

图模互补：知识图谱与大模型融合综述

黄 勃¹, 吴申奥¹, 王文广^{2†}, 杨 勇³, 刘 进⁴, 张振华¹, 陈南希⁵, 杨洪山⁶

1. 上海工程技术大学 电子电气工程学院, 上海 201620;
2. 达观数据有限公司, 上海 201203;
3. 上海宝信软件股份有限公司, 上海 201203;
4. 武汉大学 计算机学院, 湖北 武汉 430072;
5. 中国科学院 上海微系统与信息技术研究所, 上海 200050;
6. 星环信息科技(上海)股份有限公司, 上海 200233

收稿日期: 2024-03-11 † 通信联系人 E-mail: wangwenguang@datagrand.com

基金项目: 科技部科技创新 2030“新一代人工智能”重大项目(2020AAA0109300); 上海地方能力建设项目(21010501500); 上海市科技创新行动计划(21DZ1204900)

第一作者: 黄 勃, 男, 副教授, 现从事人工智能, 软件工程研究。E-mail: huangbosues@sues.edu.cn

摘 要: 大模型(LLM)的兴起在自然语言处理领域引起了广泛关注, 其涌现能力在各个垂直领域(如金融、医疗、教育等)也取得一定进展。然而, 大模型自身面临解释性不足、知识实时性差、生成结果存在虚假信息等诸多挑战。为了应对这些问题, 知识图谱与大模型的融合逐渐成为了研究热点。知识图谱作为一种结构化的知识模型, 其真实性和可靠性, 成为提高大模型解释和推理能力的有力工具。同时大模型具备语义理解能力, 为知识图谱的构建和更新提供了有力支持。因此, 知识图谱和大模型是互补的(本文称为图模互补)。本文系统性地介绍知识图谱与大模型融合的方法, 分别从 1) 大模型增强知识图谱, 2) 知识图谱增强大模型两个角度进行全面回顾和分析。最后, 本文从医学诊断预测和时间知识图谱出发, 介绍图模互补的领域应用, 并讨论图模互补未来发展的方向, 为知识图谱与大模型的进一步研究提供帮助。

关 键 词: 大模型; 知识图谱; 自然语言处理

中图分类号: TP391; TP18

文献标志码: A

文章编号: 1671-8836(2024)04-0397-16

KG-LLM-MCom: A Survey on Integration of Knowledge Graph and Large Language Model

HUANG Bo¹, WU Shen'ao¹, WANG Wenguang^{2†}, YANG Yong³, LIU Jin⁴, ZHANG Zhenhua¹, CHEN Nanxi⁵, YANG Hongshan⁶

1. School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China;
2. DataGrand Inc., Shanghai 201203, China;
3. Shanghai Baosight Software Co. Ltd, Shanghai 201203, China;
4. School of Computer Science, Wuhan University, Wuhan 430072, Hubei, China;
5. Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China;
6. Transwarp Technology Co., Ltd. Shanghai 200233, China

Abstract: The rise of the Large Language Model (LLM) has attracted wide attention in Natural Language Processing, and its emergent ability has also made some progress in various vertical fields (such as finance, healthcare, education, etc.). However, LLM faces numerous challenges, including inadequate interpretation, limited real-time knowledge, and the potential presence of

引用格式: 黄勃, 吴申奥, 王文广, 等. 图模互补: 知识图谱与大模型融合综述[J]. 武汉大学学报(理学版), 2024, 70(4): 397-412. DOI: 10.14188/j.1671-8836.2024.0040.

HUANG Bo, WU Shen'ao, WANG Wenguang, et al. KG-LLM-MCom: A Survey on Integration of Knowledge Graph and Large Language Model [J]. J. Wuhan Univ. (Nat. Sci. Ed.), 2024, 70(4): 397-412. DOI: 10.14188/j.1671-8836.2024.0040(Ch).

false information in generated outcomes. In order to address these problems, integrating Knowledge Graphs (KG) and LLM has gradually become a research hotspot. As a structured knowledge model, KG has become a powerful tool for improving the interpretation and reasoning ability of LLM due to its authenticity and reliability. At the same time, LLM has the ability to understand semantics, which supports the construction and update of knowledge graphs. Therefore, KG and LLM are mutually complementary (KG-LLM-MCom is called in this article). The article provides a systematic introduction to KG and LLM integration methods. It conducts a comprehensive review and analysis from two perspectives: 1) LLM-enhanced KG and 2) KG-enhanced LLM. Finally, the article introduces the field applications of KG-LLM-MCom from the viewpoint of Medical Diagnosis prediction and Temporal Knowledge Graphs. It discusses KG-LLM-MCom's future development direction, which can provide help for further research on KG and LLM.

Key words: large language model; knowledge graph; natural language processing

0 引言

近年来,自然语言领域出现了一项令人瞩目的技术:大模型。大模型(LLM)是指通过超大规模文本数据训练出来的神经网络模型,由预训练语言模型(PLM)发展而来,其特点是模型规模较大,参数通常在数十亿乃至万亿级别。相较于传统的预训练语言模型,大模型在处理复杂任务时表现出特殊的能力(涌现能力^[1]、推理能力^[2])。相关研究表明,大模型不仅能够理解和处理大规模的文本数据^[3],同时还具备上下文学习能力^[4]和领域泛化能力。这使得它们成为各种自然语言下游任务的通用选择,能够轻松进行少样本迁移学习^[5]。如今,AI领域中涌现出许多具有代表性的大模型,例如对话式语言大模型 ChatGPT^[6]、增强推理能力的多模态大模型 GPT-4^[7]等等。这些模型不仅在传统的自然语言处理领域(如搜索引擎^[8]和机器翻译^[9])取得一定进展,还在金融^[10]、医疗^[11]、教育^[12]等各种领域提供有效帮助。大模型对传统的自然语言处理领域带来冲击,促使学者重新思考通用人工智能的可能性^[13]。

目前,虽然大模型引起了广泛的关注,但它仍然面临着诸多挑战,包括模型内部的不可控性、缺乏解释性^[14]、无法保证知识实时性^[5]、语言数据质量的不确定性以及产生幻觉和有毒信息的潜在风险^[15]。为了应对这些挑战,学者们开始思考将知识图谱与大模型融合^[5]。知识图谱(knowledge graph)是一种用于表示和存储知识的网络图形结构,其中节点表示实体,边表示实体之间的关系^[16]。知识图谱以其数据的真实性著称,这一特点可以有效地减轻大模型产生幻觉的问题。例如 KELM 语料库^[17]、通用模型 KGPT^[18]为大模型提供基于知识图谱转化的真实文本信息。大模型作为一个“黑盒模型”其输出结果通常难以解释,而知识图谱内部的结构知识更接近人类认知。因此,知识图谱可以提供一种

解释和推理知识的手段,用于探究大模型内部复杂的工作步骤和推理过程。例如个性化知识库与大模型集成检索框架 KnowledGPT^[19],提高了处理复杂搜索和歧义的能力。此外,知识图谱还可以作为外部检索工具,帮助大模型解决公平、隐私和安全等问题^[20]。

如今知识图谱的规模越来越大,传统的图谱构建、补全技术也面临许多难题,如数据获取、实体识别、知识抽取和实体消歧等^[21]。大规模知识图谱的构建往往需要投入大量的人力、物力和时间成本,且依旧无法保证知识图谱的质量和可用性。而大模型能有效解决这些问题。大模型内部存在海量的知识信息,在处理复杂的文本数据信息时,能够迅速地进行实体识别与抽取,有效应对知识构建和补全的挑战^[22]。此外,链接预测是知识图谱推理和问答的关键步骤,在零样本和少样本学习中,大模型同样能够有效地挖掘实体间的逻辑关系。根据知识图谱和大模型的上述特点,本文认为知识图谱和大模型是相互补充的,称为“图模互补”(KG-LLM-MC),图 1 为知识图谱与大模型的优缺点总结。

知识图谱与大模型融合是一个热门研究领域^[23-25]。文献[25]提出了统一大模型与知识图谱的前瞻性路线图,总结了现有的大模型与知识图谱的先进技术,并讨论了大模型与知识图谱融合的相关挑战和发展方向。其整体路线划分与本文有所差异。本文从作用功能角度进行划分,根据知识图谱和大模型在领域中的地位,将其融合划分为两个不同的类别:大模型增强知识图谱和知识图谱增强大模型;并从增益的效果出发,将每个大类别细分为不同的小类别;最后探究图模互补的领域应用。

