

# 图模融合:人工智能系统事实表达和逻辑推理增强

杨娟<sup>1,2</sup>, 沈游人<sup>2</sup>

1. 清华大学计算机科学与技术系, 北京 100084;

2. 北京海致星图科技有限公司, 北京 100085

## 摘要

知识图谱通过图结构组织和表达实体关系, 为机器理解和推理提供了基础, 但推理能力受限于覆盖范围和人工规则。大模型展现了强大的语义理解和生成能力, 但缺乏对符号知识的有效利用和可解释性。近年来, 学术界和产业界致力于探索知识图谱与大模型的结合, 以融合两者优势, 构建更强大且可解释的人工智能系统。首先, 梳理了知识图谱与大模型融合的研究现状, 重点介绍了两种技术融合在事实表达和逻辑推理增强方面的主要研究成果, 包括基于知识图谱的预训练语言模型、基于大模型的知识图谱表示学习以及二者融合的推理方法。然后, 对产业界当前图模融合的主流技术路线和应用场景进行了介绍。最后, 对图模融合未来的发展方向进行了展望, 并且提出两者融合是人工智能的重要发展方向之一。

## 关键词

知识图谱; 大模型; 图神经网络; 机器推理; 预训练; 注意力; 图表示学习

中图分类号: TP391

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2025014

## *Graph-LLM fusion: enhancing fact representation and logical reasoning in artificial intelligence systems*

**YANG Juan<sup>1,2</sup>, SHEN Youren<sup>2</sup>**

1. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

2. Beijing Haizhi Stargraph Technology Co., Ltd., Beijing 100085, China

## *Abstract*

Knowledge graphs organize and represent entity relationships through graph structures, providing a foundation for machine understanding and reasoning, but their reasoning capabilities are limited by coverage and manual rules. Large language models demonstrate strong semantic understanding and generation abilities but lack effective utilization of symbolic knowledge and interpretability. To combine the strengths of both technologies, academic and industrial communities have devoted significant effort in recent years to exploring the integration of knowledge graphs and large language models, aiming to build more powerful and interpretable AI systems. Firstly, this paper reviews the current state of research on the fusion of knowledge graphs and large language models, with a focus on the key achievements in enhancing fact representation and logical reasoning. These achievements include pre-trained language models based on knowledge graphs, knowledge graph representation learning based on large language models, and reasoning methods

that leverage the fusion of the two approaches. Furthermore, the paper outlines the mainstream technical approaches and application scenarios of graph-model integration in the industry. Finally, future development directions of graph-model integration are discussed, and it is posited that the integration of these two technologies represents a crucial trend in the advancement of artificial intelligence.

### Key words

knowledge graph, large language model, graph neural network, machine reasoning, pre-training, attention, graph representation learning

## 0 引言

知识图谱 (knowledge graph, KG) 是一种对结构化知识进行表达和符号推理的经典技术, 具备准确、推理高效、可解释性强等特点, 但同时面临着知识获取和构建成本高、泛化能力弱等挑战<sup>[1-2]</sup>。大语言模型 (简称大模型) 则对海量语义信息具备很强的提取、推理和生成能力, 但面临幻觉问题、可解释性差、知识获取和更新困难等挑战<sup>[3]</sup>。

两种技术有各自的特点和优劣势, 如何取长补短, 实现优势互补成为当前研究的热点。当前学术界对知识图谱和大模型的协同进行了广泛的研究。协同方式主要包括两种, 一种是两种技术保持独立, 对彼此进行增强, 例如使用大模型辅助知识图谱构建或利用知识图谱提升大模型的知识储备<sup>[4]</sup>; 另一种是两种技术进行深度的融合, 实现  $1+1>2$  的效果。

本文重点介绍深度融合知识图谱与大模型的主要研究成果和状态, 并对未来可能的发展方向进行讨论。其中, 融合方式又分为增强人工智能系统事实表达和逻辑推理能力两类, 分别着重于构建对事实更加完备的描述与表达及联合推理, 以处理解决复杂的任务和问题。例如, 在事实表达方面, 可以将知识图谱中的结构化知识融入大模型的预训练过程, 从而提升模型对实体和关系的理解

能力; 而在逻辑推理方面, 可以利用大模型基于图上的数据进行推理, 使复杂的逻辑问题得到准确且可解释的回答。

本文的内容组织如下。首先, 回顾知识图谱和大模型各自的技术原理和发展历程, 并深入分析两者的优势和局限性。其次, 聚焦于知识图谱与大模型的协作方式, 着重探讨深度融合这一前沿方向。本文提出融合方法主要有两大类——知识表达能力增强和推理能力增强, 并结合具体研究案例, 对这两类方法进行详细的阐述和比较分析。然后, 讨论产业界关于图模融合技术技术路径的探索, 并以智能审计场景为例, 给出图模融合实际应用案例, 以展示图模融合的优势。最后, 本文针对图模融合现有的问题和挑战进行深入探讨, 并展望未来可能的技术突破点和研究方向。

## 1 大模型和知识图谱发展现状和特点

### 1.1 大模型发展现状和特点

2017年, Google的研究者在机器学习领域引入自注意力机制, 提出 Transformer<sup>[5]</sup> 模型, 该架构通过自注意力机制并行处理序列中的每个单词, 从而有效地捕捉单词间的长期依赖关系。在预训练阶段, 大模型通常采用掩码语言模型或下一个词预测等任务学习语言的统计特性; 然后, 通过微调来适应特定的下游任务。

Transformer 模型在处理序列数据方面比传统的循环网络更加高效,这为后续大模型的发展奠定了基础。随着 Transformer 模型的成功,研究人员开始构建越来越大的模型,如 OpenAI 的 GPT 系列、Google 的 BERT 模型等。这些模型的参数规模从数十亿迅速增长到千亿甚至万亿级别。2020 年,OpenAI 发布了 GPT-3,这是一个具有 1 750 亿参数的大模型,它在自然语言生成、理解和多任务学习方面展现了惊人的能力。

此后,国内外迅速涌现出 Anthropic、Cohere、LLAMA、ChatGLM、QWen 等大模型。当前,大模型呈现出参数规模增长迅速、多模态兴起、开源模型增多、商业应用广泛、由通用领域走向专业领域等特点。尽管大模型在多个领域展现出强大的应用潜力,但也面临着计算资源需求高昂、数据需求量大、可解释性差、知识获取和更新困难等挑战。

## 1.2 知识图谱的发展现状和特点

知识图谱是一种用于表示实体和实体间的各种关系和属性的结构化数据结构。它将现实世界中的各种实体(如人、地点、事物等)和它们之间的关系以图的形式表示出来,并通过节点和边来存储信息。它能发现数据之间的潜在联系,使数据检索变得更加高效,尤其是在处理复杂的关系查询时。知识图谱能够灵活地更新和扩展,可以随着时间的推移不断地加入新的知识和信息。同时,知识图谱的推理过程是准确的、显式的。

