面向精准教学的教育知识图谱表示学习与动态推理研究*

徐多礼[1,2] 连硕[1,2**]

1.重庆移通学院计算机学院，重庆 401420；

2.公共大数据安全技术重庆市重点实验室，重庆 401420

摘 要 随着高等教育数字化转型深入，教学数据多源异构、语义挖掘不足与个性化推荐冷启动等问题日益凸显。针对教育数据稀疏性与时序动态特性，本文构建了面向精准教学的教育知识图谱计算框架，提出基于超平面投影的深度表示学习机制，结合图注意力网络设计多跳概率推理模型，并引入元学习策略适配小样本场景，最终搭建融合 OBE 理念的层级化达成度评价体系。该方法可有效将隐性认知转化为量化指标，提升稀疏数据下的推理鲁棒性，实现可解释的个性化学习路径生成与全过程教学评价，为高校教学从经验驱动向数据驱动的精准治理转型提供了解决方案。

关键字 教育知识图谱；表示学习；动态推理；OBE；元学习

Research on Educational Knowledge Graph Representation Learning and Dynamic Reasoning for Precision Teaching *

Duoli Xu[1,2] Shuolian[1,2]

1. College of Computer Science Chongqing College of Mobile Communication, Chongqing 401420, China;

2. Chongqing Key Laboratory of Public Big Data Security Technology Chongqing 401420, China;

Abstract—As the digital transformation of higher education deepens, issues such as the multi-source and heterogeneous nature of teaching data, insufficient semantic mining, and the cold-start problem in personalized recommendations have become increasingly prominent. Addressing the sparsity and dynamic temporal characteristics of educational data, this paper constructs a computational framework for educational knowledge graphs aimed at precision teaching. It proposes a deep representation learning mechanism based on hyperplane projection, designs a multi-hop probabilistic inference model using graph attention networks, and introduces a meta-learning strategy to adapt to small-sample scenarios. Ultimately, the paper establishes a hierarchical achievement evaluation system that integrates the OBE (Outcomes-Based Education) philosophy. This method effectively transforms tacit knowledge into quantifiable metrics, enhances the robustness of inference under sparse data conditions, and enables the generation of interpretable personalized learning pathways and end-to-end instructional evaluation. It provides a solution for the transition of higher education from experience-driven to data-driven precision governance.

Keywords—Educational Knowledge Graph; Representation Learning; Dynamic Reasoning; OBE; Meta-Learning

1 引 言

在高等教育数字化转型迈向深水区的当下，高校教学环境正面临着前所未有的数据爆发与认知迷航双重挑战[1]。随着大规模在线开放课程（MOOCs）、虚拟仿真实验及各类数字化教学平台的广泛应用，海量的多源异构数据——从结构化的教学大纲、半结构化的习题库到非结构化的课堂实录与多模态教案——呈指数级增长[2]。然而，正如王威[3]在其研究中深刻指出的，这种单纯的“资源数字化”并未直接转化为教学效能的提升，反而因缺乏有效的语义组织导致了严重的“信息过载”，使得学习者在碎片化的知识海洋中面临“检索难、路径乱、认知浅”的困境。周童[4]进一步强调，现有的主流教学系统仍主要依赖基于关键词匹配的浅层统计模型，无法理解知识点之间深层的逻辑依赖与学科概念的演化规律，导致了技术与育人之间的巨大“语义鸿沟”（Semantic Gap）。在此背景下，知识图谱（Knowledge Graph，KG）[5]作为连接数据与认知的桥梁，其价值不再局限于静态的知识可视化，

*基金资助：2026 年重庆移通学院校级高等教育教学改革研究项目“知识图谱与大模型双驱的通用编程课程教学新生态建构及评价研究”（26JG2007）；2025 年重庆移通学院校级应用研究项目“知识图谱驱动下高校“教-学-评”一体化融合机制研究与探索”（KY202552）；重庆市高等教育学会 2025-2026 年度高等教育科学研究课题“知识图谱驱动下高校“教－学－评”一体化融合机制研究与探索”（cqgj25197C）

