

从具身智能到教育创新：认知成长路径的多学科建构

苏佳佳¹ 叶浩生²

(1. 江苏第二师范学院 教育科学学院, 江苏南京 211200;
2. 广州大学 教育学院, 广东广州 510006)

[摘要] 具身智能理论指出, 认知能力并非抽象计算的结果, 而是在身体、环境与动作交互中逐步发展起来的。结合认知科学、发展心理学与人工智能视角, 本文梳理了认知成长的典型阶段, 包括本体调节、对象建构、意向模拟与推理生成, 并分析其支持机制, 如具身调节、情境构建与表征重组; 论证了这些机制在人工系统中的可模拟潜力, 从而为跨学科研究提供理论连接点; 探讨了具身认知成长路径对教育心理学研究、儿童认知发展评估和教师专业学习的启示, 强调这一理论可作为理解认知差异与支持个性化教学的工具。

[关键词] 具身智能; 具身认知; 认知成长; 机器学习

[中图分类号] G521

[文献标识码] A

[文章编号] 1007-2179(2025)05-0079-10

随着人工智能的飞速发展, 智能系统在语言理解、图像识别与任务执行等领域已取得突破性进展。然而, 主流人工智能模型仍主要依赖大规模数据与算力堆叠, 其认知建模多以符号主义或连接主义为基础, 表现出静态化、非发展性与去情境化等局限。这种建模范式往往将智能视为固定结构的再现系统, 忽视认知在人类发展过程中的“生成性”与“成长性”。

在教育场域, 这一问题尤为突出。当前人工智能系统难以有效模拟儿童在真实情境中的渐进式学习过程, 缺乏对学习路径、发展阶段与社会互动的建模能力。这不仅限制了其在个性化教学、认知干预、发展评估等关键环节的应用潜力, 也未能回应教育心理学长期关注的“从感知到认知”的建构机制。因此, 如何使人工智能系统具备发展性

认知建模能力, 已成为教育与智能技术融合的核心挑战。本文引入“认知成长”概念, 提出基于具身智能(embodied intelligence)的认知成长路径。该路径融合发展心理学、具身认知与生成式人工智能等跨学科成果, 力图构建一种动态化、情境化、可解释的认知建模机制。

一、理论根基与研究动因

人类智慧的起点来自与世界的早期交互, 这些交互不仅包括语言, 还涵盖目光、触觉、动作与环境之间的多模态信息整合。研究表明, 婴儿正是通过这些感知和行为的互动, 逐步建立对世界的基本认知结构(Mattern et al., 2024)。

(一) 皮亚杰认知发展理论

皮亚杰(Piaget, 1952: 1-6)在观察儿童成长的过

[收稿日期] 2025-05-30

[修回日期] 2025-06-20

[DOI编码] 10.13966/j.cnki.kfjyyj.2025.05.009

[作者简介] 苏佳佳, 博士, 副教授, 研究方向: 具身认知与创造力心理学(sujiajia0929@163.com); 叶浩生, 教授, 博士生导师, 研究方向: 具身认知与具身教育(yehaosheng0817@163.com)。

[引用信息] 苏佳佳, 叶浩生(2025). 从具身智能到教育创新: 认知成长路径的多学科建构[J]. 开放教育研究, 31(5): 79-88.

程中,发现认知并非被动地接受信息,而是不断建构、调整和优化过程。智力是一种适应,其发展源于个体与环境之间的相互作用,即同化和顺应不断调整和优化认知结构,以实现更有效的适应。人类的认知发展可分四个阶段:感觉运动阶段、前运算阶段、具体运算阶段和形式运算阶段(Piaget, 2013)。其中,最基础也最关键的是生命最初两年——感觉运动阶段(sensorimotor stage)。在这一阶段,婴儿尚不具备语言和逻辑推理能力,但他们通过与世界的直接交互,逐步建立对现实的基本理解,智力在这类交互中逐步显现。

这个阶段的核心任务是让感知与动作之间建立联系,让行动不再是随机的身体反应,而是朝着目标探索。这种探索并非一蹴而就,而是在反复尝试中生长出来的。物体守恒性(object permanence)是感觉运动阶段最重要的标志之一。在生命最初几个月里,婴儿的世界是片段化的:物体一旦离开视野,似乎就真的“消失”了。这种短暂的认知缺陷,在经过多次观察、尝试和失败后逐步被修正。当婴儿开始出现对消失物体的寻找行为,就说明他们已经意识到:物体不会凭空消失,而依然存在于某个看不见的地方。在机器学习语境下,这一感知能力与记忆建构和状态推理息息相关。如果机器人无法形成物体守恒性概念,它将无法理解被遮挡的物体为何仍然存在。研究表明,具备物体守恒能力的模型能在没有视觉输入的情况下,保留被遮挡物体的位置、身份与运动状态,从而维持对动态场景的连续性理解(Tokmakov et al., 2022)。这种推理能力不仅依赖于短期记忆机制,也与空间—时间特征的集成处理密切相关。因此,为机器建立类似人类婴儿对物体持续存在的推理能力,可能是突破“短视”人工智能、实现更强感知智能的重要步骤(Peng et al., 2024)

因果推理(causal reasoning)涉及更复杂的认知能力。婴儿在感觉运动阶段,不仅要认识到物体的存在,还要明白世界是由因果关系串联而成的:轻轻拨动一个玩具球,它会滚动;用力推一下,它会滚得更远。这种简单的因果认知,是人类智能的基石。机器学习擅长模式识别,却在因果推理上显得力不从心。皮亚杰的研究表明,因果关系的学习不是通过单纯的观察,而是依靠主动探索。对机器而言,