1 相关介绍

本节将介绍知识图谱的定义和主要架构,然后

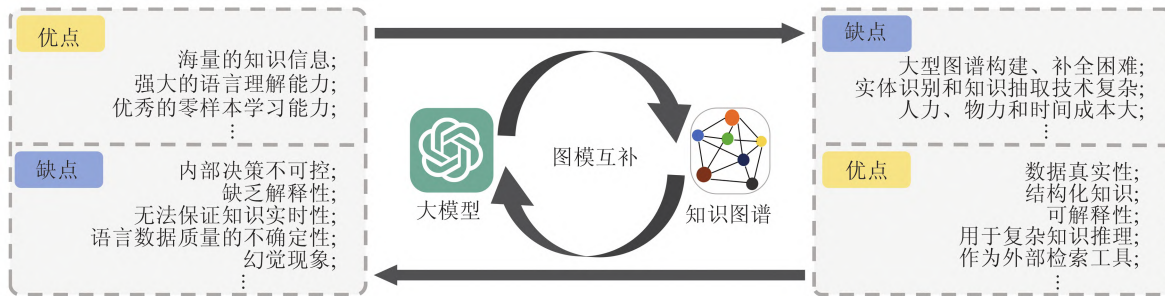


图1 知识图谱与大模型的优缺点总结

Fig. 1 Summary the Pros and Cons of KG and LLM

介绍大模型的概念和发展历程,为后续对知识图谱与大模型的领域融合提供知识基础。

1.1 知识图谱定义

知识图谱是一种用于表示和存储结构化知识的图形知识库^[26],以实体和实体之间的关系构成的三元组(头实体,关系,尾实体)为基本组成单位。知识图谱的思想来源于语义网络^[27],但知识图谱不仅包含实体和关系和属性信息,还包含实体和关系之间的语义关系。知识图谱提供了一种符号化形式来组织表示知识数据,使计算机可以通过事物以及事物之间的内在联系,查询和推理来获取有关特定实体或关系的信息。

知识图谱的形式化定义为 $G=(E,R,S)$, 其中 $E=\{e_1,e_2,e_3,\dots,e_m\}$ 表示实体名的集合,包含头实体和尾实体,数量为 m ; $R=\{r_1,r_2,r_3,\dots,r_n\}$ 表示实体之间的关系集合,数量为 n 。 S 表示事实集合,一个三元组 $(h,r,t)\in$ 一个事实。例如事实 S_1 表示“李明是一名研究生”。在知识图谱中,可以用一个三元组(李明,是,研究生)表示这个事实 S_1 。知识图谱领域存在多个大型开放知识库,如 YAGO^[28]、DBpedia^[29]、Freebase^[30]、Wikidata^[31] 等。

1.2 知识图谱架构

知识图谱的构建是将大量结构化、半结构化或者非结构化的数据通过知识抽取、知识融合、知识推理等操作处理之后加入知识库,构建过程需要反复多次迭代直至补全知识。知识图谱的构建方式包括自顶向下和自底向上,自顶向下构建方式借助现有的结构化知识库,完成知识抽取到图谱构建的过程;自底向上构建方法是从非结构化文本中提取知识,选取置信度较高的三元组,完成知识图谱的构建。目前知识图谱大多采用自底向上的方式构建,如 Google 的 Knowledge Vault^[32]。

知识抽取是从各种数据源中提取有价值的信息,涵盖了结构化和半结构化文本数据、非结构化文本数据等场景。在(半)结构化数据的知识抽取

中,模型通过简单的规则和固定的模式,从结构化数据(如关系数据库)或半结构化数据(如维基百科)中实现知识抽取。而非结构化数据的知识抽取是抽取任务的难点,一般包括实体识别、关系抽取和事件抽取,需从文本中抽取原子信息、实体间的语义关系等。例如生成网络 KEGNET^[33],在没有可观测数据的情况下进行知识抽取。

知识推理是指基于现有的事实和逻辑规则推理新的知识。在知识图谱的构建过程中,往往在知识抽取阶段存在知识缺失、识别错误等问题,需通过逻辑推理、机器学习、深度学习等方法分析知识图谱中的实体和关系,从现有的知识库中自动识别错误并推断出新的知识,从而扩展和丰富知识图谱。早期的知识图谱推理基于一定的规则和限制^[34],需要依赖规则、假设等前提条件。随着机器学习的研究不断深入,神经网络模型 Neulnfer^[35] 从主三元组和辅助描述构成的事实中进行未知元素推理。

1.3 大模型概念

大模型是指具有超大规模参数量和数据量的神经网络模型。大模型通过从大量文本或多模态数据中学习语言和模式识别能力,可以更好地理解和处理自然语言,甚至生成新的文本或内容^[36]。狭义的大模型,即大型语言模型,指具备数十亿乃至万亿参数,通过高达万亿词元数量的文本语料训练出来的深度神经网络模型。大型语言模型在语言理解和生成上取得了可观的成绩,其发展历史可以追溯到 2013 年的 Word2Vec^[37],但直到近年诸如 GPT-3^[38]、LaMDA^[39]、PaLM^[40]、PaLM-2^[41]、LLaMA^[42]、LLaMA^[42]、LLaMA-2^[43]、Claude 2^[44]、WizardMath^[45] 等大模型的出现,才使得大型语言模型的应用得到普及。广义的大模型则包含了文本、声音、图像和视频等多模态任务,其典型代表是 GPT-4^[47]、Gemini^[46] 和 LLaVA^[47],还有 OpenAI 最新发布的大模型 Sora^[48]。大模型的快速发展,得益于充分利用注意力机制^[49]进行序列建模的变换器网络(Trans-

former)架构^[50],以及诸如稀疏变换器网络^[51]的各种模型。如今,变换器网络架构最终统一了语言^[38,45]、视觉^[52]、声音^[53]和多模态^[7,46,54]的建模。大模型通过提示工程(prompt engineering)^[55]实现应用于特定任务上的情境学习,特别是 OpenAI 发布的多模态大模型 GPT-4,其性能表现超过 ChatGPT 等大模型,部分学者将其称为通用人工智能的早期不完善版本^[56]。

1.4 大模型的发展历程

在大模型早期探索阶段,OpenAI 发布两个原始 GPT 模型,GPT-1^[57]和 GPT-2^[58],该团队采用了神经网络架构(Transformer)进行模型构建,希望通过语言模型探索智能系统。GPT-1 的构建主要基于无监督预训练和有监督微调的混合方法,确立 GPT 通过自然语言文本进行建模的基本原则:模型通过预测下一个单词解决自然语言处理(NLP)问题。随着 GPT-2 的发布,数据规模增加到了十亿级别,并完全采用无监督语言建模,省去了使用标注数据进行微调的过程。同时,OpenAI 团队提出将每个自然语言文本任务视为固定化的输入、输出和任务信息三段式结构,GPT 通过分析任务信息对解决方案进行单词预测,不过 GPT-2 在处理下游任务时表现并不理想。随着不断深入的研究,GPT-3 将参数规模增加到千亿级别,这一扩展使得 GPT3 在解决复杂任务方面表现出显著的性能提升,成为大模型领域的重要进展。至此,大模型领域快速发展,如增强推理能力和代码能力的 Codex^[59]、改进人类对齐能力的 InstructGPT(GPT-3.5)^[60]、优化对话能力的 ChatGPT 以及从文本输入扩展到多模态信号的 GPT-4 等模型。

大模型的主流框架主要以 Transformer 为基础。这个架构具有出色的并行性能和扩展能力,使得语言模型能够扩展到数百亿或数千亿个参数。主流架构大致分为三种类型:编码器-解码器、因果解码器和前缀解码器。编码器-解码器架构适用于各种自然语言任务,因果解码器架构采用单向注意力掩码,而前缀解码器架构修正了掩码机制,实现对前缀 token 的双向注意力。这三种架构还能通过混合专家(MoE)扩展来提高性能。

2 大模型增强知识图谱

传统知识图谱的构建往往要面对繁琐的设置、规则解析和多次测试模型。而大模型的出现彻底改变这一格局,通过充分利用会话型大模型的语言能力,

在零样本和少样本的训练中,轻松应对知识图谱的各种挑战^[61]。本节依次介绍大模型增强知识图谱构建和补全、增强推理能力和增强问答能力的模型。

2.1 知识图谱的构建任务

2.1.1 增强知识图谱构建

在知识图谱构建方面,利用大模型零样本或少样本学习的信息提取能力,从文本或其他数据源中完成实体抽取和关系抽取任务^[62]。大模型内部海量的潜在知识,能够有效地完成图谱的构建任务,节约数据标注的时间和成本。例如 ChatIE^[63]将实体抽取、命令实体识别和事件抽取任务转化为多回合问答问题,旨在将复杂的信息抽取任务分解为多个简单的子任务,并设计提示词输入 ChatGPT,在多个信息提取数据集上取得良好效果。同样地,ChatExtract^[64]设计了一种强迁移性的数据提取方法,其核心通过构建一系列用于提取结构化数据的工程提示词,实现大模型在零样本学习的自动识别和提取数据功能,并且 ChatExtract 可以应用于各种对话式大模型,产生高质量的数据提取结果。但是上述的方法不涉及大模型微调任务,且输出结果依赖于大模型自身性能,在一些复杂的自然语言处理任务中,可能无法保证抽取结果的准确性。

