

多智能体系统研究的历史、现状及挑战

关键词：智能体 多智能体系统 人工智能

多智能体系统 (multi-agent systems) 由分布式人工智能演变而来, 其研究目的是解决大规模、复杂、实时和有不确定信息的现实问题, 而这类问题是单个智能体所不能解决的。多智能体系统通常具有自主性、分布性、协调性等特征, 并具有自组织能力、学习能力和推理能力。对多智能体系统的研究既包括构建单个智能体的技术, 如建模、推理、学习及规划等, 也包括使多个智能体协调运行的技术, 例如交互通信、协调、合作、协商、调度、冲突消解等。随着互联网技术的发展, 多智能体系统中往往会出现一些“自私”的智能体 (如电子商务市场的交易方), 因此需要引入博弈论来分析智能体的交互策略, 其研究内容包括电子商务、拍卖、机制设计、社会选择理论等。

经过近 30 年的发展, 多智能体系统已经成为国际人工智能领域的前沿和研究热点。在近年

来的 AAAI 人工智能会议 (AAAI Conference on Artificial Intelligence) 和国际人工智能联合会议 (International Joint Conference On Artificial Intelligence, IJCAI) 上,

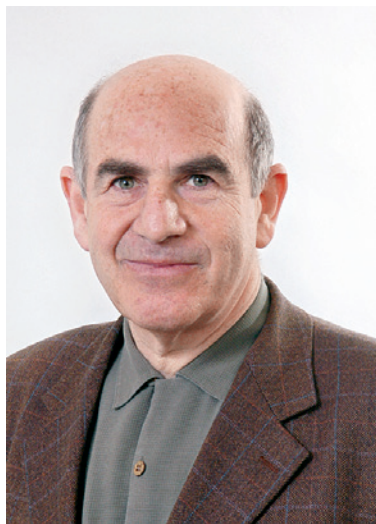


图1 多智能体系统创始人维克托·莱瑟

录用的关于多智能体系统的文章数量一直名列前两位。多智能体系统领域的创始人维克托·莱瑟 (Victor Lesser) 教授于 2009 年获得了 IJCAI 杰出研究奖 (IJCAI

安波^{1,2} 史忠植²

¹新加坡南洋理工大学

²中国科学院计算技术研究所

Award for Research Excellence)。自 1995 年以来, IJCAI 计算机与思维奖 (IJCAI Computers and Thought Award) 的获奖者有一半以上来自多智能体系统领域。

多智能体系统研究的历史

自 1956 年约翰·麦卡锡 (John McCarthy) 在著名的达特茅斯研讨会上提出“人工智能”这一概念后, “智能体”的概念便开始兴起。例如, 阿兰·图灵 (Alan Turing) 提出了用来判断一台机器是否具备人类智能的“图灵测试”。在此测试中, 测试者通过监视设备向被测试的实体 (也就是我们现在所说的智能体) 提问。根据图灵的观点, 如果测试者无法区分被测试对象是计算机还是人, 那么被测试对象就是智能的。人工智能泰斗马文·明斯基在他的《心智社会》这本书中将智能体描述为实现人类智能的

基本模块。尽管智能体的概念很早就已经出现,但在20世纪70年代之前将多个智能体作为一个功能上的整体(即能够独立行动的自主集成系统)进行研究的做法却很少。一些研究着眼于构建一个完整的智能体或多智能体系统,比如Hearsay-II语音理解系统^[1]、STRIPS规划系统^[2]、Actor模型^[3]等。1978年,美国国防部高级研究计划署(DARPA)在卡耐基梅隆大学举办分布式传感器网络研讨会,这是最早的此类会议。1980年,分布式人工智能领域的首次研讨会在麻省理工学院举办。在会议上,研究人员讨论了分布式问题求解、多智能体规划、组织控制、合同网、协商、分布式传感器网络、功能精确的协作分布式系统、大规模行为者模型以及智能体规范逻辑框架等重要的多智能体系统研究问题。自此,集成智能体构建和多智能体系统研究的各个分支领域都有了较大发展。