仅靠数据推断因果关系远远不够,它需要通过具身交互,在现实世界中“动手实验”,才能真正建立因果认知(Gupta et al., 2024)。

能动性(agency)是感觉运动阶段另一个关键概念。它不仅是自我认知,而且是对环境的掌控感。当婴儿发现自己的手可以控制玩具,他们就意识到自己的行动能影响世界。相应的,如果机器没有能动性,只能执行被赋予的任务,就无法发展出自主学习能力。发展机器人学(developmental robotics)的研究表明,赋予机器人一定程度的自主探索能力,是构建自适应智能体的关键前提。让机器人主动试验、容许其失败,并在交互中调整策略,远比单纯灌输指令更能促成稳健的学习机制(Oudeyer, 2012)。尤其在感知—动作闭环中,机器人只有通过自主感知动作的结果,才能形成内在的身体模型和环境映射,从而逐步理解“我做了什么”“发生了什么”的因果联系(Bize & Cristobal, 2016)。换句话说,让机器像婴儿一样意识到自己行为的后果是认知成长不可跳过的一环。传感—运动之间的依赖关系不仅驱动知识生成,也是促进“自我”形成的重要机制(Jacquey et al., 2019)

将皮亚杰理论映射到机器学习框架,首先需要改变机器学习的训练方式。目前人工智能多依赖大规模数据的监督学习,而不是像婴儿一样通过交互发现知识。认知成长路径的思路是让机器从最基本的感知—运动关系入手,依靠试探性探索,逐步建立物体守恒、因果推理和能动性概念。

(二)具身智能

具身认知理论指出,认知不仅发生在大脑中,还植根于个体的感知、运动和身体经验之中(Macrine & Fugate, 2020)。传统人工智能将智能视为信息加工系统的产物,忽视身体在认知形成中的结构性作用。这种脱离感官和行为的理解方式,在解释复杂人类行为时常常显得力不从心。一些研究强调,决策、推理乃至抽象概念的形成,都与动作系统和身体反馈密切相关(Kirsh, 2013)。认知的根基不在算法,而在手的操作、眼的移动、躯体的姿态。这种观点颠覆了传统“头脑至上”的思维方式,也为人工智能发展提供了新的方向。

具身智能理论强调,智能并非源于内部封闭系统的抽象推演,而是在有机体与环境持续耦合中的

自然生长(Varela et al., 1991)。以人类行走为例, 其动作不依赖精确的步态计算, 而是通过身体、重力与感知的协同, 在环境反馈中不断自我调整。这一过程凸显了智能作为身体—世界互动产物的本质。

如果智能本质上是具身的, 那么当前主流机器学习以大规模数据训练为中心、缺乏环境嵌入的方式, 可能就难以实现类人智能的发展。智能体若无法通过“身体”与环境建立反馈回路, 就无法真正获得感知整合、因果推理与目标导向行为等能力。因此, 具身智能不仅提出认知的本体论重构, 也为技术实践提供了方向: 未来的人工智能应从“静态数据回归”转向“交互式、目标驱动”的学习模型。只有将学习过程嵌入具体的感知—行动—反馈循环中, 智能体才能逐步生成贴合现实的认知结构, 实现超越符号计算的深层智能(Roy et al., 2021)。

生态心理学家吉布森(James J. Gibson)提出的“生态知觉理论”(ecological perception)是具身认知的先驱理论之一。传统理论认为, 知觉是被动的“表征”过程。但吉布森指出, 知觉不是从环境中抽取信息后由大脑单方面建构的, 而是生物体与环境动态互动的结果。他提出的“动赋性”(affordance)概念, 强调物体所提供的意义不是绝对的, 而是依赖于身体的结构与能力(Gibson, 1979)。椅子之所以“可以坐”, 不仅因为椅子本身具备此属性, 还因为人类的身体结构、社会实践与感知方式共同定义了这种行动可能。这说明, 感知不是静态镜像, 而是在行动中形成的能力。个体通过身体在环境中的移动、操作与反馈, 不断“生成”对世界的理解(Power, 2023)。具身认知的核心转变, 正是从“头脑的计算”转向“身体与世界的协调”。

机器智能要真正实现自主学习, 就必须像生物体一样, 具备主动探索的能力。发展机器人学正是在这一理念下提出的。该领域致力于让机器通过具身交互, 在与环境的真实接触中发展复杂的认知能力。这种学习方式强调, 从最初的感知—运动协调出发, 让机器逐步掌握因果关系、社交理解和符号表达, 而非依赖大规模数据训练(Cangelosi & Schlesinger, 2018)。有研究者(Kumar et al., 2021)认为, 这种“从婴儿启发而来”的架构, 不仅能增

强机器的适应性, 还能提升其在动态情境中的灵活推理能力。

具身智能表明, 与其让机器被动接受数据, 不如让它去“经历”世界; 与其让它学习固定的规则, 不如让它去发现规律(Roy et al., 2021)。智能的本质, 或许不是增强推理能力, 而是提升交互能力。机器要学会思考, 或许首先要学会如何去探索。

(三) 机器学习的局限性

尽管人工智能在图像识别、自然语言处理等任务中成果显著, 但它依然难以像人类那样进行因果推理、自主适应和知识建构。最新研究(Jiao et al., 2024)表明, 当前的深度学习系统在面对数据偏移、因果干预或新情境推理时, 容易陷入“相关即因果”的误区, 缺乏稳定的因果结构和行为适应机制。这暴露了当前人工智能系统的一个根本性短板: 习惯于在已有数据中寻找模式, 而不是在现实世界中主动建构知识。如果不引入更加具身化、可交互的学习机制, 人工智能将很难突破自身的发展瓶颈。