**通讯作者：连硕 shuo.lian@qq.com

而在于通过表示学习（Representation Learning）与动态推理（Dynamic Reasoning）技术，将复杂的教学交互转化为机器可理解的计算模型。然而，教育场景具有极高的数据稀疏性与时序动态性，江丹[6]与蔡丽[7]的研究表明，传统的通用图谱嵌入算法在处理特定学科的冷门知识点及随时间快速迭代的技术概念（如计算机学科的版本更迭）时，往往表现出鲁棒性不足与推理失效的问题。因此，本研究聚焦于“面向精准教学的教育知识图谱表示学习与动态推理方法”，旨在通过引入深度神经网络与时序演化机制，构建能够精准捕捉多模态语义与动态认知状态的计算基座，从而解决精准教学中面临的“冷启动”推荐与“不可解释”推理难题，为推动高校教学从“经验驱动”的粗放治理向“数据驱动”的精准治理转型提供坚实的算法理论支撑。

2 教育知识图谱构建的理论基础与表示学习机制

2．1 教育知识图谱的形式化定义与多模态特征融合

在面向精准教学的智能教育系统中，教育知识图谱（Educational Knowledge Graph, Edu-KG）的构建不再是简单的实体罗列，而是一个涉及多模态语义对齐、深度特征表示以及时空动态演化的复杂计算过程。首先，从形式化定义的角度来看，教育知识图谱本质上是一个具有高度异构性和语义丰富性的有向多重图。不同于传统的通用图谱，Edu-KG需要同时通过符号化与向量化两种范式来描述教学场景中的实体及其复杂关系。我们将其形式化定义为一个包含多模态特征空间的四元组 $G=(E,R,T,\mathrm{M})$。其中，$E=\{e_1,e_2,...,e_{|E|}\}$ 代表教学实体的集合，涵盖了知识点、教学资源（视频、课件）、习题及认知状态等离散对象；$R=\{r_1,r_2,...,r_{|R|}\}$ 代表实体间关系的集合，既包含显性的“前驱”、“包含”逻辑关系，也包含隐性的“相关”、“增强”语义关联；$T\subseteq E\times R\times E$ 则是由头实体、关系与尾实体构成的三元组事实集合 (h,r,t)。更为关键的是，针对李琴[8]与黄冕[9]在研究中指出的多源异构数据融合难题，必须引入多模态特征集合 M 来增强图谱的语义表达能力。在实际构建中，单一的文本模态往往无法完整刻画一个复杂的教学概念（如“电磁感应”既需要公式定义的文本描述，也需要实验现象的视频表征）。因此，本研究采用基于大模型增强的跨模态对齐理论，对于任意教学实体 $e\in E$，利用预训练编码器将视觉（Visual）、听觉（Audio）与文本（Text）信息映射到统一的语义空间中。其多模态特征聚合表示 $\mathbf{x}_e$ 定义为各模态特征向量的非线性拼接：

$$\mathbf{x}_e=\mathrm{Concat}(\phi_v(e),\phi_a(e),\phi_t(e)) \quad (1)$$

这种形式化定义为解决教育数据的“语义鸿沟”提供了数学公理基础，确保了后续的计算模型能够在一个完备且统一的高维向量空间中对教学对象进行精确操作，从而实现从非结构化数据到结构化知识的根本性转化。