20世纪80年代后期,随着《分布式人工智能》(*Distributed Artificial Intelligence*)和《分布式人工智能教程》(*Readings in Distributed Artificial Intelligence*)的出版^[4,5],分布式人工智能领域开始显著扩张,建立在博弈论和经济学概念之上的自私智能体交互的研究也逐步兴盛起来^[6]。随着协作型和自私型智能体研究的

交融,分布式人工智能逐渐演变并最终有了一个包罗万象的新名字——多智能体系统。

90年代初期,多智能体系统研究快速发展,智能体的出现催生了许多软件技术。面向智能体的编程范式(Agent-Oriented Programming, AOP)将软件的构建集中在软件智能体的概念之上。相比面向对象的编程以对象(通过变量参数提供方法)为核心的理念,AOP以外部指定的智能体(带有接口和消息传递功能)为核心。很快,一批移动智能体框架以Java包的形式开发和发布,2000年国内开展网构软件的研究^[7],2002年开展了构建智能体网格的工作^[8]。

由于90年代初期开发的KQML和KIF这两种颇具影响力的智能体通信语言一直没有被正式标准化,因此一些研究人员认为智能体技术可能会因为国际标准的缺乏而难以得到广泛认可。90年代中期,工业界对智能体系统的关注集中在标准化这一焦点上。FIPA¹在1995年开始多智能体方面的标准化工作,它的核心目标是制定一套用于智能体协作的语言。

90年代后期,研究人员开始在不断拓宽的现实领域寻求多智能体系统的新发展。机器人世界杯(RoboCup)应运而生。举行RoboCup比赛的目标是,在50

年内,产生出一支能够战胜具有世界杯水准的人类球队的机器人足球队。其出发点是,从理论上说一支出色的球队需要一系列技能,比如利用有限的带宽进行实时动态的协作,机器人球队如果能战胜人类球队,标志多智能体系统在技术和理论上有根本的突破。参加RoboCup的热度在世纪之交暴涨,定期举办的RoboCup锦标赛吸引了来自世界各地的数百支参赛队伍。2000年,RoboCup推出了一项新的活动——RoboCup营救。这项活动以1995年发生的日本神户大地震为背景,目标是建立一支能够通过相互协作完成搜救任务的机器人队伍。

在世纪之交,以智能体为媒介的电子商务成为智能体技术的最大单一应用场合,为智能体系统在谈判和拍卖领域的发展提供了巨大推动力。2000年以后,以促进和鼓励高质量的交易智能体研究为目的的交易智能体竞赛(Trading Agent Competition)推出并吸引了很多研究人员参与。

自2001年以来,多智能体系统的思想对主流计算机科学产生了深远影响,拍卖和机制设计等研究领域已经跻身理论计算机科学的主要课题。这些突破性进展首先缘于易趣(eBay)等在线拍卖行的巨大成功,但更普遍的是许多有趣的、重要的组合问题都可以表示为拍卖。对博弈论和计

¹ FIPA(The Foundation for Intelligent Physical Agents)是一个由活跃在智能体领域的公司和学术机构组成的国际组织,其目标是为异质的智能体和智能体系统之间的互操作而制订相关的软件标准。

算机科学之间关联性的研究同样高涨。2008年,国际博弈论学会特意为最佳博弈论与计算机科学交叉研究设立了一个新的奖项。在多智能体系统领域涌现并得到应用的思想还包括:社会科学领域复杂分布式系统建模的多智能体模拟技术、形式化描述分布式系统的DEC-POMDP以及个人协助技术。同时下一代分布式传感器、网络化自动驾驶车辆和机器人系统、自主计算以及网络服务等未来多智能体技术更广阔的应用场景正展现在我们面前。

2014年AAMAS年会

1995年,第一届国际多智能体系统会议(International Conference on Multi-Agent Systems, ICMAS)在美国旧金山举办^[9]。紧接着,智能体理论、架构和语言国际研讨会(International Workshop on Agent Theories, Architectures, and Languages, ATAL)和国际自主智能体会议(International Conference on Autonomous Agents, AA)也相继举办。2002年,国际智能体及多智能体系统协会(International Foundation for Autonomous Agents and Multiagent Systems, IFAAMAS²)将上述三个会议合并为智能体及多智能体系统国际会议(International Joint Conference on Autonomous