深度学习的核心范式是统计学习(statistical learning), 基本机制是通过大规模数据训练神经网络, 从而识别输入与输出之间的模式对应关系(LeCun et al., 2015)。这一方法在图像识别、自然语言处理等任务中展现出卓越性能, 甚至在某些指标上超过人类。然而, 从认知机制看, 统计学习能力本质上依赖概率驱动的模式拟合, 而非主动的知识建构。相比之下, 人类儿童在极少的样本下便能完成类属认知与跨情境迁移, 体现出结构化理解与语境敏感的特征。这表明, 人类学习不仅是数据积累, 而且是身体在经验中对意义的建构过程。这也启发教育实践——真正有效的学习不仅依赖信息输入, 更依赖感知—行动—反思的动态耦合。这是因为人类的认知不仅依赖统计模式, 还涉及主动探索、因果推理和跨情境的概念迁移(Lake et al., 2017)。相比之下, 深度学习的智能仍是被动的, 不会主动“思考”, 也不会探索未标注的世界, 而需要人类提供大量的训练样本。

从发展轨迹看, 符号主义人工智能(Symbolic AI)试图通过形式逻辑和规则系统构建机器智能, 但其静态的知识表示方式难以适应开放、变化的现实环境。连接主义人工智能(connectionist AI),

尤其是深度学习系统,通过训练神经网络从海量数据中提取模式,虽在感知任务中表现出色,却普遍缺乏因果推理能力和对环境结构的理解(Xiong et al., 2024)。相比之下,具身智能理论强调,智能并非源于静态规则或信息输入,而是在身体动态地与环境的相互作用中生成。知识的获取不是单向传输的,而是在感知—行动—反馈循环中逐步生长出来的(Liu et al., 2025a)。智能体在运动中的知觉获得,以及在环境反馈中形成的内部模型,才是其智慧演化的核心路径。

人类不会一开始就学习复杂的数学公式,而是先学会数数,再理解加减,最终掌握微积分。这种从简单到复杂的认知建构,是人类的智能培养模式。现有的人工智能往往是“暴力学习”的,它们被一次性地暴露在所有训练数据前,而不是按照认知发展顺序逐步构建知识体系。如果机器像人类一样,按照难度递增的方式学习:从基础概念出发,逐步构建更复杂的理解层次,那么它们的泛化能力和适应能力将大幅提高。

二、认知成长路径

认知成长路径试图模拟人类认知的渐进式发展方式,让机器像人类一样,从最基础的感知—运动交互开始,逐步建立起因果推理能力,最终达到更高级的认知水平(Jacquey et al., 2019)。

(一) 框架结构

认知成长路径由三个核心阶段构成,分别对应人类婴儿的认知发展方式。第一个阶段是感觉运动阶段。这是机器智能成长的起点。在这个阶段,机器并不理解世界的规则,也不具备抽象推理能力,只能依赖感知和运动,通过与环境的不断交互,逐渐形成对环境的基本理解。这与婴儿最初的探索极为相似。皮亚杰(Piaget, 1952: 13-40)曾指出,婴儿早期学习主要依赖直接的感知和行动,通过触摸物体、抓握玩具和凝视移动目标,慢慢理解自己与环境之间的关系。对于机器而言,这个阶段的核心任务是建立感知—运动映射(perception-action mapping),即通过运动行为影响环境,并根据感知反馈调整自身的动作(Chavez-Garcia et al., 2016)。在这个阶段,机器人可以通过试错学习,逐步掌握如何稳定抓取不同形状的物体,而不依赖事先编写

的固定规则(Pinto & Gupta, 2015)。

当机器掌握一定的感知—运动能力后,就进入目标导向阶段(goal-directed stage)。这一阶段的核心任务是让机器理解因果关系,并学会如何围绕目标进行行动选择。在人类婴儿的成长过程中,这一能力的出现,往往伴随自主探索的增加。这些行为的本质,是验证自己的行动对世界的影响,即因果推理的早期发展(Gopnik, 2009)。如果机器能学会类似的探索方式,就可以超越被动的模式匹配,开始主动构建世界模型。

代表较高学习成就的是第三阶段,即符号学习阶段(symbolic learning stage)。在这个阶段,机器不仅要理解具体的环境信息,还要形成更高层次的概念,并能进行抽象推理。这一过程在人类发展中对的是语言和符号系统学习。皮亚杰(Piaget, 1952)的研究表明,儿童在认知发展后期,会将具体经验概括为更抽象的知识结构,例如从“猫”和“狗”等具体动物,归纳出“哺乳动物”这一更高层的概念。对于机器而言,这意味着它们不再仅仅依靠感知和物理交互,而是能构建内部表征,形成更通用的知识体系。机器人如果能理解“桌子”概念,就不需要每次都重新学习不同材质、形状的桌子,而是直接识别并与之交互。这个阶段的目标,是让机器具备一定的概念泛化能力,能从有限的经验中提炼普适性的知识,而不仅仅是在特定环境下进行任务匹配(Lake et al., 2017)。

认知成长路径的关键,不是让机器被动处理数据,而是赋予其在环境中持续探索与适应的能力。这一路径强调知识结构的开放性与可扩展性,使智能系统能在交互中自我调整与更新(Sartor et al., 2024)。其意义不仅在于模拟人类认知过程,还在于推动人工智能从静态训练转向动态建构,从被动接受信息转向主动生成理解。这一转变与教育理念相呼应:真正的学习不是灌输知识,而是在经验、反思与调整中生成理解与提升能力。