当前知识图谱被广泛应用于搜索、风控、营销、产品推荐、客服等领域。在运用过程中,它也面临着知识获取和构建成本

## 2 图模融合的技术发展路线

知识图谱和大模型两种技术具备互补性优势。知识图谱通常包含大量的结构化信息,这些信息可以作为先验知识帮助大模型更好地理解数据,提高数据效率,减少大模型对大规模标注数据的依赖,这在需要逻辑推理和知识检索的任务中非常有用。将知识图谱与大模型相结合,可以增强模型在需要复杂推理的任务中的表现,并清晰地展示决策的依据,提高模型的可解释性和用户的信任度。知识图谱可以帮助大模型在不同的领域和任务中迁移知识,帮助大模型更快地适应新的专业领域任务。知识图谱可以较容易地更新和扩展,以包含新的知识和信息;大模型可以借助知识图谱进行持续学习,实时更新模型的知识库,保持模型的时效性。与此同时,大模型能够捕捉大量文本数据中的丰富语义信息,有很强的泛化能力,能够帮助知识图谱进行图谱的自动化抽取、生成和补全,大大提高了知识图谱的生产准确率和效率。两者的优势如图 1 所示。

知识图谱和大模型的协同包括两种方式,一种是两种技术相对独立,对彼此进行增强;另一种是两种技术进行深度融合。本文聚焦于未来两者更深层次的融合,因此将简要阐述它们如何相互增强,而不过

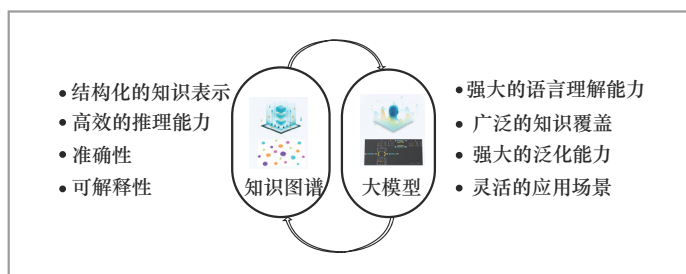


图1 知识图谱与大模型各自的优势

多探讨具体的增强细节。

## 2.1 大模型和图相互增强

### 2.1.1 大模型增强知识图谱

大模型可以对知识图谱进行多方面的增强。大模型具备比知识图谱更加广泛和全面的知识，在知识图谱构建阶段，可以以较高的效率完成实体抽取、关系抽取等关键任务。以重叠和不连续命名实体识别为例，传统方法难以处理重叠和不连续的实体，然而通过大模型的跨度式模型可以同时识别重叠和不连续的实体。Li等<sup>[6]</sup>提出可以首先通过遍历可能的文本片段来识别实体片段，然后进行关系分类以判断片段间的重叠度或连续性，从而实现更精确的实体识别。大模型可以对知识图谱中缺失的实体、关系、属性进行有效的补全，以及对实体、关系、属性进行分类等，大大减少过往知识图谱构建成本昂贵、效率低下等问题。KG-BERT<sup>[7]</sup>提出了使用预训练语言模型进行知识图谱补全的框架。该方法将知识图谱中的三元组（实体及其关系）视为文本序列，利用BERT模型来对这些三元组进行评分，从而实现三元组分类、链接预测和关系预测。实验结果表明，KG-BERT在多项基准知识图谱任务（如链接预测）中取得了最先进的性能，极大地改善了传统的知识图谱补全方法。

此外，大模型通过语义消歧的能力增强了知识图谱的构建。例如，跨文档语言建模<sup>[8]</sup>引入了跨文档学习的预训练方法，通过在多个相关文档上进行预训练，学习文档之间的关系。这一机制使大模型能够在长文本或多文档场景下更好地建模语义上下文，从而提高了知识图谱中实体的消歧能力。同样，BERT<sup>[9]</sup>在共指消解任务中的表现也得到了提升，尤其是在区分相似

但不同的实体（如总裁和首席执行官）方面效果显著。

### 2.1.2 知识图谱增强大模型

知识图谱在多个方面显著增强大模型的性能，知识图谱通过提供丰富的结构化知识，提升了大模型对复杂任务的理解和处理能力。知识图谱可以在大模型的预训练阶段作为目标注入，大模型也可以在输入层结合知识图谱的内容。

在大模型的预训练阶段，知识图谱可以作为大模型的训练目标，帮助增强大模型对结构化知识的理解。例如，GLM<sup>[10]</sup>采用了一种自回归空白填充的预训练方式，并通过引入二维位置编码，优化了大模型在自然语言理解任务中的表现。GLM通过不同的任务设计（如空白预测）将结构化的知识信息融入预训练过程，使其在多种任务上超越了BERT的表现。SKEP<sup>[11]</sup>针对情感分析任务，提出了一种将情感知识（如情感词汇和情感关联对）融入预训练的方式。通过构建情感知识预测目标，SKEP在多个情感分析任务中嵌入了情感信息，显著提升了模型在情感分类、极性预测等任务中的表现。

除了在预训练阶段融入知识图谱，知识图谱还可以在输入层引入外部知识或结构化信息以增强大模型的推理能力。例如，ERNIE 3.0<sup>[12]</sup>是一种知识增强的预训练框架，结合了自回归和自编码网络结构。ERNIE 3.0将知识图谱中的结构化知识与文本信息相结合，通过在大规模知识图谱上进行训练，显著提升了模型在自然语言理解和生成任务中的表现。通过这种方式，大模型在不需要重新训练的情况下，能够有效利用知识图谱中的结构化信息。K-BERT<sup>[13]</sup>通过将知识图谱中的三元组（实体及其关系）注入输入文本中，增强了大

模型对领域知识的理解。该模型利用软位置编码和可见矩阵来限制知识噪声,使大模型能够在不失去语义信息的情况下利用知识图谱中的领域知识。

## 2.2 大模型和图互相融合

在大模型和图的相互增强的过程中,由于两方是独立的技术组件,交互过程中一方往往只能利用另一方的部分能力,增益有限。越来越多的研究将二者进行深度融合,构建更加完善的人工智能系统,从而充分发挥两种技术的优势。目前,图模融合对人工智能系统的能力提升集中在事实表达与逻辑推理两方面。相关技术见表1。

### 2.2.1 人工智能系统事实表达能力提升

知识图谱擅长表达知识的复杂结构,大模型擅长表达文本中丰富的语义信息。两者如能充分融合,则能构建起对事实更加完备的描述与表达。神经网络是表达图的主要技术手段之一,通过节点之间的递归消息传递和聚合机制,在节点、边或图级别为不同类型的下游任务获取表达性表示。虽然图神经网络(graph neural network, GNN)擅长捕获结构信息,但

它们主要依赖于语义约束的嵌入作为节点特征,这在一定程度上限制了其对节点复杂性特征的全面表达能力。而大模型擅长对文本进行编码,但难以捕获图谱中存在的结构信息。两者相结合可以实现更全面、更强大的事实表达能力。笔者将相关研究成果划分为两类,一类是基于GNN提升人工智能系统事实表达能力的研究,另一类是基于非GNN提升人工智能系统事实表达能力的研究。

