

多智能体系统专刊序言

孙长银¹ 孙健² 杨涛³ 易新蕾⁴

近年来,随着互联网技术、通信技术和传感技术的快速发展,多智能体系统逐渐成形.多智能体系统是指由多个具有自主决策能力的智能体通过分布式网络架构组成的协同系统.每个智能体既是独立的功能单元,具备局部感知、计算和执行能力,又能通过特定通信协议与相邻节点进行信息交互,最终通过分布式协调机制实现全局优化目标.

该方向仍面临许多挑战,迫切需要融合跨学科理论、设计新型分布式算法,并通过虚实结合的验证平台推动技术落地,最终实现从“机械协同”到“自主认知”的跨越.因此,《自动化学报》组织“多智能体系统”专刊,旨在促进控制、优化、学习理论与多智能体系统的融合发展,为科研人员和其他相关从业者提供一个展示最新理论和实践成果的平台.本专刊包括综述、多智能体系统控制、多智能体系统协同优化和学习,经同行评议录用共计 28 篇,部分文章将在下期发表.

1 综述

中国科学技术大学秦家虎教授等重点讨论在微分博弈下多智能体协同问题的建模与分析,并简要总结如何应用强化学习算法求解博弈均衡.选取多移动机器人导航和电动汽车充电调度这两个典型的多智能体协同场景,介绍博弈与控制融合的思想如何用于解决相关领域的难点问题.

中南大学罗彪教授等梳理多智能体强化学习在控制与决策领域的研究,分析其面临的主要问题与

挑战,从控制理论与自主决策两个层次综述现有的研究成果与进展,并针对未来的研究方向进行展望.

华东理工大学唐漾教授等多智能体强化学习在博弈中的应用进行系统综述.介绍多智能体强化学习的基本理论,并回顾相关算法及基线测试环境的演变.针对合作、对抗和混合任务,分析提高合作效率和对抗能力的最新进展.同时,结合实际应用,讨论混合博弈的研究前沿.

北京大学谢广明教授等概述追逃博弈问题的相关研究进展,从空间环境、信息获取等五个方面介绍追逃博弈问题的各类设定;简述理论求解、数值求解等四种当下主流的追逃博弈问题研究方法.

北京航空航天大学吕金虎教授等从协同制导的发展历程和任务层次进行分类,提出四种协同制导方法:打击时间/角度约束下的协同制导、编队构型约束下的协同制导、角色与任务互补下的协同制导、博弈对抗条件下的协同制导.总结最新理论进展,并从研究方法角度分析相关成果的优缺点及适用范围.深入探讨协同制导面临的技术挑战,并对未来发展进行展望.

燕山大学闫敬教授等梳理跨域集群相关概念及其发展现状,分析其面临的挑战与关键问题;进而从控制理论和通信技术相结合的角度出发,简述跨域集群任务规划、组网传输、协同控制等关键技术的研究进展;最后结合实际发展情况和未来发展趋势,对海上无人系统跨域集群未来值得深入研究的方向进行总结与展望.

安徽大学孙长银教授等阐述具身智能自主无人系统的发展现状,详细探讨包含大模型驱动的多模态感知、面向具身任务的推理与决策、基于动态交互的机器人学习与控制、三维场景具身模拟器等具身智能领域的关键技术.最后,指出目前具身智能无人系统所面临的挑战,并展望未来研究方向.

2 多智能体系统协同控制

东北大学郭戈教授等研究高维线性时不变动力学系统构成的具有加权有向的多智能体网络拓扑变

1. 安徽大学人工智能学院 合肥 230031 2. 北京理工大学自动化学院 北京 100081 3. 东北大学流程工业综合自动化国家重点实验室 沈阳 110819 4. 同济大学电子与信息工程学院 上海 201804

1. School of Artificial Intelligence, Anhui University, Hefei 230031 2. School of Automation, Beijing Institute of Technology, Beijing 100081 3. State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819 4. College of Electronic and Information Engineering, Tongji University, Shanghai 201804

引用格式 孙长银, 孙健, 杨涛, 易新蕾. 多智能体系统专刊序言. 自动化学报, 2025, 51(3): 485–488

Citation Sun Chang-Yin, Sun Jian, Yang Tao, Yi Xin-Lei. Guest editorial of the special issue on multi-agent systems. *Acta Automatica Sinica*, 2025, 51(3): 485–488

DOI 10.16383/j.aas.c202501 CSTR 32138.14.j.aas.c202501

化可辨识性,分析异构内耦合矩阵对网络拓扑可辨识性的影响,发现网络拓扑结构的可辨识性与智能体之间的内耦合矩阵相关.当内耦合矩阵由同构变为异构时,网络拓扑的可辨识性可能发生变化,既可能由可辨识变为不可辨识,也可能由不可辨识变为可辨识.针对一般网络结构,提出一些充分和必要的条件以验证拓扑变化的可辨识性.