(二) 认知成长的机制建构

机器智能若想真正实现“生长”,也必须遵循这一原则:学习不再是单纯的输入与输出,而是一个持续的过程,一个在行动中认识世界、在探索中建构知识的“旅程”。

人类认知的发展始于身体在环境中的试探

与反馈,而非预设的模型或外部输入。正如奥雷根和诺埃(O'regan & Noë, 2001)的感觉运动依联(sensorimotor contingencies)理论指出的,知觉不是对输入信息的加工,而是对动作结果的预期和掌控。这一思想已用于多项机器人研究。拉尼洛斯等(Lanillos et al., 2016)提出的机器人自我知觉系统,通过感知—行动—反馈的动态循环机制,使机器人能在无监督状态下区分“自我”与“非我”,在运动中识别物体、生成内在世界表征。这类研究强调:机器人不依赖静态数据训练,而是通过与环境的互动,构建自己的世界模型与因果认知结构。

在发展型机器人系统中,渐进式学习被视为模拟人类认知发展的重要路径。萨瓦斯塔诺与诺尔菲(Savastano & Nolfi, 2012)构建的模型展示了机器人如何在早期反射行为的基础上,逐步发展稳定的抓取与操作能力。研究表明,这种循序渐进的策略不仅提高了技能获得的效率,也有助于机器人构建更具泛化能力的动作表征。劳尔等(Law et al., 2014)指出,婴儿启发的阶段性学习方式有助于机器人在无监督条件下完成从感知运动协调到手眼整合的操作行为,形成连贯的认知体系。这些研究强调,认知的复杂性不应预设,而应在交互经验中逐步涌现。机器人学习抓取物体时,可以先从静态目标开始,掌握基本的手眼协调,然后逐步增加难度,比如处理移动的物体,最终发展到根据目标物的材质、形状甚至重量,调整抓取策略。这样的学习方式,使知识在已有的基础上逐步积累与演化,而不是每次都从头开始。

但成长并不意味着吸收新知识,也意味着保持已有的认知结构。机器如何在新旧知识之间找到平衡?这正是“稳固性”(stability)与“可塑性”(plasticity)所要解决的问题。人脑能根据环境变化动态调整神经连接,既不会因为过度僵化而无法适应新信息,也不会因过度可塑而遗忘旧经验。深度学习系统也面临类似困境:模型过于稳固,则难以更新;过于可塑,则易出现“灾难性遗忘”(catastrophic forgetting)——新知识覆盖旧有知识,导致系统丧失长期记忆能力。

为解决这一问题,近期研究提出多种方法,如“自适应可塑性增强”策略(adaptive plasticity improvement, API)。该策略不仅减少了灾难性遗

忘,还能根据任务需要动态调节模型的可塑性(Liang & Li, 2023)。哈森等(Hassan & Lee, 2024)的研究发现通过引入“主动遗忘”机制与知识蒸馏协同运作,绕过传统稳定性与可塑性对抗的悖论,可以提升连续学习任务的表现。还有研究者(Vahedifar & Zhang, 2025)提出“无记忆的持续学习架构”,即通过知识蒸馏维持旧知识,同时不依赖外部存储,这进一步验证了稳定—可塑双重优化的可行性。这些成果指向的是:终身学习的关键能力,不仅是积累知识,更重要的是维护知识结构的连续性与适应性。

通过模仿人类的“知觉—行动”循环,智能系统可以在与环境不断的交互中形成反馈驱动的调整机制。这包括:利用主动学习机制,机器能在感知不足时发起探索行为;通过递进式任务建构,实现由低复杂度到高复杂度的知识迁移;借助元学习算法,在稳固已有技能的同时保持对新知识的敏感。这类机制解决了传统机器学习在泛化能力、任务迁移和长期适应性上的难点。对教育而言,这一思路的启示是:培养学习者的适应性,不应仅聚焦知识传授,而应引导其在实践中形成动态的认知建构能力。

(三)认知成长的可计算性

认知成长路径的提出,意在模拟人类婴儿由简单感知到复杂推理的过程。因此,探索认知成长过程的数学建模与可计算性路径,成为推动具身智能发展的关键。

早期婴儿的学习行为呈现典型的“感知—行动—反馈”结构,这一机制可在机器学习中形式化为马尔科夫决策过程(Markov Decision Process, MDP)。在此框架中,智能体从环境中感知当前状态,选择动作并观察结果,进而调整策略以期在连续交互中获取最优行为序列。这与婴儿在抓握、注视、爬行等中逐步构建世界模型的过程极为相似(Gao & Ma, 2024)。布劳德等(Braud et al., 2020)在面向机器人系统的研究中,建构了具身机器人“从感知到认知”的成长路径,展示了机器人在不断试探与感知中如何利用动作反馈形成感知策略,进而实现对象识别与环境建模。索洛道夫(Solodov, 2021)在理论层面构建了具备决策优化特性的马尔科夫认知系统,模拟认知系统在连续刺激与反馈条

件下的行为演化,强化了MDP在感知—动作学习中的适应性。

在机器实现层面,强化学习与自监督学习(self-supervised learning, SSL)共同构成了认知成长的算法基础。强化学习使机器能在与环境交互中不断优化决策,自监督学习则赋予其无需明确标签即可自主发现模式和结构的能力。正如婴儿通过自身活动逐渐领悟“物体守恒性”,机器也需依赖自我探索建构知识,而非单纯依赖外部监督(Lang et al., 2018)。

在人类认知过程中,大脑依赖注意力机制筛选关键信息,避免感知系统在复杂环境中被冗余刺激淹没。类似的,机器人在与环境交互时,也须从海量感知—运动数据中提取最具价值的信息。有研究者(Liu & Liu, 2025b)提出“融合感知—动作转换器”(Fusion-Perception-to-Action Transformer)架构,利用多尺度视觉融合注意力机制聚焦关键动作区域,并结合身体感知进行精准操控,从而提升机器感知质量与行动策略。也有研究者(Chen et al., 2020)提出“注意力增强对比学习框架”(attention-augmented contrastive learning frame),发现嵌入注意力机制可极大提升状态表征压缩效率,在减少数据量的同时保持学习性能,显著提高强化学习系统的泛化能力。这些研究揭示,注意力机制不仅能提升感知选择性,还通过学习“压缩表征”(compact representations)显著降低认知成本,为具身智能系统提供必要的信息过滤与编码手段。