基于GNN的人工智能系统事实表达能力提升的研究成果如下。ConGraT<sup>[14]</sup>提出了一种通用的自监督的方法,用于对齐大模型和GNN的在潜在空间中的表示。GNN-LM<sup>[15]</sup>是一个先用异质图的信息对大模型进行微调,再与GNN进行联合训练的框架,获取的表达既包括结构和连接信息,又包括语义信息。TAPE<sup>[16]</sup>设计了一个翻译器,将语义信息转换为特征,传递给GNN。GLEM<sup>[17]</sup>提出了一个期望值最大化的框架,在大规模的文本属性图上进行图结构和语义信息的融合。GPEFT<sup>[18]</sup>用GNN将邻居节点中的结构信息编码为一个图提示词,并将该提示插入文本序列的开头。如图2所示,JAKET<sup>[19]</sup>提出一个预训练框架对知识图谱和语言进行建模。知识模块和语言模块生成互补的基本信息,知识模块为文本中的实体生成嵌入,而语言

表1 图模融合相关技术

| 类别             | 子类别        | 相关技术                                                                              |
|----------------|------------|-----------------------------------------------------------------------------------|
| 人工智能系统事实表达能力提升 | 基于图神经网络技术  | ConGraT、GNN-LM、TAPE、GLEM、GPEFT、JAKET                                              |
|                | 基于非图神经网络技术 | ERNIE、GraphBridge、METAG、BERT-MK、CokeBERT、KEPLER、UniGLM、WalkLM、DRAGON、DeltaKG、HKLM |
| 人工智能系统逻辑推理能力提升 | 需要重新训练     | GoT、StructGPT、KD-Cot、KSL、Think-on-Graph、Graph-CoT                                 |
|                | 不需要重新训练    | JointLK、KagNet、MHGRN、GreaseLM、ReasonLM、QA-GNN、Graph Reasoning Transformers        |

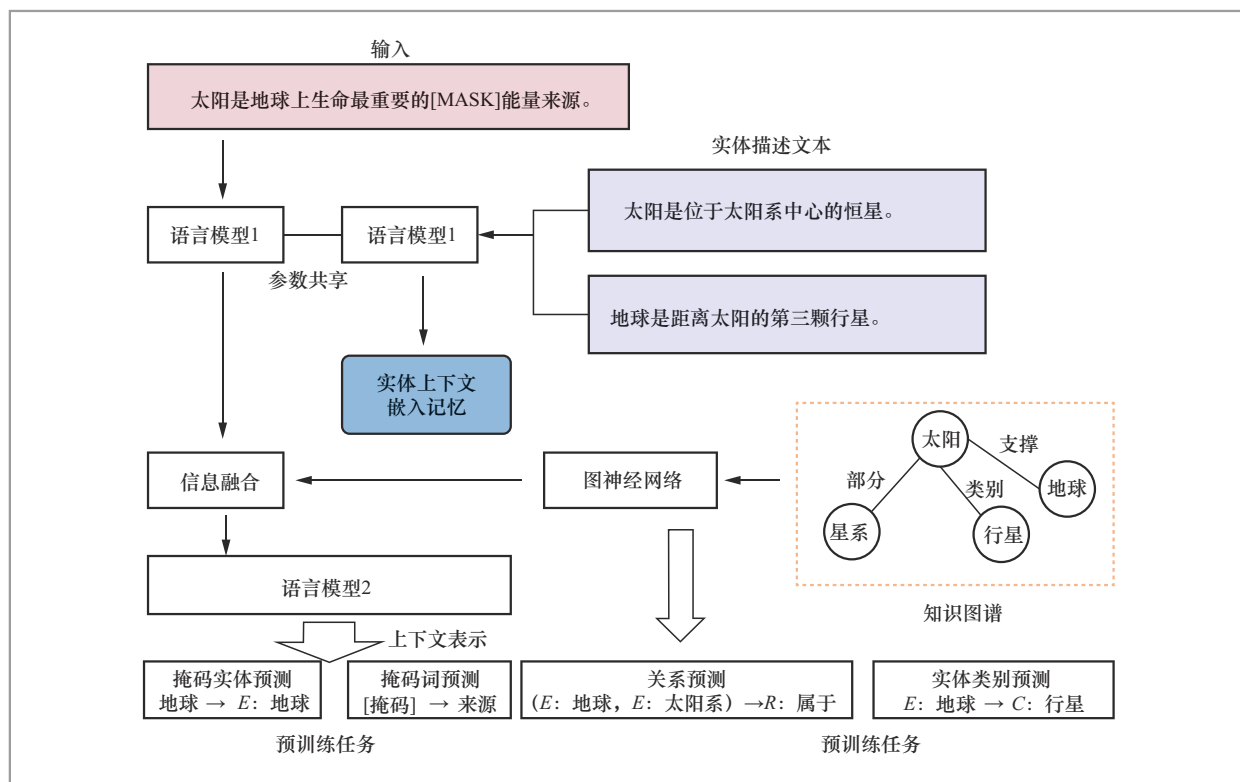


图2 JAKET的整体架构

模块为图中的实体和关系生成上下文感知的初始嵌入。GNN具有良好的表征能力，能够处理复杂的数据结构，与大模型之间具有较强的融合能力，但GNN要进行多轮消息传递和特征更新，计算开销较大，因此在处理的数据量较大时，其需要的计算资源往往会急剧增加。同时，GNN的训练过程较复杂，加大了以上基于GNN的图模融合技术在实际应用中落地的难度。

基于非GNN技术的人工智能系统事实表达能力提升的研究成果如下。ERNIE<sup>[20]</sup>利用大规模文本语料库和知识库来训练一个增强的语言表示模型，这个模型可以同时充分利用词汇、句法和知识信息。GraphBridge<sup>[21]</sup>提出了一个框架，将上下文信息和图结构信息进行联合表达，从而获得对文档的全局性视角。METAG<sup>[22]</sup>利

用一个文本编码器对跨关系的共享知识进行建模，同时通过每个关系的少量参数生成特定于关系的表示。这种设计既能有效捕获图中复杂的多路复用结构，又能在不同关系间共享知识，从而避免了单一嵌入难以表达多种语义关系的问题。BERT-MK<sup>[23]</sup>支持从知识图谱中提取任何子图，并将这些子图输入一个新的基于变换器的模型中以学习知识嵌入，并使用预先训练的语言模型来构建一个包含所学知识表示的模型。CokeBERT<sup>[24]</sup>是一种新的框架，用于根据预训练模型的文本上下文动态选择上下文知识并嵌入知识上下文，避免KG与输入文本不匹配导致的冗余和模糊知识。KEPLER<sup>[25]</sup>使用预训练模型作为嵌入对文本实体描述进行编码，然后共同优化知识表示和语言建模目标，以更好地将

事实知识整合到预训练模型中,并通过预训练模型生成有效的文本增强知识嵌入方法。UniGLM<sup>[26]</sup>可以对多个跨领域的文本属性图进行统一的学习和编码。WalkLM<sup>[27]</sup>将大模型与随机游走技术相结合,获取属性图的结构和语义信息。DeltaKG<sup>[28]</sup>构建了一个对以上的事实表达进行修改和更新的方法,在快速高效更新事实表达的同时不显著影响模型的性能。如图3所示,DRAGON<sup>[29]</sup>是一种联合训练文本和知识图谱,增强模型在知识密集型任务中的表现的自监督方法,将成对的文本片段和相关的子图作为输入,并双向融合两种模态的信息,通过统一两个自监督推理任务(即掩码语言建模和KG链接预测)来预训练该模型。HKLM<sup>[30]</sup>是一个统一的预训练语言模型,使用掩码语言模型目标来学习单词知识,使用三重分类目标和标题匹配目标分别学习实体知识和主题知识,适用于所有形式的文本,包括非结构化文本、半结构化文本和结构化良好的文本。

