



社论:

多智能体系统的体系化和组织化博弈

陆军¹, 王飞跃^{†2}, 董琦^{†‡1}, 魏庆来²

¹中国电子科学研究院, 中国北京市, 100049

²中科院自动化研究所, 中国北京市, 100049

[†]E-mail: feiyue.wang@ia.ac.cn; dongqiouc@126.com

本文编译自 Lu J, Wang FY, Dong Q, et al., 2022. Institutionalized and systematized gaming for multi-agent systems. *Front Inform Technol Electron Eng*, 23(7):991-994. <https://doi.org/10.1631/FITEE.2240000>

广泛存在于军事智能、信息网络、无人系统、智慧交通和智能电网等领域的多智能体博弈问题具有体系化和组织化特点, 它要求多智能体系统在复杂动态环境中进行感知和行动, 同时在系统内达成个体利益的平衡和群体利益的最大化。这其中存在系统结构复杂、博弈环境不确定、决策信息不完整以及结果不可解释等问题, 使得多智能体博弈的研究从传统简单博弈转变为面向高维连续复杂环境的博弈, 对组织化和体系化博弈 (institutionalized and systematized gaming, InSys gaming) 提出迫切需求。在此背景下, 针对多智能体系统的体系化和组织化博弈的研究呈现出以下几方面的研究趋势:

1. 分析多智能体系统的博弈演化规律, 建立多智能体组织化体系化博弈理论模型

组织化和体系化的多智能体系统具有有序、结构化特点, 建立系统博弈模型是必不可少的研究环节。针对政治、军事、经济等体系对抗博弈, 分析系统的内部演化特性和外部交互信息, 建立系统博弈演化模型, 研究博弈演化的要素、关系、准则, 可为该领域的体系设计、决策规划等提供理论支撑。

2. 运用并结合多种人工智能学习算法, 实现

多智能体系统的协同决策

当前主流人工智能学习方法均具备特定场景下的应用优势。在体系化、组织化博弈问题求解中, 可以运用并结合深度学习的环境表征能力与强化学习方法的决策生成能力, 通过构建数字化仿真训练环境, 设计智能体决策算法和非监督训练方法, 实现复杂未知环境下多智能体系统的协同决策生成。

3. 采取分层任务规划与决策行动架构, 降低协同决策算法复杂度随着多智能体系统规模增加, 协同决策面临的节点耦合、观测非确定、交互无序等问题凸显, 其体系化、组织化博弈问题的求解复杂度大幅提升。通过博弈任务分解, 结合长期规划与实时行动决策, 构建多智能体分层算法架构, 可有效降低协同决策算法搜索过程的复杂度, 为解决组织化与体系化博弈提供一种可行思路。

4. 建立算法模型的鲁棒性分析框架, 解决数据驱动方法与实际场景存在的模型偏差

对于数据驱动方法, 当训练数据与实际场景存在模型偏差时, 算法性能将下降。通过研究数据驱动方法的鲁棒性分析框架, 设计更为稳健的算法模型和实际数据微调方法, 减轻训练得到的算法向实际场景迁移时的性能损失, 为数据驱动方法的实际部署提供支撑。

博弈理论已成为战略政治、军事对抗、市场

[‡] 通讯作者

© 浙江大学出版社 2022

经济等领域中解决问题的基本分析框架，所分析对象具备复杂的体系化和组织化特点，得到学术界和产业界高度关注。利用多智能体系统对组织化和体系化博弈建模，结合人工智能方法解决博弈决策问题，为该领域的理论、方法、技术的发展提供新思路。

在此背景下，《信息与电子工程前沿(英文)》期刊组织了“多智能体系统的组织化和体系化博弈”专题。主要内容关于多智能体演化博弈、无人机编队控制、多智能体自主规划、多智能体协同控制、群体智能以及多智能体强化学习框架设计等。经严格评审，录用 8 篇文章，包括 2 篇观点类文章和 6 篇研究论文。

中国电科陆军等从对博弈的认识过程出发，探讨了博弈的存在与实践，阐述了多智能体博弈中的环境复杂多变、系统动态异构、单体计算及感知能力有限等难点问题，提出一种多智能体演化博弈理论框架，并以下一代预警探测系统为例，介绍了多智能体演化博弈的应用实践，对研究高维复杂环境下的组织化、体系化博弈行为具有重要意义。

海军航空大学何友等对多智能体协同与博弈的发展情况进行了概述，从任务挑战、技术方向和应用领域 3 个角度简要回顾了多智能体协同与博弈技术。其中，重点分析了多智能体系统研究中的典型问题和挑战，讨论了关于多智能体协同和博弈任务的一些有前景的研究方向，并对多智能体协同与博弈在游戏等方面的应用进行了展望。

中国科学院自动化所李洪阳、魏庆来提出一种新的输入饱和最优同步控制方法，将最优同步控制问题转化为多智能体非零和博弈问题，并求解具有非二次输入能量项的哈密顿—雅可比—贝尔曼(Hamilton–Jacobi–Bellman, HJB)方程，以实现纳什均衡。同时，提出一种新的无模型非策略强化学习方法，该方法允许迭代控制率在不考虑系统模型信息的情况下收敛到纳什均衡，为饱和输入的多智能体系统同步控制问题提供了方法支撑。

