

第五届智能体及多智能体系统研讨会

2018年8月7日·济南

08:30-08:35	开幕式	
特邀报告		
08:35-09:15	特邀报告 1: 人工智能的三重空间: 知识、数据与现实 报告人: 陈小平 教授 中国科学技术大学	主持人: 高阳、唐平中
09:15-09:55	特邀报告 2: 最优拍卖机制设计: 简单性与鲁棒性 报告人: 陆品燕 教授 上海财经大学	
09:55-10:10	茶歇	
特邀报告		
10:10~10:50	特邀报告 3: 数海淘金: 数据定价的挑战与初探 报告人: 吴帆 教授 上海交通大学	主持人: 王崇骏、蒋崑川
10:50-11:30	特邀报告 4: 面向视觉感知的大规模深度 Q 网络方法 报告人: 刘全 教授 苏州大学	
论文报告		
11:30-12:10	5 位讲者汇报最近 1-2 年发表在顶级会议或顶级刊物上的研究成果	主持人: 章宗长
12:10~13:30	午餐	
特邀报告		
13:30-14:10	特邀报告 5: 关于群体智能的基本观点和初步应用 报告人: 张伟 副教授 北京大学	主持人: 曾一锋、俞扬、郝建业
14:10-14:50	特邀报告 6: 强化学习方法在自然语言处理中的应用 报告人: 赵立 副研究员 微软亚洲研究院	
14:50-15:30	特邀报告 7: 面向海量智能体系统的深度强化学习技术 报告人: 张伟楠 助理教授 上海交通大学	
15:30-15:50	茶歇	
博士论坛		
15:50-16:20	先验相关拍卖中的信息及鲁棒性 报告人: 曾驭龙 博士 清华大学	主持人: 卓汉逵、赵登吉
16:20-16:50	面向合作竞争博弈的强化学习方法及应用 报告人: 张启超 博士 中国科学院自动化所	
论文报告		
16:50-17:30	5 位讲者汇报最近 1-2 年发表在顶级会议或顶级刊物上的研究成果	主持人: 吴锋
颁奖		
17:30-18:00	学组几个奖项的颁奖 学组发展讨论	主持人: 于剑、毛新军

论文报告（11:30-12:10）

1. 俞扬，南京大学，Stabilizing Reinforcement Learning in Dynamic Environment with Application to Online Recommendation
2. 卓汉逵，中山大学，Extracting Action Sequences from Texts Based on Deep Reinforcement Learning
3. 李斯源，清华大学，Context-Aware Policy Reuse
4. 杨尚东，南京大学，An Optimal Algorithm for the Stochastic Bandits with Knowing Near-optimal Mean Reward
5. 杨耀东，天津大学，Recurrent Deep Multiagent Q-Learning for Autonomous Brokers in Smart Grid

论文报告（16:50-17:30）

6. 潘颖慧，江西财经大学，Data-Driven Approaches to Multiagent Sequential Decision Making through Interactive Dynamic Influence Diagrams
7. 王万元，东南大学，Strategic Social Team Crowdsourcing: Forming a Team of Truthful Workers for Crowdsourcing in Social Networks
8. 蔡庆芑，清华大学，Reinforcement Mechanism Design for E-commerce
9. 程浩，南京大学，Optimal Constraint Collection for Core-Selecting Path Mechanism
10. 张雯，上海科技大学，Diffusion Mechanism Design

【报告人】：陈小平 教授



【报告题目】：人工智能的三重空间：知识、数据与现实

【报告摘要】： **

【报告人简介】：陈小平博士，现任中国科学技术大学机器人实验室主任，机器人技术标准创新基地主任，计算机学院教授，中国 RoboCup 委员会主席，国际 RoboCup 联合会理事。曾任 2015 世界人工智能联合大会机器人领域主

席、2015 和 2008RoboCup 机器人世界杯及学术大会主席。国务院政府特殊津贴获得者，中科大“杰出研究”校长奖获得者。多次获得最佳论文奖，包括年 IEEE ROBIO 2016 最佳大会论文奖等。提出基于“开放知识”的机器人智能技术路线，并在“可佳”和“佳佳”智能机器人系统中进行了持续性研究和工程实现。团队自主研发的“可佳”机器人 2015 年获国际服务机器人精确测试第一名，2014 年获国际服务机器人标准测试第一名，2013 年获第 23 届世界人工智能联合大会最佳自主机器人奖和通用机器人技能奖，2012 年入选 JHRI 全球人-机器人互动 5 个代表性大型系统之一。2005 年以来，带领中科大机器人团队在机器人世界杯上先后获得 12 项世界冠军，三次排名金牌榜第一。