在以上技术中,GraphBridge、BERT-MK、UniGLM、WalkLM、DRAGON需要将生成知识图谱作为输入,适用于图应用已经比较成熟的领域;ERNIE、METAG、CokeBERT、KEPLER、DeltaKG、HKLM不依赖于已有的知识图谱,适用于更加通用的领域,但是也对训练的资源 and 技巧提出了更高的要求。

## 2.2.2 人工智能系统逻辑推理能力提升

推理是一种认知过程,涉及从已知信息得出新结论的过程。推理是人类智能的核心特征之一,也是逻辑学和哲学研究的重要内容。在人工智能领域,推理是构建智能系统的重要部分,尤其是在机器学习、自然语言处理和专家系统中,机器通过推理来模拟人类的思考过程,处理复杂的任务和问题。常用的大模型的推理技术包括预训练、微调、对齐等。常用的知识图谱推理技术包括符号推理、路径发现等。大

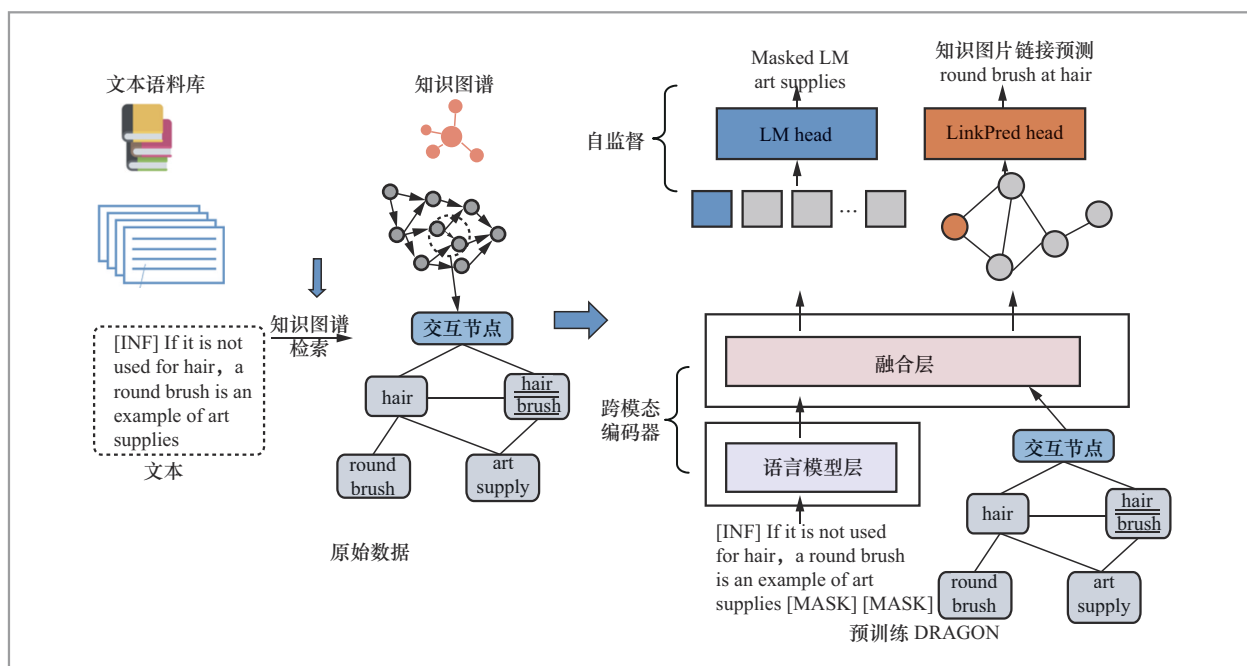


图3 DRAGON的主要架构

模型推理往往存在幻觉问题,知识图谱凭借其结构化的知识表示和精确的逻辑推理能力,为实现高效、可解释的推理提供了重要途径。然而,传统的知识图谱推理方法往往受限于知识图谱的覆盖范围和推理规则的表达能力,难以处理复杂的现实问题,例如用户多样化的自然语言查询、隐含的语义关系等。两者的融合可以有效地提升人工智能系统整体的推理能力。利用大模型的优势来增强知识图谱的推理能力,在提高推理的泛化性、灵活性和可解释性等方面具有重要意义。近年来,一系列研究工作尝试将大模型和知识图谱推理技术进行融合,提出一系列图上推理或图模联合推理等技术,并取得了不错的效果。笔者根据技术实现的难易程度以及对算力的消耗情况,将其分为不需要重新训练模型和需要重新训练模型两类。

不需要重新训练的人工智能系统逻辑推理能力提升的研究成果如下。思维图(graph of thought, GoT)<sup>[31]</sup>将大模型生

成的信息建模成图,信息单元作为节点,信息单元之间的关联作为边,通过不断变换图的结构生成新的信息聚合。StructGPT<sup>[32]</sup>是一种迭代式读取然后推理(reading-then-reasoning)框架,完成多种结构化数据(包括知识图谱)的问题回答任务,对结构化数据的访问接口进行迭代式的调用和评价。KD-Cot<sup>[33]</sup>是一种新的方法,在多个文档上创建一个知识图谱,其中节点象征着段落或文档结构,边表示段落之间或文档结构关系之间的语义/词汇相似性,作者设计了一个基于大模型的图遍历代理充当本地导航器,收集相关上下文以逐步解决问题并保证检索质量。KSL<sup>[34]</sup>提出了一种范式,称为知识求解器,设计了一个简单而有效的提示,将检索转换为多跳决策序列,使大模型能够以零样本的方式搜索知识,并提供了完整的检索路径,从而提高大模型推理过程的可解释性。如图4所示,Think-on-Graph<sup>[35]</sup>提出一种方法,使大模型迭代地在知识图谱

