



脑启智跃:人脑与类脑的协同创新范式

邬霞, 李子遇, 李晴, 李秀星, 王焯

引用本文:

邬霞, 李子遇, 李晴, 等. 脑启智跃:人脑与类脑的协同创新范式[J]. *智能系统学报*, 2026, 21(2): 337-352.

WU Xia, LI Ziyu, LI Qing, et al. Intelligent leap inspired by brain function mechanisms: a collaborative innovation paradigm between human brain and brain-inspired systems[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(2): 337-352.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202509020>

您可能感兴趣的其他文章

一致性协议匹配的跨模态图像文本检索方法

Matching with agreement for cross-modal image-text retrieval

智能系统学报. 2021, 16(6): 1143-1150 <https://dx.doi.org/10.11992/tis.202108013>

基于智能计算的脑机制研究

Brain mechanism research based on intelligent computing

智能系统学报. 2021, 16(5): 850-856 <https://dx.doi.org/10.11992/tis.202103029>

新一代人工智能十问十答

Ten questions and answers for the new generation of artificial intelligences

智能系统学报. 2021, 16(5): 828-833 <https://dx.doi.org/10.11992/tis.202103044>

人工智能中的封闭性和强封闭性——现有成果的能力边界、应用条件和伦理风险

Criteria of closeness and strong closeness in artificial intelligence——limits, application conditions and ethical risks of existing technologies

智能系统学报. 2020, 15(1): 114-120 <https://dx.doi.org/10.11992/tis.202001001>

大数据智能:从数据拟合最优解到博弈对抗均衡解

Big data intelligence: from the optimal solution of data fitting to the equilibrium solution of game theory

智能系统学报. 2020, 15(1): 175-182 <https://dx.doi.org/10.11992/tis.201911007>

泛逻辑学理论——机制主义人工智能理论的逻辑基础

Universal logic theory: logical foundation of mechanism-based artificial intelligence theory

智能系统学报. 2018, 13(1): 19-36 <https://dx.doi.org/10.11992/tis.201711033>

DOI: 10.11992/tis.202509020

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20251223.1708.004>

脑启智跃: 人脑与类脑的协同创新范式

邬霞, 李子遇, 李晴, 李秀星, 王焯

(北京理工大学 计算机学院, 北京 100081)

摘要: 在人工智能技术重塑未来的今天, 其应用成果正深刻改变着社会的各个领域。从智能机器人辅助手术, 到自动驾驶重塑交通模式, 人工智能正以前所未有的速度推动社会进步。然而, 当前人工智能仍面临静态规则、局部模拟与表层表征等挑战, 这导致其难适动态环境、难撑全局智能、难触智能本质, 进而阻碍了通用人工智能的发展进程。人脑作为自然界最复杂的智能系统, 其多模感知、深度认知与智能决策为类脑研究提供了核心启示。现有研究已在感知增强、认知升级和决策优化等层面取得进展, 但仍面临单一层次模拟局限和脑功能机制理解不足等问题。据此, 本文拟打破学科壁垒, 以人脑感知、认知、决策功能机制为纽带, 构建全链条协同创新人工智能框架, 实现从人脑到类脑的功能机制启发优化, 以及类脑反哺人脑功能机制研究的深化。未来, 人脑与类脑的协同创新范式将催生“脑启智跃”, 以脑功能机制为启迪实现智能跃迁, 最终开启人机共生的智能文明新篇章。

关键词: 人工智能; 脑功能机制; 类脑; 全链条协同; 智能跃迁; 感知; 认知; 决策

中图分类号: TP18 **文献标志码:** A **文章编号:** 1673-4785(2026)02-0337-16

中文引用格式: 邬霞, 李子遇, 李晴, 等. 脑启智跃: 人脑与类脑的协同创新范式 [J]. 智能系统学报, 2026, 21(2): 337-352.

英文引用格式: WU Xia, LI Ziyu, LI Qing, et al. Intelligent leap inspired by brain function mechanisms: a collaborative innovation paradigm between human brain and brain-inspired systems[J]. CAAI transactions on intelligent systems, 2026, 21(2): 337-352.

Intelligent leap inspired by brain function mechanisms: a collaborative innovation paradigm between human brain and brain-inspired systems

WU Xia, LI Ziyu, LI Qing, LI Xiuxing, WANG Zhuo

(School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China)

Abstract: Today, the transformative applications of artificial intelligence (AI) are revolutionizing various sectors of society. From surgical robots enhancing medical precision to autonomous vehicles redefining transportation systems, AI is accelerating social advancement at an unprecedented pace. However, current AI technologies still face significant challenges, such as reliance on static rules, limited simulation capabilities, and shallow representation models. These limitations hinder their adaptability in dynamic environments, capacity for global intelligence, and ability to capture the essence of intelligence. The human brain is the most sophisticated intelligent system in nature through its multimodal perception, advanced cognition, and intelligent decision-making. Therefore, this paper tries to transcend disciplinary boundaries by using the functional mechanisms as a bridge and construct a full-chain collaborative innovation paradigm. This paradigm seeks to realize both the optimization of brain-inspired systems guided by human brain mechanisms and the deepened understanding of brain functions through feedback from these systems. Looking ahead, the integration of the human brain and brain-inspired systems will catalyze an “Intelligent Leap Inspired by Brain Function Mechanisms,” achieving a breakthrough in intelligence and ultimately ushering in a new era of intelligent civilization characterized by human-machine symbiosis.

Keywords: artificial intelligence; brain functional mechanisms; brain-inspired systems; full-chain collaboration; intelligent leap; perception; cognition; decision-making

收稿日期: 2025-09-01. 网络出版日期: 2025-12-24.

基金项目: 国家自然科学基金青年基金(A类)项目(62325601);
国家自然科学基金重点项目(62236001).

通信作者: 李子遇. E-mail: ziyuli@bit.edu.cn.

在科技革新驱动发展的当下, 人工智能已成为推动科技变革与社会发展的核心力量。当前, 人工智能领域呈现出蓬勃发展的态势, 在图像识

别、语音交互、文本处理等单一任务上,智能系统展现出超越人类的性能。其发展已进入深度学习主导的阶段,其核心特征体现为数据驱动、算力依赖和任务专用性。以Transformer架构为代表的大规模预训练模型在自然语言处理领域取得突破性进展,GPT(generative pre-trained Transformer)系列模型参数量已达万亿级别,展现出惊人的模式识别能力。然而,现有技术路线面临三重根本性瓶颈:其一,静态规则无法适应动态环境。当前智能系统依赖预设的确定性规则或固定特征工程,在动态环境中暴露出环境适应僵化、知识更新滞后、跨任务迁移失能等矛盾,成为智能系统迈向实用化的关键阻碍^[1]。其二,局部模拟难撑全局智能体系。当前智能系统架构普遍遵循模块化设计范式,聚焦于局部功能的模拟,难以实现真正意义上的全局智能体系,人工智能的发展需要从局部功能的优化转向整体智能的协同进化^[2]。其三,表层处理难触智能本质内核。当前智能系统依赖深度学习等数据驱动范式,在基于浅层表征的模式识别层面表现卓越,缺乏对因果推理、元认知、自主意识等深层次认知机制探索,难以触及智能的本质内核,实现质的飞跃^[3]。这些技术瓶颈暴露出人工智能领域中3种重要的哲学和方法论流派功能主义(Functionalism)、结构主义(Structuralism)和行为主义(Behaviorism)各自局限:功能主义依赖预先设定的物理符号系统和规则库,难以自主更新与调整而无法适应动态环境^[4];结构主义通过人工神经网络模拟大脑结构,易陷入局部功能优化难以构建整体智能的困境^[5];行为主义聚焦刺激-反应映射,因表层处理与深层机制的割裂难以触及智能本质^[6]。即功能主义困于静态规则、结构主义陷于局部模拟、行为主义止于表层处理,三者共同构成人工智能迈向强智能的核心障碍。这迫使学界重新审视“智能”的本质定义,是否需要超越单纯的层次化模仿,转向对人脑智能本质机制的借鉴。

作为亿万年进化的产物,人脑是智能的天然范例,其独特的功能机制为人工智能的发展提供了重要的启发价值。具体而言,人脑的智能依赖于其三大独特的性质:其一,动态可塑学习机制打破静态规则限制。人脑具有很强的适应性和可塑性,能够根据环境变化调整自身结构和功能,这种能力使得人脑能够在复杂多变的环境中表现出卓越的灵活性和鲁棒性^[7]。其二,全脑协同计算范式突破模块架构桎梏。人脑的运作是高度协同的,不同区域之间通过神经连接形成一个有机

整体,这种能力使得人脑能够高效地处理复杂的任务,并实现整体智能^[8]。其三,推理思维形式引擎触及智能本质内核。人脑能够通过推理关系理解世界,对自身进行反思和调控,并主动探索和学习,这种能力使得人脑能够形成对世界本质的深刻理解,实现自我进化^[9]。人脑的动态可塑学习、全脑协同计算及推理思维能力,为突破现有智能系统的静态规则、局部模拟与表层表征三大瓶颈,构建具有环境自适应、全局涌现与本质认知的下一代智能系统提供了生物启发范式。人脑的动态可塑学习、全脑协同计算及推理思维能力是通过感知、认知和决策3个层面的脑功能机制协同作用实现的。其中,感知适配层面通过初级皮层驱动的动态可塑在线学习实现基础构建^[10],认知映射层通过高级联合皮层主导的全脑协同知识整合深化发展^[11],决策执行层通过逻辑规则与因果价值导向的推理思维过程助力综合应用^[12]。从感知适配层的多模态协同处理、认知映射层的知识抽象与推理,到决策执行层的目标策略调整,每一个层次的机制都蕴含着解决人工智能技术瓶颈的关键线索,三者协同作用可增强人工智能在复杂多变环境中的灵活性和鲁棒性,推动人工智能从“工具型智能”向“认知型智能”的历史性跨越。