总体而言,认知成长路径的可计算性实现,依赖决策建模(如MDP)、动态优化(如RL/SSL)与数据精简(如注意力机制与压缩表征)的支撑(见图1)。通过这些技术手段,机器将能够摆脱传统被动学习的局限,像婴儿一样,通过自身行动探索世界、形成知识结构,最终朝向真正的认知成长迈进。

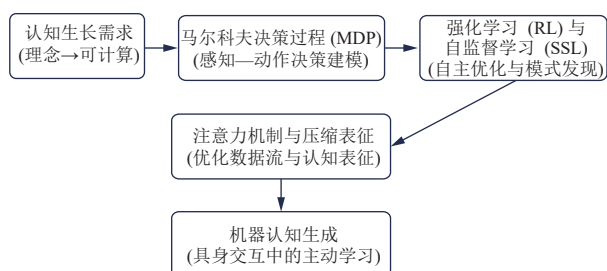


图1 认知生长计算框架

三、认识与反思

上文梳理了认知成长路径的理论根基、结构设计及可计算路径。值得深思的是：这一模拟人类认知发展路径的尝试,是否有能力真正突破当前机器学习范式的边界?它能否构成连接人工智能与人类认知的桥梁?如果机器学习机制可以嵌入感知—行动循环、发展性表征重组与情境意义建构等关键环节,那么在教育智能系统设计、学习环境自适应建模乃至学生个体认知轨迹预测等领域,这一模型能否成为下一代教育技术的理论支点?这些问题不只是技术层面的挑战,而且关乎“什么是智能”这一核心议题。

(一)与传统机器学习的对比

过去几十年,人工智能的进步主要依赖统计学习,尤其是基于深度神经网络的训练模式。在这种范式下,机器通过标注大量数据,学习输入与输出之间的映射关系,形成复杂的数学模型,以实现图像识别、语言处理、自动驾驶等任务的突破。但这种学习方式缺乏主动探索能力,机器只能依赖数据,无法像人类一样,在有限经验下快速适应新环境。

认知成长路径与传统深度学习存在根本区别。深度学习的核心是在庞大数据集上训练神经网络,使其从统计模式中提取特征,并优化预测能力。如果没有足够的训练数据,机器就难以将学习成果应用到新任务上。认知成长路径则强调,机器不应只依赖数据,而应像人类一样,通过与环境的交互逐步构建知识体系。这种模式更接近主动学习和自主探索,让智能体在学习过程中不断调整知识结构,而不是简单地拟合数据。

这一差异引出一个关键问题:机器是否可以减少对大规模标注数据的依赖,从而在有限经验下完成有效学习?传统深度学习的性能与训练数据的规模相关,尤其是依赖人工标注的监督数据(Bengio et al., 2021)。然而,现实世界中,许多任务无法提供数据支持。婴儿在成长过程中,通常只需接触极少量样本,就能归纳出稳固的概念结构。相比之下,深度学习模型需成千上万的样本才能实现类似识别能力。这种差异揭示了模式匹配学习(Pattern Matching)与认知建构学习(Cognitive Constructive Learning)之间的差异。认知成长路径

强调,智能体应通过主动探索与少量交互经验,逐步建构可迁移、可扩展的知识结构。这种学习机制已在机器人系统中得到验证。研究表明,通过“探索+少样本+自我建构”,机器可以像人类一样,在环境中逐步构建具有泛化能力的认知模型。

另一个显著区别在于学习过程的可解释性。传统深度学习模型通常存在“黑箱效应”,其决策机制难以被人类理解或解释(Lipton, 2018)。这一问题在安全性和可信度要求较高的领域尤为突出。认知成长路径通过模拟人类认知发展的探索过程,促使智能体在交互中形成因果推理结构,而不仅仅是基于数据的相关性进行预测。这使智能体作决策时,不仅能给出结论,还能追溯推理路径,从而提高系统的透明度和可解释性。

认知成长路径不是否定深度学习已有的成就,而是提出一种补充性学习方式——通过强调主动探索、知识建构、因果推理和可解释性,为人工智能的发展提供新的理论框架。未来的智能体如果能减少对大规模标注数据的依赖,在有限经验中建构知识,同时具备因果推理与自主适应能力,那么人工智能的发展将更接近人类认知的本质。

(二)教育启示

认知成长路径不仅是人工智能的建模策略,也为教育研究提供了可计算的模拟平台。柯尔泰(Kirtay, 2021)指出,发展型机器人模型可将皮亚杰与维果斯基的建构主义学习理论具象化,模拟儿童通过感知、注意与模仿逐步建构知识结构。毕竟这类系统不仅借鉴了认知发展的分层机制,还通过行为生成过程对教育变量(如动机、反馈、情境)进行计算建模,从而为教学机制研究提供新工具。拉焦利与坎杰洛西(Raggioli & Cangelosi, 2022)构建的认知机器人系统再现了婴儿在语言学习中如何通过具身注意力聚焦目标对象、建立词—物映射。这种“感知—注意—符号”耦合机制为教育过程的建模提供了解释性强、可调试的实验架构。

此外,该路径强调的内在动机机制,是教育心理学“自主性驱动学习”理念的工程化表达。斯里瓦斯塔瓦与辛格(Srivastava & Singh, 2025)提出的“自目的强化学习”框架模拟学生在无外部奖励的自主目标设定行为;阮(Nguyen, 2024)的“社会引导型内在动机系统”则模拟学生在模仿与探