在处理知识密集型任务时,为了解决大模型无法捕捉到复杂的关系动态,AutoKG^[65]提出一种轻量级和高效的自动化知识图构建方法。与传统的通过语义相似性搜索方式不同,AutoKG 使用预训练大模型构建简化的知识图谱。AutoKG 构建的知识图与传统的知识图结构不同,它以关键字作为节点,记录各个节点间的相关性权重,AutoKG 会根据关键词之间的关联程度建立知识图中的边。但 AutoKG 仅仅在外部知识库的提取关键词阶段,利用大模型的信息抽取能力,没有在后继丰富知识图阶段结合大模型。同样的,通用大模型在特定领域的实体关系抽取能力还存在不足,而大模型自身可以通过生成标注样本,作为示例或微调大模型方案构建知识图谱。文献[66]探索大模型在医疗保健领域实体抽取的效果,通过大模型生成大量高质量带标签的训练数据,并将其作为提示信息局部微调大模型。实验结果表明,微调大模型相较于通用大模型,性能有较好的提升。

2.1.2 增强知识图谱补全

知识图谱的构建往往是不完备的,传统的知识图补全任务利用知识图谱的结构信息预测三元组中缺失的实体,但是在解决长尾实体时需要付出很

高的代价。大模型可作为额外知识库提取可信知识,完成知识图谱的补全。一种做法是直接将大模型应用在知识图谱补全任务中,将三元组分类、关系预测和实体(链接)预测转化为提示文本,通过提示大模型生成预测结果。例如文献[67]构建了以三元组的实体和关系描述作为提示,在微调大模型 KG-ChatGLM-6B 和 KG-LLaMA(7B 和 13B)上进行实验,在补全任务中取得了较好的性能。而另一种补全方法则是间接利用提示大模型。例如 KIC-GPT^[68]集成大模型与传统的结构感知模型(知识补全检索器),检索器对链接预测任务中缺失三元组进行查询,根据分数排序生成候选实体列表,然后以结构知识编码作为大模型情景学习示例,重新对候选实体列表进行排序。与之相似,文献[69]通过提示大模型,为三元组生成出高质量的上下文描

述,之后借助传统的补全方案训练模型,间接通过大模型补全知识图谱。

然而简单的问答方式无法利用大模型的推理能力,缺乏对知识图谱结构中实体和关系的描述,大模型可能无法理解知识图谱的结构信息。为实现更有效和准确的图谱补全,学者们将知识图谱中的实体、关系等结构信息融入大模型,使大模型具有结构感知推理能力。文献[70]提出了一种名为 KoPA 的知识前缀适配器,将知识图谱结构信息整合到大模型中。KoPA 的整体框架如图 2 所示, KoPA 模型将知识图谱结构信息进行知识嵌入,并投射到文本空间。借助 KoPA 生成的虚拟知识令牌(Token)增强输入提示序列微调大模型,使得大模型能够在结构感知状态下解码指令的答案,提高了大模型在知识图谱补全任务中的性能。

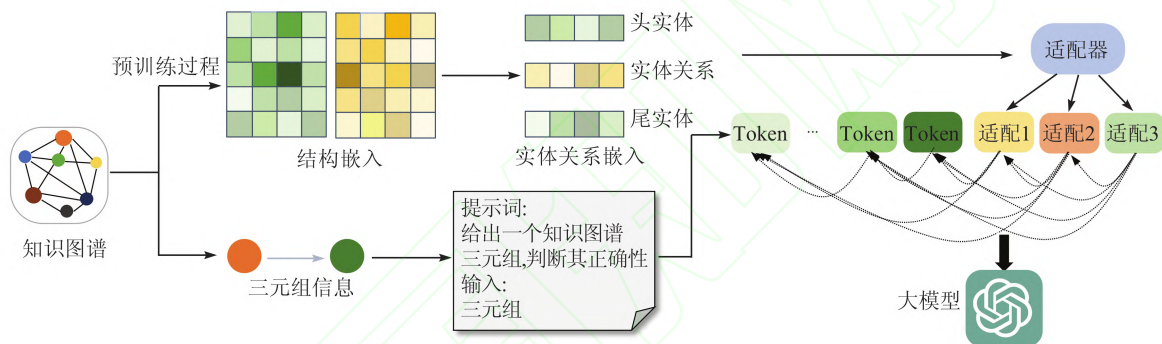


图2 KoPA 概述图^[70]

Fig. 2 An overview of the knowledge prefix adapter (KoPA)^[70]

2.2 增强知识图谱推理

以往的知识图谱推理任务中,研究人员利用知识计算进行显式的推理,借助深度学习模型将实体和关系嵌入到高维向量空间实现推理^[71]。不过该方法依赖于知识图谱自身的知识,缺乏“世界知识”支撑,大模型的到来使得通识知识和知识图谱联合推理成为了可能。LLM-ARK^[72]不访问模型参数直接利用大模型作为代理,将多跳推理任务表示为强化学习序列决策问题,通过全文本环境提示聚合多尺度输入,使大模型获得丰富信息状态的嵌入表示。具体步骤为:LLM-ARK 将知识图谱推理表述为一个由六元组(S, O, A, T, R, γ)描述的马尔可夫决策过程(MDP)。其中, S 描述环境的状态集合; O 用于观察环境的完整状态; A 表示下一步可能发生的行动; T 记录状态的更新; R 表示奖励信号; γ 表示奖励折扣因子,并且引入近端策略优化(PPO)强化学习算法,使模型在推理任务中不断学习。实验表明,模型 LLaMA-2-7B-ARK 在二跳推理任务中表现出优越的性能。LARK^[73]也利用大模型引导知识图

谱的复杂推理,不同的是它将复杂的知识图谱推理转化为上下文知识图谱搜索和逻辑查询推理的两个步骤组合,将多操作复杂逻辑查询分解为大模型更擅长的单操作基本查询,最后将检索到的邻域和分解的查询转换为大模型提示,获取输出结果;同时利用大模型评估最终答案集,LARK 利用图提取算法和大模型双方的优势实现高性能的复杂推理。

在传统演绎推理方向,推理规则挖掘存在资源成本较大、逻辑规则缺乏可扩展性、密集型规则搜索效果不佳等问题,很难扩展到大规模的知识图谱。而大模型能够理解自然语言文本,利用内化到深度学习模型中的语义知识和结构信息生成有意义的规则,结合知识图谱的知识,实现辅助挖掘知识图谱推理的规则及评估方案。例如规则生成器 ChatRule^[74],框架如图3所示。ChatRule 以知识图中的规则实例作为大模型提示,生成一组粗糙的候选规则,设计逻辑规则评分器过滤不合要求的规则,最后采用思维链(CoT)^[75]强化推理能力的大模型验证规则的正确性,消除大模型潜在的虚假规则。

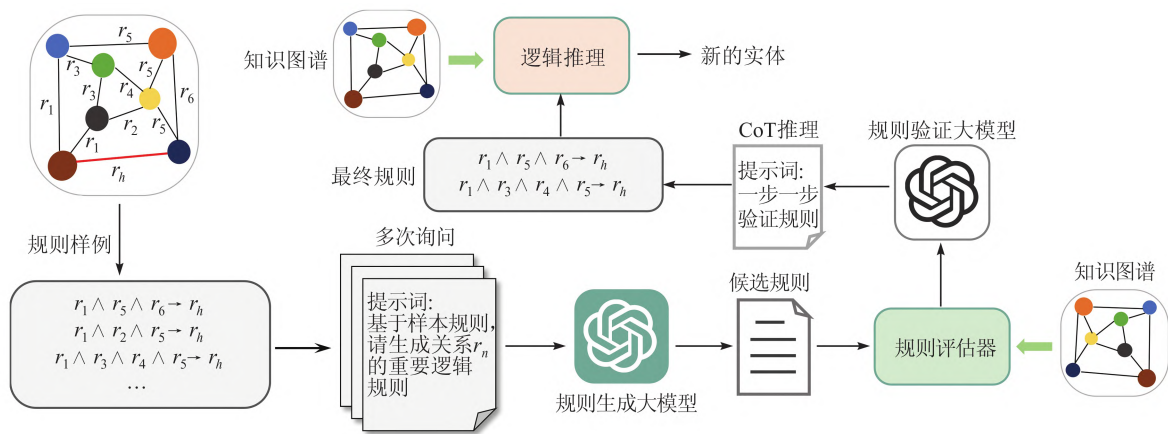


图3 ChatRule 总体框架^[74]