本文旨在建立生物智能与机器智能的统一解释框架,推动脑启智跃。脑启智跃是本研究提出的核心范式,指通过解码人脑“感知-认知-决策”全链条功能机制的底层逻辑,构建生物智能与机器智能的双向赋能、协同进化体系,最终实现从专用智能向通用智能的质性跃迁。这一框架将为揭示生物智能与机器智能在信息处理、知识表达、决策制定等方面的共性与差异,构建起统一的理论体系,理解智能的本质、探索智能的起源与发展提供全新视角,促进认知科学、神经科学与人工智能等多学科的交叉融合。此外,本文致力于为强人工智能提供可解释、可进化的实现路径,该路径不仅能够突破现有人工智能技术瓶颈,更可能揭示智能涌现的本质规律。这种从人脑到类脑的创新正在构建“生物-机器”共生演化的新生态,为人工智能进化提供新的借鉴方向。

1 脑认知机制的研究进展

人脑功能机制在感知适配、认知映射与决策执行3个层面呈现出精密且层次分明的协同运作体系。在感知适配层,初级皮层凭借动态可塑特性,通过神经连接与功能的适应性改变,实现对

单模态信息的精准处理及多模态信息的跨脑区整合,为后续认知加工提供结构化信息基础;认知映射层以前额叶、颞叶和顶叶等高级联合皮层为核心,将分散的感知信息、经验记忆、抽象概念等进行知识整合,形成结构化知识体系,构建全局、连贯的认知表征;决策执行层则是基于规则的逻辑推理和基于价值的因果推理的动态平衡,以前额叶为枢纽,基底神经节、边缘系统等协同作用,实现推理思维与目标导向决策。三大层面层层递进、紧密关联,共同构成了人脑从环境感知、知识构建到行为决策的完整智能链条,为理解人类智能本质与推动类脑智能发展提供了重要参照。

1.1 感知适配层:初级视听感皮层驱动的动态可塑在线学习机制

感知适配层是人脑信息处理的基础层次,主要负责对来自外部环境的原始信息进行接收、处理和初步解析^[10]。在感知适配层,初级皮层构成了与外界环境交互的基础神经架构,凭借其高度动态可塑的特性,实现对环境信息的高效在线学习。

以视觉系统为例,初级视皮层(V1区)作为视觉信息处理的首站,接收来自视网膜的原始视觉信号。当个体处于持续变化的视觉环境中,如频繁切换不同光照强度、物体布局的场景时,V1区与外侧膝状体之间的神经连接会发生适应性改变,通过皮层可塑性机制调整信息传递效率^[13]。研究发现,长期进行精细视觉任务的个体,其V1区的激活范围和强度显著高于普通人,表明该脑区通过结构与功能的重塑,增强对特定视觉特征的处理能力^[14]。在听觉感知中,初级听觉皮层(A1区)同样展现出强大的适应性。例如,当个体暴露于嘈杂的环境中,A1区与听觉丘脑的协同活动模式会发生改变,通过增强对目标频率的响应、抑制背景噪声相关神经活动,实现听觉信息的精准提取^[15]。更进一步,通过初级皮层之间的跨脑区协作可实现多模态感知整合,初级视皮层、初级听觉皮层和初级体感皮层通过胼胝体、前连合等神经纤维束建立联系^[10]。例如,婴儿在探索环境时,手部触摸物体的触觉信号由初级体感皮层处理,与视觉皮层接收到的物体形状信息相互关联,通过皮层间的同步激活,逐步建立起多模态感知表征^[16]。

现代神经影像学技术进一步揭示,初级皮层在学习过程中呈现出动态的功能重组。例如,当个体初次接触虚拟现实环境,初级视皮层与前庭皮层的功能连接在数小时内显著增强,这种快速的神经可塑性变化使得个体能够迅速适应虚拟空

间中的运动感知需求^[17]。此外,长期的感知学习还会引发初级皮层的结构改变,例如,经过专业训练的音乐家,其初级躯体感觉皮层(S1区)、初级听觉皮层(A1)与初级运动皮层(M1)之间形成的超模态整合网络,其功能连接强度显著高于非音乐家^[18]。

感知适配层作为人脑与外界交互的“先锋阵地”,通过初级皮层的动态可塑特性完成在线学习,持续优化感知策略,快速适应环境变化。其不仅为认知映射层提供经过初步处理的、结构化的信息素材,是高级认知功能得以展开的前提;也为人工智能突破感知瓶颈、实现多模态融合感知提供了生物原型,启发类脑感知系统的设计与构建。

1.2 认知映射层:高级认知皮层主导的全脑协同知识整合体系

认知映射层是人脑信息处理的核心层次,负责对感知适配层提供的信息进行深度加工和理解,形成知识和概念,并进行推理和联想^[19]。认知映射层以高级皮层为核心,构建多脑区协同的知识整合网络。

前额叶皮层、颞叶皮层和顶叶皮层通过弓状束、钩束等白质纤维束形成紧密的神经环路,实现感知信息的深度融合与抽象加工^[20]。以大脑语言理解过程为例,枕叶视觉皮层将文字信息传递至颞叶联合皮层的视觉词形区进行字形解析,随后信息经弓状束传输至位于前额叶皮层的布洛卡区进行句法分析,同时颞叶皮层的韦尼克区完成语义理解,这些脑区通过神经振荡实现同步活动,形成“语言理解网络”,协同完成从文字符号到语义概念的转换^[21]。

这一过程的知识整合也高度依赖于高级皮层与记忆相关脑区的协同运作。海马体作为情景记忆的关键脑区,与前额叶皮层存在双向投射。当个体学习历史事件时,海马体将事件的时间、空间和情节信息编码,前额叶皮层则对这些信息进行逻辑分析与关联整合,通过持续的神经活动,将分散的记忆片段转化为结构化知识图谱^[22]。神经心理学研究发现,患有海马体损伤的患者,虽然保留基础感知能力,但在知识整合与推理任务中表现显著下降,凸显海马体-前额叶皮层通路在知识构建中的重要性^[23]。

最后,高级皮层的协同能力还体现在复杂问题解决过程中。主要表现在,顶叶皮层负责空间关系与数据整合,前额叶皮层进行逻辑推理与策略规划,颞叶皮层提取相关领域知识,多个脑区通过默认模式网络和任务正激活网络的动态切

换,实现信息的高效交互^[24]。例如,在数学证明过程中,研究者的前额叶背外侧皮层与顶叶下小叶的功能连接强度与推理效率呈正相关,表明全脑协同程度直接影响认知效能^[25]。这种多脑区协同机制使人类能够超越感知信息的表层特征,构建抽象概念体系,推动认知边界的扩展。

认知映射层作为感知与决策间的关键枢纽,衔接了感知适配层处理后的多模态信息,通过多脑区同步神经振荡和网络动态切换,实现知识的结构化整合,完成从具体感知素材到抽象认知的跨越,也为复杂推理和决策奠定了认知基础。这种承上启下的特性,让认知映射层成为实现智能跃迁的核心桥梁,也为类脑智能的研究指明了从信息处理到知识构建的关键路径。

1.3 决策执行层:逻辑规则与因果价值导向的推理思维决策过程

决策执行层是人脑信息处理的最高层次,负责根据认知映射层提供的知识和信息,制定行动计划并做出决策^[26]。在这一过程中,逻辑推理侧重于规则和概念的理性推导,因果推理则更强调对事件因果关系的理解 and 价值判断,二者相互补充,为复杂问题的决策提供强大的工具。

逻辑推理是人类基于规则和概念进行理性思考的重要能力,主要依赖大脑前额叶皮层、顶叶皮层以及相关脑区的协同活动。前额叶皮层,尤其是背外侧前额叶皮层,在逻辑推理过程中发挥着核心作用^[26]。负责对信息进行工作记忆的维持、组织和操作,以及对推理策略的规划与执行。顶叶皮层则在逻辑推理中参与空间关系处理和数字运算等功能^[27]。当个体进行逻辑推理任务时,颞叶率先发挥作用,作为语义信息处理中心,从长时记忆中快速提取与任务相关的概念、知识和语义规则^[28]。背外侧前额叶皮层接收来自颞叶的语义信息后,利用工作记忆对其进行暂存和组织,同时制定推理策略^[26]。在此过程中,背外侧前额叶皮层会与顶叶建立紧密的神经连接。背外侧前额叶皮层则根据顶叶提供的空间信息,结合颞叶的语义知识,对推理步骤进行规划和调整,确定逻辑推导的方向和顺序,得出结论^[29]。

因果推理是人类理解世界、预测事件和做出合理决策的核心能力,其依赖于前额叶皮层和边缘系统的紧密协同^[30]。前额叶皮层能够处理复杂的逻辑推理、规划和目标导向的行为。通过整合来自认知映射层多个脑区的信息,评估不同行动方案的潜在后果,并选择最优的行动方案^[26]。边缘系统主要负责处理情感和本能反应,包括杏仁

核、扣带回皮层和基底神经节等结构,这些结构在情感调节、记忆形成和习惯行为的形成中起重要作用^[31]。杏仁核作为情感处理的关键脑区,与前额叶皮层的腹内侧区域形成双向连接,实时评估决策选项的情感价值^[32]。前扣带回皮层则在决策过程中扮演错误监测与策略调整的重要角色^[33]。当实际决策结果与预期不符时,前扣带回皮层与纹状体、前额叶皮层形成的神经环路被激活,通过调节多巴胺能神经传递,促使前额叶皮层重新评估决策规则。