Agents and Multi-Agent Systems, AAMAS)。第13届AAMAS于2014年5月5~9日在法国巴黎举行,吸引了全世界800多名研究人员参会。

AAMAS 2014设置了一个主要议题——多智能体系统研究,4个特别议题——创新应用、机器人、虚拟智能体和创新思维。会议共收到投稿709篇,其中60%的论文第一作者是学生,会议最终录用了169篇全文,录取率为23%。从录用论文的关键词来看,博弈论以及相关领域的研究最受关注,紧随其后的是智能体合作和智能体推理。

AAMAS 2014共设28个专题研讨会,既包括相对传统的多智能体优化、增强学习、合作博弈论、电子商务、多智能体模拟、序贯决策、自动协商等,也包括最近兴起的面向可持续发展的多智能体系统、人与智能体交互等热点问题。

ACM/SIGART智能体研究奖(ACM/SIGART Autonomous Agents Research Award)是多智能体系统领域最重要的学术成就奖。本届会议的获奖者是美国密歇根大学的迈克尔·韦尔曼(Michael Wellman)教授,以表彰他将经济学应用到多智能体系统研究的开创性贡献。

IFAAMAS维克托·莱瑟杰出博士论文奖(Distinguished Dis-

sertation Award)旨在奖励全世界多智能体系统领域最优秀的博士论文。2013年杰出博士论文奖得主是美国南加州大学的曼尼什·简(Manish Jain)博士,表彰其在安全博弈论方向的研究及应用。

IFAAMAS有影响力论文奖(Influential Paper Award)于2006年创立,旨在奖励10年以前发表的对多智能体系统理论及应用产生巨大影响的论文。今年该奖颁发给了IBM以色列研究院的安·史哈瑞(Onn Shehory)博士和巴伊兰大学的克劳斯·萨里塔(Sarit Kraus)教授,以表彰他们于1998年在《人工智能》上发表的关于组织形成(coalition formation)的论文。

研究热点

算法博弈论

算法博弈论是博弈论和算法设计的交叉领域。算法博弈论中算法的输入通常来自许多分布的参与者,这些参与者对输出有不同的个人偏好。智能体参与者可能会因为个人利益而隐瞒真实的输入信息。除了经典算法设计理论要求的多项式运算时间及较高近似度之外,算法设计者还需要考虑约束智能体的动机。显然,算法博弈论的研究有两个关键。

1. 分析:参照现有算法,利

² 一个以提供统一的、高质量的,并广受国际关注的智能体和多智能体系统理论和实践研究论坛,以促进人工智能、智能体与多智能体系统领域科技发展的非盈利机构。

用博弈论工具对其进行计算均衡、证明均衡性质、无序代价(price of anarchy)、最优动态反应等方面的分析。

2. 设计: 设计具备良好均衡和算法属性的博弈, 即算法机制设计。

自1999年尼桑(Nisan)和罗恩(Ronen)在美国计算机学会计算理论会议(ACM Symposium on Theory of Computing, STOC)^[10]上将理论计算机学术界的目光引向针对自私用户的算法设计上, 算法博弈论便开始兴起。雅虎、谷歌以及百度等从搜索广告拍卖获得的巨额收入掀起了对算法博弈论研究的高潮。算法博弈论当前的研究热点包括算法机制设计、计算社会选择理论、(社交)网络分析、均衡分析等。

安全博弈论及其应用

保护关键公共基础设施和目标是各国安全机构肩负的一项极具挑战性的任务。有限的安全资源使得安全机构不可能在任何时候都提供全面的安全保护。此外, 安全机构的对手可以通过观察来发现安全机构保护策略的固定模式和弱点, 并据此来选择最优的攻击策略。一种降低对手观察侦查能力的方式是随机调度安全机构的保护行为, 这种方法广泛用于警察巡逻、行李检测、车辆检查以及其他安全程序。然而, 安全机构在进行有效的随机安全策略调度时面临许多困难^[11], 特别是有限的安全资源无法无处不在或每时每刻提供安全保护。安

