

《智能控制与机器人》

—期末大作业

本次期末大作业占本次课 60% 的成绩， 每名学生可以在编程和写报告中二选一， 完成后将资料打包成 zip 文件， 发送至邮箱 yuanyang@smail.nju.edu.cn， 截止提交时间为 2019.12.29， 下午 11:59。

选择一：Python 编程

要求使用 Python 语言完成以下两项任务， 并附上编程报告， 说明主要的实施步骤和算法效果图。

1. 寻找一个案例（比如经典迷宫导航问题）， 复现离散状态动作空间的经典 Q-learning 算法；
2. 寻找一个案例（比如连续状态动作空间下的导航问题、Mujoco 机器人运动控制问题）， 复现经典的 Policy Gradient 算法， 即 REINFORCE 算法。

选择二：写报告

要求完成一份综合报告， 内容包括两点几点。

1. 强化学习综述： 主要的几类算法， 各自的特点与优缺点， 以及每种算法之间的联系与区别；
2. 结合自己的研究领域（比如机械、控制、智能电网、管理科学、工业工程、金融工程等）， 在已有文献的基础上， 思考潜在的可以与强化学习相结合的地方， 如何将强化学习应用到自己的领域。