清华大学何潇教授等设计基于比例-积分结构的容错控制律.考虑到传统的比例型控制律无法消除加性干扰影响下的稳态误差,在一致性控制律中融入状态的积分项,用于改善多智能体系统一致性过程的稳态性能.针对领航者输入不为零的情况,设计非线性的一致性控制律,并借助黎卡提方程以及 Lyapunov 函数,进行多智能体系统在故障情况下的一致性分析和控制律设计.

东南大学王庆领教授等研究严格反馈多智能体系统的最优一致性问题,提出一种新的分布式比例积分变量,使得经典的控制技术能够处理复杂的多智能体系统.此外,结合所提出的分布式 PI 变量和预设性能控制,设计一类基于 PI 调节的最优一致性算法,用以解决带有死区输入非线性和有界扰动的严格反馈的最优一致性问题.

上海大学张宪霞教授等提出一种新的融合动态模糊神经网络 (DFNN) 和自适应动态规划算法的无模型自适应控制方法.类似于强化学习中执行者-评论家结构,DFNN 和神经网络分别逼近控制策略和性能指标.每个智能体的 DFNN 执行者从零规则开始,通过在线学习,与其局部邻域的智能体交互而生成和合并规则.

中国地质大学 (武汉) 陈鑫教授等设计了基于同胚分布式控制协议的无模型方法.通过将输出反馈线性化理论与自适应动态规划相结合,实现非线性智能体的线性化.通过设计双层分布式控制结构和无模型反馈线性化方法,在微分同构空间层利用线性控制方法进行分布式共识控制.

中南财经政法大学王巍副教授等将低增益反馈方法与 Q 学习相结合,提出采用后向时间迭代的模型无关控制方法.将执行器饱和的有限时域一致性控制问题的求解转变为执行器饱和的单智能体有限时域最优控制问题的求解,证明可以通过求解修正的时变黎卡提方程以实现有限时域最优控制.此外,引入参数化时变 Q 函数,并提出基于 Q 学习的模型无关后向时间迭代算法,可以更新低增益参数,同时实现逼近求解修正的时变黎卡提方程.

华中科技大学曾志刚教授等提出一种新的基于常微分-偏微分方程 (ODE-PDE) 的分析方法,以解决稀疏-密集混合分布的大规模异构 MASs 编队问题.通过设计特定的通信协议,并将原始大量的异构 MASs 的 ODE 动力学模型转化为由一个 PDE 和少数几个 ODE 耦合而成的 ODE-PDE 模型,拓扑权值被规定为半马尔科夫切换的,且稀疏分布和密集分布智能体遵循不一致的切换规则.

安徽大学何舒平教授等研究距离-变权重通信拓扑下非线性多智能体系统固定时间防碰防离编队控制问题,充分考虑通信路径损耗所引起的拓扑变化的不确定性和距离相关性、系统中未知非线性动力学特性以及固定时间收敛的控制性能要求等.基于人工势场原理,设计一套新的预设时间防碰防离策略,以确保每个智能体在预设时间内离开碰撞与离群预警区,避免碰撞与离群现象.

重庆大学苏晓杰教授等以分布式网络中的移动机器人为研究对象,构建基于视线和速度的低触发势能函数,实现邻近编队机器人近距排斥作用下的避碰保持;引入驱动机器人绕过障碍物的安全加速度包络,并复合近距排斥的弱能量、低触发势能,避免与非合作障碍物的碰撞.此外,嵌入复合自适应辅助动态系统,平滑躲避过程中触发的嵌套运动饱和及安全加速度约束引起的轨迹振荡.

东南大学温广辉教授等提出一种新型分布式固定时间二分编队跟踪控制协议.将命令滤波引入反推控制中,进而分别设计虚拟控制协议与真实控制协议.此外,为估计未知参数设计了参数复合学习律,利用在线记录的数据和即时数据产生预测误差,并利用跟踪误差和预测误差更新参数估计.

东北大学高伟男教授等针对离散时间多智能体系统的协同最优输出调节问题,在不依赖多智能体系统矩阵的条件下提出分布式数据驱动自适应控制策略.基于自适应动态规划和分布式自适应内模,通过引入值迭代和策略迭代两种强化学习算法,利用在线数据学习最优控制器,实现多智能体系统的协同输出调节.

浙江工业大学何德峰教授等提出一种新的多耦合分布式经济模型预测控制 (EMPC) 策略.由于耦合非线性系统的经济性能函数的非凸性和非正定性,引入关于经济最优平衡点的正定辅助函数和相应的辅助优化问题.利用辅助函数的最优值函数构造原始分布式 EMPC 的一类隐式收缩约束.然后建立状态耦合分布式 EMPC 的递推可行性和闭环

系统关于最优经济平衡点的输入-状态稳定性结论.

西安工业大学高嵩教授等提出一种基于变分稀疏高斯过程回归的分布式感知与围捕算法. 针对传统高斯过程回归在处理大量数据时的不足, 研究引入分离超平面的质心维诺划分算法, 用于机器人动态规划任务区域, 并考虑障碍物的影响. 通过多机器人在任务区域内的移动探索获取环境信息, 并利用变分自由能方法近似模型的后验分布, 实现对未知环境的感知.