全领域资源分配的关键问题是设计出配置有限安全资源的最优方案, 以获取最佳的安全保护效果。

博弈论提供了一个恰当的数学模型来研究对有限安全资源的配置, 以最大限度地提高资源分配的有效性。安全博弈模型基于20世纪30年代的斯塔克尔伯格(Stackelberg)博弈模型, 自科尼策(Conitzer)和桑德霍尔姆(Sandholm)2006年的经典论文^[12]发表后迅速发展起来。安全博弈论初期研究的主要参与者包括南加利福尼亚大学米兰·丹部(Milind Tambe)教授领导的TEAMCORE研究小组。现在越来越多的学者参与到此项研究中, 近100篇论文发表在人工智能领域的顶级会议上, 如AAMAS, AAAI和IJCAI。过去几年, 博弈论在安全领域的资源分配及调度方面的理论——安全博弈论逐渐建立, 并在机场检查站的设置及巡逻调度、空中警察调度、海岸警卫队的巡逻、机场安保、市运输系统安全等领域得到成功应用。但其依然面临很多挑战, 包括提高现有的安全博弈算法的可扩展性、提高安全资源分配策略的鲁棒性、对恐怖分子的行为建模、协同优化、多目标优化等^[13]。

多智能体学习

多智能体学习(Multi-Agent Learning, MAL)将机器学习技术引入多智能体系统领域, 研究如何设计算法来创建动态环境下

的自适应智能体。强化学习(reinforcement learning)是多智能体学习领域广泛研究的技术。单智能体的强化学习在马尔科夫决策过程(Markov Decision Processes, MDP)的框架内能被较好地描述。一些独立的强化学习算法(如Q-learning)在智能体所处环境满足马氏性且智能体能够尝试足够多行动的前提下会收敛至最优的策略。尽管马尔科夫决策过程为单智能体学习提供了可靠的数学框架, 但对多智能体学习却并非如此。在多个自适应智能体相互作用的情况下, 一个智能体的收益通常依赖于其他个体的行动, 学习环境不是静态的。此时每个智能体面临一个目标不断变化的问题——单个智能体需要学习的内容依赖于其他智能体学到的内容, 并随之改变。因此, 有必要对原有的马尔科夫决策过程框架做相应的扩展, 包括马尔科夫博弈和联合行动学习机等^[14,15]。在扩展中, 学习发生在不同智能体的状态集和行动集的积空间上。因而当智能体、状态或行动的数量太大时, 这些扩展面临积空间过大的问题。此外, 共享的联合行动空间也未必可用。比如在信息不完全的情况下, 智能体未必能观察到其他智能体的行动。如何处理复杂的现实问题, 如何高效地处理大量的状态、大量的智能体以及连续的策略空间, 已经成为目前多智能体学习研究的首要问题。多智能体学习需要建立在可扩展规模的理论之上。在可

扩展的框架下,多智能体学习算法能够适应各种规模的智能体系统。到目前为止,多智能体学习领域通常处理的都是比较简单的问題,或采用初级的实验性场景,或取自博弈论并且只包含很少的(一般为两个)学习智能体。此外,多智能体学习领域需要关注更加复杂、更加实际的应用,如地面和空中交通管控、分布式监测、电子市场、机器人营救和机器人足球赛、智能电网等一系列实际应用场合都是多智能体学习的用武之地。

分布式问题求解

作为多智能体系统的一个子领域,分布式问题求解着眼于让多个智能体共同解决一个问题。这些智能体通常都是合作性的。在众多分布式问题求解模型中,分布式约束推理(Distributed Constraint-Reasoning, DCR)模型,如分布式约束满足问题(Distributed Constraint-Satisfaction Problem, DCSP)和分布式约束优化问题(Distributed Constraint-Optimization Problem, DCOP)的使用和研究较为广泛。分布式约束推理模型历史悠久,在各种分布式问题上都有应用,包括分布式会议安排和分布式传感器任务分配。自20世纪90年代中期以来,分布式约束满足问题和分布式约束优化问题算法设计(包括完全的和非完全的)得到学术界的广泛关注。根据搜索策略(最佳优先搜索还是深度优先分支定