Fig. 3 The overall framework of ChatRule^[74]

2.3 增强知识图谱问答

传统基于知识图谱的问答系统(KBQA),如检索-推理结构^[76]、基于语义解析框架^[77],这些模型面临需要大量训练数据作支撑、构建过程对人类专家过度依赖,以及专业系统泛化能力不足等问题。而图模互补为知识图谱问答系统创造新的机会,现今增强问答模型通过微调技术或直接应用大模型实现相关操作^[78],更有学者探讨了ChatGPT在SPARQL查询任务中的潜力^[79]。文献[80]提出了一种利用大模型在少样本情况下回答学术知识图谱问题的方法。具体步骤为:对于一个目标问题,模型通过分析训练集中的问题,找到与测试问题相似的问题;以这些相似问题作为提示输入大模型,生成目标问题的SPARQL查询,最终通过查询知识图谱获取答案。此过程不涉及大模型预训练,减少了计算资源和时间的消耗。研究表明,该模型在SciQA-one挑战基准中获取较好的成绩。ChatKBQA^[81]则使用指令调优技术对开源大模型进行微调,使大模型获得令人满意的语义解析能力,利用大模型将新的自然语言问题转换为根据候选的逻辑形式,最后通过检索转换逻辑形式中的实体信息,生成相对应的SPARQL查询。ChatKBQA将大模型强大的语义解析功能和知识图谱可解释性相结合,为图模互补引入了一种新的思想图查询(GQoT)范式,更好地利用外部知识提高问答的可解释性,同时减少大模型的幻觉现象。

与大模型对话增强路径推理和检索的方式不同,KnowPAT^[82]提出了一个用于处理领域中特定问题回答的偏好对齐框架,通过知识偏好增强知识图感知问答的全新方法。KnowPAT认为大模型应该利用领域知识来生成可靠的答案,KnowPAT构建了风格偏好集和知识偏好集。此外,KnowPAT

设计了一个新的对齐目标,将大模型偏好与人类偏好对齐,为真实场景问答领域训练一个更可靠和友好的问答系统。不过KnowPAT模型无法适用于某些领域,如在医学领域、计算机科学领域,与原始模型相比KnowPAT的能力出现明显的下降。BYOKG^[83]设计一个通用的问题回答系统,旨在提高模型的通用性和便捷性,并且无需任何人工标注数据,就能够在任何知识图上运行。该系统受人类好奇心驱动学习的启发,首先采用无监督学习方式,通过图遍历学习未知知识图谱。并借助大模型生成自然语言问题补充探索语料库,最后使用检索增强推理实现问答预测。

表1为大模型增强知识图谱总结。

3 知识图谱增强大模型

尽管大模型在处理自然语言下游任务时,表现出色的性能,但是大模型仍存在一些致命缺陷问题尚未解决,例如大模型理解问题需要依赖上下文信息,逻辑推理能力不足、输出存在幻觉和编造的现象^[84]。这些问题导致大模型输出结果存在虚假性,可能对社会和道德产生负面影响^[15]。本文从多个角度分别介绍知识图谱增强大模型的自身性能、推理能力、检索和问答,并介绍减轻幻觉现象的方法。

3.1 增强大模型自身性能

3.1.1 知识图谱构建预训练语料库

大模型幻觉的缺陷是语言模型的通病,主要在于生成文本时过度依赖语言模型的概率分布,并且大模型自身不考虑文本是否与现实世界或常识一致,以及预训练数据可能包含错误、过时、有偏见或有争议的信息。研究表明^[85],使用结构化的数据可以提高大模型的性能,有助于大模型确认事实性问

表 1 大模型增强知识图谱总结

Table 1 Survey of LLM-enhanced KG methods

类别	名称	年份	原理
增强图谱构建	ChatIE ^[63]	2023	通过提示工程抽取信息
	ChatExtract ^[64]	2023	通过提示工程提取提示问题集
	AutoKG ^[65]	2023	通过预训练大模型提取关键词
	Tang 等 ^[66]	2023	利用 ChatGPT 生成标注样本
增强图谱补全	Yao 等 ^[67]	2023	知识图谱内嵌大模型
	KoPA ^[70]	2023	知识图谱内嵌大模型
	KICGPT ^[68]	2024	大模型作为额外知识库
	Li 等 ^[69]	2024	大模型作为训练数据生成器
增强图谱推理	LLM-ARK ^[72]	2023	大模型引导知识图谱推理
	LARK ^[73]	2023	大模型引导知识图谱推理
	ChatRule ^[74]	2023	大模型作为规则生成器
增强图谱问答	Taffa 等 ^[80]	2023	通过提示工程生成 SPARQL 查询
	ChatKBQA ^[81]	2023	微调大模型生成 SPARQL 查询
	KnowPAT ^[82]	2023	微调知识偏好的大模型
	BYOKG ^[83]	2023	大模型作为补充语料库

题,从模型本身的角度避免幻觉与编造。因此在大模型预训练阶段,可将知识图谱中的结构化信息(实体、关系、链接路径)作为训练数据,增强大模型自身涌现能力。例如文献[17]设计 TEKGEN 语言化管道模型将实体子图转换为高质量的语料库(KELM)。该管道包括启发式对齐器、三元组转换为文本的生成器、实体子图创建器、删除低质量输出的后处理过滤器四个部分。通过使用序列到序列模型(T5模型^[86])对训练语料进行微调,生成高质量的自然语言文本。KGPT^[18]是一种知识增强的预训练语言模型,利用自动对齐知识图谱和文本构建基于知识的语料库 KGTEXT,并选择高度语义重叠的对齐策略去噪。同时模型本身具有极强的泛化能力,在零样本和少样本学习上表现出了较好的性能。除此之外,预训练语料库还包括增强大模型类比推理能力的 ANALOGYKB 语料库^[87]。

3.1.2 知识图谱内嵌大模型

事实上,在 BERT^[88]和 GPT 为代表的预训练模型发布不久,不少学者已经开始研究知识增强型预训练模型^[89,90]。例如将知识图谱的三元组作为领域知识注入到句子的 K-BERT^[89],以及实体链接模型 KnowBERT^[90]。KnowBERT 设计一种使用集成实体链接来检索相关的实体嵌入的方法,以词到实体的关注形式更新上下文词的表示,并将多个知识库嵌入到大规模模型中实现知识增强。

如今知识内嵌大模型通常是采用对齐技术将知识图谱与自然语言相关联。例如 ERNIE^[91]和 ERNIE 3.0^[92]。ERNIE 是一种增强语言表示模型,通过构造

结构化知识编码模块,将知识纳入语言理解,显著提高知识驱动的性能。ERNIE 3.0 则是对其进行改进,同时融合自回归网络和自编码网络,使用大量纯文本和大规模知识图谱进行训练。不同于上述模型,SKILL^[93]设计一种直接在知识图谱的事实三元组上训练 T5 模型的方法,避免了知识图模型之间的差异,使模型能够轻易学习内嵌的事实化知识,应用在各种行业领域的问答工作中。此外,知识增强型大模型还包括集成实体间的细粒度关系的预训练语言模型 KLMo^[94]、统一知识嵌入和预训练语言表示模型 KEPLER^[95]、将知识图数据转换为自然语言的嵌入模型^[96]以及基于 ChatGPT 的黑箱知识注入方法 KnowGPT^[97]。知识增强大模型在 AI4Science 领域取得了一定的成果,文献[98]提出一种基于功能提示的知识图增强分子对比学习(KANO)方法,利用元素导向的图增强对比学习框架实现化学领域专业知识图谱的嵌入,同时知识图谱还用于生成功能提示,帮助大模型在微调过程中更好地理解任务相关知识,并为预测结果提供合理的化学解释。

3.2 增强大模型推理

大模型在处理结构化推理方面(如解决数学问题^[99])表现不佳,目前增强大模型推理能力的技术主要有思维链(CoT)、思维树(ToT)^[100]、思维图(GoT)^[101]等等。思维链是一种通过少样本示例提示来增强大型模型推理任务的方法,它能够通过生成中间推理步骤执行复杂的推理,例如 CohortGPT^[102]采用链式思维采样策略辅佐领域知识图增强大模型在医学领域推理能力。然而,思维链模型在使用