决策执行层是大脑将感知信息与认知成果转化为实际行动的“最后一环”,其承接认知映射层整合形成的知识体系,通过前额叶皮层提供的逻辑分析和理性评估以及边缘系统对决策选项的情感价值反馈,综合理性与感性因素做出决策。决策执行层为类脑智能系统探索如何融合情感计算与因果推理,实现更具人类智慧特征的决策提供了生物蓝本。

2 人工智能的发展现状与变革必然性

2.1 人工智能的变革必然性

在科技高速发展的当下,人工智能正处于关键的变革节点。从经济、哲学和政策等多视角审视,人工智能向类脑智能的变革已成为不可阻挡的趋势,这一转变关乎技术发展的可持续性、人类对智能本质的探索以及社会发展的战略需求。

2.1.1 经济视角:智能摩尔定律与人脑能效竞赛

从经济视角来看,人工智能发展正面临算力-能效比的终极制约。传统的摩尔定律描述了集成电路上可容纳的晶体管数目大约每 18 个月便会增加 1 倍,性能也将提升 1 倍^[34]。为了追赶人脑的效率,人工智能领域提出了“智能摩尔定律”这一量化标尺,人工智能系统的能效必须每 18 个月提升 10 倍。而现有的人工智能系统能耗巨大,例如,训练 GPT-3 耗费 128 万美元电力,相当于人类大脑 10 年代谢能量。并且随着人工智能在医疗、交通、金融等领域应用的逐渐深入,对智能系统的实时性、准确性和通用性要求的不断提高,若人工智能无法实现跨越式提升,将导致各行业的发展陷入瓶颈。因此,推动人工智能的变革,实现人工智能与人脑的竞争力平衡,促进其在各行业的深度融合,可创造更大的经济价值。

2.1.2 哲学视角:认知之轮隐喻与符号困境批判

从哲学视角来看,人工智能的发展面临着深刻的理论困境。哲学家 Danieli^[35]提出的“认知之轮”(Cognitive Wheels)隐喻,深刻揭示了当前智能

系统的本质局限: 依赖脆弱符号系统的智能如同坐在轮椅上的巨人, 虽能执行特定任务, 但难以自主跨越认知鸿沟, 无法像人脑一样理解世界的本质与意义。现有的人工智能大多基于符号系统和数据驱动, 无论是早期的专家系统, 还是当下流行的深度学习模型, 都依赖于预先设定的规则或从大量数据中提取的模式。人脑则截然不同, 它能够在与环境的交互中, 通过感知、认知和决策, 构建起对世界的理解体系。人工智能的符号困境, 本质上是哲学层面上对智能本质认知的缺失。要突破这一困境, 就必须深入研究人脑的功能机制, 向具有真正认知能力的类脑智能迈进, 实现从“工具智能”到“认知智能”的转变。

2.1.3 政策视角: 中国类脑计算战略与全球格局

从政策视角来看, 中国《新一代人工智能发展规划》明确将类脑计算列为重点发展方向, 并制定了 2025 年的具体目标, 这一战略部署深刻体现了人工智能变革的必然性与紧迫性。从全球竞

争格局来看, 世界各国都在积极布局人工智能与类脑计算领域。美国国防部的 DARPA、空军研究实验室等耗时近 20 年资助 IBM 研发类脑推理芯片 NorthPole; 欧盟继续推进“Human Brain Project”第二阶段, 整合神经科学数据、开发 EBRAINS 平台、推动脑疾病模型、类脑计算理论以及类脑机器人研究。政策层面的导向对人工智能的发展起着关键的推动作用, 在这样的国内和国际背景加持下, 可进一步加速人工智能的变革进程, 推动我国在这一前沿领域实现跨越式发展。

2.2 主流人工智能范式的技术突破与瓶颈

人工智能领域的发展始终在探索与突破中前行, 结构主义、功能主义和行为主义作为主流范式, 深刻影响着技术演进方向, 同时也面临着难以逾越的瓶颈。其在各自领域的辉煌成就与深层局限, 不仅揭示了当前人工智能发展的困境, 也为后续技术突破指明了方向。三大人工智能范式对比分析如表 1 所示。

表 1 功能主义、结构主义、行为主义三大人工智能范式对比分析

Table 1 Comparative analysis of three major artificial intelligence paradigms: functionalism, structuralism and behaviorism

维度	功能主义	结构主义	行为主义
核心思想	符号操作与逻辑推理, 依赖预设规则系统	神经网络结构仿生, 通过连接权重重建功能	环境刺激-行为反应的关联优化, 强调试错学习
代表技术	专家系统、知识图谱、逻辑编程 (Prolog)、符号推理引擎	人工神经网络 (卷积神经网络/循环神经网络/Transformer)、脉冲神经网络、神经形态芯片	强化学习 (Q-learning、近端优化策略)、进化算法、多智能体系统
核心优势	可解释性强、逻辑严谨、适合规则明确的封闭系统	感知能力强、自适应学习、擅长模式识别	环境适应性好、无需预设知识、适合动态决策
核心瓶颈	静态规则僵化、知识获取困难、无法处理不确定性	缺乏全局协同、可解释性差、知识迁移能力弱	依赖奖励设计、缺乏深层认知、难以因果推理
与人脑差距	缺少动态可塑性、无法自主更新知识体系	未实现全脑区协同、停留在局部功能模拟	缺乏内在价值体系、仅停留在表层行为模仿

2.2.1 功能主义: 预设规则体系的僵化困局

功能主义以物理符号系统为基础, 通过将领域知识编码为规则和逻辑表达式, 实现对问题的推理与决策。早期的专家系统是功能主义的典型应用, 这些系统在规则明确、边界清晰的场景中, 展现出高效且准确的决策能力, 为特定领域的自动化和智能化发展提供了有力支持^[36]。

然而, 现实世界的复杂性与动态性远超功能主义预设规则体系的处理能力。首先, 规则的获取与更新面临巨大挑战。领域知识的获取往往依赖专家经验的人工总结与编码, 这一过程耗时耗力且容易出现知识遗漏或错误。其次, 功能主义系统对不确定性和模糊性信息的处理能力极为有限。现实中的许多问题无法用精确的规则描述, 自然语言的语义模糊性、复杂环境中的噪声干扰、未定义场景的出现等, 都会导致系统决策失

效。最后, 功能主义系统的推理过程严格依赖预设规则, 缺乏自主学习与创新能力。当遇到规则覆盖之外的新问题时, 系统无法通过类比、归纳等方式进行推理, 只能输出错误结果或无法响应。这种僵化的决策模式, 使得功能主义在应对动态变化的环境时, 逐渐暴露出严重的局限性, 亟需向更具适应性和灵活性的智能范式转型。

2.2.2 结构主义: 固定架构的局部模拟困境

结构主义以人工神经网络为核心载体, 通过对大脑皮层神经网络结构的仿生, 试图复刻人类智能的信息处理模式。其核心逻辑在于构建多层次、多节点的网络架构, 从输入层接收原始数据, 经隐藏层的逐层特征提取与变换, 最终在输出层产生决策结果^[37]。

然而, 这种局部模拟的方式在构建全局智能体系时面临着根本性困境。首先, 人工神经网络

的架构在训练完成后便趋于固定,缺乏动态调整与自适应能力。其次,不同类型的神经网络往往针对单一任务设计,各模块间缺乏有效的跨模态协同机制,由于缺乏统一的表征空间和交互接口,难以实现深度融合,表现出明显的割裂感。最后,结构主义智能系统在知识迁移与泛化能力上的不足,由于每个网络架构针对特定任务优化,其学习到的特征与知识具有强任务依赖性,难以在不同任务或领域间复用。这种局部模拟的局限性,使得结构主义人工智能始终徘徊在专用智能阶段,难以跨越到能够处理复杂现实问题的通用智能水平。

2.2.3 行为主义: 动态环境适应的表象局限

行为主义以强化学习为代表技术,通过智能系统与环境的交互和反馈机制,实现行为策略的优化与环境适应。其基本原理是定义奖励函数,智能系统在环境中执行动作后,根据获得的奖励信号调整策略,以最大化长期累积奖励^[38]。

尽管在动态环境适应方面表现出色,但行为主义仅停留在行为层面的反馈调节,难以触及智能的本质内核。首先,智能系统缺乏对环境的深度理解和抽象认知能力,其行为策略的形成仅依赖于奖励信号的驱动,而非对环境内在规律的理解。其次,行为主义智能系统不具备因果推理、元认知等高级认知能力,难以像人类一样通过因果分析预测行为的长期后果,反思自身决策过程并进行策略调整。最后,行为主义对奖励函数的高度依赖带来了严重的局限性。奖励函数的设计直接决定了智能系统的行为目标,不合理的奖励设计可能导致智能系统学习到错误或短视的行为策略。这种表层的适应方式,使得行为主义智能始终处于“刺激-反应”的低层次循环,无法像人类智能一样,基于对世界本质的理解、价值判断和目标规划,实现具有深度和创造性的智能行为。

这三大主流人工智能范式,沿用传统物质科学的“分而治之”策略,缺乏有效沟通与协同机制,如同“信息孤岛”般独立运作。结构主义的神经网络难以直接与功能主义的符号系统进行信息交互,行为主义的强化学习策略也无法融入前两者的知识体系,使得人工智能系统在面对复杂现实任务时,无法整合多方面能力,陷入局部最优而难以实现全局智能。这种碎片化的研究现状,使得人工智能难以形成通用理论。三大范式的割裂现状,凸显了“脑启智跃”所倡导的全链条协同创新的必要性。因此,亟需构建新型信息科学范式,整合人脑的“感知-认知-决策”全链条功能机

制,实现信息在各环节之间的无缝流转与协同处理,构建基于“信息转换与智能创生”的通用理论,为人工智能的发展开辟新的道路。

3 脑功能机制启发的人工智能优化

在人工智能发展历程中,借鉴脑功能机制有望成为突破现有技术瓶颈,推动人工智能向更高智能水平迈进的关键路径。从感知适配层、认知映射层到决策执行层,人脑独特的信息处理模式为人工智能算法的优化提供了丰富灵感,推动了人工智能在感知效率、认知深度与决策合理上的突破。