北京航空航天大学段海滨等针对干扰对分布式无人机编队控制的影响，提出一种具有外界干扰和障碍物的分布式无人机编队控制博弈策略。

该策略包括分布式模型预测控制(model predictive control, MPC)架构和基于莱维飞行的鸽子启发式优化(Levy flight based pigeon inspired optimization, LFPIO)。首先，提出一个非奇异的快速终端滑动膜观测器；然后，建立干扰 MPC 架构，每个无人机只与它的邻居交换信息；接着，将无人机编队问题转化为博弈问题并设计成本函数；最后，设计了 LFPIO 并进行数值仿真，通过对比实验验证了基于 LFPIO 的分布式 MPC 的有效性，为无人机编队控制问题的解决提供了新的研究思路。

中国科学技术大学赵鉴等针对仿真场景和现实之间的差距为多智能体强化学习在实际运用中带来的挑战，提出一个具有意外崩溃的合作多智能体强化学习系统的形式化概念。设计了一套虚拟教练员辅助的多智能体强化学习框架，进一步模拟实际操作中智能体可能在协调过程中意外“崩溃”的现象，为解决多智能体系统模拟与现实之间存在差异的问题提供了研究思路。

北京航空航天大学董希旺等针对多智能体微分博弈问题及其在协同同步控制中的应用，提出一种多智能体微分博弈的系统化设计和分析方法，并给出了基于强化学习技术的数据驱动方法。通过分解博弈问题和定义最佳响应，推导出替代的局部纳什解，构建了使用相邻交互数据的非策略强化学习算法，该算法可在无模型情况下更新控制器，且具有稳定性和鲁棒性。此外，通过修改耦合指数函数的微分博弈配置，可实现全局纳什均衡。仿真结果说明了学习过程的有效性和同步控制的稳定性。

北京科技大学宋睿卓等针对离散二阶多智能体系统在外界干扰约束下的有限时间领导—跟随一致性问题，利用自适应滑模控制理论设计了一种新的一致性控制方案。该方案采用传统的滑模趋近律设计自适应控制器，能够有效减少抖振并保持对干扰的不变性，同时利用离散李雅普诺夫函数证明了有限时间稳定性，为多智能体有限时间下一致性问题的解决提供了理论支撑。

清华大学芦维宁等针对多智能体协同规划中的问题，将图神经网络与面向任务的知识融合抽样方法结合，提出一种新的协同规划体系结构，能够对任意数量智能体协同规划问题建立通用模

型。同时设计了一种面向任务的抽样方法，用于从特定方向聚合可用知识，为在未知复杂环境下的多智能体博弈协同规划问题提供了研究框架。

总体而言，本专题涵盖了从多智能体演化博弈研究框架、多智能体协同控制与规划到多智能体强化学习框架等方面的内容。我们衷心希望这一专题能够激发对这些课题感兴趣的研究人员，进一步促进该领域的发展。

最后，我们要特别感谢作者和审稿人对本专题的大力支持和宝贵贡献，感谢主编潘云鹤院士、卢锡城院士以及期刊编辑。



陆军，中国工程院院士，中国电子科技集团公司首席科学家。1985年在东南大学获学士学位，1988年在中国科学技术大学获硕士学位。我国电子信息领域专家、预警机信息系统领域学术带头人。主持完成首型国产预警机和出口型预警机等，建立了我国预警机信息系统的技术体制，在网信体系和综合电子信息系统领域做出突出贡献，获国家科技进步特等奖。研究兴趣为综合电子信息系统理论与开发。



王飞跃，研究员，中国科学院自动化研究所复杂系统管理与控制国家重点实验室主任。1990年获美国伦塞利尔理工学院博士学位。IEEE/AAAS/ASME/IFAC/INCOSE 会士，获 IEEE ITS 杰出研究奖、ASME MESA 杰出成就奖、诺伯特·维纳奖、国家自然科学基金二等奖等。2021年成为 IFAC Pavel J. Nowacki 杰出讲师。研究兴趣为平行智能，社会计算，复杂系统的建模、分析和控制，等。



董琦，高级工程师，中国科学院研究院智能博弈团队负责人。2017年在天津大学获控制科学与工程博士学位，中国电子学会优博获得者，2018年入选中国科协青托工程，2019年获吴文俊人工智能优秀青年奖。研究兴趣包括智能博弈、自主规划、多智能体系统控制等。



魏庆来，研究员，中国科学院自动化研究所复杂系统管理与控制国家重点实验室副主任。分别于2002、2005和2008年获东北大学自动化学士学位、控制理论与控制工程硕博学位。*Neurocomputing* 副主编，自2021年起担任 IEEE 计算智能学会北京分会副主席。研究兴趣包括自适应动态规划、基于神经网络的控制、最优控制、非线性系统及其工业应用。