偏见特征时可能会导致推理结果受到影响^[103],改变其原本的方向。通常情况下,大模型被视为黑箱模型,即使是开源模型也很难显式判断其内部推理链和决策过程,为了克服大型模型在结构化推理方面的不足,可将知识图谱结构化推理、支持可解释的预测的能力与大型模型结合起来。

JointLK^[104]和 QA-GNN^[105]则采用图神经网络(GNN)和知识图谱来提高模型推理能力。相较于以往文本与知识独立的模式,QA-GNN将问题上下文与检索到的知识连接起来,构成一个联合图。实验表明,相较于比RoBERTa^[106],QA-GNN表现出更好的效果。JointLK模型则是在QA-GNN模型上的优化,由于QA-GNN仅将QA上下文作为一个额外节点附加到知识图,无法完成双向交互。而JointLK通过密集的双向注意力模块实现语言模型和知识图谱的多步联合推理。具体来说,JointLK模型将任务文本信息与外部知识图谱数据这两种不同模态的信息结合起来,设计了一个联合推理模块,在每个问题标记和每个知识图谱节点之间生成细粒度的双向注意映射,实现不同模态信息的融合。同时,JointLK模型设计了一个动态图裁剪模块,通过移除无关的图节点进行去噪,以确保模型正确地使用完整和适当的证据进行推理。实验结果表明,JointLK在解决带有否定词的复杂推理问题方面表现出色。

DRAGON^[107]模型在文本和知识图的深度双向联合(QA-GNN)的基础上,采用自监督学习策略。通过统一两个自我监督的推理任务,包括掩码语言建模(MLM)和链接预测,DRAGON模型实现了对文本和知识图的全面预训练。这种自监督学习策略使得模型能够更好地理解文本和知识图之间的关系,从而更准确地进行推理。

3.3 增强大模型检索

知识图谱通过检索增强大模型是自然语言处理领域备受关注的研究方向之一。一种代表性的方法是检索增强生成(retrieval-augmented generation, RAG)^[108]。大模型自身无法精确处理知识密集型任务,也无法探究信息的来源和知识的更新,RAG旨在通过外部真实知识向量索引来解决大模型无法自我更新知识的问题。RAG是一种包含查询编码器(Query Encoder)、预训练检索器(Retriever)和预训练的生成式模型相结合的端到端训练方法。具体步骤为通过文档检索器检索潜在信息,将潜在文本作为输入信息的附加上下文执行大模型,最终获得文本目标序列。与之类似,LaMDA^[39]模型包括LaMDA-Base和LaMDA-Research模型,在执行过程中,首先调用LaMDA-Base模型生成输出,但是

输出结果可能存在不真实信息,模型会继续调用LaMDA-Research与信息检索系统进行多次交互,直到LaMDA-Research响应用户输出,将无法验证的结果进行替换。相较于RAG,LaMDA的检索范围较小,但检索结果的精确度相对较高。

为了解决大模型内存开销过大的问题,RETRO^[109]通过从大型语料库中检索相似文档块来增强语言模型,在处理下游知识密集任务时,小参数模型性能达到了GPT-3的表现效果。RAG^[108]和RETRO^[109]二者都需单独训练检索模型,并且在面对超大规模和更新性强的外部文档时,可能会导致计算成本的增加。在检索的基础上,KaLMA^[110]构建了一个基于检索、重新排序和生成的管道,检索增强大模型的知识感知属性,同时提出“Conscious Incompetence”的设置,即当语言模型意识到自己无法提供某些知识,会在文本中插入“[NA]”标记,表示该部分知识无法被验证。实验结果表明,KaLMA^[110]在提高大模型的引用生成能力和检索准确性方面具有一定的潜力。与上述不同,KMLM^[111]是一种直接基于多语言三元组的知识增强型预训练方法,将三元组信息以代码形式转化为多语言文本,同时在预训练过程中附加文本的结构信息,实验结果表明,KMLM提高了隐性知识的推理能力,在跨语言知识密集型任务(事实知识检索)中表现出显著的性能改进。

3.4 增强大模型可解释性

大模型因其异常复杂的结构、庞大的参数量、对大规模文本数据的高依赖性等问题,往往很难提供具备逻辑推理过程的解释,在许多领域(如生物医疗、军事和金融等)是无法接受的。并且大模型内部工作机制难以理解,加之大模型输出的不确定性,导致相同输入可能得到不同输出,即使目前性能比较出色的大模型GPT-4也存在不可解释性的问题。

传统解决模型可解释性问题的方法集中于模型内在和事后的解释,如可解释模型结构和事后特征选择。然而,这些方法在解释模型决策过程方面存在不足。近期研究通过多实例学习、注意矩阵和外部知识结构等手段提供了文本解释,但仍未完全理解模型的推理过程。LMExplainer^[112]设计了一个知识增强的解释模块,将知识图谱与大模型相结合,采用知识图谱和图注意力网络(GAT)^[113]提取大模型的关键决策信号,提供了全面、清晰、可理解的文本解释。其过程如图4所示,通过大模型生成输入语言的嵌入,同时从知识图谱中检索到相关知识以构建子图,以语言嵌入和子图作为图神经网络的输入,通过图注意力网络来获取注意力分数,生成

最终的预测结果和决策的解释过程。实验结果显示,LMExplainer不仅提高了模型性能,并且可以更准确地解释模型推理过程。同样的,XplainLLM^[114]是首个捕捉大模型推理元素并通过人类可理解的解释呈现决策过程的数据集,满足大模型在决策过程中透明度、可解释性和可理解性的需求。通过结

合知识图和图注意力网络,构建一个问题-答案-解释(QAE)三元组,将大模型推理过程与知识图谱中实体和关系相连接的。评估结果显示,使用解释后的大模型性能提高2.4%,并且在问答任务中具有更出色的可解释性和理解效果。

表2为知识图谱增强大模型总结。

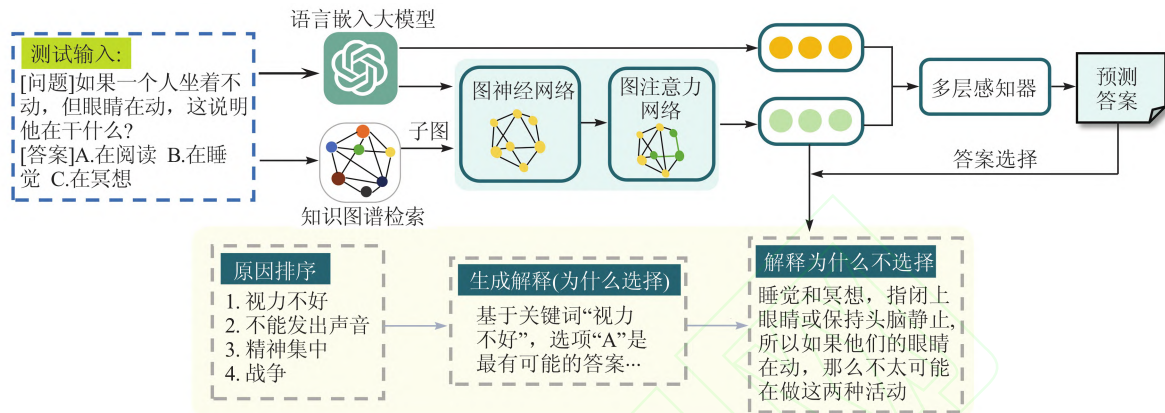


图4 LMExplainer总体框架

Fig. 4 The overall framework of LMExplainer

表2 知识图谱增强大模型总结

Table 2 Survey of KG-enhanced LLM methods

类别	名称	年份	原理
构建预训练语料库	KELM 语料库 ^[17]	2020	大模型作为生成器
	KGTEXT 语料库 ^[18]	2020	大模型作为生成器
	ANALOGYKB 语料库 ^[87]	2023	大模型作为生成器
知识图谱内嵌大模型	K-BERT ^[89]	2020	知识注入大模型
	KnowBERT ^[90]	2019	知识注入大模型
	ERNIE ^[91]	2019	模块化内嵌大模型
	ERNIE 3.0 ^[92]	2019	知识图谱训练大模型
	SKILL ^[93]	2022	知识图谱训练大模型
	KLMO ^[94]	2021	知识图谱与大模型链接表示
	KEPLER ^[95]	2019	实体嵌入大模型
	Lu 等 ^[96]	2022	知识注入大模型
	KnowGPT ^[97]	2023	知识注入大模型
	Fang 等 ^[98]	2023	知识图谱训练大模型
增强大模型推理	JointLK ^[104]	2021	图神经网络+知识图谱
	QA-GNN ^[105]	2021	图神经网络+知识图谱
	DRAGON ^[107]	2022	知识图谱训练大模型
增强大模型检索	LaMDA ^[39]	2022	知识图谱训练大模型
	RAG ^[108]	2020	知识图谱作为大模型的检索文档
	RETRO ^[109]	2021	知识图谱作为大模型的检索文档
	KaLMA ^[110]	2023	知识图谱作为大模型的检索文档
	KMLM ^[111]	2021	知识图谱训练大模型
增强大模型可解释性	LMExplainer ^[112]	2023	知识图谱作为检索器
	XplainLLM ^[114]	2023	将大模型推理与知识图谱关联