界搜索)、智能体间同步类型(同步还是异步)、智能体间通讯方式(约束图邻居间点对点传播方式还是广播方式)以及主要通讯拓扑结构(链式还是树形)的不同,可以对众多分布式约束推理算法进行归类。例如,ADOPT^[16]采用的是最佳优先搜索、异步同步、点对点的通讯和树形通讯拓扑结构。

近年来,研究人员在多个方面延伸了分布式约束推理模型和算法,从而能够更精确地建模和求解现实问题^[17]。使用分布式约束推理来对分布式会议安排这类问题建模的目的是隐私保护。例如当两名用户参加的不是同一个会议时,他们应当无权知悉对方对于会议时间的偏好。然而,在实际生活中事实并非如此,用户隐私经常被泄露。为此,研究人员引入一些指标去衡量分布式约束推理的隐私泄露程度,并设计新算法来更好地保护隐私。分布式约束推理模型可以扩展为动态分布式约束推理模型,如动态分布式约束满足问题和动态分布式约束优化问题。一个典型的动态分布式约束推理问题由一连串静态的分布式约束推理问题组成,问题彼此不完全相同。其他相关的扩展还包括智能体以截止时间决定价值的连续时间模型、智能体对所处环境认知不完全的模型。研究人员还将分布式约束优化问题扩展为多目标分布式约束优化问题和资源受限的分布式约束优化问题。

分布式规划

不确定性是多智能体系统研究面临的最大挑战。即便在单智能体系统中,行动的输出也会存在不确定性(比如行动有失败的几率)。此外,在许多问题中,环境的状态也会因噪声或传感器能力所限而具有不确定性。在多智能体系统中,这些问题更加突出。一个智能体只能通过自己的传感器来观测环境状态,因此智能体预知其他智能体动向的能力十分有限,合作也因此变得复杂。如果这些不确定性不能得到很好的处理,各种糟糕的情况都有可能出现。理论上,智能体通过相互交流和同步信念协调行动。然而,由于带宽的限制,智能体不可能将必要的信息广播至其他所有智能体。此外,在许多实际场景中,通讯是不可靠的,对其他智能体行动的感知也不可避免地具有不确定性。

近年来,大量的研究关注于寻求处理多智能体系统不确定性的规范方法,建立了各种模型和求解方法。分布式马尔科夫决策过程(Decentralized Markov Decision Process, Dec-MDP)和分布式部分可观测马尔科夫决策过程(Decentralized Partially Observable Markov Decision Process, Dec-POMDP)是在不确定性情形下多智能体规划建模最为常用的两种模型。不幸的是,求解Dec-POMDP(即计算最佳规划)通常是很难的(NEXP-complete)^[18],

即便是计算绝对误差界内的解也是 NEXP-complete 的问题^[19]。特别要强调的是,联合策略的数量随智能体数量和观察次数呈指数增长,随问题规模呈双重指数增长。尽管很难找到对所有问题都行之有效的办法,但开发更好的 Dec-POMDP 优化求解方法依然是一个至关重要的问题。

未来展望:多智能体“图灵测试”

在 AAMAS'13 上,美国工程院院士、哈佛大学芭芭拉·格罗兹 (Barbara Grosz) 教授作了题为《多智能体系统的“图灵挑战”》(A Multi-Agent Systems “Turing Challenge”) 的报告,她提出了“多智能体图灵测试”问题:一群智能体是否能够在动态、不确定的环境中表现得像一群人一样。此问题可以看作经典图灵测试的扩展。尽管过去几十年已经有很多关于多智能体协同、合作以及处理对抗性或策略性环境的研究成果,但当前的多智能体技术还难以应对复杂情况下的新挑战。我们需要更善于与人一起工作的智能体,并且这样的需求越来越强烈和普遍。未来的多智能体系统研究将继续关注规划、学习、协调、机制设计、人机交互等众多理论问题。同时,未来的研究将面向安全、可持续发展、医疗、老龄化等现实挑战。■