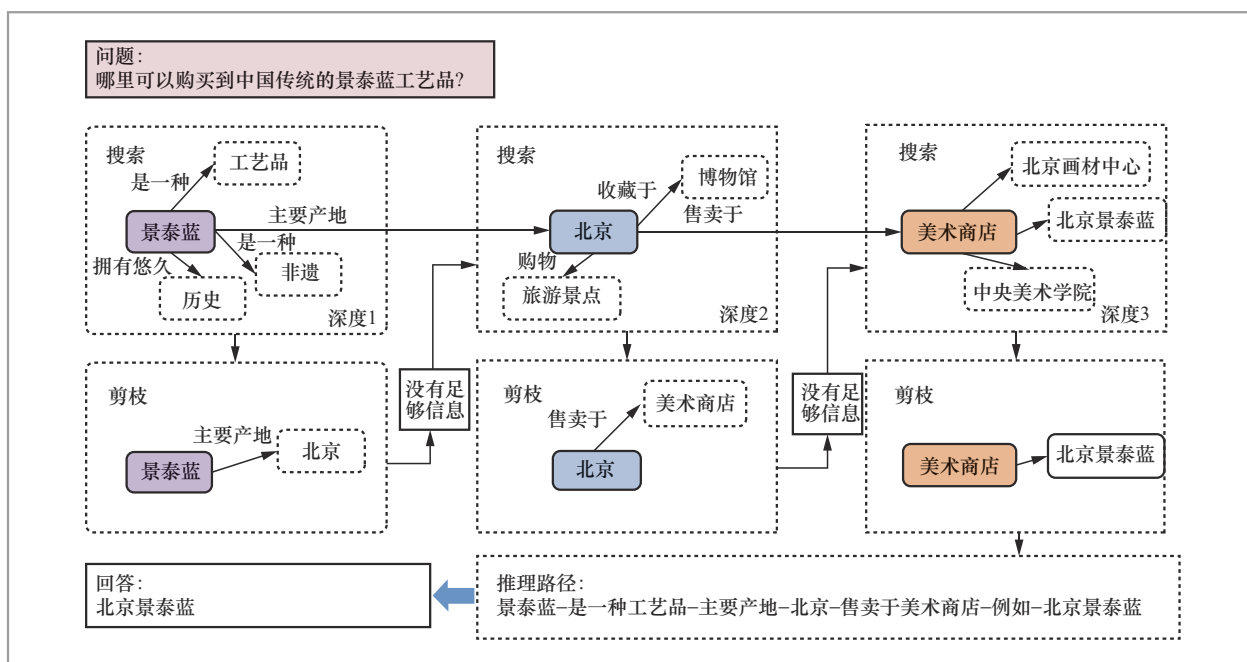


图4 ToG的执行示例

上执行波束搜索,发现最有希望的推理路径,从而返回最可能的推理结果。Graph-CoT<sup>[5]</sup>提出了一个图上推理的评测数据集 GRBench,同时给出一个框架,帮助大模型在图上进行迭代式推理,包括大模型推理、图模交互、图上执行3个阶段。以上技术的共同特点是将大模型的推理问题转换为图上的检索和跳转问题,不断迭代直至收敛,有效地增强了推理过程的逻辑性和可解释性。在此过程中,多轮迭代的效率和最终的收敛条件成为增强大模型与知识图谱融合效果的核心挑战。

需要创新训练的人工智能系统逻辑推理能力提升的研究成果如下。如图5所示,JointLK<sup>[36]</sup>中有一个双向注意力的模块,每个问题令牌都关注知识图谱节点,每一个知识图谱的节点也都关注问题令牌,两种模态的表示通过多步交互来进行融合和更新。KagNet<sup>[37]</sup>是一种用于回答常识问题的文本推理框架,将问答对从语义空间扩展到基于知识的符号空间,再用一种新型

知识感知图网络模块表示模式图,最后用图表示对答案进行评分。MHGRN<sup>[38]</sup>是一种新的知识感知方法,为预训练语言模型配备了一个多跳关系推理模块,从外部知识图谱中提取的子图进行多跳、多关系推理,将基于路径的推理方法与图神经网络相结合,以实现更好的可解释性和可扩展性。GreaseLM<sup>[39]</sup>是一种新的模型,在多层模态交互操作中融合了来自预训练语言模型和图神经网络的编码表示,允许语言上下文表示以结构化的世界知识为基础,并允许上下文中的语言细微差别(如否定、对冲)为知识图谱提供信息。ReasonLM<sup>[40]</sup>基于一种子图感知的自注意力机制来模仿GNN进行结构化的推理,并合成了20 000个子图及相应的问题对大模型的参数进行适应性调整。QA-GNN<sup>[41]</sup>使用大模型来估计大模型节点相对于给定QA上下文的重要性,给出相关性评分,并将QA上下文和KG连接起来形成一个联合图,并通过GNN更新它们的表示。

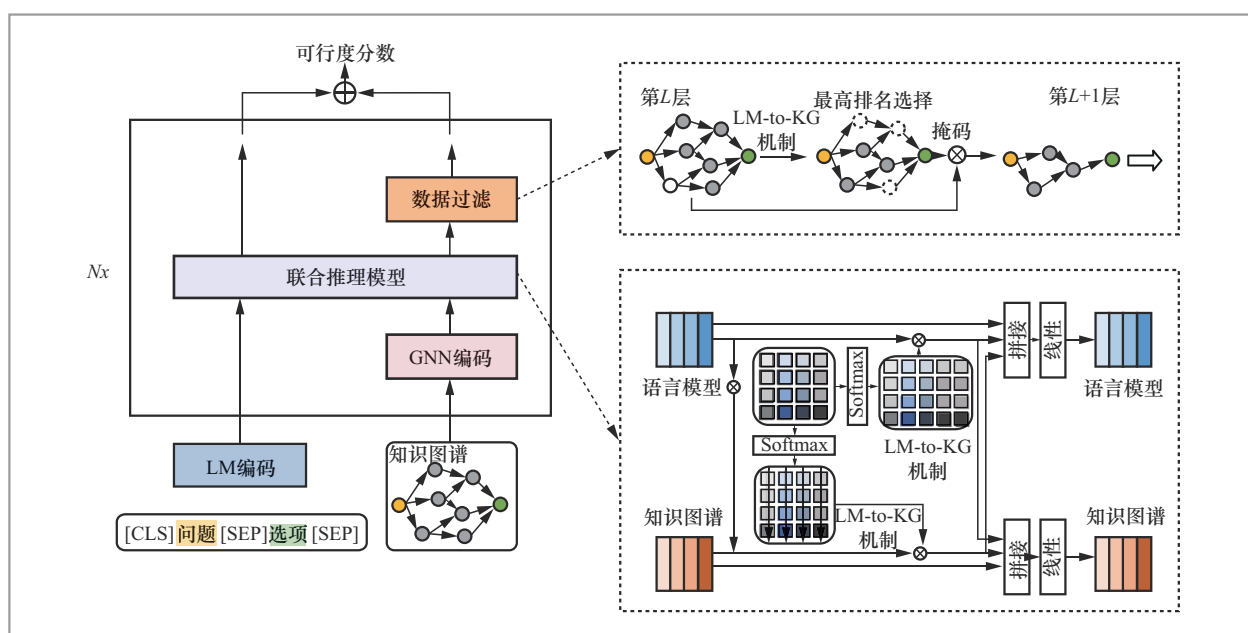


图5 JointLK的整体架构

Graph Reasoning Transformers<sup>[42]</sup>使用一个三元组级别的图编码器来捕捉图特征，然后构建了一个跨模态的信息混合层，使用注意力机制来进行文本信息和图信息的充分混合。进行重新训练的图模融合方式可以对图和大模型中的知识进行充分的细粒度混合，具有良好的理论效果，但是在实际应用中其对于算力和算法提出了较高要求，同时面临知识更新方面的挑战。

### 3 图模融合的产业界探索和实践

随着人工智能技术的迅猛发展，图模融合不仅在学术界取得了丰硕的研究成果，同时也在产业界引发了广泛的关注与应用。各大企业和技术公司逐渐认识到，将知识图谱的结构化知识与大模型的强大生成能力相结合能够为多种实际场景提供更智能化、精准化的解决方案。本节将重点介绍产业界在图模融合领域的技术探索与实践，分析其在实际应用中的成功案例。

