

基于深度强化学习的平行企业资源计划

秦蕊^{1,2,3} 曾帅^{1,2} 李娟娟^{1,2} 袁勇^{1,2}

摘要 传统的企业资源计划 (Enterprise resource planning, ERP) 采用静态化的业务流程设计理念, 忽略了人的关键作用, 且很少涉及系统性的过程模型, 因此难以应对现代企业资源计划的复杂性要求. 为实现现代企业资源计划的新范式, 本文在 ACP (人工社会 (Artificial societies)、计算实验 (Computational experiments)、平行执行 (Parallel execution)) 方法框架下, 以大数据为驱动, 融合深度强化学习方法, 构建基于平行管理的企业 ERP 系统. 首先基于多 Agent 构建 ERP 整体建模框架, 然后针对企业 ERP 的整个流程建立序贯博弈模型, 最后运用基于深度强化学习的神经网络寻找最优策略, 解决复杂企业 ERP 所面临的不确定性、多样性和复杂性.

关键词 企业资源计划, 深度强化学习, ACP 理论, 平行管理, 多 Agent 建模

引用格式 秦蕊, 曾帅, 李娟娟, 袁勇. 基于深度强化学习的平行企业资源计划. 自动化学报, 2017, 43(9): 1588–1596

DOI 10.16383/j.aas.2017.c160664

Parallel Enterprises Resource Planning Based on Deep Reinforcement Learning

QIN Rui^{1,2,3} ZENG Shuai^{1,2} LI Juan-Juan^{1,2} YUAN Yong^{1,2}

Abstract Traditional enterprise resource planning (ERP) usually adopts static business processes design and does not take the key role of “human” into consideration. It rarely involves the systematic process modeling, which makes it impossible to tackle the management complexity of modern enterprises. Considering the big data driven environment of modern enterprises, we utilize the ACP (Artificial societies, computational experiments, parallel execution) theory integrated with deep reinforcement learning approaches to establish a parallel management system for modern ERP management. We first propose a framework for ERP systems based on multi-agent technology where a sequential game model is included. Then, we seek for the optimal strategy using a deep reinforcement learning based neural network. Our proposed framework and approaches can well deal with uncertainty, diversity and complexity of modern ERP systems.

Key words Enterprise resource planning (ERP), deep reinforcement learning, ACP theory, parallel management, multi-agent technology

Citation Qin Rui, Zeng Shuai, Li Juan-Juan, Yuan Yong. Parallel enterprises resource planning based on deep reinforcement learning. *Acta Automatica Sinica*, 2017, 43(9): 1588–1596

企业是国民经济的微单元, 是推动社会经济进步的主导力量. 高效的企业资源计划 (Enterprise resource planning, ERP) 则是企业持续健康运转的源动力^[1–2]. 企业 ERP 是引导并规范企业这

条列车向正确的目标高效前进的轨道. 在现代企业 ERP 中, 由于对市场环境大数据缺乏全面、准确和实时的获取能力和技术手段, 以及对企业内部各类数据缺乏实时高效的汇聚和融合技术, 实际的企业 ERP 目前仍严重依赖于知识型工作者的经验, 这就导致人和社会环境因素对企业 ERP 动态变化趋势的影响被忽略. 随着信息技术和网络技术的发展, 人类社会逐渐迈入知识经济时代, 而知识社会必然要求实现知识自动化, 满足人类对知识快速获取和应用的需求^[3]. 面对互联网和大数据中的海量信息, 人们更需要借助知识自动化的方法“弥补”智能上的不足, 进而完成各种层出不穷的不定、多样、复杂任务. 如何实现大数据、知识与三者之间的动态闭环反馈以及实时交互, 成为现代企业 ERP 面临的重大挑战.

在企业中, 人员行为与关系网络的复杂性、管理过程的主观性等, 导致企业资源计划过程的动态性、不可控性和突变性巨增. 此外, 随着社会化和网络化趋势的日益增强, 企业已全面深度地融入网络 (Cyber) 环境, 演变成为“现实物理世界–网络虚拟

收稿日期 2016-09-14 录用日期 2016-11-28
Manuscript received September 14, 2016; accepted November 28, 2016

国家自然科学基金 (71702182, 71472174, 71232006, 61533019, 61233001, 71402178), 复杂系统管理与控制国家重点实验室优秀人才基金 (Y6S9011F4E, Y6S9011F4H) 资助

Supported by National Natural Science Foundation of China (71702182, 71472174, 71232006, 61533019, 61233001, 71402178) and the Early Career Development Award of State Key Laboratory of Management and Control for Complex Systems (Y6S9011F4E, Y6S9011F4H).

本文责任编辑 王飞跃

Recommended by Associate Editor WANG Fei-Yue

1. 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 北京 100190 2. 青岛智能产业技术研究院 青岛 266109 3. 中国科学院自动化研究所 北京市智能化技术与系统工程研究中心 北京 100190

1. The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 2. Qingdao Academy of Intelligent Industries, Qingdao 266109 3. Beijing Engineering Research Center of Intelligent Systems and Technology, Institute of Automation, Chinese Academy of Sciences, Beijing 100190

世界”耦合空间中的现代新型企业. 由此可见, 现代企业资源计划目前已呈现出显著的快速性、开放性、交互性、突变性和数据海量等特点.

企业 ERP 的最终目标是将其面临的 UDC, 即不确定性 (Uncertainty)、多样性 (Diversity) 和复杂性 (Complexity) 转化为完成特定管理任务和目标的 AFC, 即敏捷 (Agility)、聚焦 (Focus) 和收敛 (Convergence) 等特性^[4], 而单纯依靠物理空间的资源无法有效地实现企业 ERP 的这种转变. 综合计算、网络和物理环境的多维复杂系统—物理信息系统 (Cyber-physical systems, CPS) 已不能满足企业 ERP 应用的需求, 亟需加入社会化维度和人的因素, 构建新的社会物理信息系统 (Cyber-physical-social systems, CPSS)^[5], 并在此框架下为企业决策支持提供必要参考, 使得企业 ERP 能够迈向自动化和智能化.