2．2 基于深度神经网络的超平面投影表示学习机制

在建立完备的形式化定义之后，核心挑战在于如何将离散的符号化知识映射为连续的低维稠密向量，即表示学习（Representation Learning）。传统的平移距离模型（如TransE）假设头实体向量加上关系向量等于尾实体向量，这种浅层线性模型在处理教育数据普遍存在的稀疏性（Sparsity）与复杂关系（1-N, N-N）时表现出严重的欠拟合。为此，本研究引入江丹[6]提出的基于深度神经网络的表示学习框架，并通过肖昌义[10]的超平面投影理论进行改进。该理论认为，教育图谱中的关系并非单一维度的平移，而是发生在特定关系子空间中的投影变换。例如，实体“Python语言”在“属于”关系下指向“编程语言”类别，而在“包含”关系下则指向“列表”、“元组”等具体知识点，这意味着同一实体在不同关系语境下应具有不同的语义投影。为了数学化这一机制，我们将实体 h,t 和关系 r 映射到低维向量空间 $\mathbb{R}^d$，并引入关系特定的投影矩阵 $\mathbf{w}_r$，将头尾实体投影到关系超平面上，记为 $\mathbf{h}_\perp$ 和 $\mathbf{t}_\perp$。在此基础上，构建基于L2范数的能量评分函数 $f_r(h,t)$，用于衡量三元组在语义空间中的合理性。为了进一步优化模型在稀疏数据上的表现，我们采用基于间隔的排序损失函数（Margin-based Ranking Loss）作为训练目标，旨在最大化正样本（真实教学关联）与负样本（随机噪声关联）之间的距离。该优化过程通过随机梯度下降算法驱动深度神经网络参数的更新，使得模型能够自动学习到知识点之间隐含的层级结构与逻辑依赖。其评分函数与损失函数的数学表达分别如下：

$$\mathbf{h}_\perp=\mathbf{h}-\mathbf{w}_r^{\text{晦}}\mathbf{h}\mathbf{w}_r,\quad \mathbf{t}_\perp=\mathbf{t}-\mathbf{w}_r\,\mathbf{t}\mathbf{w}_r \quad (2)$$

$$f_r(h,t)=-\|\mathbf{h}_\perp+\mathbf{r}-\mathbf{t}_\perp\|_2^2 \quad (3)$$

$$L=\sum_{(h,r,t)\in S}\sum_{(h',r,t')\in S'}\max(0,\gamma+f_r(h',t')-f_r(h,t)) \quad (4)$$

2.3 融合图拓扑聚合与时序演化的动态推理模型

这一理论框架不仅克服了浅层模型无法建模“多对多”复杂关系的缺陷，更通过深度非线性变换显著提升了知识表示的鲁棒性，为精准教学推荐提供了高精度的向量化依据。然而，静态的向量表示尚不足以支撑复杂的认知推理与动态知识更新，必须进一步融合图拓扑结构聚合机制与时序演化理论。针对知识点之间深层的逻辑依赖与认知传递效应，本研究结合武月佳[11]提出的图注意力网络（Graph Attention Network, GAT）模型，设计了邻域感知的信息聚合机制。在这一机制中，一个知识点的语义表示不再是孤立的，而是由其邻居节点（即前修知识点或相关概念）加权聚合而成。通过引入自适应注意力系数 α_{ij}，模型能够自动识别并赋予关键前修知识点更高的权重，模拟人类学习过程中的“联想”与“迁移”机制。同时，考虑到计算机等学科知识的高频迭代特性，蔡丽[7]的时序知识图谱（Temporal Knowledge Graph, TKG）理论被引入以解决“知识过时”问题。我们将时间维度 τ 映射为复数向量空间中的旋转算子，使得知识表示随时间推移发生动态位移。例如，随着Python版本的更新，某些语法规则的有效性会随时间衰减或改变，这种变化通过旋转向量 $\mathbf{rot}(\tau)$ 作用于实体向量来实现。这种时空联合建模方法，使得教育知识图谱具备了“动态认知”的能力，既能通过图卷积捕捉全局的学科结构，又能通过时序旋转追踪知识的演化轨迹。其图卷积聚合公式、注意力系数计算及包含时间维度的动态评分函数分别表示为：

$$\mathbf{h}_i^{(l+1)} = \sigma\left(\sum_{j \in \mathrm{N}_i} \alpha_{ij} \mathbf{W}^{(l)} \mathbf{h}_j^{(l)} \right) \tag{5}$$

$$\alpha_{ij} = \frac{\exp(\mathrm{LeakyReLU}(\mathbf{a}^{\cdot}\,[\mathbf{W}\mathbf{h}_i \,\|\, \mathbf{W}\mathbf{h}_j]))}{\sum_{k \in \mathrm{N}_i} \exp(\mathrm{LeakyReLU}(\mathbf{a}^{\cdot}\,[\mathbf{W}\mathbf{h}_i \,\|\, \mathbf{W}\mathbf{h}_k]))} \tag{6}$$

$$f_r(h,t,\tau) = \| (\mathbf{h} \circ \mathbf{rot}(\tau)) \circ \mathbf{r} - (\mathbf{t} \circ \mathbf{rot}(\tau)) \| \tag{7}$$