3.1 人脑多模感知机制启发的人工智能感知优化

人脑的感知系统是一个高度复杂且精妙的体系,初级视听感皮层在其中扮演着至关重要的角色,负责对视觉、听觉、躯体感觉等多模态信息进行初步处理与特征提取。受此启发,人工智能领域致力于开发基于初级视听感皮层启发的智能算法,以提升智能系统的感知能力。

在视觉感知方面,人脑视觉系统以层级化和模块化著称^[39],初级视觉皮层(V1)神经元依据自身的感受野特性,对视觉刺激进行初步筛选与分析,提取出基础视觉特征。随后,这些特征信息被逐级传递至高级视觉皮层,如V2、V4和颞下回等区域。在高级视觉皮层中,神经元的感受野逐渐扩大,对复杂物体和场景的识别能力不断增强。受此启发,人工智能视觉模型中引入层级化结构与特征提取机制。以卷积神经网络为例,其通过多层卷积核的设计,模拟人脑视觉皮层神经元从低级特征到高级特征的逐步提取过程^[40]。谷歌的Inception系列模型便是这一思路的杰出代表^[41-43],通过精心设计的多尺度卷积模块,模拟人脑对不同视觉信息的并行处理,极大提升了智能算法对复杂图像的感知与分类能力。

在听觉感知方面,人脑听觉系统以时序敏感性和空间定位性为核心,实现场景化语义解析。具体来说,人脑对声音的时序依赖通过听觉皮层的层级化时间滤波器实现,受此启发,循环神经网络^[44]、长短期记忆网络^[45]以及Transformer^[46]等架构应时而生,用于捕捉时序序列的相关性以及长程依赖信息。此外,人脑听觉系统具有空间定位功能,下丘中的神经元按照一定的拓扑结构对不同空间方位的声音进行编码,德国海德堡大学基于此提出一种模仿生物听觉系统下丘神经元拓扑映射原理的创新型芯片,在实时3D声源定位方面表现出色,且具有低功耗的优势^[47]。

此外,多模态感知融合也是感知适配层优化的重要方向。在日常生活中,人脑能够轻松地将视觉、听觉、触觉等多种感官信息融合,形成对周围环境和事物的全面、准确认知。受此启发,人工智能领域的多模态感知融合研究取得了诸多进展^[48-52]。例如,Wang等^[48-49]提出一种受神经元多样性启发的多感官识别模型,模拟大脑初级视听皮层单感官神经元以及多感官神经元之间的相互作用关系,设计特殊的连接约束来调节不同神经元之间的信息传递,以促进不同模态数据之间的交互融合。Li等^[50]基于哺乳动物丘脑和皮层整合多模态感觉信息的工作原理构建脑启发模型,分别模拟听觉皮层、视觉皮层和背侧高级听觉丘脑进行跨模态信息整合。He等^[51]提出一种逆效能驱动的多模态融合策略,模拟大脑在单一感觉信号弱时增强多感官整合效果,信号强时降低融合依赖的方式,使智能系统能够动态调整多模态信息融合强度,在大幅降低计算成本的同时,提高了模型在视听分类、持续学习和问答任务上的性能。

感知适配层的研究正从“单模态独立处理”向“脑启发的多模态融合”演进,视觉感知启发的模型通过层级化与选择性机制提升特征表达;听觉感知启发的模型借助时序建模与空间感知增强场景理解;多感知启发的模型则以跨模态统一表征为核心,推动人工智能向“更接近生物感知”的智能系统发展。

3.2 人脑知识整合机制启发的人工智能认知升级

人脑的认知系统主导着全脑协同知识整合体系,负责将感知信息转化为抽象知识,并建立知识之间的关联与推理网络,为人类的逻辑推理、抽象思维和知识迁移等能力奠定坚实基础。在人工智能领域,受此启发的研究致力于开发能够模拟高级联合皮层中知识表征、抽象思维以及知识迁移功能的算法,以实现人工智能认知能力的升级。

在知识表征方面,人脑通过神经元之间的连接模式对知识进行编码存储^[53]。不同的概念、事实和经验在大脑中并非孤立存在,而是通过关联记忆形成复杂的关联网络。这种分布式、关联性的知识表征方式,使人类能够快速检索和运用知识,知识图谱技术正是对这一机制的模拟与应用。知识图谱以图结构的形式,将实体如人、物、概念等表示为节点,实体之间的关系如所属、因果等表示为边,构建起一个庞大的语义网络^[54]。当用户输入问题时,搜索引擎能够通过知识图谱快速定位相关实体和关系,理解问题的语义,并

给出准确的答案,给智能系统注入人类般的知识理解与运用能力。

在抽象思维方面,人脑能够从具体的事物和现象中提炼出一般性的概念和规律^[55]。在这一过程中,大脑可基于大量已有经验的观察和分析,对复杂信息进行筛选、归纳和概括,抽象出事实规律。为实现抽象思维能力,现有智能算法主要包括基于元学习和基于生成对抗网络两大类。基于元学习的方法让模型在多个不同的任务和领域中进行学习,从而掌握解决问题的通用策略和模式^[56]。生成对抗网络由生成器和判别器组成,生成器尝试生成符合特定概念的样本,判别器则判断生成的样本是否真实。通过二者的对抗训练,生成器能够逐渐学习到抽象概念的本质特征,并生成高质量的样本,实现抽象概念的学习与生成^[57]。

在知识迁移方面,大脑可对不同知识之间的相似性和关联性进行识别,实现高效的多任务迁移^[58-59]。为了模拟大脑的知识迁移能力,知识迁移学习被提出,旨在解决智能算法在不同任务和领域之间的泛化问题。迁移学习方法通常先在一个大规模的源任务上进行预训练,使模型学习到通用的知识和特征表示,然后将这些知识迁移到一个或多个目标任务上进行微调^[60]。例如,在自然语言处理中,许多预训练语言模型如基于Transformer的双向编码器(bidirectional encoder representations from Transformer, BERT)、GPT等在大规模文本数据上进行预训练,学习到丰富的语言知识和语义表示^[61]。当应用于具体的下游任务,如文本分类、情感分析时,只需在少量目标任务数据上进行微调,模型就能快速适应新任务,利用已学到的知识提高任务性能,实现知识的跨任务迁移。此外,为了提升智能系统的动态适应性,增量学习、持续学习等模仿人脑的终身学习机制,进一步增强了预训练模型的灵活性。例如,Zhou等^[62]结合预训练模型的泛化能力与适应性调整,通过预训练特征与微调特征融合和原型分类器更新,在类别增量学习中实现高效知识整合。Han等^[63-64]借鉴人脑海马体的模式分离机制与长短时记忆动态更新策略,提出了一种新型的持续学习模型以提升其在多任务学习过程中的稳定性与可塑性。

认知映射层的研究核心聚焦于将感知信息转化为可推理、可交互的结构化知识,未来,认知映射模式向共生智能的转变,成为连接感知和决策的智能核心枢纽,为实现跨领域通用智能奠定基础。

3.3 人脑推理思维机制启发的人工智能决策革新

人脑的决策过程由推理思维机制支撑,涉及逻辑推理和因果推理协作,融合了情感价值判断、理性逻辑分析以及对潜在风险的评估,这为人工智能决策算法的革新提供了重要启示。现有研究尝试将人脑决策机制融合人工智能系统,使其具备更合理、高效的决策能力。

在逻辑推理方面,演绎推理是人工智能领域实现逻辑推理的重要方式。基于规则的专家系统是早期演绎推理的典型应用,通过预先设定的规则和知识库,系统能够对输入的事实进行匹配和推理,得出结论。然而,这种方法在面对复杂多变的现实问题时,灵活性和扩展性有限^[65]。近年来,深度学习与逻辑推理的结合成为研究热点。以神经符号推理为例,神经符号推理将神经网络强大的学习能力和符号逻辑系统的推理能力相结合。神经网络负责从数据中提取特征和模式,生成符号化的知识表示,然后利用符号逻辑推理引擎进行推理运算。例如,在知识图谱推理任务中,Tang等^[66]通过训练神经网络对知识图谱中的实体和关系进行嵌入表示,使其能够捕捉到语义信息,再结合逻辑规则进行推理,预测知识图谱中缺失的关系,显著提升了推理的准确性和效率。在智能系统路径规划任务中,Jia等^[67]从两个维度模拟大脑逻辑推理机制,结构层面仿照大脑对逻辑链条的可靠性筛选设计了智慧边缘选择器,过程层面仿照大脑关系记忆结合感官观察和关系结构设计了记忆感知预测器,显著提升了规划性能。

在因果推理方面,因果推理旨在让机器能够理解事件之间的因果关系,从而做出更具解释性和可靠性的决策。传统的机器学习方法大多基于相关性进行预测,难以挖掘数据背后的因果机制。近年来,研究人员开始借鉴人脑因果推理机制,发展出基于因果贝叶斯网络和结构因果模型等的人工智能算法^[68]。其中,因果贝叶斯网络模拟大脑杏仁核与前额叶在风险评估中,会构建因果链这一特性,在贝叶斯网络中引入因果方向,模拟大脑对因果层次的判断。结构因果模型则通过构建因果图来表示变量之间的因果关系,利用干预和反事实推理等方法,模拟人脑在因果判断中的思维过程。例如,Matthieu等^[69]开发了一款强化学习智能体——人工智能临床决策系统。该系统从海量患者数据中提取隐性临床知识,通过分析大量以非最优方案为主的临床治疗决策,习得脓毒症的最优治疗策略。另一方面,基于深度学习的因果表征学习成为研究热点,因果表征