2016 年 3 月, 谷歌围棋人工智能 AlphaGo 与韩国棋手李世石进行的人机大战引起了各界广泛关注^[6]. 最终 AlphaGo 以 4:1 获胜, 体现了人工智能在围棋领域的重大突破, 这种突破为面临 UDC 的复杂系统的管理与决策带来了重要启示^[7]: 1) AlphaGo 为复杂系统优化中面临的高计算复杂度提供了一个现实的解决方案. 在理论最优解无法求出的情况下, 探索工程解是求解这一类复杂问题的重要突破. 可以大胆推断: 所有“合理”的复杂性问题, 包括 NP-Complete 甚至 NP-Hard 问题, 都可以用 AlphaGo 方法与思路处理^[8]. 2) AlphaGo 的建设思路是首先根据历史数据进行“人工棋”建模, 其次通过自博弈强化学习在人工围棋系统中进行计算实验, 最后的人机大战则是典型的虚实互动^[9]. 这与王飞跃提出的基于 ACP (人工社会 (Artificial societies)、计算实验 (Computational experiments)、平行执行 (Parallel execution)) 方法^[10-11] 的平行管理不谋而合. ACP 方法是将实验经济学方法与经济计算实验方法结合, 融合人工系统、计算实验和平行执行, 形成的一套研究和解决复杂社会经济问题的计算理论和方法体系^[12].

本文安排如下: 第 1 节阐述研究问题和研究思路, 第 2 节提出企业 ERP 3.0 的理念, 并基于企业 ERP 流程进行建模, 第 3 节给出基于深度强化学习的求解算法, 第 4 节对本文的工作进行总结.

1 研究问题与思路

1.1 研究问题

企业 ERP^[13-14] 是指建立在信息技术基础上, 以系统化的管理思想整合企业内部业务流程, 为企业决策层及员工提供决策运行手段的管理平台. 企业 ERP 的发展^[15] 经历了以传统的 ERP、ERP-II 和 e-ERP 等系统为代表的 ERP 1.0 时代和 Web

2.0 技术下的 ERP 2.0 时代. ERP 1.0 虽然实现了对采购、生产、成本、库存、分销、运输、财务、人力资源的综合管理, 但由于采用静态化的业务流程设计理念, 将企业的工程管理需求放在第一位, 仅关注对工程复杂性的管理, 缺少对企业人员行为的管理, 从而忽略了人员行为的关键作用^[16]. ERP 2.0 虽然解决了如何使用新的社会化软件平台, 从而可以使组织及人员交流更顺畅的问题, 但未考虑人的行为, 且存在一定的安全、技术和机制方面的问题^[17].

在现代企业 ERP 中, 企业面临着内部和外部双重挑战. 在企业内部, 计算机技术的兴起和广泛应用极大减轻了企业人员的劳动负担, 但同时也增加了 ERP 系统的复杂性. 此外, 人员行为的复杂性、管理过程的主观性等致使企业 ERP 过程的动态性、不可控性、突变性巨增; 在企业外部, 由于同一个市场中往往存在着多个相互竞争的企业, 企业 ERP 决策面临着复杂、动态、竞争性的决策环境.

由于受到内部和外部多种影响因素的复杂作用, 企业 ERP 呈现出显著的快速性、开放性、交互性、突变性和数据海量等特点, 给企业 ERP 带来了前所未有的困难. 显然, 传统的 ERP 1.0 和 ERP 2.0 系统已难以应对现代企业 ERP 面临的巨大挑战.

1.2 研究思路

为构建现代企业 ERP 的新范式, 本文基于 AlphaGo 建设思路, 在 ACP 方法框架下, 融合深度强化学习理论, 提出基于平行管理^[18-19] 的 ERP 建设新思路和新方法, 为企业的复杂系统管理与决策提供一套行之有效的解决方案.

在 ACP 方法框架下, 以大数据为驱动, 运用人工智能方法, 构建基于平行管理的企业 ERP 系统, 实现现代企业 ERP 从现实物理空间向虚拟网络空间的扩充, 在人工企业建模的基础上, 融合运作管理和资源管理, 以计算实验和虚实互动模式解决复杂企业 ERP 问题^[20], 具体研究思路如图 1 所示.

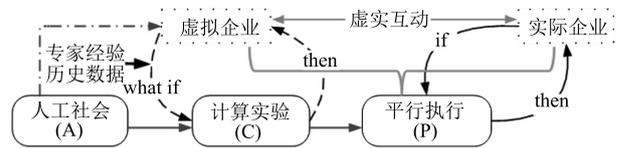


图 1 平行企业 ERP 思路

Fig. 1 Basic idea of parallel ERP

1) 基于多 Agent^[21-22] 方法对企业进行人工建模, 并建立与实际企业“等价”的虚拟企业. 在人工企业建模过程中, 给出企业 ERP 决策输入以及相应的评价指标体系. 与传统企业建模方法相比, 不仅对企业的整个生产流程进行 Agent 建模, 并且在此基础上构建不完全信息动态环境下的人工企业序贯博弈模型.

2) 针对虚拟企业进行计算实验, 计算复杂虚拟企业在不同策略下的行为、反应及目标实现, 并针对不同决策的效果进行评估, 作为选择和支持管理决策的依据. 在计算实验之前, 虚拟企业已经根据专家经验和历史数据通过监督学习和强化学习获得一定的管理决策规则, 给出可能输入, 可获得相应的输出, 并通过估值网络对该决策输出进行评估 (实现图 1 中 what if 到 then 的过程). 在计算实验中, 将先进的人工智能方法融入已有的涌现方法及其他计算方法中, 运用深度强化学习神经网络寻找最优策略, 并通过自博弈手段进行策略的自我进化.

3) 通过虚拟企业和实际企业的互动反馈形成对实际企业的管理与控制. 在这一过程中, 虚拟企业逐步逼近实际企业, 对实际企业中的行为决策及所产生的决策收益进行预估和判断, 从而为确定实际企业中的有效决策方案或改进当前决策提供必要的依据 (实现图 1 中 if 到 then 的过程). 进一步, 通过实际企业与虚拟企业的相互连接, 对二者之间的行为进行对比和分析, 观察相应的误差反馈, 并基于此对虚拟企业的评估方法和参数进行修正, 完成对各自未来状况的“借鉴”和“预估”, 从而相应地调节各自的管理与控制方式, 达到实施有效解决方案以及学习和培训的目的.

基于 ACP 方法的平行企业 ERP 系统针对企业全流程建模, 不局限于特定的企业 ERP 场景, 也不过多地依赖于专业领域知识, 因此建模及仿真优化均可以根据具体应用场景进行相应的扩展, 可应用于典型企业的 ERP 研究和实践. 企业 ERP 研究的落脚点应该在决策支持上, 任何理论上的创新最终都应该为实践服务. 在大多数企业仿真和企业建模中, 由于求解搜索空间太大, 导致其沦为概念模型, 无法为企业决策优化提供有效的支持. 为了避免这类问题, 本文采用基于深度强化学习的神经网络来求解所建立的平行企业 ERP 模型.

2 企业 ERP 建模

本节主要针对企业 ERP 系统进行建模, 首先提出基于平行管理的 ERP 3.0, 然后构建基于多 Agent 的 ERP 3.0 建模框架, 在此基础上, 建立基于企业 ERP 全流程的不完全信息动态博弈模型.