【报告人】：陆品燕 教授



【报告题目】：最优拍卖机制设计：简单性与鲁棒性

【报告摘要】：设计最优拍卖机制是微观经济学中一个非常重要的课题，特别是网络经济的发展让这个课题更加具有现实意义。经济学中关于这个课题最重要的结果就是著名的 Myerson 最优拍卖理论，Myerson 因此而获得了诺贝尔经济学奖。但这个漂亮的经济学理论在现实中的使用却非常少，主要有两个原因：一是这个最优拍卖机制比较复杂；二是这个机制的最优性有很严格的数学假设条件，这些条件不一定在现实中满足。针对这些问题，最近十几年里在理论计算机界有一系列关于最有拍卖机制设计的工作，主要突出机制的简单性与鲁棒性。本报告会综述这方面的工作并展望未来的方向。

【报告人简介】：陆品燕，上海财经大学信息学院教授，副院长，理论计算机科学研究中心主任。2009 年 1 月于清华大学计算机系获博士学位后加入微软亚洲研究院，历任理论组副研究员，研究员，主管研究员。2015 年 12 月全职加盟上海财经大学，领衔组建理论计算机科学研究中心，经过两年时间的建设，他的研究中心在 CS Rankings 上算法与复杂性、计算经济学两个方向已经排到亚洲第一名、世界第十五名。他的主要研究方向是理论计算机，并注重与其它学科的交叉，包括自然科学中的统计物理以及社会科学中的经济学与社会选择理论等。有 60 余篇科研论文在 STOC、FOCS、SODA、EC 等顶级计算机理论及博弈论的国际会议和杂志发表，荣获 ICALP 2007、FAW 2010、ISAAC 2010 等重要国际会议最佳论文奖。2010 年曾受丘成桐先生邀请在第五届国际华人数学家大会 (ICCM) 上作 45 分钟的大会报告。担任 FAW-AAIM 2012、WINE 2017、FAW 2018 等国际会议程序委员会联合主席，以及多次担任 STOC, FOCS, ICALP 等顶级国际会议的程序委员会委员。曾荣获上海市拔尖青年 (2017)、中国计算机学会青年科学家 (2014)、微软金星员工奖 (2010)、微软学者 (2008)、清华大学特等奖学金 (2007) 等荣誉。

【报告人】：吴帆 教授



【报告题目】：数海淘金：数据定价的挑战与初探

【报告摘要】：被喻为 21 世纪新金矿的大数据，其内在价值已经得到了广泛的关注，并已经成为世界各国竞相发展的关键领域。然而，现有的数据大都被其拥有者内部分析和使用，缺乏流通、共享，形成了大量的数据孤岛，成为大数据产业发展的瓶颈。因此需要开放的数据交易平台来促进数据在互联网上的交易和流通，进一步挖掘大数据的经济价值，发现各类数据背后的应用潜力。本报告将从计算机科学研究的角度介绍数据商品展现出的新特性，以及对数据商品进行定价的技术挑战和开放问题。最后以感知数据为例介绍数据商品的构造方法和动态适应市场供需变化的数据定价机制。

【报告人简介】：吴帆博士，上海交通大学计算机科学与工程系教授、博导、副系主任。2004 年获南京大学学士学位，2009 年获美国纽约州立大学布法罗分校博士学位，2009 至 2010 年在美国伊利诺伊大学香槟分校担任博士后研究员，2010 年加入上海交通大学计算机科学与工程系。在无线网络与移动计算、博弈论算法与应用等领域取得了一系列科研成果，已发表学术论文 150 余篇，论文发表在 JSAC、TON、TMC、TPDS、TOC 等国际著名期刊，以及 MOBICOM、MobiHoc、CoNEXT、INFOCOM、ICDE、VLDB 等重要国际会议。曾获教育部自然科学奖一等奖（排名第三）、国家自然科学基金优秀青年科学基金、ACM 中国新星奖、CCF-腾讯犀牛鸟卓越奖、CCF-Intel 青年学者提升计划等。