#### 3.1 产业界图模融合的技术路径探索

产业界图模融合技术的探索主要分为两大路径：一类是以大模型为主体，将知识图谱作为增强模块融入大模型中；另一类则是以知识图谱为主体，大模型作为知识图谱应用的一部分进行协同工作。以下将重点介绍以知识图谱为主的图模融合技术在工业界的应用实践。

##### 3.1.1 大模型驱动的图模融合技术探索

百度提出的 ERNIE 3.0 是一个多范式统一的大规模预训练框架，旨在克服传统预训练模型无法充分利用常识知识和世界知识的局限性。与仅基于纯文本数据进行

预训练的传统大模型不同，ERNIE 3.0 通过引入知识图谱类数据，将知识与语言模型有机结合起来，从而提升了模型在多种任务上的表现，尤其是在常识推理和复杂关系推理任务中的能力。在知识图谱与模型的融合方面，ERNIE 3.0 采用了一种词-实体对齐的预训练目标。ERNIE 3.0 将文本中的句子和对应的实体同时输入大模型中，并训练模型预测文本词元与知识图谱中实体之间的对齐关系。这种对齐机制使模型能够在语言理解的过程中，将文本内容与外部知识图谱中的实体信息关联起来，从而增强模型对常识性和结构化知识的感知能力。

智谱 AI 是一家由清华大学计算机科学与技术系的技术成果转化而来的公司，致力于构建新一代认知智能通用模型。在其产品研发中，智谱 AI 成功实现了超大规模预训练模型和高精度知识图谱的深度融合，形成了数据与知识双轮驱动的认知引擎。这一融合大大增强了模型的推理和生成能力，尤其是在复杂的认知任务中展现了优异的表现。智谱 AI 的代表性模型 GLM-130B 是一个双语千亿级预训练模型，与高精度知识图谱协同工作。通过引入知识图谱，模型能够在文本生成和理解任务中更加准确地利用背景知识，提升对复杂语义和逻辑关系的处理能力。基于这一融合技术，智谱 AI 推出了一个智能对话系统 ChatGLM，其能够为用户提供自然、上下文相关且知识丰富的交互体验。

OpenAI 最新发布的 OpenAI o1 大幅提升了大模型的推理能力，尽管其具体实现方法尚未公开，但根据部分公开资料推测，极有可能是通过强化学习技术与蒙特卡罗树搜索技术相结合，将链式推理（chain of thought, CoT）能力内化。相比 CoT 技术，GoT 技术更适合处理复杂的逻辑推

理问题,因为它能够更好地处理多步推理和复杂的关系网络。因此,未来知识图谱的推理能力很有可能进一步被吸纳进产业界的大模型中,以增强大模型的逻辑推理能力。

微软提出了 GraphRAG 框架,该框架旨在通过知识图谱增强大模型的推理和生成能力。GraphRAG 不仅将知识图谱作为外部辅助系统来丰富大模型的语义理解,还通过图结构的组织形式提升生成内容的准确性和一致性。通过这一框架,知识图谱成为大模型推理过程中的重要参考,不仅提高了生成的质量,还降低了大模型生成错误或无关内容的可能性。

### 3.1.2 知识图谱驱动的图模融合技术创新

数据库厂商如 Neo4j、TigerGraph 和 PingCap,以及知识图谱技术服务商如海致星图和 Palantir,正积极探索以知识图谱为核心的图模融合技术。这些公司通过将大模型的能力整合到知识图谱的构建与应用中,提升了知识查询、复杂推理和智能生成的效率与精度。

具体而言,这些系统首先利用大模型来构建企业内部的知识图谱。大模型从非结构化的数据源(如文本、图像、表格等)中抽取出实体和关系,并将其与企业现有的结构化知识图谱进行融合。在这一融合过程中,大模型技术还被用于处理实体消歧、指代消解以及实体链接等任务,以确保知识图谱的一致性和准确性。

在构建出企业内部的知识图谱之后,系统通过算法模块结合大模型与知识图谱来回答用户的问题。具体操作流程是,系统根据用户的输入在知识图谱中搜索相关资料,并将这些资料返回给大模型。大模型再基于返回的知识,生成更丰富的答案,

或决定是否需要进一步在知识图谱中进行检索,从而确保回答的全面性和准确性。一些产品,例如 Neo4j 和 AtlasGraph,不仅在图数据库中存储知识图谱,还存储向量数据,使系统可以通过模糊搜索查找相关信息,进一步增强搜索的灵活性和效率。

在知识图谱和大模型的融合搜索之上,相关系统通常还提供智能体应用的编排功能。这使用户可以根据不同的业务需求,构建灵活的检索流程和应用逻辑。最终,用户获得的是一个基于知识图谱与大模型融合的应用构建平台,能够满足复杂的知识检索和推理需求。

上述以知识图谱为主的图模融合具备可解释性强、图谱的数据模型贴近实际应用场景、技术路径成熟等特点。因此,这一融合技术已成为大模型应用落地的重要方向,在多个领域展现出广泛的应用潜力。

### 3.2 图模融合的产业界实践案例

本文以智能审计场景为例,提出了一种基于图模融合技术的智能审计方案生成系统,如图 6 所示。该系统的核心任务是根据审计规章制度与审计目标的业务指标,自动生成适配的审计方案。审计规章制度数据为非结构化数据,而审计目标的业务指标数据被构建为结构化的知识图谱。因此,生成审计方案时,需要同时结合非结构化的审计规章制度与结构化的业务指标,综合分析两者提供的信息,才能生成符合实际需求的智能审计方案。

智能审计系统的构建过程涉及多个数据源和步骤。审计方法、审计制度、报告模板、过往审计资料等均属于非结构化数据,而来自业务指标的具体明细则被构建为结构化的知识图谱。为了整合这些异构数据并生成符合业务需求的审计方案,本

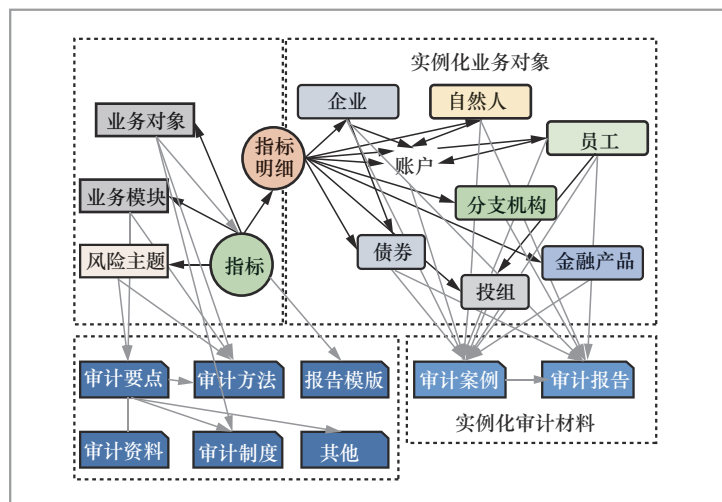


图6 智能审计方案生成系统

文提出了基于大模型与知识图谱融合的技术方案，具体包括以下6个步骤。

- 指标接入：将构建好的结构化业务指标（如企业财务数据、风险主题、业务模块等）接入系统。这些结构化数据为后续的审计方案生成提供了基础的分析维度。

- 文档接入：将审计规章制度、审计方法、报告模板等非结构化文档接入系统。这些文档包含了生成审计方案所需的法律法规、行业标准以及过往的审计经验。

- 图谱融合：利用大模型对接入的非结构化规章制度文本进行语义解析，自动抽取关键实体（如审计方法、审计要点），并将其与结构化的业务指标知识图谱进行融合，形成统一的图谱表示。这一步骤确保了非结构化和结构化数据的有机结合。

