

调研 - 动态环境的解决方法 - 强化学习 RL

传统算法的缺陷原因：

传统的路径规划方法通常**依赖于对环境的精确地图信息**，这些方法包括经典的图搜索算法，如A*算法和Dijkstra算法等。

这些算法能够在已知的环境中找到最短或最优的路径，但它们往往**需要事先对环境有完整的了解**，并且对环境变化的适应性较差。

针对动态环境的解决方法 - 深度强化学习：

随着深度学习（Deep Learning, DL）和强化学习（Reinforcement Learning, RL）技术的迅速发展，深度强化学习（Deep Reinforcement Learning, DRL）算法开始在移动机器人路径规划和避障领域展现出其强大的潜力。深度强化学习结合了**深度学习的感知能力**和**强化学习的决策能力**，使得机器人能够在未知或动态变化的环境中进行有效的路径规划。

学习了解：

- 强化学习：概率论基础、基本/核心概念、Q-Learning
 - <https://www.bilibili.com/video/BV1dN4y1Z7qU>
 - 模仿学习：概念方法、GAN的概念和区别
 - <https://blog.csdn.net/caozixuan98724/article/details/103765605>
-

论文选择：

- 强化学习相关
 - [多智能体路径规划 python 多智能体路径规划 碰撞]https://blog.51cto.com/u_13354/6969236
 - Multi-agent navigation based on deep reinforcement learning and traditional pathfinding algorithm
将传统的A star算法选择路径作为强化学习算法中的action（相当于用A*进行保底）
 - PRIMAL: Pathfinding via Reinforcement and Imitation Multi-Agent Learning（经典）
总结：结合了强化学习和模仿学习，能用于多个智能体（1024）；分布式算法
设计了expert centralized MAPF planner，各个agent的决策都是有益于整体的；agent不需要显式通信，但可以在路径规划中表现出隐式协同。
单agent的决策靠的是RL + 模仿集中专家
(有点类^①似于：集中式训练，分布式执行)
-

TODO:

- 要用的话，怎么保证训练情况表现良好
- 怎么在船舶的环境下进行训练
- 看的东西的关联性、用脑图组织整理
总结出文章的问题、解决方案
- 理解总结个架构