东北大学孙佳月教授等提出基于混合双端事件触发机制的模糊跟踪控制策略, 构建一种灵活可调的阈值设计方法以满足系统实时性需求; 改进的状态触发机制将状态估计值作为触发信号, 可有效降低现存机制的保守性并提高阈值设计的灵活性.

临沂大学邱建龙教授等研究异构不确定二阶非线性多智能体系统的事件触发状态趋同控制问题. 为每个智能体设计参数观测器用以估计不确定参数; 此外, 为每个跟随智能体设计分布式参数观测器渐近估计领导智能体不确定参数, 基于估计的参数和邻居状态, 提出完全不依赖智能体间连续信息传输的事件触发趋同算法.

东北大学杨涛教授等考虑量化通信下多智能体系统的协同最优输出调节问题, 设计分布式量化观测器, 使得每个跟随者对外部系统状态的估计误差渐近收敛至零. 此外, 在系统动态未知的情况下, 提出基于自适应动态规划的数据驱动算法在线学习次优控制策略, 解决协同最优输出调节问题, 保证每个跟随者的输出信号渐近跟踪参考信号, 并抑制由外部系统产生的干扰信号.

3 多智能体系统协同优化和学习

安徽大学柳文章等在集中训练和分散执行框架的基础上, 考虑从优势函数输入端注入扰动量来提升优势函数的泛化能力, 提出一种新的基于优势函数输入扰动的多智能体近端策略优化方法, 从而提升多无人艇协同策略的探索效率.

同济大学李修贤教授等研究分布式在线复合优化场景中的几种反馈延迟, 包括梯度反馈、单 Bandit 反馈和两点 Bandit 反馈, 在分布式场景下, 研究每个智能体具有不同的时变延迟, 基于近端梯度下降算法, 分别设计上述三种延迟反馈的分布式在线复合优化算法, 并且对动态遗憾上界进行分析.

南京理工大学张保勇教授等研究多智能体时变网络上基于 Bandit 反馈的分布式在线鞍点问题,

在 Bandit 反馈下, 结合单点梯度估计方法和预测映射技术, 提出一种分布式在线 Bandit 鞍点优化算法. 对于一般的凸-凹损失函数, 算法实现动态遗憾上界的次线性收敛. 进一步设计了近似算法, 并分析近似精确度对算法遗憾上界的影响.

中国科学院大学赵冬斌教授等提出熵引导的极小极大值分解强化学习方法, 在线学习队内合作和队间对抗的策略. 提出基于极小极大值分解的执行器-评估器框架, 在不限动作空间的、高采样成本的任务中, 通过极小极大值分解提升优化效率和博弈性能. 此外, 引入最大熵使智能体可以更充分地探索状态空间, 避免在线学习过程收敛到局部最优.

湖南大学张辉教授等提出一种基于分层仿生神经网络的多机器人协同区域搜索算法. 将仿生神经网络和不同分辨率下的区域栅格地图结合, 构建分层仿生神经网络信息模型. 在分层仿生神经网络信息模型基础上引入分布式模型预测控制框架, 并设计多机器人分层协同决策机制.



孙长银 安徽大学人工智能学院教授. 2004 年获得东南大学电子工程专业博士学位. 主要研究方向为智能控制, 飞行器控制, 模式识别和优化理论. E-mail: 20168@ahu.edu.cn

(SUN Chang-Yin Professor at the School of Artificial Intelligence, Anhui University. He received his Ph.D. degree in electrical engineering from Southeast University in 2004. His research interest covers intelligent control, flight control, pattern recognition, and optimal theory.)



孙 健 北京理工大学教授. 2007 年获得中国科学院自动化研究所控制理论与控制工程专业博士学位. 主要研究方向为网络化控制系统, 信息物理融合系统安全性, 多智能体系统.

E-mail: sunjian@bit.edu.cn

(SUN Jian Professor at Beijing Institute of Technology. He received his Ph.D. degree in control theory and control engineering from Institute of Automation, Chinese Academy of Sciences in 2007. His research interest covers networked control systems, security of cyber-physical-systems, and multi-agent systems.)



杨 涛 东北大学流程工业综合自动化国家重点实验室教授. 主要研究方向为工业人工智能, 信息物理系统, 分布式协同控制和优化.

E-mail: yangtao@mail.neu.edu.cn

(**YANG Tao** Professor at the State Key Laboratory of Synthetical Auto-

mation for Process Industries, Northeastern University. His research interest covers industrial artificial intelligence, cyber physical system, distributed collaborative control and optimization.)



易新蕾 同济大学准聘教授. 2020 年获得瑞典皇家理工学院博士学位. 主要研究方向为分布式优化, 在线优化, 元学习和图神经网络.

E-mail: xinleiyi@tongji.edu.cn

(**YI Xin-Lei** Tenure-track professor of Tongji University. He re-

ceived his Ph.D. degree from KTH Royal Institute of Technology, Stockholm, Sweden, in 2020. His research interest covers distributed optimization, online optimization, meta-learning and graph neural networks.)