2.1 基于平行管理的 ERP 3.0

针对企业 ERP 复杂过程中面临的重大需求以及企业 ERP 中人员行为多样化和数据海量等问题, 本节对传统的企业 ERP 1.0 和 ERP 2.0 的概念进行升级, 提出基于平行管理的 ERP 3.0 的整体思路, 并建立企业 ERP 3.0 系统, 以数据为驱动, 通过将现实物理空间向虚拟网络空间扩充, 在人工企业组织的基础上, 融合运作管理和资源管理, 以计算实验和虚实互动模式来解决复杂管理问题.

ERP 3.0 采用平行管理理论, 在工程复杂性的传统闭环管理基础上, 将企业的制度规范和人员行为纳入到企业 ERP 范畴, 实现社会与工程一体化的大闭环管理, 通过虚实系统的平行执行, 实现对实际系统的双闭环管理, 并从根本上考虑人员行为在运作管理中的核心地位和作用. 其基本思想是将企业的业务流程看作一个紧密连接的供应链, 将供应商和企业内部的采购、生产、销售以及客户紧密联系起来, 便于对供应链上的所有环节进行有效管理, 实现对企业的动态控制和各种资源的集成与优化.

企业 ERP 3.0 系统在原有的 ERP 系统各环节的基础上, 引入数据采集和实验决策环节, 动态采集与监测企业内部数据和外部市场环境, 并据此进行实时决策. 企业 ERP 3.0 系统的各个环节如图 2 所示.

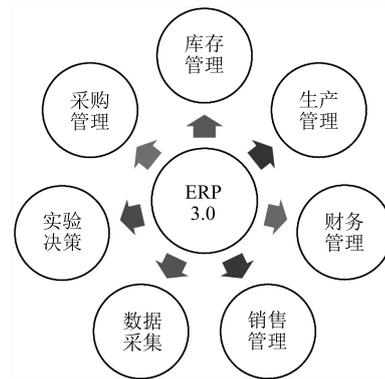


图 2 企业 ERP 3.0 的系统构成

Fig. 2 Composition of ERP 3.0

2.2 基于多 Agent 的企业 ERP 3.0 建模框架

本节将采用多 Agent 建模方法^[23], 构建各个环节的 Agent, 从而实现对企业 ERP 3.0 系统的建模. 由于 Agent 具有高度的自治性、智能性、协作性、交互性和适应性等特点, 多 Agent 技术为企业 ERP 3.0 系统的建模提供了一种崭新的研究思路.

根据企业 ERP 3.0 系统的核心组成, 基于多 Agent 技术的 ERP 系统主要包含: 采购管理 Agent、库存管理 Agent、生产管理 Agent、财务管理 Agent、销售管理 Agent、数据采集 Agent 和实验决策 Agent, 各 Agent 的功能如下:

1) 采购管理 Agent: 主要负责原料采购计划的制定, 并与供应商进行交互, 进行采购计划的实施.

2) 库存管理 Agent: 主要负责剩余产品和原料的库存管理, 以适应不断变化的市场需求.

3) 生产管理 Agent: 根据市场需求制定产品的生产计划和原料的采购需求, 并进行生产计划的实施.

4) 财务管理 Agent: 控制企业 ERP 流程中各个环节中的资金流和财务状况.

5) 销售管理 Agent: 根据整个市场中客户的产品需求和企业自身状况, 制定销售计划, 并将产品销售给客户.

6) 数据采集 Agent: 采集企业 ERP 整个流程中的数据信息, 并监测外部环境信息.

7) 实验决策 Agent: 根据数据采集 Agent 所提供的数据, 控制企业 ERP 整个流程中各 Agent 的交互.

基于以上定义, 图 3 给出了企业 ERP 系统 Agent 建模的流程图, 具体过程可描述如下:

1) 销售管理 Agent 根据整个市场中的客户需求、企业自身状况和外部环境 (其他竞争企业的情况) 等因素确定每一时刻的订单, 通过与库存管理 Agent 交互, 查询当前时刻的库存状态, 并根据库存状态和订单需求确定需要生产产品的数量, 之后与生产 Agent 交互, 将相关需求信息和生产信息传递给生产 Agent. 在该环节中, 销售管理 Agent 需要通过实验决策 Agent 与库存管理 Agent、生产管理 Agent 交互, 确定销量和生产量, 并将所有信息传递给数据采集 Agent.

2) 生产管理 Agent 根据销售 Agent 所传递的生产信息、企业自身的状况 (如企业的财务状况、产品的库存容量、单位时间的生产能力、产品的库存成本等信息) 和外部环境等信息制定生产计划, 确定各原料的需求量, 并与库存管理 Agent 交互, 确定需要采购原料的数量, 然后将这些信息传递给采购管理 Agent. 在该环节中, 生产管理 Agent 需要通过实验决策 Agent 与库存管理 Agent、采购管理

Agent 交互, 确定生产量和原料采购量, 并将所有信息传递给数据采集 Agent.

3) 采购管理 Agent 根据生产管理 Agent 传递的原料采购信息、企业自身的状况 (如企业的财务状况、原料的库存容量、原料的库存成本等信息), 与财务管理 Agent 和库存管理 Agent 交互, 制定采购计划. 在该环节中, 采购管理 Agent 需要通过实验决策 Agent 与生产管理 Agent、库存管理 Agent 交互, 确定原料采购量、原料采购费用、剩余原料的入库, 并将所有信息传递给数据采集 Agent.

4) 生产管理 Agent 利用从库存管理 Agent 和采购管理 Agent 中获得的原料进行生产, 并根据销售管理 Agent 传递的需求量将生产的产品传递给销售管理 Agent, 剩余的产品则传递给库存管理 Agent. 在该环节中, 生产管理 Agent 还需要通过实验决策 Agent 与财务管理 Agent 交互, 确定产品的生产成本和库存成本, 并将所有信息传递给数据采集 Agent.

5) 销售管理 Agent 将从生产管理 Agent 和库存管理 Agent 获得的产品销售给客户, 并从客户获得产品收益. 在该环节中, 销售管理 Agent 需要通过实验决策 Agent 与财务管理 Agent 交互, 将产品收益传递给财务管理 Agent, 并将该过程中的所有信息传递给数据采集 Agent.

