

多智能体系统与强化学习

主讲人：高阳、杨林、杨天培

<https://reinforcement-learning-2025.github.io/>

课后实验一

实验背景：

假设你是一个机器人控制系统的设计师，你的任务是在一个网格世界中导航。网格世界是一个 5×5 的方格，其中包含：

起点：网格的左上角（位置 $(0, 0)$ ）。

目标：网格的右下角（位置 $(4, 4)$ ）。

障碍物：某些格子是不可通过的（例如 $(2, 2)$ 和 $(3, 1)$ ）。

奖励：每移动一步的奖励是 -1 ，到达目标的奖励是 $+10$ ，碰到障碍物的奖励是 -10 。

实验目标：

你的目标是设计一个强化学习算法，让机器人从起点出发，找到一条最优路径到达目标，同时避开障碍物。

课后实验二

实验背景：

CartPole 是一个经典的强化学习环境，智能体需要通过左右移动小车来保持杆子的平衡。智能体的目标是通过学习，使得杆子尽可能长时间地保持直立状态。

实验目标：

1. 使用深度强化学习方法，分别实现基于价值的方法（如DQN）和基于策略的方法（如Policy Gradient）来解决CartPole问题。
2. 可视化智能体在学习过程中的表现。

课后实验三

实验背景：

在 OpenAI Gym 的 CarRacing-v2 离散环境中实现深度 Q 网络 (DQN) 算法

实验目标：

1. 把CarRacing-v2环境中的连续动作空间离散化。
2. 构建一个合适的DQN神经网络模型。
3. 实现DQN算法的训练流程，包含经验回放和目标网络更新。
4. 训练智能体并绘制训练过程中的奖励曲线。