4 图模互补应用与展望

本节论述由知识图谱和大模型协同构建的专业领域系统,从医疗诊断预测和时间知识图谱领域应用出发,介绍组件之间互相配合、场景丰富度高的智能系统,对图模互补领域的未来进行展望。

4.1 图模互补应用

在图模互补的系统中,建立有效的反馈机制以动态调整和优化知识图谱与大模型之间的互动至关重要。构建这类系统的关键在于:① 大模型和知识图谱的反馈交互;② 适用于不同任务和场景的通用知识图谱;③ 各种知识增强型大模型的应用(例如通过 ChatExtract 和 AutoKG 工具完成领域知识图谱的构建)。总的来说,知识图谱和大模型相互协作构建,有利于提升系统的可靠性、可解释性和智能程度。

临床诊断案例、电子健康记录等各类患者的医疗记录是计算机医疗领域的重要知识来源。大模型能够利用这些冗余的记录自动构建医学知识图谱^[115],同时,图模互补系统还能够推理预测临床案例,推进医疗诊断决策系统的发展^[116]。如 Graph-Care^[117]通过提示工程从丰富的临床知识中抽取知识,为患者构建个性化医疗知识图谱,并利用双注意增强(BAT)图神经网络(GNN)模型进行下游任务预测。对于更复杂实体关系抽取,GPT-RE^[118]采取任务感知检索和金标签(gold label)诱导推理的方法,实现关系抽取的情境学习。如图 5(a)所示,GPT-RE 采用了两种任务感知检索方法,通过编码表示强调文本中的实体与关系信息,接着通过金标诱导推理方法(类似思维链)注入推理逻辑,获取大量输入与标签的对齐演示案例,最后通过提示大模

型实现关系抽取的高精确性和可解释性。此外,为了避免幻觉现象,REALM 模型^[119]提出一种检索增强生成(RAG)驱动框架,能够从各种非结构化数据(临床记录和电子健康记录)中提取实体,并与外部专业知识图谱进行匹配,以确保模型输出结果的一致性和准确性。

除医学领域外,大模型在与时间相关的特殊知识图谱方面(如事件抽取^[120]、时间关系探索^[121]和时间关系提取^[122])也取得一定进展。但是由于时间知识图谱(TKG)是一个复杂时态多关系图数据结构,大模型无法直接理解结构化的时间关系数据,在时间知识图谱的预测方面具有一定潜力^[123]。如 GenTKG^[124]将指令调优大模型引入时间知识图谱,采用基于时间规则的检索策略实现时间关系的预测,且性能优于传统的规则式和嵌入式方案。同时 GenTKG 通过轻量级的指令调优技术,极大程度地降低了训练成本。而大模型在时间知识推理方面^[125]也面临着一些问题,大模型的时间推理能力会因时间信息的变化而产生偏差。为了解决这一问题,文献[126]利用时间知识图谱和大模型融合设计一个复杂时间推理的系统,该系统通过学习时间推理的上下文,实现对于未来事件发生的可解释预测。同时,该文献提出一个可解释时间推理(ExpTime)的多源指令调整数据集,其构架过程如图 5(b)所示,借助时间知识图谱数据集及其时间推理路径完成构建任务。实验表明,该数据集能够提高大模型(LlaMA2)的事件预测和解释推理能力。

4.2 展望

4.2.1 增强大模型自身性能

大模型如今表现出的理解能力已经赶超人类,未来的研究将更加关注如何有效地将结构化、高质

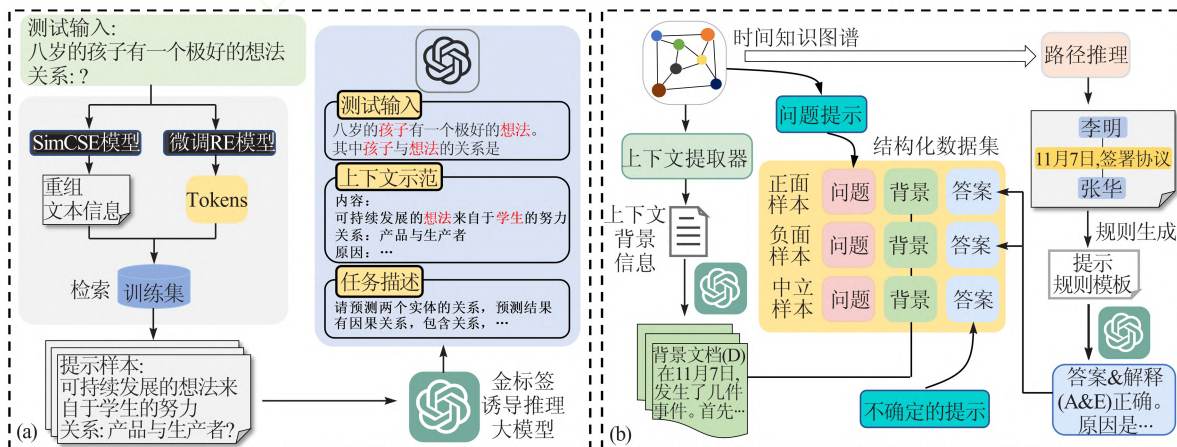


图 5 GPT-RE 执行过程和 ExpTime 构建过程

Fig. 5 The execution process of GPT-RE and the construction process of ExpTime

量的知识注入到大模型之中,以及如何使大模型能够更好地理解结构化数据。如今核心问题就是如何让大模型学习知识图谱信息,一种方式是将知识图谱直接转化为语言文本,但是这破坏了知识图谱的结构信息。另外一种利用深度学习从知识图中提取相关信息^[97]。因此在未来需要采用更先进的知识编码策略,如图神经网络,以更好地捕捉知识图谱中的关系和语义信息。

4.2.2 图模互补深度融合

知识图谱与大模型融合已经取得一定的成果,如 JointLK^[104] 和 QA-GNN^[105] 通过图神经网络(GNN)将知识图谱与大模型联立起来,DRAGON^[107]更是在联合基础上增加自监督学习策略,使得大模型能够理解知识图谱中的结构信息,提高大模型的推理能力。在未来可以结合先进的强化学习技术^[72],构建更完善的检索架构^[19],进一步探索图模互补更高效的交互机制,以及更有效的微调和更新策略。

4.2.3 缓解大模型幻觉现象

目前大模型普遍存在的一个问题:幻觉问题,在一些高精度的领域(如医疗领域、数学领域等)幻觉问题是不能接受的。如今已有不少研究通过设置逻辑中间推理步骤,以此来提高大模型自身的问题解决能力,进而缓解幻觉现象。例如,思维链(CoT)通过外部知识图谱实现让大模型逐步思考,使大模型更好地理解问题,减少了幻觉的发生。因此在未来研究中,大模型可以借助外部工具如可靠的、最新的知识图谱。知识图谱可以作为强大的信息检索工具,为大模型提供更加准确的信息。未来的工作将关注如何将知识图谱与语言模型的检索和推理有机结合,以提供更精确和真实的答案。

4.2.4 可解释型大模型

大型模型面临着可解释性的挑战,由于大模型的参数量过大,对于模型结构和数据的影响尚不明确,学者还未发现涌现能力的根源;大模型自身缺乏统一的解释标准和可靠的解释评估方法,这为可解释型大模型带来巨大挑战。因此需要探索大模型内部的工作流程,如 LMEExplainer^[112]通过知识图谱和图注意网络对大模型的关键决策信号进行探索。研究人员还需要进一步探索大模型的推理过程,理解模型自身决策原理,使模型能够解释其决策和推理过程,借助知识图谱来增强大模型的可解释性。这将有助于增加人们对大模型决策的信任,特别是在敏感领域和决策制定中的应用。

5 结 语

本文对知识图谱与大模型领域融合进行全面阐述,希望未来能有更多的研究工作,以推动这一领域的不断前进。首先,全面回顾了知识图谱与大模型领域的最新研究进展;然后介绍知识图谱和大模型的基础知识;接着从大模型增强知识图谱和知识图谱增强大模型两个角度,详细论述相关模型;最后介绍图模互补应用并提出展望。