2.3 竞争环境中多企业 ERP 3.0 的序贯博弈模型

本节基于企业 ERP 3.0 系统 Agent 建模框架, 采用数学表达式将每个 Agent 的行为与策略符号化,

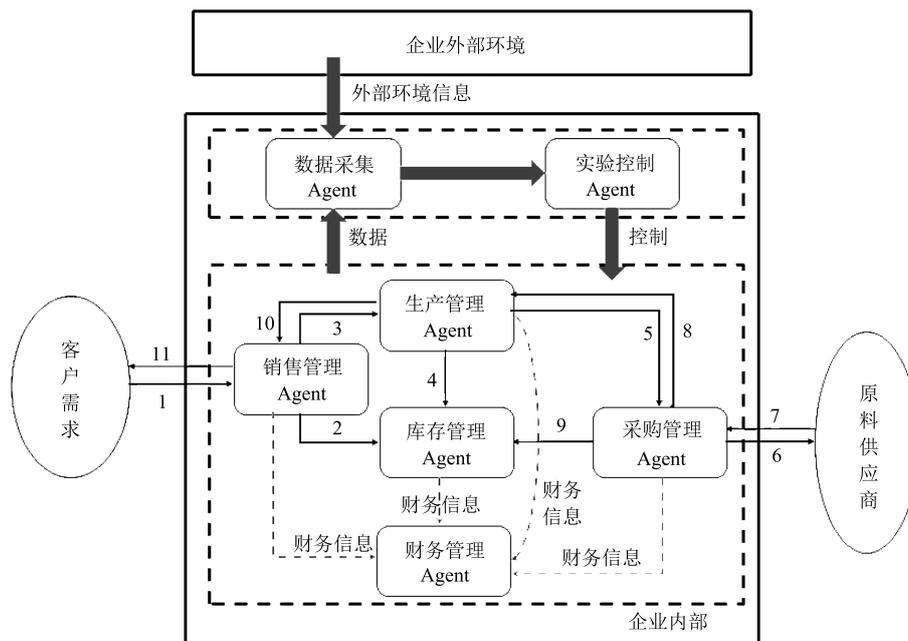


图 3 企业 ERP 3.0 系统 Agent 建模流程图

Fig. 3 Agent modeling framework for ERP 3.0

并采用序贯博弈方法^[24], 建立竞争环境中的多企业 ERP 3.0 系统的序贯博弈模型.

2.3.1 符号介绍

假设市场中有相互竞争的 K 个企业, 这些企业生产同质的产品, 具有相同的客户和供应商. 假设在一个固定的时间周期中包含 T 个阶段, 各阶段整个市场的产品需求随时间而变化, 每个企业需要根据市场需求、企业自身状况和竞争对手的决策来制定自己的决策. 由于每个企业的优势不同, 假设这些企业进行决策时是有一定顺序的, 即按企业 1, 企业 2, \dots , 企业 K 的顺序进行决策. 此外, 由于各企业规模的不同, 不同企业生产同一产品具有不同的的生产成本, 并且每个企业的库存成本也不相同.

1) 产品及原料假设: 假设企业 1, 2, \dots , K 均生产 N 种产品, 记为 B_1, B_2, \dots, B_N , 生产这些产品的原料共有 M 种, 记为 G_1, G_2, \dots, G_M . 生产每种产品 B_i 所需的原料可以用向量

$$\mathbf{l}_i = (l_{i,1}, l_{i,2}, \dots, l_{i,M}) \quad (1)$$

表示. 其中, $l_{i,j}$ 表示生产产品 B_i 时需要原料 G_j 的数量.

2) 库存状态: 企业 k 在阶段 t 的库存状态为

$$\mathbf{s}_{k,t} = (\mathbf{p}_{k,t}, \mathbf{q}_{k,t}) = (p_{k,1,t}, p_{k,2,t}, \dots, p_{k,N,t}, q_{k,1,t}, q_{k,2,t}, \dots, q_{k,M,t}) \quad (2)$$

其中, $p_{k,i,t}$ 表示成品 B_i 的数量, $q_{k,j,t}$ 表示原料 G_j 的数量.

3) 库存成本: 企业 k 每个成品 B_i 的库存成本为 $c_{1,k,i}$, 每个原料 G_j 的库存成本为 $c_{2,k,j}$. 则在阶段 t 所有成品和原料的总库存成本为

$$Z_{k,1,t} = \sum_{i=1}^N c_{1,k,i} p_{k,i,t} + \sum_{j=1}^M c_{2,k,j} q_{k,j,t} \quad (3)$$

4) 库存容量假设: 假设企业 k 中成品的最大库存容量为 $C_{k,1}$, 原料的最大库存容量为 $C_{k,2}$, 则有

$$\sum_{i=1}^N p_{k,i,t} \leq C_{k,1}, \quad \sum_{j=1}^M q_{k,j,t} \leq C_{k,2} \quad (4)$$

5) 生产能力与采购能力: 企业 k 在阶段 t 生产每个产品 B_i 的最大能力为 $\alpha_{k,i,t}$, 采购每个原料 G_j 的最大能力为 $\beta_{k,j,t}$.

6) 生产成本与定价: 企业 k 生产每个产品 B_i 的成本为 $c_{3,k,i}$ (此处指除原料之外的成本), 在阶段 t 的定价为 $\omega_{k,i,t}$, 采购每个原料 G_j 的成本为 $c_{4,k,j}$.

7) 环境及状态变量: 整个市场中存在 K 个相互竞争的企业, 因此每个企业 k 进行决策时不仅需

要考虑企业自身的内部环境, 还要考虑其他企业所构成的外部环境对自身企业的影响. 企业 k 在阶段 t 的内部环境用状态变量 $\mathbf{s}_{k,t}$ 表示, 外部环境用状态变量 $\mathbf{s}_{-k,t}$ 表示. 企业内部状态变量 $\mathbf{s}_{k,t}$ 表示企业自身在阶段 t 的状态, 由于市场的隐私性, 每个企业并不能获得其竞争对手的所有信息, 因此企业外部状态变量 $\mathbf{s}_{-k,t}$ 表示企业 k 能够获得的市场中其他企业的部分相关信息.

8) 市场需求及需求比例选择: 设 t 阶段每个产品 B_i 的市场总需求为 $d_{i,t}$, $i = 1, 2, \dots, N$, $t = 1, 2, \dots, T$. 每个企业 k 所选择的比例分别为 $\lambda_{k,i,t}$, 则有

$$\sum_{k=1}^K \lambda_{k,i,t} \leq 1 \quad (5)$$

9) 行动 (Action): 企业 k 在 t 阶段需要根据企业当前的内部环境和外部环境, 采取决策行动, 制定每个产品 B_i 的生产量 $x_{k,i,t}$, 每种原料 G_j 的采购量 $y_{k,j,t}$, 满足市场需求的比例 $\lambda_{k,i,t}$, 每个产品 B_i 的定价等决策, 用 $\mathbf{a}_{k,t} = (\mathbf{x}_{k,t}, \mathbf{y}_{k,t}, \boldsymbol{\lambda}_{k,t}, \boldsymbol{\omega}_{k,t})$ 表示.