学习依赖因果推理的核心概念如因果结构、干预、反事实,通过因果发现、干预模拟构建和优化表征实现有效决策,其典型代表包括因果图生成模型和因果图神经网络^[70]。其中,因果图生成模型将因果图视为生成过程的先验,用生成对抗网络或变分自编码器学习数据的生成机制,并推断背后的因果结构。因果图神经网络利用图神经网络的图结构建模能力,将变量间的依赖关系编码为图节点的交互,通过消息传递机制学习因果方向。例如,Wang等^[71]提出类脑因果感知对比学习的对比式强化学习框架,旨在实现高效的表征学习与类人自动驾驶决策。

此外,强化学习与逻辑/因果推理机制的结合近来受到了研究者的广泛关系^[72-73]。强化学习通过不断尝试动作(探索),记录每个动作对应的环境反馈(奖励),并基于统计规律(如蒙特卡洛采样、时序差分法)更新策略,学习状态-动作-奖励的概率映射关系,实现与环境的交互^[74]。可以看出,强化学习的核心范式与人类的“经验试错学习”类似,依赖数据关联而非先验的因果理解或逻辑规则。将因果推理和逻辑推理引入强化学习,不仅能够促进知识在不同任务间的迁移泛化,而且可以增强可解释性表征,让决策过程从“黑箱”变为“白箱”,显著增强智能系统在复杂场景中的决策可靠性。逻辑推理型强化学习系统分为逻辑正则化强化学习和神经符号强化学习两大类。其中,逻辑正则化强化学习将符号逻辑规则转化为数值约束,融入强化学习的损失函数中。例如,Fan等^[75]通过挖掘训练文本中的逻辑规则,在文档级关系抽取任务中引入一致性正则化,惩罚违反规则的预测,提升全局推理能力。神经符号强化学习用神经网络处理感知,用逻辑引擎处理高层决策。Jia等^[76]借鉴逻辑推理过程中大脑的并行推理机制和多策略协同机制,设计了并行仲裁模块有效平衡探索与利用的权衡,开发了多策略协调学习模块指导智能系统动态优化训练缓冲区,进一步提升了高维连续环境中的运动规划效率。因果推理型强化学习系统^[77-79]根据因果信息是否先验,可划分为以下两大类:基于先验因果信息和未知因果信息的强化学习。其中,基于先验因果信息的强化学习通常假定环境模型中的因果结构已知,利用已有的因果知识来指导智能系统学习最优策略,优化决策过程。例如, Ji等^[77]首次将因果分析引入熵正则化,通过评估动作维度与奖励的因果效应,动态调整探索优先级;并进一步引入休眠引导重置机制,防止过度依赖局

部最优动作, 显著提升了多个连续控制任务的采样效率。基于未知因果信息的强化学习通常需要额外地进行因果结构学习、因果表征学习等, 智能系统需要从数据中发现因果结构, 或者学习有效的因果表征, 以便更好地理解环境并做出决策。例如, Cao 等^[78]利用观测数据学习环境的因果转移模型, 捕捉变量间的直接因果效应; 通过最大化未来状态与动作的互信息选择干预动作, 优先探索高控制潜力的节点; 并使用探索数据更新因果模型, 结合内在好奇心奖励平衡因果学习与工业机器人路径规划目标。通过与逻辑与因果推理的交叉融合, 强化学习可突破“缺乏显式知识约束”和“仅知相关, 不知因果”的局限, 向更具解释性、泛化性的智能决策系统演进。未来, 逻辑强化学习和因果强化学习可能成为连接机器学习与认知推理的关键路径, 推动智能系统在复杂场景中的决策能力向人类水平靠近。

可以看出, 脑功能机制在人工智能优化的各个层面提供了丰富且有效的借鉴思路, 从感知适配到认知映射, 再到决策执行, 基于脑功能机制启发的算法研究正逐步推动人工智能向更智能、更接近人类智能水平的方向发展。然而, 现有基于脑功能机制启发的人工智能研究仍然存在局限性, 一方面, 现有大多数研究仅聚焦于单一层次机制的模拟, 未能实现感知、认知与决策层面的全链条优化; 另一方面, 人类对脑机制的理解尚不充分, 使得人工智能只能进行表层的结构和功能模仿, 难以把握其本质。因此, 亟需构建脑功能机制启发的全链条协同创新范式, 以实现人工智能质的飞跃。

4 脑功能机制启发的全链条协同创新

人脑作为已知宇宙中最复杂的智能系统, 其运行机制为人工智能的发展提供了终极“蓝本”。脑功能机制启发的全链条协同创新, 本质是通过解码脑功能机制的底层规律, 将生物智能的“软件逻辑”与机器智能的“硬件能力”深度融合, 构建自适应并且可进化的智能系统。这一创新范式可打破“生物智能”与“机器智能”的边界, 实现二者的相互促进, 并最终推动智能跃迁。

4.1 全链条协同创新框架

“脑启智跃”范式具有双向性(人脑启发类脑优化, 类脑反哺脑科学研究)、全链性(突破单一层次模仿, 实现感知-认知-决策闭环协同)和涌现性(通过动态耦合产生 $1+1>2$ 的智能涌现效应)。脑功能机制启发的全链条协同创新, 以实

现生物智能与机器智能深度融合为核心目标, 致力于构建具备自适应、可进化能力的智能系统, 框架如图 1 所示。

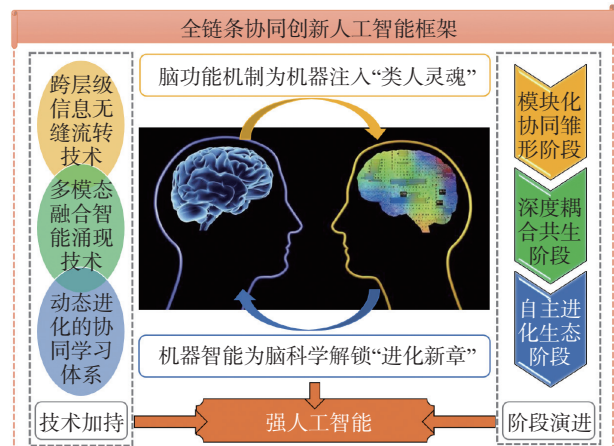


图 1 全链条协同创新人工智能框架

Fig. 1 Full-chain collaborative innovation artificial intelligence framework

人脑作为自然界最复杂的智能系统, 其感知环境、认知整合和决策优化等方面蕴含着巨大的智慧。而机器智能则具有强大的计算和数据处理能力, 二者的深度融合, 旨在使机器智能摆脱“局部模仿”的局限, 向通用智能迈进。同时, 借助机器智能的无限拓展性, 反哺人脑实现前所未有的进化跃升, 最终达成生物智能与机器智能的完美共生。

4.1.1 人脑→机器: 脑功能机制为机器注入“类人灵魂”

人类的感知、认知、决策能力, 本质上是亿万年进化形成的“生物算法”。未来, 通过解码这些算法的底层逻辑, 机器将从“工具”升级为“伙伴”, 具备类人的信息处理、理解与行动能力。

具体来说, 在感知层面, 机器智能不再依赖固定的模型参数, 而是通过“人脑顶叶皮层启发的显著性检测”, 动态调整感知焦点; 且其多模感知并非独立运作的, 通过构建跨模态感知共鸣网络, 让不同的感官信息进行语义对话。在认知层面, 可复刻人脑高级联合皮层与记忆皮层的神经连接模式, 前者像“临时工作台”可处理当前任务, 而后者作为“知识图书馆”支撑长期理解, 二者共同助力构建会生长的智能系统; 在决策层面, 将深度融合人脑逻辑推理与因果推理的协同决策, 不仅是计算最优解, 更是“价值导向”的选择, 使得智能系统能够自主决策, 具备类人的价值判断能力。

4.1.2 机器→人脑: 机器智能为脑科学解锁“进化新章”

机器不仅是“智能的载体”, 更是“脑科学的实

验室”。未来,机器将通过强大的计算能力与数据模拟,帮助人类突破“脑机制研究”的瓶颈,甚至主动优化人脑的功能。

具体而言,在感知层面,机器可构建“虚拟人脑模型”,模拟不同注意力策略对决策的影响,为个体定制“注意力画像”。此外,借助脑机接口等智能辅助设备,人脑的感官将得到无限扩展,如可“看到”电磁波传递,以全新的视角感知世界。在认知层面,机器智能可成为人类大脑的“超级外挂”,协助构建复杂的虚拟模型进行推演;并辅助人类突破思维定式,激发潜在的创造力。在决策层面,机器智能可模拟决策可能引发的所有后果,以全息投影等直观方式呈现给人类;同时分析人类的情感波动与认知偏差,提供客观且富有洞察力的建议。人类可结合自身的经验与价值观,做出更加明智、更具前瞻性的决策。

此外,全链条协同创新的终极实现,依赖两大硬件支柱,脑机接口构建“神经信号↔机器指令”的通道,类脑芯片提供“全链条协同”的硬件加速。二者共同构成生物智能与机器智能的物理融合层。其终极目标不是让机器取代人类,而是让人类与机器成为共生的智能体。通过人脑机制的赋能与机器能力的反哺,将突破生物智能与机器智能的边界,创造出更高效、更温暖、更具生命力的智能形态,引领人类文明走向更加辉煌的未来。

4.2 全链条协同创新阶段

智能系统的发展始终遵循“从简单到复杂、从孤立到协同、从被动执行到主动进化”的底层逻辑。脑功能机制启发的全链条协同创新也不例外,其演进过程可划分为模块化协同雏形阶段、深度耦合共生阶段、自主进化生态阶段。这3个阶段为“脑启智跃”范式的演进形态,既体现了技术能力的阶梯式突破,也反映了智能系统与物理世界、社会系统交互深度的质变。

