

课程大纲

COURSE SYLLABUS

1.	课程代码/名称 Course Code/Title	强化学习 Reinforcement Learning
2.	课程性质 Compulsory/Elective	专业选修课 Major elective course
3.	开课单位 Offering Dept.	计算机科学与工程系 Department of Computer Science and Engineering
4.	课程学分/学时 Course Credit/Hours	3/64
5.	授课语言 Teaching Language	中文 Chinese
6.	授课教师 Instructor(s)	袁博, 助理教授, 计算机科学与工程系, yuanb@sustech.edu.cn Bo Yuan, Assistant Professor, Department of Computer Science and Engineering, yuanb@sustech.edu.cn
7.	开课学期 Semester	秋季 Fall
8.	是否面向本科生开放 Open to undergraduates or not	是 Yes
9.	先修要求 Pre-requisites	(如面向本科生开放, 请注明区分内容。 If the course is open to undergraduates, please indicate the difference.) CS303 人工智能 Artificial Intelligence
10.	教学目标 Course Objectives	(如面向本科生开放, 请注明区分内容。 If the course is open to undergraduates, please indicate the difference.) 为了实现能够学会做出正确决策的自治系统, 强化学习是一类强大的方法, 它与众多应用相关。本课程将为强化学习领域提供扎实的介绍, 学生将学习核心挑战和方法, 包括马尔可夫决策过程、动态规划、无模型预测和控制、值函数近似、策略梯度、学习与规划、深度强化学习和多智能体强化学习等主题。本课程的目的是让学生了解和掌握强化学习的一些基本思想、算法和工具, 以便于为学生将来进一步的工作和研究打下基础。 In order to realize autonomous systems that can learn to make correct decisions, reinforcement learning is a powerful paradigm that is related to many tasks. This course will provide a solid introduction to the field of reinforcement learning, and the students will learn the core challenges and methods, including Markov decision process, dynamic programming, mode-free prediction and control, value function approximation, policy gradient, learning and planning, deep reinforcement learning and multi-agent reinforcement learning. The goal is to provide the basic ideas and methods in reinforcement learning, and the students will be able to pursue advanced study and research in the field if desired.
11.	教学方法 Teaching Methods	(如面向本科生开放, 请注明区分内容。 If the course is open to undergraduates, please indicate the difference.) 理论课 32 学时 + 实验课 32 学时 Lectures 32 credit hours + Lab 32 credit hours
12.	教学内容 Course Contents	(如面向本科生开放, 请注明区分内容。 If the course is open to undergraduates, please indicate the difference.)
	Section 1	介绍 Introduction
	Section 2	马尔可夫决策过程 MDP

Section 3	动态规划 Dynamic Programming
Section 4	无模型预测 Model-Free Prediction
Section 5	无模型控制 Model-Free Control
Section 6	值函数近似 Value Function Approximation
Section 7	策略梯度 Policy Gradient
Section 8	学习与规划 Learning and Planning
Section 9	探索与利用 Exploration and Exploitation
Section 10	深度学习基础 Deep Learning
Section 11	深度强化学习价值方法 Deep Reinforcement Learning I
Section 12	深度强化学习策略方法 Deep Reinforcement Learning II
Section 13	基于模型的深度强化学习 Deep Reinforcement Learning III
Section 14	多智能体强化学习一 Multi-agent Reinforcement Learning I
Section 15	多智能体强化学习二 Multi-agent Reinforcement Learning II
Section 16	多智能体强化学习前沿 MARL Advanced Topic
13. 课程考核 Course Assessment	
	(①考核形式 Form of examination; ②分数构成 grading policy; ③如面向本科生开放, 请注明区分内容。 If the course is open to undergraduates, please indicate the difference.) 课程项目 60%+期末考试 40% Projects 60% + Final Exam 40%
14. 教材及其它参考资料 Textbook and Supplementary Readings	
	[Book] An Introduction to Reinforcement Learning [Book] Algorithms for Reinforcement Learning