2.3.2 模型建立

每个企业在每个阶段 t 都要进行正常的生产和采购活动, 并且每个阶段所选择的需求必须被完全满足. 此外, 考虑到未来需求的不确定性和动态变化性, 以及企业自身的生产能力和库存能力, 当前阶段生产和采购的数量可以超过当前的需求量, 以保证将来能够更好地满足客户需求. 生产的数量超过需求量的部分, 以及未用完的原料均要存入库存中, 并且产生相应的库存成本. 企业的目标是制定一个合理的每个阶段的生产、采购、需求比例选择和定价的计划, 使得在整个周期的总收益最大化.

企业内部状态变量 $\mathbf{s}_{k,t}$ 中主要包括企业自身在阶段 t 的库存信息 (即每个产品 B_i 的库存 $p_{k,i,t}$, 每种原料 G_j 的库存 $q_{k,j,t}$), 即

$$\mathbf{s}_{k,t} = (\mathbf{p}_{k,t}, \mathbf{q}_{k,t}) \quad (6)$$

其中,

$$\begin{aligned} \mathbf{p}_{k,t} &= (p_{k,1,t}, p_{k,2,t}, \dots, p_{k,N,t}) \\ \mathbf{q}_{k,t} &= (q_{k,1,t}, q_{k,2,t}, \dots, q_{k,M,t}) \end{aligned} \quad (7)$$

阶段 t 所采取的行动信息 $\mathbf{a}_{k,t}$ (即每个产品 B_i 的生产量 $x_{k,i,t}$, 每种原料 G_j 的采购量 $y_{k,j,t}$, 满足市场需求的比例 $\lambda_{k,i,t}$, 每个产品 B_i 的定价) 为

$$\begin{aligned} \mathbf{a}_{k,t} &= (x_{k,1,t}, \dots, x_{k,N,t}, y_{k,1,t}, \dots, y_{k,M,t}, \\ &\lambda_{k,1,t}, \dots, \lambda_{k,N,t}, \omega_{k,1,t}, \dots, \omega_{k,N,t}) \end{aligned} \quad (8)$$

而阶段 t 企业 k 的信息状态 $S_{k,t}$ 主要包括企业 k 前 t 阶段的状态信息和前 $t-1$ 阶段的行动信息, 即

$$\mathbf{S}_{k,t} = \{\mathbf{s}_{k,1}, \mathbf{a}_{k,1}, \mathbf{s}_{k,2}, \mathbf{a}_{k,2}, \dots, \mathbf{a}_{k,t-1}, \mathbf{s}_{k,t}\} \quad (9)$$

用 $\mathbf{f}_{k,t}(\cdot) = (f_{k,1,t}(\cdot), f_{k,2,t}(\cdot), f_{k,3,t}(\cdot), f_{k,4,t}(\cdot))$ 表示由当前阶段 t 的内部状态 $\mathbf{s}_{k,t}$ 和外部状态 $\mathbf{s}_{-k,t}$ 映射为 t 阶段行动的策略函数, 则有

$$\mathbf{a}_{k,t} = \mathbf{f}_{k,t}(\mathbf{s}_{k,t}, \mathbf{s}_{-k,t}) \quad (10)$$

在下一阶段 $t+1$, 状态变为

$$\mathbf{s}_{k,t+1} = (\mathbf{p}_{k,t+1}, \mathbf{q}_{k,t+1}) \quad (11)$$

其中,

$$\begin{aligned} \mathbf{p}_{k,t+1} &= \mathbf{g}_{k,t}(\mathbf{s}_{k,t}, \mathbf{a}_{k,t}) \\ \mathbf{q}_{k,t+1} &= \mathbf{g}'_{k,t}(\mathbf{s}_{k,t}, \mathbf{a}_{k,t}) \end{aligned} \quad (12)$$

$\mathbf{g}_{k,t}(\cdot)$ 和 $\mathbf{g}'_{k,t}(\cdot)$ 分别表示 t 阶段 $\mathbf{p}_{k,t}$ 和 $\mathbf{q}_{k,t}$ 的状态转移函数.

企业 k 在阶段 t 生产 B_i 的数量 $x_{k,i,t}$ 为

$$x_{k,i,t} = f_{k,1,t}(\mathbf{p}_{k,t-1}, \mathbf{q}_{k,t-1}, \mathbf{s}_{-k,t-1}) \quad (13)$$

并且满足

$$x_{k,i,t} \leq \alpha_{k,i,t}, \quad i = 1, 2, \dots, N \quad (14)$$

其总的生产成本为

$$z_{k,t}^{(1)} = \sum_{i=1}^N c_{3,k,i} x_{k,i,t} \quad (15)$$

阶段 t 每个企业 k 采购每个原料 G_j 的数量为 $y_{k,j,t}$, 则 $y_{k,j,t}$ 为

$$y_{k,j,t} = f_{k,2,t}(\mathbf{p}_{k,t-1}, \mathbf{q}_{k,t-1}, \mathbf{s}_{-k,t-1}) \quad (16)$$

并且满足

$$y_{k,j,t} \leq \beta_{k,j,t}, \quad j = 1, 2, \dots, M \quad (17)$$

其总的采购成本为

$$z_{k,t}^{(2)} = \sum_{j=1}^M c_{4,k,j} y_{k,j,t} \quad (18)$$

阶段 t 每个企业 k 满足市场需求的比例 $\lambda_{k,j,t}$ 为

$$\lambda_{k,j,t} = f_{k,3,t}(\mathbf{p}_{k,t-1}, \mathbf{q}_{k,t-1}, \mathbf{s}_{-k,t-1}) \quad (19)$$

阶段 $t+1$ 每个产品 B_i 和每个原料 G_j 的库存状态分别为

$$p_{k,i,t+1} = g_{k,1,t}(\mathbf{s}_{k,t}, \mathbf{a}_{k,t}), \quad i = 1, 2, \dots, N \quad (20)$$

和

$$q_{k,j,t} = g'_{k,1,t}(\mathbf{s}_{k,t}, \mathbf{a}_{k,t}), \quad j = 1, 2, \dots, M \quad (21)$$

并且满足

$$\sum_{i=1}^N p_{k,i,t+1} \leq C_{k,1}, \quad \sum_{j=1}^M q_{k,j,t+1} \leq C_{k,2} \quad (22)$$

其对应的库存成本分别为

$$z_{k,t}^{(3)} = \sum_{i=1}^N p_{k,i,t} c_{1,k,i} \quad (23)$$

和

$$z_{k,t}^{(4)} = \sum_{j=1}^M q_{k,j,t} c_{2,k,j} \quad (24)$$

阶段 t 企业 k 的收益为

$$r_{k,t} = h(\mathbf{s}_{k,t}, \mathbf{a}_{k,t}) \quad (25)$$

其中, $h(\cdot)$ 是支付函数.

