

浙江大学 2025 暑期多智能体强化学习实训

课程简介

本课程讲述强化学习、智能体的理论与应用，涉及人工智能强化学习基本理论、深度强化学习算法、多智能体强化学习以及前沿应用。课程分为讲授和实训两个部分。

- 课程共 12 天，形式为授课 + 练习作业 + 大作业。课程内容涵盖了动态规划理论、基于表格的基础强化学习算法、Eligibility Trace 应用、基于神经网络的深度强化学习算法、基于价值/基于策略的算法、基于有/无模型的算法、多智能体强化学习算法、以及在 Gym、SMAC 和具身智能物理仿真环境中的实践应用。
- 优先推荐表现优秀的同学到国家人工智能学院实习。

课程目标

通过实训，学生将学习和实践一系列强化学习理论与算法。从最基础的基于表格的 Q-learning 出发，到人形机器人运控、多智能体 AI 的开发，最终大作业是完成多智能体博弈算法开发。

时间安排

6 月 25 日至 7 月 8 日，12 天全天课程，每天上午知识讲授，下午实训。

时间：全天课程 9:30-12:00 14:00-17:00 || 每天上午知识讲授，下午实训。大作业提交截止时间：8 月 18 号。

助教：1-2 个

课程大纲

课程的设计采用讲授与实训结合的方式，每日上午知识讲授结束后，下午安排对应的实训，让学生完成相应的练习。

时间	上午	下午
6 月 25 日	第一讲 动态规划理论，Bellman 公式，策略评估与策略优化理论	配置强化学习环境，安装 Gym，完成案例 CartPole 环境的运行、开始 Issac Gym 环境配置
6 月 26 日	第二讲 基于表格的 Q-learning 算法，SARSA 算法，Eligibility Trace 的应用	基于 Gym，测试 Q-learning 和 SARSA 算法的表现与区别

6 月 27 日	第三讲 深度学习基础，包含损失函数、梯度回传等知识	安装 PyTorch，并完成基础的回归、分类任务的网络训练
6 月 28 日	第四讲 强化学习与具身智能	人形机器人平地行走实验
6 月 29 日	休息	休息
6 月 30 日	第五讲 深度强化学习 I：基于价值的 DQN 与基于 DQN 的算法	基于 Gym，完成 DQN 算法实现，并在 Freeway 等环境测试
7 月 1 日	第六讲 深度强化学习 II：基于策略的 Reinforce 算法、AC 算法的讲解、以及 A2C、A3C 等算法	在 Gym 环境中实现基础 AC 算法
7 月 2 日	第七讲 深度强化学习 III：先进算法 TRPO、PPO、DDPG、SAC 等算法原理与实现	人形机器人复杂地形行走实验（一）
7 月 3 日	第八讲 基于模型的强化学习方法与基于监督学习的行为克隆方法	人形机器人复杂地形行走实验（二）
7 月 4 日	休息	休息
7 月 5 日	第九讲 多智能体强化学习 I：基于价值的 QMIX 算法以及改进算法	在 SMAC-Hard 环境中实现 QMIX 算法
7 月 6 日	第十讲 多智能体强化学习 II：基于策略的 MADDPG、MAPPO 算法实现	在 SMAC-Hard 环境中实现 MAPPO 算法
7 月 7 日	第十一讲 多智能体系统研究展望	大作业布置
7 月 8 日	人形机器人行走技术方案汇报	大作业答疑