- 图谱检索：系统根据用户的审计需求，结合业务对象（如企业、金融产品、员工等）进行图谱检索。大模型通过解析用户需求，提取出相关的关键词，并在图谱中检索匹配的审计规章和业务指标。

- 内容召回：根据图谱检索的结果，系统从规章制度和业务指标中召回与审计目标相关的内容片段。例如，系统可能会

返回适用于某类企业的审计方法、报告模板，以及对应的风险评估指标。

- 智能生成审计方案：大模型基于召回的内容，生成一个符合用户需求的审计方案。该方案不仅包括审计流程和方法，还结合了历史审计经验、业务指标及合规要求，以确保方案的全面性和准确性。

在上述场景中，知识图谱与大模型的融合实现了高效的知识管理与智能生成。大模型从非结构化文档中自动抽取信息，构建结构化的知识图谱，提供了精准的检索与推理支持。知识图谱作为基础信息载体，增强了系统在复杂场景中的处理能力。与此同时，大模型依托其强大的自然语言处理和生成能力，基于知识图谱的查询结果，生成准确、上下文相关的方案。最终，这种融合技术有效满足了规章制度复杂性和多维度数据需求，显著提高了智能审计方案的准确度和丰富度。

## 4 未来展望

知识图谱和大模型作为各自领域的尖端技术，二者融合是人工智能发展的重要趋势。现实世界充满了复杂的关系网络，而知识图谱结构为这些关系提供了有效的表征方式。将知识图谱与大模型相结合，能够显著增强大模型的理解、推理和生成能力，从而促进其在更广泛的实际应用中落地。然而，尽管图谱融合具有广阔的前景，仍存在诸多挑战和机遇。

### （1）增强模型对知识图谱的理解能力

当前，大模型在处理复杂的知识图谱结构，特别是包含丰富语义信息的图谱时，其理解能力依然有限。例如，GPT4Graph一文中<sup>[43]</sup>研究显示，大模型在知识图谱的理解测试中的表现还需提高。因此，未来的

研究方向应包括探索更有效的模型架构和训练策略。例如知识图谱神经网络与大模型的深度融合,将图神经网络作为大模型的重要组成部分,在预训练和微调阶段引入知识图谱信息,以增强模型的图谱表征能力。基于知识图谱的预训练任务设计开发专门用于知识图谱的预训练任务,如链接预测、节点分类等,使模型能够更好地捕捉图谱结构和语义信息。

(2) 提升模型对复杂知识图谱的构建能力

大模型可以利用其强大的文本处理能力,从海量文本数据中自动抽取实体、关系等信息,进而构建大规模知识图谱。未来需要进一步关注如下方面:基于大模型的知识抽取方法,如利用大模型的语义理解能力提高实体识别、关系抽取等任务的准确率和效率;知识图谱的自动构建和更新,如探索利用大模型自动构建和动态更新知识图谱的方法,以适应不断变化的现实世界。

(3) 发展模型对复杂知识图谱的知识表示和推理能力

大模型需要具备对复杂知识图谱进行逻辑推理、路径查找、知识发现等的高级认知能力。未来需要关注:图神经网络与符号推理的结合,如将图神经网络的强大表征能力与符号推理的逻辑性相结合,实现更精准和可解释的图推理;基于大模型的知识问答和推理,开发基于大模型的知识问答系统,使其能够理解复杂问题,并在知识图谱上进行推理,给出准确答案。

(4) 探索多模态大模型与知识图谱的融合

现实世界的信息往往以多模态的形式存在,如文本、图像、视频等。未来,需要研究如何将多模态信息融入知识图谱和大模型中。例如,多模态知识图谱的构建

和应用,构建包含多模态信息知识图谱,并研究其在大模型中的应用;基于图的多模态推理和生成,利用图结构关联不同模态的信息,实现跨模态推理和生成。

(5) 应用侧的落地展望

随着上述技术的发展,知识图谱与大模型的融合将在各个领域催生出颠覆性的应用,如更智能的AI智能体。融入图谱推理能力的AI智能体将具备更强的逻辑思维和决策能力,在虚拟助理、游戏AI等领域发挥重要作用。神经符号推理与图推理的结合利用神经符号推理技术自动生成图推理代码,可以处理更加复杂的问题,在图谱深度分析等领域具有巨大潜力。

经过数十年的发展,知识图谱的应用在各行各业已有了广泛的基础。当前,大模型作为新兴技术,正在向各行各业快速渗透。两者作为人工智能的不同流派,具有各自的技术特点。如何将其充分地进行融合,解决人工智能在各个领域落地应用的实际问题,已经成为学术领域的焦点之一。

参考文献:

[1] JI S X, PAN S R, CAMBRIA E, et al. A survey on knowledge graphs: representation, acquisition, and applications[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(2): 494-514.

[2] HOGAN A, BLOMQVIST E, COCHEZ M, et al. Knowledge graphs[J]. ACM Computing Surveys (Csur), 2021, 54(4): 1-37.

[3] MINAEES, MIKOLOV T, NIKZADN, et al. Large language models: a survey[EB]. arXiv preprint, 2024, arXiv: 2402.06196.