每个企业 k 在 t 阶段的目标是最大化其在阶段 t 到阶段 T 所获得的总的折现收益, 即

$$\max R_{k,t} = \sum_{t'=t}^T \gamma^{t'-t} r_{k,t'}, \quad k = 1, 2, \dots, K \quad (26)$$

其中, γ 是折现因子.

3 基于深度强化学习的求解算法

在企业 ERP 3.0 中, 由于每个企业面临着市场的动态性、竞争性、难以预测性、竞争对手信息的不完全性等复杂的决策环境, 模型 (26) 很难用传统的方法进行求解. 由于深度强化学习^[25-26] 融合了深度学习^[27]、强化学习^[28]、蒙特卡洛树搜索等方法, 因此在解决复杂决策问题时具有强大的优势, 能够高效地解决在规模巨大的解空间中搜索最优策略的复杂问题. 基于此, 本节将基于深度强化学习^[29], 建立虚拟自我对局模型, 求解模型 (26) 的近似最优反应.

3.1 基本思路

假设在时刻 t , 企业观察到环境 ξ 的部分状态 s_t (即企业 Agent 建模的输入信息), 选择行为 $a_t \in A$ (A 为博弈中可选的行为集合), 在下一时刻 $t+1$, 它获得收益 r_{t+1} , 且状态改变为 s_{t+1} . 定义折现因子 γ , 则从 t 到博弈终止时刻 T 的总收益可定义为

$$R_t = \sum_{t'=t}^T \gamma^{t'-t} r_{t'} \quad (27)$$

考虑到实际企业 ERP 场景中往往面临不确定因素, R_t 一般为随机变量, 因此企业的目标是最大

化总收益的期望值. 以 π 表示将状态 s 映射为行为 a 的策略函数:

$$a = \pi(s) \quad (28)$$

$Q^\pi(s, a)$ 表示以 s 为初始状态, 在策略 π 下采取行为 a 可获得的期望总收益, 则有

$$Q^\pi(s, a) = E(r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots | s_t = s, a_t = a, \pi) \quad (29)$$

令 $Q^*(s, a)$ 表示最优期望收益, 则有

$$Q^*(s, a) = \max_{\pi} E(R_t | s_t = s, a_t = a, \pi) \quad (30)$$

由 Bellman 方程, 若已知 $t+1$ 时刻状态为 s' , 且最优收益为 $Q^\pi(s', a')$, 则 $Q^*(s, a)$ 可表示为 $r + \gamma Q^\pi(s', a')$, 即

$$Q_t^*(s, a) = E_{s'}(r + \gamma \max_{a'} Q_{t+1}^*(s', a') | s, a) \quad (31)$$

2015 年, Mnih 等^[30] 提出深度强化学习, 用深度神经网络对 Q 函数进行拟合^[31], 即

$$Q(s, a; \theta) \approx Q^*(s, a) \quad (32)$$

其中, θ 为该 DQN (Deep Q-network) 的参数. 在深度神经网络的训练过程中, 通常采用均方误差 (Mean squared error)、对数均方误差 (Mean squared logarithmic error)、交叉熵 (Cross entropy) 等损失函数, 以及随机梯度下降方法 (Stochastic gradient descent)、内存受限的 BFGS 算法 (Limited-memory BFGS)、共轭梯度法 (Conjugate gradient method) 等优化方法.

由于企业在 ERP 过程中面临着竞争环境, 因此企业不仅要考虑自己的行为, 还要考虑竞争对手的决策. 考虑多个企业参与的不完全信息博弈问题, 在 t 时刻, 每个企业只能获得自己的信息状态 s_t . 每个局中人 (企业) 选择行为策略 (Behavioral strategy) 将信息状态 s_t 映射为可选行为 a 的概率分布, 以最大化自己的期望收益. 定义策略组合 $\Pi = \{\pi^1, \pi^2, \dots, \pi^n\}$, 其中 π^i 是局中人 i 的策略, π^{-i} 表示 Π 中除 π^i 以外的其他策略. 对于一个固定的策略组合 π^{-i} , 使得局中人 i 获得最大收益的策略称为 π^{-i} 的最佳反应 (Best response). 如果在策略组合中每个策略对于其他策略都是最佳反应, 则这个策略组合称为纳什均衡 (Nash equilibrium).

由于最佳反应和纳什均衡的要求非常严格, 在实际应用中往往很难达到, 通常情况下, 只需求解次优的最佳反应, 从而获得近似的纳什均衡. 虚拟对局 (Fictitious play)^[32] 是从自我对局 (Self-play) 中进行学习的博弈论模型. 按照虚拟对局理论, 如果局中人按照对手的平均策略选择最佳反应, 在某些类型

的博弈 (例如二人零和博弈和多人势力场博弈) 中可以收敛到纳什均衡. 推广的弱化自我对局不仅可以保证收敛, 还允许有近似最佳反应和扰动的平均策略更新. 2016 年, Heinrich 等^[33] 提出虚拟自我对局 (Fictitious self play, FSP), 分别用强化学习和监督学习来替换最优反应计算和平均策略更新. 我们将企业之间的虚拟对局经验 $(s_t, a_t, r_{t+1}, s_{t+1})$ 用于深度强化学习, 以训练出能预测行动期望值的贪婪策略网络; 企业自己的行为经验 (s_t, a_t) 用于监督学习, 以训练拟合智能体平均策略的网络. 在虚拟对局中, 假定虚拟竞争对手都采取平均策略, 求解企业的近似最佳反应.

3.2 求解算法

基于上一节的分析, 本节给出求解模型 (26) 的算法, 主要包含以下步骤:

步骤 1. 训练策略网络. 由专家知识或历史数据得到信息状态和行为数据 (s, a) , 训练平均策略网络, 即监督学习 (Supervised learning, SL) 网络. 根据 SL 网络, 可获得企业在给定信息状态下, 可选行为的平均概率分布.

结合企业 ERP 的需求, 采用图 4 所示的深度神经网络框架构建 SL 网络:

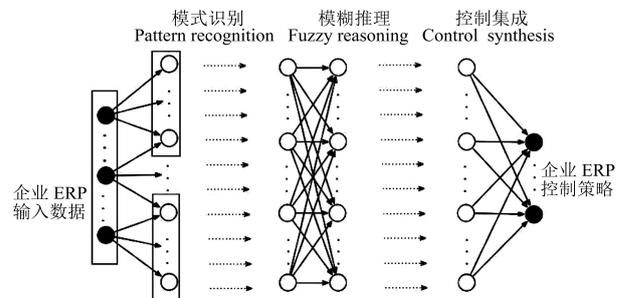


图 4 基于深度神经网络框架的 SL 网络

Fig. 4 SL network based on deep neural network

输入. 企业 ERP 建模的输入数据 s_t .