图模互补研究领域具有广阔的前景和潜力。随着知识图谱和大模型的不断发展与完善,相信图模互补的深度融合在人工智能领域中能发挥越来越重要的作用。

参考文献:

- [1] WEI J, TAY Y, BOMMASANI R, *et al.* Emergent abilities of large language models[EB/OL]. [2022-10-26]. <https://doi.org/10.48550/arXiv.2206.07682>.
- [2] HUANG J, CHANG K C C. Towards reasoning in large language models: A survey[EB/OL]. 2022; arXiv: 2212.10403. <http://arxiv.org/abs/2212.10403>.
- [3] Chen Y B, Guo S R, Liu K, *et al.* Large language models and knowledge graphs[EB/OL]. [2022-10-26]. <https://aclanthology.org/2023.ccl-2.6/>.
- [4] DONG Q X, LI L, DAID M, *et al.* A survey on In-context learning[EB/OL]. 2022; arXiv: 2301.00234. <http://arxiv.org/abs/2301.00234>.
- [5] ZHAO W X, ZHOU K, LI J Y, *et al.* A survey of large language models [EB/OL]. 2023; arXiv: 2303.18223. <http://arxiv.org/abs/2303.18223>.
- [6] OpenAI. Introducing chatgpt [EB/OL]. [2022-11-30]. OpenAI Blog. <https://openai.com/blog/chatgpt>.
- [7] ACHIAM J, ADLER S, AGARWAL S, *et al.* GPT-4 Technical Report[EB/OL]. [2024-03-04]. <https://arxiv.org/abs/2303.08774>.
- [8] WANG L, YANG N, HUANG X L, *et al.* Large search model: Redefining search stack in the era of LLMs[J]. *ACM SIGIR Forum*, 2023, **57**(2): 1-16. DOI: 10.1145/3642979.3643006.
- [9] ZHU J H, XIA Y C, WU L J, *et al.* Incorporating BERT into neural machine translation [EB/OL]. 2020; arXiv: 2002.06823. <http://arxiv.org/abs/2002.06823>.
- [10] WU S J, IRSOY O, LU S, *et al.* BloombergGPT: A large language model for finance[EB/OL]. 2023; arXiv: 2303.17564. <http://arxiv.org/abs/2303.17564>.
- [11] SINGHAL K, AZIZI S, TU T, *et al.* Large language models encode clinical knowledge [J]. *Nature*, 2023, **620**: 172-180. DOI: 10.1038/s41586-023-06291-2.

- [12] GAN W S, QI Z L, WU J Y, *et al.* Large language models in education: Vision and opportunities [C]//2023 IEEE International Conference on Big Data (BigData). New York: IEEE Press, 2023: 4776-4785. DOI: 10.1109/BigData59044.2023.10386291.
- [13] 车万翔, 窦志成, 冯岩松, 等. 大模型时代的自然语言处理: 挑战、机遇与发展[J]. 中国科学: 信息科学, 2023, **53**(9): 1645-1687.
- CHE W X, DOU Z C, FENG Y S, *et al.* Towards a comprehensive understanding of the impact of large language models on natural language processing: Challenges, opportunities and future directions[J]. *Scientia Sinica (Informationis)*, 2023, **53**(9): 1645-1687 (Ch).
- [14] LIU Y, YAO Y S, TON J F, *et al.* Trustworthy LLMs: A survey and guideline for evaluating large language models' alignment[EB/OL]. 2023: arXiv: 2308.05374. <http://arxiv.org/abs/2308.05374>.
- [15] BANG Y J, CAHYAWIJAYA S, LEE N, *et al.* A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and interactivity [EB/OL]. 2023: arXiv: 2302.04023. <http://arxiv.org/abs/2302.04023>.
- [16] 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. DOI: 10.1109/TNNLS.2021.3070843.
- [17] AGARWAL O, GE H M, SHAKERI S, *et al.* Knowledge graph based synthetic corpus generation for knowledge-enhanced language model pre-training[EB/OL]. 2020: arXiv: 2010.12688. <http://arxiv.org/abs/2010.12688>.
- [18] CHEN W H, SU Y, YAN X F, *et al.* KGPT: Knowledge-grounded pre-training for data-to-text generation[EB/OL]. 2020: arXiv: 2010.02307. <http://arxiv.org/abs/2010.02307>.
- [19] WANG X T, YANG Q W, QIU Y T, *et al.* KnowledGPT: Enhancing large language models with retrieval and storage access on knowledge bases[EB/OL]. 2023: arXiv: 2308.11761. <http://arxiv.org/abs/2308.11761>.
- [20] YAO Y F, DUAN J H, XU K D, *et al.* A survey on large language model (LLM) security and privacy: The good, the bad, and the ugly[J]. *High-Confidence Computing*, 2024, **4**(2): 100211. DOI: 10.1016/j.hcc.2024.100211.
- [21] PENG C Y, XIA F, NASERIPARSA M, *et al.* Knowledge graphs: Opportunities and challenges[J]. *Artificial Intelligence Review*, 2023, **56**(11): 13071-13102. DOI: 10.1007/s10462-023-10465-9.
- [22] ZHU Y Q, WANG X H, CHEN J, *et al.* LLMs for knowledge graph construction and reasoning: Recent capabilities and future opportunities [EB/OL]. 2023: arXiv: 2305.13168. <http://arxiv.org/abs/2305.13168>.
- [23] ZHANG B W, SOH H. Extract, define, canonicalize: An LLM-based framework for knowledge graph construction [EB/OL]. 2024: arXiv: 2404.03868. <http://arxiv.org/abs/2404.03868>.
- [24] LIU P F, QIAN L, ZHAO X W, *et al.* Joint knowledge graph and large language model for fault diagnosis and its application in aviation assembly[J]. *IEEE Transactions on Industrial Informatics*, 2024, PP(99): 1-10. DOI: 10.1109/TII.2024.3366977.
- [25] PAN S R, LUO L H, WANG Y F, *et al.* Unifying large language models and knowledge graphs: A roadmap[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2024, PP(99): 1-20. DOI: 10.1109/TKDE.2024.3352100.
- [26] 徐增林, 盛泳潘, 贺丽荣, 等. 知识图谱技术综述[J]. 电子科技大学学报, 2016, **45**(4): 589-606. DOI: 10.3969/j.issn.1001-0548.2016.04.012.
- XU Z L, SHENG Y P, HE L R, *et al.* Review on knowledge graph techniques[J]. *Journal of University of Electronic Science and Technology of China*, 2016, **45**(4): 589-606. DOI: 10.3969/j.issn.1001-0548.2016.04.012(Ch).
- [27] 刘峤, 李杨, 段宏, 等. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, **53**(3): 582-600. DOI: 10.7544/j.issn1000-1239.2016.20148228.
- LIU Q, LI Y, DUAN H, *et al.* Knowledge graph construction techniques [J]. *Journal of Computer Research and Development*, 2016, **53**(3): 582-600. DOI: 10.7544/j.issn1000-1239.2016.20148228 (Ch).
- [28] SUCHANEK F M, KASNECI G, WEIKUM G. YAGO: A large ontology from Wikipedia and WordNet[J]. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2008, **6**(3): 203-217. DOI: 10.1016/j.websem.2008.06.001.
- [29] BIZER C, LEHMANN J, KOBILAROV G, *et al.* DBpedia: A crystallization point for the Web of Data [J]. *Journal of Web Semantics*, 2009, **7**(3): 154-165. DOI: 10.1016/j.websem.2009.07.002.
- [30] BOLLACKER K, COOK R, TUFTS P. Freebase: A shared database of structured general human knowledge [DB/OL]. [2024-02-02]. <https://cdn.aaai.org/AAAI/2007/AAAI07-355.pdf>.
- [31] WMF. Wikidata[EB/OL]. [2015-11-11]. https://www.wikidata.org/wiki/Wikidata:Main_Page.
- [32] DONG X, GABRILOVICH E, HEITZ G, *et al.* Knowledge vault: A web-scale approach to probabilistic knowledge fusion [C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. New