4.2.1 模块化协同雏形阶段

模块化协同雏形阶段是全链条协同创新的起点,其核心目标是通过标准化技术手段,将分散的感知、认知、决策模块从“独立运行”转向“初步协同”。这一阶段,各模块多源自独立的研究领域,基于现有技术栈开发。因此,模块化协同的首要任务是打破技术壁垒,通过标准化接口与协议,实现多模块的信息互通与规则联动,形成简单的协同工作流程。此时,各模块相对独立运行,仅在数据与指令层面交互。感知模块负责采集并初步编码原始数据,经标准化处理后传输给认知模块;认知模块对数据进行分析处理,提取

特征并构建知识表征;决策模块依据预设规则与认知模块输出,生成决策指令。此外,模块化协同雏形阶段的硬件支撑以传统 GPU 与外置脑机接口为核心,二者处于互不耦合的独立运行状态。传统 GPU 承担感知数据处理、简单认知计算等基础算力任务,外置脑机接口仅负责单向采集脑电信号并传输至计算模块,无反向反馈与协同优化机制,仅为全链条协同搭建基础硬件雏形。这种协同模式虽能完成从数据采集到决策的基本闭环,但缺乏深度信息融合与协同优化能力,系统灵活性与适应性有限,面对复杂多变的实际情况时,响应迟缓、决策僵化,不过为后续更高级别的协同创新奠定了基础。

4.2.2 深度耦合共生阶段

当模块化协同完成技术底座的整合,智能系统开始面临单一任务的动态性、多任务的交叉性以及环境的不确定性等更复杂的挑战。此时,模块化协同的“规则驱动”模式已难以满足需求,系统需要从“被动执行”转向“主动适应”,深度耦合共生阶段应运而生。在深度耦合模式下,感知、认知、决策模块不再是独立的“黑箱”,而是通过数据流、控制流的双向互动形成“感知→认知→决策→反馈→再感知”的闭环;系统可根据任务需求动态分配计算资源、调整模块参数,通过模块间的协同交互,系统整体能力超越单一模块的简单叠加,形成“1+1>2”的智能涌现效应。此外,深度耦合系统的架构不再严格区分“感知层-认知层-决策层”,而是采用混合架构,感知模块嵌入轻量级认知模型,提前过滤无效数据;决策模块集成实时感知反馈等,实现从分层解耦到逻辑融合的模式创新。并且,深度耦合共生阶段采用 NorthPole 等异构类脑芯片与脑机接口的半集成架构。异构芯片整合脉冲神经网络与传统计算单元,为感知-认知-决策的动态协同提供高效算力,能效比极大提升。脑机接口系统实现神经信号高精度双向传输,既向芯片输入认知状态数据,又接收优化后的神经反馈信号。二者通过定制化协议实现功能联动,打破硬件壁垒,支撑闭环协同。

4.2.3 自主进化生态阶段

当深度耦合系统具备高效的任务执行与环境适应能力,智能的发展开始触及“意识”与“目标”的本质,人工智能迈向自主进化生态阶段。在此终极形态下,基于全流程协同创新框架的智能系统突破传统局限,系统不再依赖人类预设的规则或目标,形成具有自我意识与目标驱动能力的自

主进化生态,实现从被动执行任务到主动规划、自主学习、自我进化的质变。该自主进化系统能够感知自身状态,具备“目标发现”能力,并通过主动规划实现多目标优化平衡短期收益与长期价值。系统不再是孤立的个体,也不再依赖单一模块的性能提升,而是通过复杂的自组织,与外部环境、其他智能体形成共生关系,通过交互涌现出超越个体能力的群体智能。自主进化生态阶段的硬件核心是混合生物-硅芯片与闭环脑机接口的完全融合架构。混合芯片将生物神经元与硅基电路无缝集成,具备自主学习与算力动态重构能力,完美适配系统进化需求。闭环实现神经信号与机器指令的实时双向交互,不仅精准解码人脑意图,更能动态反馈优化脑功能。二者深度嵌入全链条协同体系,硬件与算法、生物智能与机器智能无界融合,支撑系统自主进化与目标自驱。这种自主进化系统能够广泛应用于科学探索、社会治理等领域,从主动发现科学新问题、探索研究新路径,到实现社会治理从被动应对到主动预防、从局部优化到全局协同的转变。

4.3 全链条协同创新关键技术

近来,以 Darwin3 为代表的大规模神经形态芯片搭载全新指令集架构,支持灵活的神经元模型编程,以及本地学习规则的自定义设计受到了广泛关注^[80]。该芯片可作为全链条协同创新的先验,指导其打破传统智能系统中“感知-认知-决策”层级的割裂,实现信息、知识与能力的全链路贯通与动态进化。跨层级信息无缝流转技术、多模态融合智能涌现技术与动态进化的协同学习体系,作为支撑这一创新的关键技术,从信息传输、数据融合与系统进化3个维度,为提升智能系统在复杂环境中的高适应性和强泛化能力提供了技术基石。

4.3.1 跨层级信息无缝流转技术

信息的高效流转是实现全链条协同创新中各层级模块协同工作的基础。传统智能系统常面临“层级壁垒”困境,因此实现全链条协同创新的首要需求是开发跨层级信息无缝流转技术,其核心目标是构建“感知-认知-决策”层间的“信息高速公路”也即新型信息编码与传输协议,实现从原始数据采集到决策执行的全链条无障碍贯通,同时形成“数据-信息-知识-决策-优化数据”的反馈闭环,提升智能系统的协同效率和适应能力。新型信息编码与传输协议的设计可借鉴人脑神经递质传递机制。人脑神经元通过释放神经递质,实现电信号到化学信号再到电信号的转换,精准调

控神经活动。类似地,可设计“智能信号分子”将感知数据在各层级之间精准转化为可处理的认知符号与决策指令。具体来说,该“智能信号分子”不仅能够对感知到的数据进行高效编码,确保信息在传输过程中不失真;而且能够在不同层级之间进行无缝转换,避免语义偏差;还能够根据任务需求动态调整信息的优先级和处理方式。该“智能信号分子”也可以助力各层级间反馈回路的构建,当决策执行层下达的指令在实际执行中未达到预期效果时,相关反馈信息将以“智能信号分子”形式,快速传递至认知映射层。认知映射层接收到反馈后,会重新评估当前的知识体系与推理模型,调整认知策略;同时,该反馈还会进一步传递至感知适配层,促使其优化数据采集参数与方式。此外,为保障信息在跨层级流转中的安全性与可靠性,还可给“智能信号分子”添加数字签名和加密标识,防止层间信息的篡改,并记录信息流转全过程,便于排查和溯源。基于脑功能机制启发的跨层级信息无缝流转技术,可为人工智能全链条协同创新构建坚实的信息纽带,支撑系统实现高效、稳定的运行。

4.3.2 多模态融合智能涌现技术

在复杂的现实环境中,单一模态的感知和处理往往难以满足智能系统的需求。传统智能系统多依赖单一模态或简单叠加多模态数据,导致信息利用率低、复杂场景适应能力差。人脑每天接收约1100万比特的信息,其中80%来自视觉,10%来自听觉,其余来自触觉、嗅觉等多模态输入。人脑可基于初级皮层协调视觉、听觉、触觉等相关脑区的活动,实现信息的深度整合与理解。类似地,可模拟人脑的多模态信息处理机制,通过注意力机制动态聚焦关键信息,构建统一语义空间支持跨模态推理。具体来说,在感知阶段,可借鉴人脑在处理多模外界刺激时的视觉编码、语义加工、感官整合顺序功能,聚焦关键模态信息,动态分配计算资源;在认知阶段,可模拟联合皮层对多模态信息的抽象整合功能将来自不同类型初级皮层的信息转化为统一的语义表征,挖掘模态间潜在关联,形成对任务的综合认知;在决策阶段,可基于认知阶段挖掘的各模态潜在关联,结合情感、记忆等因素,评估不同模态信息对决策结果的影响权重,并根据环境动态调整,实现高效可靠的决策。基于脑功能机制启发的多模态融合智能涌现技术,可为人工智能全链条协同创新搭建灵活的交互桥梁,支撑系统实现多任务协同处理。

4.3.3 动态进化的协同学习体系

智能系统需要具备强大的适应性和泛化能力,以应对各种动态场景。传统智能系统的学习是“静态”的,模型基于历史数据训练完成后,难以适应环境变化,导致模型滞后问题。因此,动态进化的协同学习体系的核心目标是构建全流程协同进化的学习框架,使智能系统能够不断自我优化。人脑的学习与适应能力依赖海马体、前额叶皮层和杏仁核等脑区的协同作用,海马体负责记忆编码与存储,前额叶皮层主导决策与策略调整,杏仁核参与情感评估与动机调控,三者协同实现人脑的持续学习与动态适应,这为构建动态进化的协同学习人工智能体系提供了灵感源泉。感知层是智能系统与外部环境交互的第一道防线,通过在线学习适应环境变化;认知层基于新知识动态更新知识库与推理规则,模拟人脑海马体将新知识整合到已有知识体系,并由前额叶皮层调用知识进行推理;决策层根据执行反馈调整策略,模拟杏仁核评估决策结果的情感价值与潜在风险,影响决策倾向。此外,可进一步模拟人脑内部神经元间的竞争与协作机制,对不同层级模块设置对抗机制,通过感知层与认知层、认知层与决策层之间的反馈与评估,实现相互促进、协同进化,推动智能系统整体性能提升。

5 结束语

人工智能全链条协同创新的发展历程,不仅是技术迭代的过程,更是研究范式与方法论革新的探索之旅。从模块化协同雏形到自主进化生态的演进,深刻揭示了人工智能迈向通用智能的必由之路,其背后的范式创新与方法论价值,为该领域的未来发展指明了方向。