输出. 企业 ERP 决策 $a_t = (x_{k,t}^*, y_{k,t}^*, \theta_{k,t}^*)$.

处理. 由模式识别、模糊推理、控制集成三部分构成. 其中模式识别部分从低层次的原始输入数据中提取高层次、高维向量形式的潜在特征; 模糊推理部分基于模糊逻辑, 由特征学习出“IF-THEN”规则隶属度; 控制集成部分综合规则隶属度, 最终输出控制策略.

步骤 2. 虚拟对局. 构建贪婪策略网络, 即强化学习 (Reinforcement learning, RL) 网络, 将 SL 网络作为 RL 网络的初始化. 在计算实验中, 构建一个虚拟企业及其多个虚拟竞争对手. 在每轮迭代中展开虚拟企业及其竞争对手之间的博弈, 令虚拟竞争对手按照之前迭代得到的 RL 网络采取策略, 虚拟企业则依据当前的 RL 网络采取贪婪策略. 记录虚拟

企业每次采取行动所获的收益,直到博弈结束.然后追溯整个博弈过程,根据 t 时刻虚拟企业状态、行动及最终总收益 (s_t, a_t, R_t) ,更新RL网络参数.

步骤3. 训练价值网络.在计算实验中,构建一个虚拟企业及其多个虚拟竞争对手.令虚拟竞争对手按照SL网络采取平均策略,虚拟企业则先后依据SL网络与RL网络采取策略.

随机初始化信息状态 s_0 和 L (L 为大于0的整数),假定虚拟企业首先依据SL网络,每次选择概率最高的行动,直到第 $L-1$ 步;第 L 步时,虚拟企业按照均匀概率分布从可选策略集合 A 中随机选择行动;从第 $L+1$ 步到终止时刻 T ,虚拟企业依据RL网络采取行动.记录最终虚拟企业的收益,由 L 步时的信息状态 s_L 和最终收益 R_L 训练价值网络(Value network),用于预测给定信息状态下企业的期望总收益.

步骤4. 平行执行.给定状态 s_T ,分别输入SL网络、RL网络和价值网络,得到平均策略 a_{SL} ,贪婪策略 a_{RL} 及预测总收益 V_k^* .企业以一定概率从 a_{SL} 和 a_{RL} 中随机选择进行执行.此外,记录上一时刻的状态、行为和当前时刻的收益、状态 $(s_{t-1}, a_{t-1}, r_t, s_t)$,对RL网络进行更新.如果企业执行的是 a_{RL} ,则以 (s_T, a_{RL}) 作为新样本对SL网络进行更新.

通过以上4个步骤,即可根据企业的初始输入数据获得近似的纳什均衡策略.

4 结束语

以人工智能为代表的新时代已经到来,企业ERP需要新的“智慧”解决面临的更多复杂性问题.本文在ACP方法框架下,构建基于平行管理的企业ERP 3.0系统,基于Agent方法进行人工企业建模,融合深度强化学习等人工智能方法对人工系统进行计算实验,以虚实互动的思想解决复杂企业ERP中面临的问题.

实现基于ACP方法的虚实互动“平行企业”是建设“智能企业”的基础,也是未来企业ERP的发展趋势.本文即为在这个方向上的一个初步探索.可以预见,在未来的企业ERP中,人工虚拟的系统、工厂、城市将成为现实,大数据真正成为原料,数字化的经验、案例、预演都将成为生产力,计算实验与计算优化将成为首要方法,而虚拟与现实的平行执行将会是企业ERP的“新常态”.

References

- Umble E J, Haft R R, Umble M M. Enterprise resource planning: implementation procedures and critical success factors. *European Journal of Operational Research*, 2003, **146**(2): 241–257
- Zhou Yu-Qing, Liu Bo-Ying, Zhou Qiang. *ERP and Enterprise Management: Theory, Method, System*. Beijing: Tsinghua University Press, 2005.

- (周玉清, 刘伯莹, 周强. ERP 与企业管理: 理论、方法、系统. 北京: 清华大学出版社, 2005.)
- Wang Fei-Yue. Software-defined systems and knowledge automation: a parallel paradigm shift from Newton to Merton. *Acta Automatica Sinica*, 2015, **41**(1): 1–8
(王飞跃. 软件定义的系统与知识自动化: 从牛顿到默顿的平行升华. *自动化学报*, 2015, **41**(1): 1–8)
- Wang Fei-Yue. Intelligence 5.0: parallel intelligence in parallel age. *Journal of the China Society for Scientific and Technical Information*, 2015, **34**(6): 563–574
(王飞跃. 情报 5.0: 平行时代的平行情报体系. *情报学报*, 2015, **34**(6): 563–574)
- Wang F Y. The emergence of intelligent enterprises: from CPS to CPSS. *IEEE Intelligent Systems*, 2010, **25**(4): 85–88
- Chen J X. The evolution of computing: AlphaGo. *Computing in Science and Engineering*, 2016, **18**(4): 4–7
- Wang F Y, Zhang J J, Zheng X H, Wang X, Yuan Y, Dai X X, Zhang J, Yang L Q. Where does AlphaGo go: from church-turing thesis to AlphaGo thesis and beyond. *IEEE/CAA Journal of Automatica Sinica*, 2016, **3**(2): 113–120
- Wang Fei-Yue. Milestone to future: from alphaGo to parallel intelligence. *Science and Technology Review*, 2016, **34**(7): 72–74
(王飞跃. 从 alphaGo 到平行智能: 启示与展望. *科技导报*, 2016, **34**(7): 72–74)
- Wang Fei-Yue. Complexity and intelligence: from Church-Turning thesis to AlphaGo thesis and beyonds (1). *Journal of Command and Control*, 2016, **2**(1): 1–4
(王飞跃. 复杂性智能化: 从 Church-Turning thesis 到 AlphaGo thesis 及其展望 (1). *指挥与控制学报*, 2016, **2**(1): 1–4)
- Wang Fei-Yue. Artificial societies, computational experiments, and parallel systems: a discussion on computational theory of complex social-economic systems. *Complex Systems and Complexity Science*, 2004, **1**(4): 25–35
(王飞跃. 人工社会、计算实验、平行系统——关于复杂社会经济系统计算研究的讨论. *复杂系统与复杂性科学*, 2004, **1**(4): 25–35)
- Wen D, Yuan Y, Li X R. Artificial societies, computational experiments, and parallel systems: an investigation on a computational theory for complex socioeconomic systems. *IEEE Transactions on Services Computing*, 2013, **6**(2): 177–185
- Wang Fei-Yue, Wang Xiao, Yuan Yong, Wang Tao, Lin Yi-Lun. Social computing and computational societies: the foundation and consequence of smart societies. *Chinese Science Bulletin*, 2015, **60**(5–6): 460–469
(王飞跃, 王晓, 袁勇, 王涛, 林懿伦. 社会计算与计算社会: 智慧社会的基础与必然. *科学通报*, 2015, **60**(5–6): 460–469)
- Ragowsky A, Somers T M. Enterprise resource planning. *Journal of Management Information Systems*, 2002, **19**(1): 11–15
- Al-Mashari M, Al-Mudimigh A, Zairi M. Enterprise resource planning: a taxonomy of critical factors. *European Journal of Operational Research*, 2003, **146**(2): 352–364
- Jacobs F R, Ted' Weston Jr F C. Enterprise resource planning (ERP) — a brief history. *Journal of Operations Management*, 2007, **25**(2): 357–363
- McAfee A P. Enterprise 2.0: the dawn of emergent collaboration. *IEEE Engineering Management Review*, 2006, **34**(3): 38–38
- Li Rui. The development of enterprise 2.0 and the related problems. *Modern Information*, 2008, **28**(11): 170–174
(李睿. 企业 2.0 的发展及存在的问题. *现代情报*, 2008, **28**(11): 170–174)