3 基于动态推理的精准教学方法实现与策略设计

3.1 基于多跳推理的可解释性路径规划策略

在精准教学系统的核心逻辑设计中，如何为不同认知基础的学生生成一条既符合学科逻辑又适应个体差异的学习路径，是本研究要解决的首要问题。本研究不再采用传统的基于相似度匹配的协同过滤推荐，而是设计了一种基于概率图模型的多跳推理机制（Multi-hop Reasoning Mechanism）。该机制的设计思想是将学习路径的生成过程转化为知识图谱上的最大概率路径搜索问题。具体而言，系统首先利用第二章构建的深度表示向量，计算当前学生认知状态节点 e_{start} 到教学目标节点 e_{target} 之间的所有潜在连通路径。为了保证推荐结果具备教育学意义上的可解释性，本文结合武月佳[11]的推理框架，设计了包含“前修铺垫”、“核心攻坚”与“拓展应用”的三段式推理策略。对于每一条候选路径 P，系统基于图注意力网络（GAT）计算其合理的转移概率。路径 P 的综合置信度分数 $Score(P)$ 不仅取决于节点间的语义相似度，还受到教学关系权重（如“前驱”关系的权重高于“相关”关系）的约束。其路径概率计算公式设计如下：

$$Score(P) = \prod_{k=0}^{n-1} \sigma(\mathbf{h}_{e_k}^{\cdot} \mathbf{M}_{r_{k+1}} \mathbf{h}_{e_{k+1}}) \cdot \beta^{len(P)} \tag{8}$$

其中，$\mathbf{M}_r$ 是关系特定转移矩阵，$\beta \in (0,1)$ 是路径长度惩罚因子，用于避免生成过于冗长的无效路径。基于此计算策略，系统针对不同学情的学生实施差异化导航：对于认知状态向量显示为“基础薄弱”的学生，算法自动锁定短路径并高权值激活包含“概念定义”类实体的补救路径；而对于“高阶思维”学生，系统则通过降低长度惩罚因子，引导其探索跨学科的远距离推理路径（如物理学中的电磁感应链接到工程学中的传感器设计），从而在逻辑层面实现黄厚伟[12]所倡导的分层教学策略。

3.2 面向冷启动场景的模型训练与参数优化设计

在实际教学应用中，面对新开设课程或新入学学生交互数据匮乏的“冷启动”场景，常规深度学习模型往往因无法收敛而失效。为此，本研究在模型训练环节创新性地引入了基于元学习（Meta-Learning）的参数初始化策略。该策略的设计核心不在于让模型“记忆”所有学生的历史数据，而在于训练模型“快速适应”新任务的能力。结合梁逸[13]的小样本学习理论，我们将训练过程设计为两个阶段：元训练阶段（Meta-training）与微调适应阶段（Fine-tuning）。在元训练阶段，模型在具有丰富数据的源域课程（如《计算机基础》）上进行训练，学习通用的认知推理模式（例如：掌握了A通常就能掌握B）；在适应阶段，当面对目标域新课程（如《Python高级编程》）时，模型仅需利用极少量的支持集样本（Support Set，例如学生对前3个知识点的答题反馈），即可通过梯度下降快速更新参数。为了防止模型在少样本下发生过拟合，我们设计了基于梯度的元学习损失函数，旨在寻找一组对新任务极

其敏感的初始化参数θ。其优化目标函数设计为：

$$\theta^{*}=\arg\min_{\theta}\sum_{T_i\sim p(T)}L_{T_i}(f_{\theta-\alpha\nabla L_{T_i}(\theta)}) \quad (9)$$

该公式的物理含义是：寻找一组最优参数θ，使得该参数经过一步或多步梯度更新（α为学习率）后，在特定任务T_i上的损失函数L_{T_i}最小化。通过这种训练策略的设计，系统能够在新学生仅产生少量交互行为时，迅速捕捉其认知特征分布，实现从“零数据”到“精准推荐”的平滑过渡，从算法设计层面解决了精准教学系统的冷启动难题。