基于信息科学范式的脑功能机制启发研究框架的建立,成为打破传统人工智能研究长期受物质科学“分而治之”范式的主导,结构主义、功能主义和行为主义等主流范式各自为战,导致人工智能研究呈现碎片化这一困境的关键。该框架借鉴大脑功能机制,以“感知-认知-决策”全链条协同为核心,革新人工智能研究。该范式从信息产生、传输、处理和利用的全过程出发,将智能视为信息转换与处理的高级形态,感知不再局限于数据采集,而是信息的初步编码与特征提取;认知过程是对感知信息的深度解析、关联与知识构建;决策则是基于认知结果进行价值判断与行动规划。各环节紧密相连、信息双向流动,实现了从单一功能模块研究向整体智能体系构建的转变,

推动人工智能实现“整体智能跃迁”,为攻克通用人工智能难题提供了全新的理论框架与研究视角。

此外,“感知-认知-决策”闭环方法论的形成,体现了“显要素”与“隐要素”的辩证统一,为通用人工智能研究开辟了“源头创新”的思维路径。在这一闭环中,算法作为“显要素”,是人工智能实现功能的直接载体,不断革新推动着人工智能性能的提升。然而,单纯的算法优化无法触及智能的本质,只有结合“隐要素”人脑功能机制,才能赋予人工智能真正的智能属性。脑认知机制作为“隐要素”,为人工智能算法设计提供了灵感源泉。感知适配层借鉴初级视听感皮层的动态可塑机制,使人工智能感知更具适应性;认知映射层参考高级联合皮层的知识整合体系,提升人工智能的认知能力;决策执行层模仿人脑的推理与决策过程,优化人工智能的决策水平。“显要素”与“隐要素”相互依存、相互促进,算法在脑功能机制的启发下不断优化,而算法实践又反过来加深对脑功能机制的理解与应用。这种辩证统一的关系,打破了以往仅从技术层面改进算法的局限,促使研究者从源头探索智能的本质。

但是,本研究提出的全链条协同创新框架,在实际落地过程中仍面临多重技术瓶颈亟待突破。其一,脑信号解码精度与实时性的双重制约,难以满足动态场景下人机协同的实时交互需求,尤其在高频决策任务中易引发响应滞后问题。其二,人脑多模态信息天然具备时序一致性与语义关联性,而类脑系统融合多模态数据存在噪声干扰与语义鸿沟。其三,人脑的感知-认知-决策机制具有 lifelong 自适应特性,而类脑系统的协同学习体系难以快速适配任务切换与环境突变,在跨场景知识迁移中易出现模型性能衰减,且缺乏有效的动态资源调度机制平衡算力消耗与协同效率。其四,双主体协同模式下,人脑与类脑系统的双向数据交互可能涉及隐私神经信号泄露,且自主进化阶段的智能系统可能出现决策偏差,亟需建立适配全链条协同的伦理规范与风险管控机制。

未来,随着脑科学与信息科学的深度融合,“脑启智跃”范式将在关键应用领域实现具象化突破。在医疗领域,基于全链条协同框架的智能诊疗系统将实现闭环赋能,通过高精度脑信号解码捕捉早期神经疾病的细微异常,结合多模态医学数据(影像、基因、临床症状)的深度融合,辅助医生完成阿尔茨海默病、癫痫等疾病的早诊早治;同时,脑机接口与类脑决策系统的协同将助力瘫

疾患者实现运动功能重建,通过神经信号与机械假肢的实时协同,实现精细化动作控制。在教育领域,个性化智能教学体系将依托全链条协同机制精准适配个体认知规律,通过感知层捕捉学习者的注意力状态、情绪波动等实时数据,认知层构建个性化知识图谱与学习路径,决策层动态调整教学内容与交互方式,实现“因材施教”的智能化落地,尤其在特殊教育领域,可通过多模态协同感知为自闭症、注意力缺陷患者提供定制化干预方案。在工业领域,人机协同智能系统将突破复杂场景作业瓶颈,操作人员的脑电信号与机器视觉、传感器数据实时融合,实现高危环境下的远程精准操控,类脑决策系统则基于全链条知识整合优化生产流程,推动工业制造向“柔性化、智能化”转型。此外,随着脑机制解码技术的突破与多模态融合算法的迭代,“脑启智跃”范式将进一步拓展至航天、安防等领域,构建人机共生的智能应用新生态。

总而言之,随着对脑功能机制的深入探索与信息科学技术的持续发展,基于此人工智能全链条协同的范式与方法论研究将以“脑”功能机制为“启”迪之源,开启通往“智跃”的大门,推动人类向通用人工智能的目标稳步迈进,为通用人工智能提供可解释、可进化的实现路径,为社会发展与人类文明进步带来深远影响。

参考文献:

- [1] BENGIO Y, COURVILLE A, VINCENT P. Representation learning: a review and new perspectives[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(8): 1798–1828.
- [2] DINO P, LUCA P, EMANUELE F, et al. Human-AI coevolution [J]. *Artificial intelligence*, 2024: 104244.
- [3] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. *Nature*, 2015, 521(7553): 436–444.
- [4] 张钹, 朱军, 苏航. 迈向第三代人工智能. 中国科学: 信息科学[J]. 2020, 50(9): 1281–1302.
ZHANG Ba, ZHU Jun, SU Hang. Toward the third generation of artificial intelligence (in Chinese)[J]. *Scientia Sinica Informations*, 2020, 50: 1281–1302.
- [5] HINTON G, DENG Li, YU Dong, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups[J]. *IEEE signal processing magazine*, 2012, 29(6): 82–97.
- [6] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. *Nature*, 2016, 529(7587): 484–489.
- [7] BAVELIER D, GREEN C S, POUGET A, et al. Brain plasticity through the life span: learning to learn and action video games[J]. *Annual review of neuroscience*, 2012, 35: 391–416.
- [8] BERNARD J B. A cognitive theory of consciousness[M]. Cambridge: Cambridge University Press, 1993.
- [9] ANDY C. Surfing uncertainty: prediction, action, and the embodied mind[M]. Oxford: Oxford University Press, 2015.
- [10] STEIN B E, STANFORD T R. Multisensory integration: current issues from the perspective of the single neuron [J]. *Nature reviews neuroscience*, 2008, 9(4): 255–266.
- [11] ALEKSANDR R L. Higher cortical functions in man[M]. Cham: Springer Science & Business Media, 2012.
- [12] MILLER E K, COHEN J D. An integrative theory of prefrontal cortex function[J]. *Annual review of neuroscience*, 2001, 24: 167–202.
- [13] BENUCCI A, SALEEM A B, CARANDINI M. Adaptation maintains population homeostasis in primary visual cortex[J]. *Nature neuroscience*, 2013, 16(6): 724–729.
- [14] ESPINOSA J S, STRYKER M P. Development and plasticity of the primary visual cortex[J]. *Neuron*, 2012, 75(2): 230–249.
- [15] DOWNER J D, RAPONE B, VERHEIN J, et al. Feature-selective attention adaptively shifts noise correlations in primary auditory cortex[J]. *The journal of neuroscience*, 2017, 37(21): 5378–5392.
- [16] GORI M, CAMPUS C, SIGNORINI S, et al. Multisensory spatial perception in visually impaired infants[J]. *Current biology*, 2021, 31(22): 5093–5101.
- [17] MAO Yurong, CHEN Peiming, LI Le, et al. Virtual reality training improves balance function[J]. *Neural regeneration research*, 2014, 9(17): 1628–1634.
- [18] PANTEV C, ENGELIEN A, CANDIA V, et al. Representational cortex in musicians: plastic alterations in response to musical practice[J]. *Annals of the New York academy of sciences*, 2001, 930: 300–314.
- [19] MARTIN A. The representation of object concepts in the brain[J]. *Annual review of psychology*, 2007, 58: 25–45.
- [20] PASCUAL-LEONE A, AMEDI A, FREGNI F, et al. The plastic human brain cortex[J]. *Annual review of neuroscience*, 2005, 28: 377–401.
- [21] HICKOK G, POEPEL D. The cortical organization of speech processing[J]. *Nature reviews neuroscience*, 2007, 8(5): 393–402.
- [22] DUDAI Y, KARNI A, BORN J. The consolidation and transformation of memory[J]. *Neuron*, 2015, 88(1): 20–32.