- 18 Wang F Y. Toward a paradigm shift in social computing: the ACP approach. *IEEE Intelligent Systems*, 2007, **22**(5): 65–67
- 19 Wang F Y. Parallel system methods for management and control of complex systems. *Control and Decision*, 2004, **19**(5): 485–489
(王飞跃. 平行系统方法与复杂系统的管理和控制. *控制与决策*, 2004, **19**(5): 485–489)
- 20 Wang Fei-Yue. From parallel universes to parallel management systems PMS. *Money China: Management Scientists*, 2007, **10**: 48–51
(王飞跃. 从平行宇宙到平行管理系统 PMS. *财经界·管理学家*, 2007, **10**: 48–51)
- 21 Balasubramanian S, Maturana F P, Norrie D H. Multi-agent planning and coordination for distributed concurrent engineering. *International Journal of Cooperative Information Systems*, 1996, **5**(2–3): 153–179
- 22 Van Liedekerke M H, Avouris N M. Debugging multi-agent systems. *Information and Software Technology*, 1995, **37**(2): 103–112
- 23 Jiang Li-Juan. Research and Application of ERP System based on Multi-Agent [Master dissertation], Central South University, China, 2008.
(蒋丽娟. 基于多 Agent 的 ERP 系统研究与应用 [硕士学位论文], 中南大学, 中国, 2008.)
- 24 Yuan Yong, Wang Fei-Yue. Sequential equilibrium analysis and computational experiments of a bargaining game with incomplete information. *Acta Automatica Sinica*, 2016, **42**(5): 724–734
(袁勇, 王飞跃. 不完全信息议价博弈的序贯均衡分析与计算实验. *自动化学报*, 2016, **42**(5): 724–734)
- 25 Kulkarni T D, Narasimhan K R, Saeedi A, Tenenbaum J B. Hierarchical deep reinforcement learning: integrating temporal abstraction and intrinsic motivation [Online], available: <http://arxiv.org/abs/1604.06057>, May 31, 2016
- 26 Mnih V, Badia A P, Mirza M, Graves A, Lillicrap T P, Harley T, Silver D, Kavukcuoglu K. Asynchronous methods for deep reinforcement learning [Online], available: <http://arxiv.org/abs/1602.01783>, June 16, 2016
- 27 Duan Yan-Jie, Lv Yi-Sheng, Zhang Jie, Zhao Xue-Liang, Wang Fei-Yue. Deep learning for control: the state of the art and prospects. *Acta Automatica Sinica*, 2016, **42**(5): 643–654
(段艳杰, 吕宜生, 张杰, 赵学亮, 王飞跃. 深度学习在控制领域的研究现状与展望. *自动化学报*, 2016, **42**(5): 643–654)
- 28 Chen Xing-Guo, Yu Yang. Reinforcement learning and its application to the game of Go. *Acta Automatica Sinica*, 2016, **42**(5): 685–695
(陈兴国, 俞扬. 强化学习及其在电脑围棋中的应用. *自动化学报*, 2016, **42**(5): 685–695)
- 29 Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M. Playing Atari with deep reinforcement learning [Online], available: <http://arxiv.org/abs/1312.5602>, December 19, 2013
- 30 Mnih V, Kavukcuoglu K, Silver D, Rusu A A, Veness J, Bellemare M G, Graves A, Riedmiller M, Fidjeland A K, Ostrovski G, Petersen S, Beattie C, Sadik A, Antonoglou I, King H, Kumaran D, Wierstra D, Legg S, Hassabis D. Human-level control through deep reinforcement learning. *Nature*, 2015, **518**(7540): 529–533
- 31 Silver D, Huang A, Maddison C J, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D. Mastering the game of Go with deep

neural networks and tree search. *Nature*, 2016, **529**(7587): 484–489

32 Heinrich J, Lanctot M, Silver D. Fictitious self-play in extensive-form games. In: Proceedings of the 32nd International Conference on Machine Learning, Lille. France: JMLR, 2015. 805–813

33 Heinrich J, Silver D. Deep reinforcement learning from self-play in imperfect-information games [Online], available: <http://arxiv.org/abs/1603.01121>, June 28, 2016



秦蕊 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助理研究员. 主要研究方向为商务智能, 计算广告学, 知识自动化与企业平行管理. 本文通信作者. E-mail: rui.qin@ia.ac.cn
(**QIN Rui** Assistant professor at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. Her research interest covers business intelligence, computational advertising, knowledge automation, and parallel management. Corresponding author of this paper.)



曾帅 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助理研究员. 主要研究方向为社会计算和策略优化. E-mail: shuai.zeng@ia.ac.cn
(**ZENG Shuai** Assistant professor at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. Her research interest covers social computing and strategy optimization.)



李娟娟 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助理研究员. 主要研究方向为商务智能, 计算广告学, 知识自动化与企业平行管理. E-mail: juanjuan.li@ia.ac.cn
(**LI Juan-Juan** Assistant professor at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. Her research interest covers business intelligence, computational advertising, knowledge automation, and parallel management.)



袁勇 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研究员. 主要研究方向为商务智能与计算广告学. E-mail: yong.yuan@ia.ac.cn
(**YUAN Yong** Associate professor at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers business intelligence and computational advertising.)