的环境为简化版 (五维观测, 包括机器人的位姿和目标的位置信息), 源自我们组织的 2022 IEEE CoG RoboMaster Sim2Real 竞赛的第一个赛道, 详细可参考 <https://github.com/DRL-CASIA/COG-sim2real-challenge> 个赛道是机器人的第一视角图像输入, 可作为本次作业的挑战任务。

MindSpore RL 算法库: <https://gitee.com/mindspore/reinforcement/tree/master>。