索间自主切换的学习策略。这些研究启示教育者可通过新奇性反馈、个体控制权等变量提升学生学习主动性。

认知成长路径在知识建构机制上的模拟,也为“从做中学”“建构主义课程设计”等教育理念提供了参照。该路径强调知识通过感知—动作循环逐步形成,强调理解世界的结构性而非频率归纳,具有很强的教育启发性。

早期的儿童研究表明,学习者可在无明确指导下通过与环境互动获取因果关系。发展型机器人系统正尝试将这一能力建模为“自监督因果学习”过程。有研究者(Huang & Tseng, 2023)提出基于自监督策略的视觉因果学习框架(Momentum based Siamese Network, MoSN),使机器人在无标签条件下建立光源与开关关系的因果模型。在另一项研究中,研究者(Brawer et al., 2020)展示了机器人如何通过自发交互构建结构化因果模型,在工具使用过程中自动识别动作与结果之间的因果链条。这些研究为无监督学习环境下学生的因果推理训练提供了实验模型,也提示我们“反馈—调整—建构”可成为教学设计的重要逻辑。

认知成长路径提供了具有高可控性与可重复的“虚拟学习者”平台。伊斯拉姆等(Islam et al., 2020)开发的 SEDRo 环境可复现婴儿阶段的典型发展场景,并允许动态调整模型参数以模拟不同学习状态;蒙多尔等(Mondol et al., 2020)进一步引入社会交互模块,模拟母婴互动过程对认知发展的影响。这些模型为教育干预、发展评估与个别化教学提供可计算的模拟工具,尤其适用于婴幼儿及特殊儿童教育领域。

这一路径以跨界方式打通了教育理论预设与智能系统建模之间的联系,它不仅提供了教学设计的新范式,也促进了教育中的认知理论的工程化检验。认知成长路径让人类对自身学习机制的理解从描述走向生成,从结构走向过程,其意义不仅在于提升教学效果,更重要的是推动教育科学与技术智能的深度融合。

(三)优势与局限

认知成长路径作为模拟人类认知发展的人工智能路径,其核心不在于性能最优化,而在于学习过程的结构化、生长性与解释力。与强调效率和

准确率的统计学习不同, 该路径关注的是机器如何在持续交互中建立认知结构。这种转向不仅体现为技术策略的调整, 而且具有方法论上的意义。它提供了从结构性、阶段性和主动性角度理解机器学习的新路径。

认知成长路径的一个突出优势是, 它提供了更符合人类认知机制的学习模式。传统机器学习依赖固定的数据输入和训练目标, 而认知成长路径引入“由简到繁、由感知到推理”的渐进式学习结构, 使机器的知识建构方式更接近人类的认知发展。这种阶段式设计不仅提升了学习过程的合理性, 也为智能系统的稳定发展奠定了结构基础。该路径强调从低级感知出发, 逐步发展目标控制、因果推理与符号表达能力。这一顺序避免了将学习过程简化为“输入即推理”, 从而构建出更具有层次性和适应性的认知结构。

另一个优势在于, 认知成长路径强调主动性与环境适应能力。多数现有模型在面对任务切换或环境变化时, 需要重新训练或显著调整参数, 表现出适应能力的不足。而认知成长路径在设计上鼓励机器主动试探、持续调整策略, 通过与环境互动积累经验。其学习过程不依赖外部明确标签, 而是以系统内部的反馈机制驱动模型更新。这种“内源驱动”机制不仅有助于减少对人工标注的依赖, 也为开放式、动态环境中的机器学习提供了更稳健的基础。

认知成长路径还具有较强的结构可解释性。由于模型的学习过程具有明确阶段划分和构建逻辑, 研究者可追踪每一阶段模型内部状态的演变路径。这种可分解结构, 使智能体的行为具有可分析性, 有助于诊断学习瓶颈、优化策略设计, 并为后续的知识迁移提供基础条件。相比深度神经网络难以解释的中间层, 在任务失败或行为偏差出现时, 认知成长路径有助于识别问题所在, 有利于推动人工智能系统的透明发展。

然而, 这一路径并非没有局限。在当前的技术条件下, 认知成长路径面临显著的计算资源压力。这一路径强调与环境的持续交互和阶段性学习, 因此需要长时间运行、反复训练, 并动态调整模型参数。这种学习方式难以通过传统的离线数据处理完成, 必须在实时环境中逐步积累经验。相比一次

性处理完毕的大规模训练, 认知成长的过程更似“在线成长”, 这对计算平台的实时性、并发性与资源调度能力都提出了更高要求(Dromnelle et al., 2022)。

从理论角度看, 认知成长路径还面临心理与生理机制映射不充分的问题。尽管其设计理念来源于人类认知发展, 但当前的机器学习仍主要基于抽象算法, 缺乏对神经系统或生理结构的有效建模。这意味着该路径只能模拟认知的功能过程, 难以深入解释认知如何在生物体中发生和维持。这种缺口在研究自我意识、情绪调节或复杂社会行为等更高层认知时尤为明显。

具身智能的“认知成长路径”突破了传统人工智能静态认知模拟的局限, 强调智能的生成性、情境性与发展性特征。这一模型不仅为人工智能系统的自适应学习与认知演化提供了可解释的设计方向, 更重要的是, 它也为教育研究领域提供了新的理论工具。该模型将发展心理学关于儿童认知形成的机制与人工智能技术实现相结合, 为理解学习过程中的身体、情境与表征之间的关系提供了可操作的建模框架; 为教师教育与教学设计提供了启示——有效学习并非来自信息灌输, 而是来自“做中学”与“境中学”的交互历程。因此, 在智能教育、认知干预、教师培养等多个方向, 具身智能的认知成长路径有望作为连接心理理论与教学实践的桥梁, 推动教育心理学与人工智能的深度融合。