【报告人】：刘全 教授



【报告题目】：面向视觉感知的大规模深度 Q 网络方法

【报告摘要】：深度强化学习是机器学习领域中一个新的研究热点。它以一种通用的形式将深度学习的感知能力与强化学习的决策能力相结合，并通过端对端的方式学习从原始输入到动作输出的一个映射。在许多基于视觉感知的大规模决策任务中，深度强化学习方法已经取得突破性的进展。其中深度 Q 网络方法在解决一类视频游戏任务时表现出了和人类玩家相媲美的水平。然而在一些现实场景下的复杂问题中，深度 Q 网络会面临奖赏的稀疏和延迟、部分状态可观察、收敛速度慢、性能不稳定等一系列问题。报告针对上述问题，在目前深度强化学习研究的基础上，从训练算法和模型架构等方面对深度 Q 网络方法进行了改进和完善，并提出基于优先级采样深度 Q 学习算法、基于视觉注意力机制的深度循环 Q 网络模型、基于混合目标 Q 值的深度确定性策略梯度方法等高效的深度强化学习方法，并在实验效果上进行了对比。

【报告人简介】：刘全，男，1969 年 10 月生，教授，博士生导师。毕业于吉林大学计算机软件与理论专业，获博士学位，南京大学软件新技术国家重点实验室博士后。近年来，共主持了“石油勘探开发分布式数据库建设”、“基于 tableau 的逻辑强化学习研究”等国家级项目 8 项，主持了“采油工程信息可视化系统”、“油田信息处理网络系统”、“基于核方法的强化学习应用研究”等省部级和市（局）级科研项目 10 多项，获省部级科技进步奖 2 项，市（局）级科技进步奖 8 项。先后在国内外学术会议及《计算机学报》、《软件学报》等国内核心期刊和统计源期刊上发表论文 80 余篇，其中 30 篇被 SC1 检索，40 篇被 EI 检索。目前主要研究方向为：深度学习、强化学习、统计软件工程、分布式数据库检索及修正、网络信息安全及处理等。

2012 年获江苏省教工委优秀共产党员称号。2011 年、2012 年入选江苏省“六大人才”、江苏省“333”人才培养计划。目前为《通信学报》编委。E-mail: quanliu@suda.edu.cn

【报告人】：张伟 副教授



【报告题目】：关于群体智能的基本观点和初步应用

【报告摘要】：每一个人类个体都是有价值的。人类文明的进一步发展不能建立在适者生存这种原始性的淘汰机制上。互联网技术的持续发展正在对人类社会的各个方面产生深远的影响。具体而言，互联网为人类个体之间的交互和协作提供了一种全新的基础设施，促成了各种群体智能现象在信息空间的蓬勃发展。如何有效地利用群体的智慧和能力进行复杂问题的求解，对于中国这样一个人口大国的持续发展、对于人类文明的持续发展，都具有重要的理论和现实意义。本报告将主要介绍报告人对于互联网环境下人类群体智能的一些基本思考以及对群体智能的初步探索性应用，希望能够对相关领域的研究产生一定的启发作用。

【报告人简介】：张伟，男，2006 年在北京大学获“计算机软件与理论”专业理学博士学位，毕业后在北京大学留校

工作至今，现为北京大学信息科学技术学院副教授。前期研究方向主要涉及软件复用、软件需求工程，近几年主要关注基于群体智能的软件开发方法。研究工作得到了多项国家自然科学基金项目和国家重点基础研究发展计划项目的资助。

【报告人】：赵立 副研究员



【报告题目】：强化学习方法在自然语言处理中的应用

【报告摘要】：在自然语言处理的各种应用中，我们会经常遇到各种显式或者隐式的结构。自然语言处理中的结构，通常是离散的，难以通过连续的方法去优化的，因此难以用传统的方法求解。随着 AlphaGo 的成功，基于强化学习特别是深度强化学习的应用工作也受到越来越多的关注，基于强化学习的自然语言处理也成为了一个研究热点。本报告将在几个不同的自然语言处理的任务上，介绍我们是如何利用强化学习来解决自然语言处理应用中的结构优化问题。

【报告人简介】：赵立，微软亚洲研究院机器学习组副研究员。2011 年 6 月在清华大学计算机科学与技术系获得学士学位，2016 年 6 月在清华大学计算机科学与技术系获得博士学位。