致谢:感谢维克托·莱瑟教

授及迈克尔·伍尔德里奇 (Michael Wooldridge) 教授提供多智能体系统发展的历史资料,同时感谢甘家瑞对文献的整理。



安波

CCF会员。新加坡南洋理工大学南洋助理教授,中科院计算所副研究员。主要研究方向为多智能体系统、博弈论等。boan@ntu.edu.sg



史忠植

CCF会士。中国科学院计算技术研究所研究员。主要研究方向为智能科学、人工智能。shizz@ics.ict.ac.cn

参考文献

- [1] Fennell, R. D., and Lesser, V. R.. Parallelism in AI problem solving: a case study of HEARSAY-II. *IEEE Transactions on Computers - Special Issue on Multiprocessors*, 1977; C-26, 2, 98~111.
- [2] Fikes, R., and Nilsson, N. J. STRIPS: a new approach to the application of theorem proving to problem solving. *Artif. Intell.*, 1971; 2(3/4): 189~208.
- [3] Hewitt, C. Viewing control structures as patterns of passing messages. *Artif. Intell.*, 1977; 8(3): 323~364.
- [4] Huhns, M. N.. *Distributed artificial intelligence*. Pitman Publishing Ltd.. London, England.
- [5] Bond, A. H., and Gasser. *Readings in distributed artificial*

intelligence. Morgan Kaufmann Publishers. San Mateo, CA.

- [6] Rosenschein, J. H., and Genesereth. Deals among rational agents. *IJCAI* 1985: 91~99.
- [7] 杨芙清, 梅宏, 吕建等. 浅论软件技术发展. 电子学报, 2002; 30(12A): 1901~1906.
- [8] Zhongzhi Shi, Mingkai Dong, Haijun Zhang, et al.. Agent-based grid computing. *International Symposium on Distributed Computing and Applications to Business, Engineering and Science*, 2002.
- [9] Lesser, V. R., and Gasser, L.. The First International Conference on Multiagent Systems. San Francisco, California, USA. The MIT Press.
- [10] Nisan, Noam; Ronen, Amir. Algorithmic mechanism design. *Proceedings of the 31st ACM Symposium on Theory of Computing (STOC '99)*, 129~140.
- [11] Tambe, M.. Security and Game Theory: Algorithms, deployed systems, lessons learned. Cambridge University Press. 2011
- [12] Conitzer, V., and Sandholm, T. Computing the optimal strategy to commit to. *Proceedings of the ACM Conference on Electronic Commerce (ACM-EC)*, 2006: 82~90.
- [13] Tambe, M.; Jiang, A.; An, B.; et al... Computational game theory for security: Progress and challenges. AAAI Spring Symposium on Applied Computational Game Theory, 2014.
- [14] Littman, M. Markov games as a framework for multiagent reinforcement learning. *Proceedings of the Eleventh International Conference on Machine Learning*, 1994: 157~163.
- [15] Claus, C., and Boutilier, C.

The dynamics of reinforcement learning in cooperative multiagent systems. *Proceedings of the 15th National Conference on Artificial Intelligence*, Menlo Park, CA: AAAI Press, 1998:746~752.

- [16]Modi, P. J.; Shen, W.; Tambe, M.; and et al.. An asynchronous complete method for distributed constraint optimization.

Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems. ACM Press, 2003:161~168.

- [17]Yeoh, W. and Yokoo, M.. Distributed problem solving. *AI Magazine*, 2012; 33(3): 53~65.
- [18]Bernstein, D. S.; Givan, R.; Immerman, N.; and et al.. The complexity of decentralized

control of Markov decision processes. *Mathematics of Operations Research*, 2002;27 (4): 819~840.

- [19]Rabinovich, Z.; Goldman, C. V.; and Rosenschein, J. S. The complexity of multiagent systems: the price of silence. *The International Conference on Autonomous Agents and Multi Agent Systems*.