3．3 基于 OBE 理念的达成度量化计算与评价体系构建

为了将精准教学的实施效果转化为可量化的教育产出，本研究构建了基于成果导向教育（OBE）理念的自动化评价体系。该体系的设计打破了传统仅依赖期末卷面成绩的评价模式，转而建立了一种基于图谱结构的全过程动态计算模型。结合惠林虎[14]的OBE落地策略，我们设计了“毕业要求指标点（Outcome）—课程目标—微观知识点（Knowledge）”的层级映射矩阵$\mathbf{W}_{K-O}$。在该矩阵中，每一个微观知识点的掌握情况都以特定权重支撑着宏观能力指标点的达成。系统通过实时采集学生在各个知识点上的交互数据（作业得分、阅读时长、实验操作），利用加权聚合算法自动计算各项能力指标的达成度$A(O_j)$。同时，对于田学泽[15]提出的课程思政建设需求，本设计进一步扩展了评价维度，通过实体关系抽取技术挖掘专业知识点与“国家安全”、“工程伦理”等思政元素之间的隐性关联。在计算策略上，系统不仅统计显性的知识得分，还通过计算学生在思政关联路径上的注意力驻留权重来量化情感价值观的达成情况。OBE达成度的核心计算公式定义为：

$$A(O_j)=\frac{\sum_{i=1}^{|K|}w_{ij}\cdot s_i}{\sum_{i=1}^{|K|}w_{ij}\cdot s_{total}}+\lambda\cdot\mathrm{Attention}(Path_{moral}) \quad (10)$$

其中，w_{ij}为图谱定义的知识点对指标点的支撑权重，s_i为学生实测得分，λ为思政维度的调节因子。这一计算策略的设计，使得教学评价不再是滞后的总结性判断，而是伴随教学过程实时生成的诊断性反馈，实现了从“经验定性评价”向“数据定量评价”的方法论转型，确保了教学过程与培养目标的精准对齐。

4 基于动态推理的编程类课程精准教学场景设计

4．1 基于小样本推理的冷启动与学情诊断机制

编程类课程（如C语言、Python等）在开课初期普遍面临学生先修基础差异巨大的“冷启动”难题。部分学生具备较强的逻辑思维或编程基础，而另一部分则为完全零基础。为解决这一问题，系统直接引入基于元学习的小样本推理机制实施敏捷学情诊断。具体而言，系统向新生推送一组极少量的微型测试题（即支持集，如基础的变量赋值与逻辑判断题）。模型通过提取学生在这些基础交互中的元知识，结合时序动态权重，快速计算出其对后续高阶概念（如“面向对象”、“高阶函数”）的潜在掌握概率。基于这一推理结果，系统自动生成差异化的学习导航：针对零基础学生，高权重推送底层的“控制流图解”等铺垫性实体资源；针对有基础的学生，则直接解锁跨节点的挑战性任务。这一场景设计从根本上规避了传统教学系统依赖海量历史数据的弊端，实现了学习初期的精准导学。

4．2 基于多跳概率推理的代码调试与伴随干预模型

在编程课程的实践与上机实验环节，学生面临的最大障碍是代码报错后的“调试（Debug）迷航”。传统的集成开发环境（IDE）仅提供表层的语法错误提示（如抛出异常代码），缺乏教育学意义上的指导。为突破此瓶颈，本环节部署了基于图注意力网络（GAT）的多跳推理策略，设计出伴随式 AI 干预机制。以 Python 教学为例，当学生在修改字符串或元组时触发了异常，系统不再仅仅翻译报错信息，而是启动图谱推理引擎。算法沿着“异常表象→语法规则→底层原理”的逻辑链条进行反向多跳搜索，最终输出具有极强可解释性的干预路径：“系统检测到类型错误，因为你试图修改元组（当前行为），而元组在内存中属于不可变对象（核心原理），建议回顾‘可变与不可变数据类型’的底层区别（前修知识点）。”这种基于最大概率路径搜索的干预模式，模拟了专家教师的循循善诱，实现了知识点的即时纠偏与认知重构。