- [23] STUSS D T, LEVINE B. Adult clinical neuropsychology: lessons from studies of the frontal lobes[J]. *Annual review of psychology*, 2002, 53: 401–433.
- [24] ADDIS D R, PAN Ling, VU M A, et al. Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering[J]. *Neuropsychologia*, 2009, 47(11): 2222–2238.
- [25] MARREIROS A, STEPHAN K, FRISTON K. Dynamic causal modeling[J]. *Scholarpedia*, 2010, 5(7): 9568.
- [26] KOEHLIN E, SUMMERFIELD C. An information theoretical approach to prefrontal executive function[J]. *Trends in cognitive sciences*, 2007, 11(6): 229–235.
- [27] WERTHEIM J, RAGNI M. The neural correlates of relational reasoning: a meta-analysis of 47 functional magnetic resonance studies[J]. *Journal of cognitive neuroscience*, 2018, 30(11): 1734–1748.
- [28] BINDER J R, DESAI R H, GRAVES W W, et al. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies[J]. *Cerebral cortex*, 2009, 19(12): 2767–2796.
- [29] HONEY C J, SPORNS O, CAMMOUN L, et al. Predicting human resting-state functional connectivity from structural connectivity[C]//Proceedings of the National Academy of Sciences of the United States of America. Washington: PNAS, 2009, 106(6): 2035–2040.
- [30] CORBETTA M, SHULMAN G L. Control of goal-directed and stimulus-driven attention in the brain[J]. *Nature reviews neuroscience*, 2002, 3(3): 201–215.
- [31] PAPEZ J W. A proposed mechanism of emotion[J]. *Archives of neurology and psychiatry*, 1937, 38(4): 725.
- [32] SEGERSTROM S C, MILLER G E. Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry[J]. *Psychological bulletin*, 2004, 130(4): 601–630.
- [33] BOTVINICK M M, BRAVER T S, BARCH D M, et al. Conflict monitoring and cognitive control[J]. *Psychological review*, 2001, 108(3): 624–652.
- [34] LUNDSTROM M. Moore’s law forever[J]. *Science*, 2003, 299(5604): 210–211.
- [35] DANIEL C D. Cognitive wheels: the frame problem of AI[J]. *The philosophy of artificial intelligence*, 1990, 147: 1–16.
- [36] STUART J R, PETER N. Artificial intelligence: a modern approach[M]. Loden: Pearson, 2016.
- [37] ROSENBLATT F. The perceptron: a probabilistic model for information storage and organization in the brain[J]. *Psychological review*, 1958, 65(6): 386–408.
- [38] RICHARD S S, ANDREW G B. Reinforcement learning: an introduction [M]. Cambridge: MIT press, 1998.
- [39] RIESENHUBER M, POGGIO T. Hierarchical models of object recognition in cortex[J]. *Nature neuroscience*, 1999, 2(11): 1019–1025.
- [40] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the acm*, 2017, 60(6): 84–90.
- [41] CHIRSTIAN S, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015.
- [42] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 2818–2826.
- [43] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning[C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. San Francisco: AAAI, 2017.
- [44] PASCANU R, MIKOLOV T, BENGIO Y. On the difficulty of training recurrent neural networks[EB/OL]. (2012–11–21)[2025–09–01]. <https://arxiv.org/abs/1211.5063>.
- [45] GRAVES A. Generating sequences with recurrent neural networks[EB/OL]. (2013–08–04)[2025–09–01]. <https://arxiv.org/abs/1308.0850>.
- [46] ASHISH V, NOAM S, NIKI P, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems. Long Beach: NIPS, 2017: 6000–6010.
- [47] CRAMER B, BILLAUDELLE S, KANYA S, et al. Surrogate gradients for analog neuromorphic computing[C]//Proceedings of the National Academy of Sciences of the United States of America. Washington: National Academy of Sciences, 2022, 119(4): e2109194119.
- [48] WANG Qixin, FAN Chaoqiong, JIA Tianyuan, et al. ND-MRM: neuronal diversity inspired multisensory recognition model[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Vancouver: AAAI, 2024, 38(14): 15589–15597.
- [49] WANG Qixin, LI Ziyu, LI Xiuxing, et al. BrainyHGNN: brain-inspired memory retrieval and cross-modal interaction for emotion recognition in conversations[J]. *IEEE transactions on circuits and systems for video technology*, 2025, 35(10): 10264–10277.
- [50] LI Kai, XIE Fenghua, CHEN Hang, et al. An audio-visu-

- al speech separation model inspired by cortico-thalamo-cortical circuits[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2024, 46(10): 6637–6651.
- [51] HE Xiang, ZHAO Dongcheng, LI Yang, et al. Incorporating brain-inspired mechanisms for multimodal learning in artificial intelligence[EB/OL]. (2025-05-15)[2025-09-01]. <https://arxiv.org/abs/2505.10176>.
- [52] SHEN Jiangrong, XIE Yulin, XU Qi, et al. Spiking neural networks with temporal attention-guided adaptive fusion for imbalanced multi-modal learning[C]//Proceedings of the 33rd ACM International Conference on Multimedia. Dublin: ACM, 2025: 11042–11051.
- [53] KEITH J H. Parallel distributed processing: Explorations in the microstructure of cognition[J]. *Science*, 1987, 236: 992–997.
- [54] BOLLACKER K, EVANS C, PARITOSH P, et al. Freebase: a collaboratively created graph database for structuring human knowledge[C]//Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data. Vancouver: ACM, 2008: 1247–1250.
- [55] KOUNIOS J, HOLCOMB P J. Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory[J]. *Journal of experimental psychology Learning, memory, and cognition*, 1994, 20(4): 804–823.
- [56] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks[EB/OL]. (2017-03-09)[2025-09-01]. <https://arxiv.org/abs/1703.03400>.
- [57] IAN J G, JEAN P-A, MEHDI M, et al. Generative adversarial nets[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: NIPS, 2014: 2672–2680.
- [58] EICHENBAUM H. Memory: organization and control[J]. *Annual review of psychology*, 2017, 68: 19–45.
- [59] EMANUELE M, GIANPAOLO B, ELISA F, et al. Neuro-symbolic continual learning: Knowledge, reasoning shortcuts and concept rehearsal[EB/OL]. (2023-02-02)[2025-09-01]. <https://arxiv.org/abs/2302.01242>.
- [60] HUANG Jiayuan, SMOLA A J, GRETTON A, et al. Correcting sample selection bias by unlabeled data[M]//Advances in Neural Information Processing Systems 19. Cambridge: The MIT Press, 2007: 601–608.
- [61] DEVLIN J, CHANG Mingwei, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis: Minnesota Association for Computational Linguistics, 2019: 4171–4186.
- [62] ZHOU Dawei, CAI Ziwen, YE Hanjia, et al. Revisiting class-incremental learning with pre-trained models: generalizability and adaptivity are all you need[J]. *International journal of computer vision*, 2025, 133(3): 1012–1032.
- [63] HAN Yuyang, LI Xiuxing, WANG Qixin, et al. A brain-inspired distributed long-term memory guided online continual learning method[C]//Human Brain and Artificial Intelligence. Singapore: Springer, 2025: 331–343.
- [64] HAN Yuyang, LI Xiuxing, JIA Tianyuan, et al. A novel sleep mechanism inspired continual learning algorithm[J]. *Guidance, navigation and control*, 2024, 4(3): 2441003.
- [65] LENAT D B. The role of heuristics in learning by discovery: three case studies[M]//Machine Learning. Amsterdam: Elsevier, 1983: 243–306.
- [66] TANG Xiaojuan, ZHU Songchun, LIANG Yitao, et al. Rule: knowledge graph reasoning with rule embedding [EB/OL]. (2022-10-24)[2025-09-01]. <https://arxiv.org/abs/2210.14905>.
- [67] JIA Tianyuan, LI Ziyu, LI Qing, et al. BrainyMP: enhancing motion planning using graph neural network inspired by brain spatial relational memory[J]. *IEEE transactions on intelligent transportation systems*, 2025, 26(6): 8880–8893.
- [68] VANESSA D, IRIS P. Causality: models, reasoning, and inference[M]. Cambridge: Cambridge University Press, 2001.
- [69] KOMOROWSKI M, CELI L A, BADAWI O, et al. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care[J]. *Nature medicine*, 2018, 24(11): 1716–1720.
- [70] JUDEA P, DANA M. The book of why: the new science of cause and effect[J]. *Science*, 2018, 361(6405): 855.
- [71] WANG Chengyu, WANG Luhan, LU Zhaoming, et al. Autonomous driving via brain-inspired causality-aware contrastive learning with time-frequency prediction[J]. *IEEE Internet of Things journal*, 2025, 12(14): 26371–26386.
- [72] LYU Daoming, YANG Fangkai, LIU Bo, et al. Logic-based sequential decision-making[C]//The Thirty-Third AAAI Conference on Artificial Intelligence. Honolulu: AAAI, 2019, 33(1): 9995–9996.
- [73] BAREINBOIM E, PEARL J. Causal inference and the data-fusion problem[C]//Proceedings of the National Academy of Sciences of the United States of America. Washington: PNAS, 2016, 113(27): 7345–7352.
- [74] SHAKYA A K, PILLAI G, CHAKRABARTY S. Rein-

- forcement learning algorithms: a brief survey[J]. *Expert systems with applications*, 2023, 231: 120495.
- [75] FAN Shengda, MO Shasha, NIU Jianwei. Boosting document-level relation extraction by mining and injecting logical rules[C]//Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Abu Dhabi: Emirates Association for Computational Linguistics, 2022: 10311–10323.
- [76] JIA Tianyuan, FAN Chaoqiong, LI Qing, et al. A brain-inspired harmonized learning with concurrent arbitration for enhancing motion planning in fuzzy environments[J]. *IEEE transactions on fuzzy systems*, 2025, 33(2): 631–643.
- [77] JI Tianying, LIANG Yongyuan, ZENG Yan, et al. ACE: off-policy actor-critic with causality-aware entropy regularization[EB/OL]. (2024–02–22)[2025–09–01]. <https://arxiv.org/abs/2402.14528>.
- [78] CAO Hongye, FENG Fan, FANG Meng, et al. Towards empowerment gain through causal structure learning in model-based RL[EB/OL]. (2025–02–14)[2025–09–01]. <https://arxiv.org/abs/2502.10077>.
- [79] HE Xiangkun, WU Jingda, HUANG Zhiyu, et al. Fear-neuro-inspired reinforcement learning for safe autonomous driving[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2024, 46(1): 267–279.
- [80] MA De, JIN Xiaofei, SUN Shichun, et al. Darwin3: a large-scale neuromorphic chip with a novel ISA and on-chip learning[J]. *National science review*, 2023, 11(5): 102.

作者简介:



郭霞, 教授, 博士生导师, 博士, 北京理工大学计算机学院脑机接口与类脑智能研究中心主任, 主要研究方向为人工智能与脑科学, 获吴文俊人工智能自然科学一等奖、教育部自然科学二等奖、茅以升北京青年科技奖等。E-mail: wuxia@bit.edu.cn。



李子遇, 助理研究员, 博士, 主要研究方向为脑信号智能分析与类脑智能。E-mail: ziyuli@bit.edu.cn。



李晴, 副研究员, 主要研究方向为脑机接口与类脑智能。E-mail: liqing@bit.edu.cn。

[责任编辑: 刘冰洁]