[参考文献]

- [1] Bengio, Y., Lecun, Y., & Hinton, G. (2021). Deep learning for AI[J]. *Communications of the ACM*, 64(7): 58-65.
- [2] Bize, S., & Cristobal, S. (2016). Self-organisation of internal models in autonomous robots. *PhD Dissertation*. <http://hdl.handle.net/1842/20981>.
- [3] Braud, R., Giagkos, A., Shaw, P., Lee, M. H., & Shen, Q. (2020). Robot multimodal object perception and recognition: Synthetic maturation of sensorimotor learning in embodied systems. [C]// *IEEE Transactions on Cognitive and Developmental Systems*, 13(3): 416-428.
- [4] Brawer, J., Qin, M., & Scassellati, B. (2020). A causal approach to tool affordance learning. [C]// 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS): 8394-8399.
- [5] Cangelosi, A., & Schlesinger, M. (2018). From babies to robots: The contribution of developmental robotics to developmental psychology [J]. *Child Development Perspectives*, 12(3): 183-188.

- [6] Chavez-Garcia, R. O., Luce-Vayrac, P., & Chatila, R. (2016). Discovering affordances through perception and manipulation [C]// 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 3959-3964). IEEE.
- [7] Chen, H., Liu, Y., Zhou, Z., & Zhang, M. (2020). A2C: Attention-augmented contrastive learning for state representation extraction[J]. *Applied Sciences*, 10(17): 5902.
- [8] Dromnelle, R., Renaudo, E., Chetouani, M., Maragos, P., Chatila, R., Girard, B., & Khamassi, M. (2022). Reducing computational cost during robot navigation and human-robot interaction with a human-inspired reinforcement learning architecture[J]. *International Journal of Social Robotics*, 15: 1297-1323.
- [9] Gao, H., & Ma, B. (2024). Online learning strategy induction through partially observable Markov Decision Process-Based Cognitive Experience Model[J]. *Electronics*, 13(19): 3858.
- [10] Gibson, J. J. (1979). The ecological approach to visual perception[M]. Houghton Mifflin.
- [11] Gopnik, A. (2009). The philosophical baby: What children's minds tell us about truth, love, and the meaning of life[M]. New York: Random House: 15-18.
- [12] Gupta, T., Gong, W., Ma, C., Pawlowski, N., Hilmkil, A., Scetbon, M., Famoti, A., Llorens, A., Gao, J., Bauer, S., Kragic, D., Schölkopf, B., & Zhang, C. (2024). The essential role of causality in foundation world models for embodied AI [J]. *arXiv preprint arXiv: 2402.06665*.
- [13] Hassan, M. A., & Lee, C. -G. (2024). Forget to Learn (F2L): Circumventing plasticity-stability trade-off in continuous unsupervised domain adaptation[J]. *Pattern Recognition*, 159: 111139.
- [14] Huang, Y., & Tseng, Y. -C. (2023). A self-supervised learning solution with momentum-based Siamese network for robotic visual task [C]//IEEE Access, 11: 112764-112775.
- [15] Islam, S. M., Mondol, M. A. R., Pothula, A., & Park, D. (2020). An open-world simulated environment for developmental robotics[J]. *arXiv preprint arXiv: 2007.09300*.
- [16] Jacquey, L., Baldassarre, G., Santucci, V. G., & O'Regan, J. K. (2019). Sensorimotor contingencies as a key drive of development: From babies to robots[J]. *Frontiers in Neurobotics*, 13: 98.
- [17] Jiao, L., Wang, Y., Liu, X., Li, L., Liu, F., Ma, W., . . . & Hou, B. (2024). Causal inference meets deep learning: A comprehensive survey[J]. *Research*, 7: 0467.
- [18] Kirsh, D. (2013). Embodied cognition and the magical future of interaction design[J]. *ACM Transactions on Computer-Human Interaction*, 20(1): 1-30.
- [19] Kirtay, M., Chevalère, J., Lazarides, R., & Hafner, V. (2021). Learning in social interaction: Perspectives from psychology and robotics [C]//2021 IEEE International Conference on Development and Learning (ICDL): 1-8.
- [20] Kumar, S., Giagkos, A., Shaw, P., Braud, R., Lee, M., & Shen, Q. (2021). Discovering schema-based action sequences through play in situated humanoid robots [C]//IEEE Transactions on Cognitive and Developmental Systems, 14(6): 1021-1035.
- [21] Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people[J]. *Behavioral and Brain Sciences*, 40: e253.
- [22] Lang, C., Schillaci, G., & Hafner, V. V. (2018). A deep convolutional neural network model for sense of agency and object permanence in robots[C]// Joint IEEE 8th international conference on development and learning and epigenetic robotics (ICDL-EpiRob) (pp. 257-262). IEEE.
- [23] Lanillos, P., Dean-Leon, E., & Cheng, G. (2016). Yielding self-perception in robots through sensorimotor contingencies[J]. *IEEE Transactions on Cognitive and Developmental Systems*, 9(2): 100-112.
- [24] Law, J., Shaw, P., Earland, K., Sheldon, M., & Lee, M. (2014). A psychology based approach for longitudinal development in cognitive robotics[J]. *Frontiers in neurobotics*, 8: 1.
- [25] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning[J]. *Nature*, 521(7553): 436-444.
- [26] Liang, Y., & Li, W. -J. (2023). Adaptive plasticity improvement for continual learning. [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 7816-7825.
- [27] Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery[J]. *Queue*, 16(3): 31-57.
- [28] Liu, H., Guo, D., & Cangelosi, A. (2025a). Embodied Intelligence: A synergy of morphology, action, perception and learning[J]. *ACM Computing Surveys*, 57(7): 2-36
- [29] Liu, Y., Liu, S., Chen, B., Yang, Z., & Xu, S. (2025b). Fusion-perception-to-action transformer: Enhancing robotic manipulation with 3-D visual fusion attention and proprioception[J]. *IEEE Transactions on Robotics*, 41(4): 1553-1567.
- [30] Macrine, S., & Fugate, J. M. B. (2020). Embodied cognition [J]. In *Oxford Research Encyclopedia of Education*: 1-24. <https://doi.org/10.1093/acrefore/9780190264093.013.885>.
- [31] Mattern, D., Schumacher, P., López, F. M., Raabe, M. C., Ernst, M., Aubret, A., & Triesch, J. (2024). MIMo: A multimodal infant model for studying cognitive development[J]. *IEEE Transactions on Cognitive and Developmental Systems*, 16: 1291-1301.
- [32] Mondol, M. A. R., Pothula, A., & Park, D. (2020). Modeling social interaction for baby in simulated environment for developmental robotics[J]. *arXiv preprint arXiv: 2012.14842*.
- [33] Nguyen, S. M. (2024). The intrinsic motivation of reinforcement and imitation learning for sequential tasks[J]. *arXiv preprint arXiv: 2412.20573*.
- [34] O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness[J]. *Behavioral and Brain Sciences*, 24(5): 939-973.
- [35] Oudeyer, P. -Y. (2012). Developmental mechanisms for autonomous life-long learning in robots. [M]//ECAI 2012. IOS Press: 39-43.
- [36] Peng, S., Wang, M. X., Shah, J. A., & Figueroa, N. (2024). Object permanence filter for robust tracking with interactive robots[C]//2024 IEEE International Conference on Robotics and Automa-