4．3 驱动 OBE 理念的量化评价与思政隐性融合路径

教学评价是检验教学成效与落实成果导向教育（OBE）的核心阵地。为实现精确的数据度量，评价系统直接依托“微观知识点—宏观指标点”层级映射矩阵$\mathbf{W}_{K-O}$，对学生的认知状态进行动态量化。在任务发布阶段，系统根据推理出的个体能力短板自动组卷，实现“千人千题”。更为重要的是，针对理工科课

程思政融入生硬的痛点，系统应用了图谱中的关系抽取技术，将专业操作与工程伦理实现了深度的隐性关联。例如，在Python“网络爬虫”或“数据分析”章节的实践中，图谱自动将“requests库应用”等技术节点与“数据安全法”、“隐私保护”等思政节点进行连通。系统通过追踪学生在处理敏感数据相关实验时的行为轨迹与停留权重，利用达成度计算公式动态输出其在“工程伦理”这一素质指标上的得分。该场景设计使得教学评价不再是单一的总结性断言，而是兼顾了技术能力与价值引领的全过程数据刻画。

5 结束语

本文针对高校精准教学中数据语义挖掘不足的现状，归纳了知识图谱在教育领域的应用现状与发展趋势，并重点探讨了面向教育场景的表示学习与动态推理方法。在技术手段上，文章整合了深度神经网络与时序演化理论，构建了适应教育数据稀疏性的图谱构建框架，并提出了基于元学习的小样本推理策略，旨在缓解新课程与新用户的“冷启动”难题。在应用策略上，设计了基于概率多跳推理的可解释性路径规划机制，以及基于OBE理念的达成度量化计算体系，为实现差异化导学与全过程评价提供了具体的实施思路。未来研究将进一步关注大语言模型与知识图谱的深度协同，以期为构建“人机协同”的智慧教育新生态提供更具实效的参考。

参 考 文 献

[1] 李向军,刘伯成,张坚林,等. 数字化转型背景下实战型网络安全人才培养探索与实践[J].计算机技术与教育学报,2024,12(2):77-82.

[2] 郭小英,闫涛.以数字化教学助力高质量发展的数字图像处理课程范式[J].计算机技术与教育学报,2025,13(7):139-144.

[3] 王威.基于知识图谱的推荐方法及应用研究[D].华中师范大学,2024.

[4] 周童,陈德春,齐宁,等.基于知识图谱的高校课程教学模式创新与实践[J].大学,2025,(32):172-175.

[5] 甘书灵,史东辉,王园园.基于知识图谱技术的计算机教育学情内涵探索[J].计算机技术与教育学报,2023,11(2):39-45.

[6] 江丹.基于深度神经网络的知识图谱表示学习研究[D].合肥工业大学,2023.

[7] 蔡丽.时序知识图谱表示学习[D].华东师范大学,2024.

[8] 李琴,宋钰.数字化教育背景下数据库原理课程知识图谱构建[J].电脑知识与技术,2025,21(33):130-132.

[9] 黄冕.多模态知识图谱赋能教师教育的三重逻辑[J].教育科学论坛,2025,(34):57-62.

[10] 肖昌义.知识图谱表示学习理论与方法研究[D].中国科学技术大学,2024.

[11] 武月佳.基于图神经网络的知识图谱推理研究[D].内蒙古大学,2024.

[12] 黄厚伟,邱晓燕.基于知识图谱的大学物理AI+课程个性化教学实践[J/OL].大学物理,1-6[2026-01-15].

[13] 梁逸.基于小样本数据的知识图谱链接预测算法的研究[D].北京邮电大学,2023.

[14] 惠林虎,李永彬,常永虎.基于OBE与AI辅助教学融合的计算机组成原理课程教学改革[J].西部素质教育,2025,11(21):145-148.

[15] 田学泽,李文涵,陈奕霖,等.基于知识图谱的课程思政建设研究热点及趋势分析[J].江苏经贸职业技术学院学报,2025,(06):76-80.