[4] SHEN T, MAO Y, HE P C, et al. Ex -

- exploiting structured knowledge in text via graph-guided representation learning[EB]. arXiv preprint, 2020, arXiv: 2004.14224.
- [5] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2017: 6000–6010.
- [6] LI F, LIN Z C, ZHANG M S, et al. A span-based model for joint overlapped and discontinuous named entity recognition[EB]. arXiv preprint, 2021, arXiv: 2106.14373.
- [7] YAO L, MAO C S, LUO Y. KG-BERT: BERT for knowledge graph completion[EB]. arXiv preprint, 2019, arXiv: 1909.03193.
- [8] CACIULARU A, COHAN A, BELTAGY I, et al. CDLM: cross-document language modeling[EB]. arXiv preprint, 2021, arXiv: 2101.00406.
- [9] JOSHI M, LEVY O, WELD D S, et al. BERT for coreference resolution: baselines and analysis[EB]. arXiv preprint, 2019, arXiv: 1908.09091.
- [10] DU Z X, QIAN Y J, LIU X, et al. GLM: general language model pretraining with autoregressive blank infilling[EB]. arXiv preprint, 2021, arXiv: 2103.10360.
- [11] TIAN H, GAO C, XIAO X Y, et al. SKPE: sentiment knowledge enhanced pre-training for sentiment analysis[EB]. arXiv preprint, 2020, arXiv: 2005.05635.
- [12] SUN Y, WANG S H, FENG S K, et al. ERNIE 3.0: large-scale knowledge enhanced pre-training for language understanding and generation[EB]. arXiv preprint, 2021, arXiv: 2107.02137.
- [13] LIU W J, ZHOU P, ZHAO Z, et al. K-bert: Enabling language representation with knowledge graph[C]//Proceedings of the AAAI Conference on Artificial Intelligence. California: AAAI Press, 2020: 2901–2908.
- [14] BRANNON W, KANG W, FULAY S, et al. ConGraT: self-supervised contrastive pretraining for joint graph and text embeddings[EB]. arXiv preprint, 2023, arXiv: 2305.14321.
- [15] IOANNIDIS V N, SONG X, ZHENG D, et al. Efficient and effective training of language and graph neural network models[EB]. arXiv preprint, 2022, arXiv: 2206.10781.
- [16] HE X X, BRESSON X, LAURENT T, et al. Harnessing explanations: LLM-to-LM interpreter for enhanced text-attributed graph representation learning[EB]. arXiv preprint, 2023, arXiv: 2305.19523.
- [17] ZHAO J N, QU M, LI C Z, et al. Learning on large-scale text-attributed graphs via variational inference[EB]. arXiv preprint, 2022, arXiv: 2210.14709.
- [18] ZHU Q, ZHENG D, SONG X, et al. Parameter-efficient tuning large language models for graph representation learning[EB]. arXiv preprint, 2024, arXiv: 2404.18271.
- [19] YU D, ZHU C, YANG Y, et al. Jaket: joint pre-training of knowledge graph and language understanding[C]//Proceedings of the AAAI Conference on Artificial Intelligence. California: AAAI Press, 2022: 11630–11638.
- [20] ZHANG Z Y, HAN X, LIU Z Y, et al. ERNIE: enhanced language representation with informative entities[EB]. arXiv preprint, 2019, arXiv: 1905.07129.
- [21] WANG Y K, ZHU Y, ZHANG W Q, et al. Bridging local details and global context in text-attributed graphs[EB]. arXiv preprint, 2024, arXiv: 2406.12608.
- [22] JIN B W, ZHANG W T, ZHANG Y, et al. Learning multiplex representations

- on text-attributed graphs with one language model encoder[EB]. arXiv preprint, 2023, arXiv: 2310.06684.
- [23] HE B, ZHOU D, XIAO J H, et al. Integrating graph contextualized knowledge into pre-trained language models[EB]. arXiv preprint, 2019, arXiv: 1912.00147.
- [24] SU Y S, HAN X, ZHANG Z Y, et al. CokeBERT: Contextual knowledge selection and embedding towards enhanced pre-trained language models[J]. AI Open, 2021, 2: 127-134.
- [25] WANG X Z, GAO T Y, ZHU Z C, et al. KEPLER: a unified model for knowledge embedding and pre-trained language representation[J]. Transactions of the Association for Computational Linguistics, 2021, 9: 176-194.
- [26] FANG Y, FAN D Z, DING S R, et al. UniGLM: training one unified language model for text-attributed graphs[EB]. arXiv preprint, 2024, arXiv: 2406.12052.
- [27] TAN Y, ZHOU Z, LV H, et al. WalkLM: a uniform language model fine-tuning framework for attributed graph embedding[J]. Advances in Neural Information Processing Systems, 2023, 36: 13308-13325.
- [28] CHENG S Y, ZHANG N Y, TIAN B Z, et al. Editing language model-based knowledge graph embeddings[C]//Proceedings of the AAAI Conference on Artificial Intelligence. California: AAAI, 2024: 17835-17843.
- [29] YASUNAGA M, BOSSELU T A, REN H, et al. Deep bidirectional language-knowledge graph pretraining[C]//Proceedings of the 36th International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2024: 37309-37323.
- [30] ZHU H Y, PENG H, LYU Z H, et al. Pre-training language model incorporating domain-specific heterogeneous knowledge into a unified representation[J]. Expert Systems with Applications, 2023, 215: 119369.
- [31] BESTA M, BLACH N, KUBICEK A, et al. Graph of thoughts: Solving elaborate problems with large language models[C]//Proceedings of the AAAI Conference on Artificial Intelligence. California: AAAI Press, 2024, 38(16): 17682-17690.
- [32] JIANG J H, ZHOU K, DONG Z C, et al. StructGPT: a general framework for large language model to reason over structured data[EB]. arXiv preprint, 2023, arXiv: 2305.09645.
- [33] WANG Y, LIPKA N, ROSSI R A, et al. Knowledge graph prompting for multi-document question answering[C]//Proceedings of the AAAI Conference on Artificial Intelligence. California: AAAI Press, 2024, 38(17): 19206-19214.
- [34] FENG C, ZHANG X Y, FEI Z C. Knowledge solver: teaching LLMs to search for domain knowledge from knowledge graphs[EB]. arXiv preprint, 2023: 2309.03118.
- [35] SUN J S, XU C J, TANG L, et al. Think-on-graph: deep and responsible reasoning of large language model on knowledge graph[EB]. arXiv preprint, 2023, arXiv: 2307.07697.
- [36] SUN Y Q, SHI Q, QI L, et al. JointLK: joint reasoning with language models and knowledge graphs for commonsense question answering[EB]. arXiv preprint, 2021, arXiv: 2112.02732.
- [37] LIN B Y, CHEN X Y, CHEN J, et al. KagNet: knowledge-aware graph networks for commonsense reasoning[EB]. arXiv preprint, 2019, arXiv: 1909.02151.

[38] FENG Y L, CHEN X Y, LIN B Y, et al. Scalable multi-hop relational reasoning for knowledge-aware question answering [EB]. arXiv preprint, 2020, arXiv: 2005.00646.

[39] ZHANG X K, BOSSELUT A, YASUNAGA M, et al. GreaseLM: graph REASoning enhanced language models for question answering[EB]. arXiv preprint, 2022, arXiv: 2201.08860.

[40] JIANG J H, ZHOU K, ZHAO W X, et al. ReasoningLM: enabling structural sub-graph reasoning in pre-trained language models for question answering over knowledge graph[EB]. arXiv preprint, 2023: 2401.00158.

[41] YASUNAGA M, REN H Y, BOSSELUT A, et al. QA-GNN: reasoning with lan-guage models and knowledge graphs for question answering[EB]. arXiv preprint, 2021, arXiv: 2104.06378.

[42] ZHAO R, ZHAO F, HU L, et al. Graph reasoning transformers for knowledge-aware question answering[C]//Proceed-ings of the AAAI Conference on Artifi-cial Intelligence. California: AAAI Press, 2024, 38(17): 19652-19660.

[43] GUO J Y, DU L, LIU H Y, et al. GPT4 Graph: can large language models under-stand graph structured data? an empirical evaluation and benchmarking[EB]. arXiv preprint, 2023, arXiv: 2305.15066.

作者简介



杨娟（1979-），女，北京海致星图科技有限公司CEO，清华大学计算机科学与技术系博士生，主要研究方向为知识图谱、图数据库、数据分析与数据挖掘。



沈游人（1993-），男，北京海致星图科技有限公司图技术专家，主要研究方向为图数据库、数据分析与数据挖掘。

收稿日期: 2024-09-01  
通信作者: 杨娟, yangjuan@stargraph.cn  
基金项目: 国家重点研发计划基金项目(No.2022YFB2404200)  
Foundation Item: National Key Research and Development Program of China(No.2022YFB2404200)