tion (ICRA): 4909-4915.

[37] Piaget, J. (1952). The origins of intelligence in children[M]. New York: International Universities Press: 1-6, 13-40, 81-146.

[38] Piaget, J. (2013). The construction of reality in the child[M]. New York: Routledge: 81-149.

[39] Pinto, L., & Gupta, A. (2015). Supersizing self-supervision: Learning to grasp from 50K tries and 700 robot hours[C]//2016 IEEE international conference on robotics and automation (ICRA). IEEE: 3406-3413.

[40] Power, F. (2023). Thinking with my hands: Embodied cognition in practice[J]. Journal of Textile Design Research and Practice, 11(1-2): 111-126.

[41] Raggioli, L., & Cangelosi, A. (2022). Embodied attention in word-object mapping: A developmental cognitive robotics model [C]//2022 IEEE International Conference on Development and Learning (ICDL). IEEE: 156-163.

[42] Roy, N., Posner, I., Barfoot, T., Beaudoin, P., Bengio, Y., Bohg, J., . . . & Van de Panne, M. (2021). From machine learning to robotics: Challenges and opportunities for embodied intelligence[J]. arXiv preprint arXiv: 2110.15245.

[43] Sartor, G., Oddi, A., Rasconi, R., Santucci, V. G., & Meo, R. (2024). Synthesizing Evolving Symbolic Representations for Autonomous Systems[J]. arXiv preprint arXiv: 2409.11756.

[44] Savastano, P., & Nolfi, S. (2012). Incremental learning in a 14 dof simulated icub robot: Modeling infant reach/grasp development [C]//Conference on Biomimetic and Biohybrid Systems. Berlin, Heidelberg: Springer Berlin Heidelberg: 250-261.

[45] Solodov, A. A. (2021). Optimal Poisson Cognitive System with Markov Learning Model[J]. Open Education, 25(6): 45-52.

[46] Srivastava, P., & Singh, J. (2025). Autotelic reinforcement learning: Exploring intrinsic motivations for skill acquisition in open-ended environments[J]. arXiv preprint arXiv: 2502.04418.

[47] Tokmakov, P., Jabri, A., Li, J., & Gaidon, A. (2022). Object permanence emerges in a random walk along memory[J]. arXiv preprint arXiv: 2204.01784.

[48] Vahedifar, M. A., & Zhang, Q. (2025). No forgetting learning: memory-free continual learning[J]. arXiv preprint arXiv: 2503.04638.

[49] Varela, F. J., Thompson, E., & Rosch, E. (1991). The embodied mind: Cognitive science and human experience [M]. Cambridge, Massachusetts: MIT Press: 13.

[50] Xiong, H., Wang, Z., Li, X., Bian, J., Xie, Z., Mumtaz, S., & Barnes, L. E. (2024). Converging paradigms: The synergy of symbolic and connectionist AI in LLM-empowered autonomous agents [J]. ArXiv preprint arXiv: 2407.08516.

(编辑: 李学书)

A Cognitive Growth Model of Embodied Intelligence: Exploring Dynamic Cognition in Educational Contexts

SU Jiajia¹ & YE Haosheng²

(1. School of Educational Science, Jiangsu Second Normal University, Nanjing 211200, China;

2. School of Education, Guangzhou University, Guangzhou 510006, China)

Abstract: Embodied intelligence research emphasizes that cognition does not emerge from abstract computation alone, but through dynamic interactions among the body, environment, and sensorimotor experiences. Integrating perspectives from cognitive science, developmental psychology, and artificial intelligence, this study outlines a staged pathway of cognitive development, ranging from bodily regulation and object construction to intention simulation and inferential reasoning. The study identifies and discusses key enabling mechanisms, such as embodied regulation, situated construction, and representational reorganization. By introducing computational principles, the study also demonstrates the theoretical plausibility of simulating these mechanisms in artificial systems. Finally, the study proposes a framework that offers conceptual tools for understanding cognitive variability and practical implications for educational psychology, cognitive development assessment, and teacher learning support.

Key words: embodied intelligence; embodied cognition; cognitive growth; machine learning