主要研究内容包括：深度学习，强化学习，自然语言处理，情感分析。在国际顶级会议上发表了多篇论文（包括：IJCAI, AACL, EMNLP, CIKM 等）。

【报告人】：张伟楠 助理教授



【报告题目】：面向海量智能体系统的深度强化学习技术

【报告摘要】：近年来，机器学习的落地场景有两个发展方向，一是从预测到决策的范式拓展，另一个则是从单智能体到多智能体的场景推广。由此，面向多智能体系统的深度强化学习开始越来越受到学术界和工业界的关注。本次报告，我将从多智能体深度强化学习的几个落地任务切入，由此引入它的基本数学定义和几个经典解决方法。之后我将深入海量智能体场景下的不同场景，探讨在海量智能体的情况下，传统多智能体强化学习方法的不足，并深入介绍基于平均场理论的强化学习和基于因子分解模型的强化学习算法。最后，我将介绍 **MAgent**，一个专为海量智能体场景提供模拟实验的平台，并展示上述算法在该平台上的初步实验效果。

【报告人简介】：张伟楠现于上海交通大学计算机系和约翰·霍普克罗夫特研究中心担任助理教授，研究方向为深度强化学习、无监督学习及其在数据挖掘问题中的应用。他于 2011 年毕业于上海交通大学计算机系 ACM 班，于 2016 年获得英国伦敦大学学院计算机系博士学位，研究成果在国际一流的会议和期刊上发表 50 篇论文，其中 5 次以第一作者身份在 ACM 国际数据科学会议 KDD 上发表；2016 年获得由微软研究院评选的“全球 SIGKDD Top 20 科研新星”称号；2017 年获得 ACM 国际信息检索会议 SIGIR 的最佳论文提名奖；2017 年获得上海 ACM 新星奖。他曾在 KDD-Cup 用户个性化推荐大赛获得全球季军，在全球大数据实时竞价展示广告出价算法大赛获得最终冠军。此外他也曾在谷歌硅谷总部、微软剑桥研究院、微软亚洲研究院做人工智能和大数据挖掘方向的研究实习。

【报告人】：曾驭龙 博士



【报告题目】：先验相关拍卖中的信息及鲁棒性

【报告摘要】：单物品拍卖是计算经济学领域最为基础和重要的课题之一。其中，贝叶斯设定是拍卖理论的重要元素。它假定每个出价者的类型（对物品的估值）服从某个先验分布。过去的四十年里对这方面的研究取得了大量成就，其中迈尔森拍卖，也叫单物品最优拍卖，被广泛的应用于实践中。但另一方面，正如威尔逊原理所指出的，一个好的机制应当尽可能少的依赖于先验类型分布。此报告研究关联实践的情形下，作为出价者如何利用先验信息来增加他的收益。报告提出了三个定量的模型，用于分析在传统先验相关的模型中，如迈尔森拍卖，出价者的最优策略。这些结果也能够反映出此类拍卖的鲁棒性。

【报告人简介】：曾驭龙，2018 年毕业于清华大学交叉信息研究院，现在在星云链公司担任研究员。目前已在国际顶级会议如 EC, AAI, AAMAS 上以第一作者身份发表多篇论文。目前主要研究方向为机制设计、计算机经济学中的拍卖理论。

【报告人】：张启超 博士



【报告题目】：面向合作竞争博弈的强化学习方法及应用

【报告摘要】：近年来，通过与神经网络逼近器的结合，一些强化学习方法在求解复杂未知环境的单个体优化控制问题方面取得了良好的效果。值得注意的是，目前大多数的现代控制系统往往包含两个或多个控制单元甚至是多个子系统，这类优化控制问题可看作为非线性的合作或竞争博弈问题，需要求解相应的纳什均衡解。针对零和博弈、非零和博弈和完全合作博弈问题如何设计相应的强化学习算法以提高数据的利用率，节省通信资源，减轻计算负担是值得深入探讨的研究问题。随着 AlphaGo 的成功，基于强化学习特别是深度强化学习的应用工作也受到越来越多的关注，基于强化学习的智能车的横纵向控制也成为了一个研究热点。本报告将介绍面向合作竞争博弈的强化学习优化控制方法以及在智能驾驶决策控制领域的初步尝试。

【报告人简介】：张启超，中国科学院自动化所，助理研究员，2017年毕业于中科院自动化所控制理论与控制工程专业获博士学位，同年留所工作。目前已在国际顶级期刊和会议如 IEEE TNNLS、IEEE TCYB、IEEE TSMC:A、IEEE CIM 上以第一作者身份发表十多篇论文，一篇论文为 ESI 高被引和热点论文，相关成果已申请发明专利 3 项，获得中科院院长优秀奖、北京市优秀毕业生，IEEE CIS Student Research Grant (同年全球 5 位)等荣誉。目前主要研究方向为强化学习、博弈论机器在智能驾驶领域的应用。