- York : ACM, 2014: 601-610. DOI: 10.1145/2623330.2623623.
- [33] YOO J, CHO M, KIM T, *et al.* Knowledge extraction with no observable data[J]. *Advances in Neural Information Processing Systems*, 2019, **32**:2705-2714.
- [34] CHEN X J, JIA S B, XIANG Y. A review: Knowledge reasoning over knowledge graph[J]. *Expert Systems with Applications*, 2020, **141**: 112948. DOI: 10.1016/j.eswa.2019.112948.
- [35] GUAN S P, JIN X L, GUO J F, *et al.* NeuInfer: knowledge inference on N -ary facts[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, 2020: 6141-6151. DOI: 10.18653/v1/2020.acl-main.546.
- [36] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning [M]. Cambridge: MIT Press, 2016.
- [37] MIKOLOV T, SUTSKEVER I, CHEN K, *et al.* Distributed representations of words and phrases and their compositionality [DB/OL]. [2022-02-10]. https://proceedings.neurips.cc/paper_files/paper/2013/file/9aa42b31882-ec039965f3c4923ce901b-Paper.pdf.
- [38] BROWN T, MANN B, RYDER N, *et al.* Language models are few-shot learners[J]. *Advances in Neural Information Processing Systems*, 2020, **33**: 1877-1901.
- [39] THOPPILAN R, DE FREITAS D, HALL J, *et al.* LaMDA: Language models for dialog applications [DB/OL]. [2022-02-10]. <https://doi.org/10.48550/arXiv.2201.08239>.
- [40] CHOWDHERY A, NARANG S R, DEVLIN J, *et al.* PaLM: Scaling language modeling with pathways [EB/OL]. 2022; arXiv: 2204.02311. <http://arxiv.org/abs/2204.02311>.
- [41] ANIL R, DAI A M, FIRAT O, *et al.* PaLM 2 technical report [EB/OL]. 2023; arXiv: 2305.10403. <http://arxiv.org/abs/2305.10403>.
- [42] TOUVRON H, LAVRIL T, IZACARD G, *et al.* LLaMA: Open and efficient foundation language models [EB/OL]. 2023; arXiv: 2302.13971. <http://arxiv.org/abs/2302.13971>.
- [43] TOUVRON H, MARTIN L, STONE K, *et al.* LLaMA 2: Open foundation and fine-tuned chat models [EB/OL]. 2023; arXiv: 2307.09288. <http://arxiv.org/abs/2307.09288>.
- [44] Anthropic. Claude's Constitution [EB/OL]. [2023-07-08]. <https://www.anthropic.com/index/claude-constitution>.
- [45] LUO H P, SUN Q F, XU C, *et al.* WizardMath: Empowering mathematical reasoning for large language models via reinforced evol-instruct [EB/OL]. 2023; arXiv: 2308.09583. <http://arxiv.org/abs/2308.09583>.
- [46] TEAM G, ANIL R, BORGEAUD S, *et al.* Gemini: A family of highly capable multimodal models [EB/OL]. 2023; arXiv: 2312.11805. <http://arxiv.org/abs/2312.11805>.
- [47] LIU H T, LI C Y, LI Y H, *et al.* Improved baselines with visual instruction tuning [EB/OL]. 2023; arXiv: 2310.03744. <http://arxiv.org/abs/2310.03744>.
- [48] BROOKS T, PEEBLES B, HOLMES C, *et al.* Video generation models as world simulators [EB/OL]. [2024-02-15]. <https://openai.com/research/video-generation-models-as-world-simulators>.
- [49] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate [EB/OL]. 2014; arXiv: 1409.0473. DOI: 10.48550/arXiv.1409.0473.
- [50] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need [EB/OL]. [2024-02-15]. <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf>
- [51] CHILD R, GRAY S, RADFORD A, *et al.* Generating long sequences with sparse transformers [EB/OL]. 2019; arXiv: 1904.10509. <http://arxiv.org/abs/1904.10509> [LinkOut]
- [52] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, *et al.* An image is worth 16x16 words: Transformers for image recognition at scale [EB/OL]. 2020; arXiv: 2010.11929. <http://arxiv.org/abs/2010.11929>.
- [53] GONG Y, CHUNG Y A, GLASS J. AST: Audio spectrogram transformer [EB/OL]. [2021-07-08]. <https://doi.org/10.48550/arXiv.2104.01778>.
- [54] AKBARI H, YUAN L Z, QIAN R, *et al.* VATT: Transformers for multimodal self-supervised learning from raw video, audio and text [EB/OL]. 2021; arXiv: 2104.11178. <http://arxiv.org/abs/2104.11178>.
- [55] SHIN T, RAZEGHI Y, LOGAN R L, *et al.* Auto-prompt: Eliciting knowledge from language models with automatically generated prompts [EB/OL]. [2020-11-07]. <https://doi.org/10.48550/arXiv.2010.15980>.
- [56] BUBECK S, CHANDRASEKARAN V, ELDAN R, *et al.* Sparks of artificial general intelligence: Early experiments with GPT-4 [EB/OL]. [2023-03-13]. <https://doi.org/10.48550/arXiv.2303.12712>.
- [57] RADFORD A, NARASIMHAN K, SALIMANS T, *et al.* Improving language understanding by generative pre-training [EB/OL]. [2018-06-25]. <https://www.mikecaptain.com/resources/pdf/GPT-1.pdf>.
- [58] RADFORD A, WU J, CHILD R, *et al.* Language models are unsupervised multitask learners [J]. *OpenAI blog*, 2019, **1**(8): 9.

- [59] CHEN M, TWOREK J, JUN H, *et al.* Evaluating large language models trained on code[EB/OL]. 2021: arXiv: 2107.03374. <http://arxiv.org/abs/2107.03374>.
- [60] OUYANG L, WU J, XU J, *et al.* Training language models to follow instructions with human feedback[EB/OL]. 2022: arXiv: 2203.02155. <http://arxiv.org/abs/2203.02155>.
- [61] WADHWA S, HASSANZADEH O, BHATTACHARJYA D, *et al.* Distilling event sequence knowledge from large language models[EB/OL]. 2024: arXiv: 2401.07237. <http://arxiv.org/abs/2401.07237>[LinkOut]
- [62] AGRAWAL M, HEGSELMANN S, LANG H, *et al.* Large language models are few-shot clinical information extractors[C]//Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2022: 1998-2022. DOI: 10.18653/v1/2022.emnlp-main.130.
- [63] WEI X, CUI X Y, CHENG N, *et al.* Zero-shot information extraction via chatting with ChatGPT[EB/OL]. 2023: arXiv: 2302.10205. <http://arxiv.org/abs/2302.10205>.
- [64] POLAK M P, MORGAN D. Extracting accurate materials data from research papers with conversational language models and prompt engineering[J]. *Nature Communications*, 2024, **15**(1): 1-11. DOI: 10.1038/s41467-024-45914-8.
- [65] CHEN B H, BERTOZZI A L. AutoKG: efficient automated knowledge graph generation for language models[C]//2023 IEEE International Conference on Big Data (BigData). New York: IEEE Press, 2023: 3117-3126. DOI: 10.1109/BigData59044.2023.10386454.
- [66] TANG R X, HAN X T, JIANG X Q, *et al.* Does synthetic data generation of LLMs help clinical text mining[EB/OL]. 2023: arXiv: 2303.04360. <http://arxiv.org/abs/2303.04360>.
- [67] YAO L, PENG J Z, MAO C S, *et al.* Exploring large language models for knowledge graph completion[EB/OL]. 2023: arXiv: 2308.13916. <http://arxiv.org/abs/2308.13916>.
- [68] WEI Y B, HUANG Q S, ZHANG Y, *et al.* KICGPT: large language model with knowledge in context for knowledge graph completion[C]//Findings of the Association for Computational Linguistics: EMNLP 2023. Stroudsburg: Association for Computational Linguistics, 2023: 8667-8683. DOI: 10.18653/v1/2023.findings-emnlp.580.
- [69] LI D W, TAN Z, CHEN T L, *et al.* Contextualization distillation from large language model for knowledge graph completion[EB/OL]. 2024: arXiv: 2402.01729. <http://arxiv.org/abs/2402.01729>.
- [70] ZHANG Y C, CHEN Z, GUO L B, *et al.* Making large language models perform better in knowledge graph completion[EB/OL]. 2023: arXiv: 2310.06671. <http://arxiv.org/abs/2310.06671>.
- [71] 王文广. 知识图谱: 认知智能理论与实战[M]. 北京: 电子工业出版社, 2022.
- WANG W G. Knowledge Graph[M]. Beijing: Publishing House of Electronics Industry, 2022(Ch).
- [72] HUANG Y X, SHI L D, LIU A Q, *et al.* Evaluating and enhancing large language models for conversational reasoning on knowledge graphs[EB/OL]. 2023: arXiv: 2312.11282. <http://arxiv.org/abs/2312.11282>.
- [73] CHOUDHARY N, REDDY C K. Complex logical reasoning over knowledge graphs using large language models[EB/OL]. 2023: arXiv: 2305.01157. DOI: 10.48550/arXiv.2305.01157.
- [74] LUO L H, JU J X, XIONG B, *et al.* ChatRule: Mining logical rules with large language models for knowledge graph reasoning[EB/OL]. 2023: arXiv: 2309.01538. <http://arxiv.org/abs/2309.01538>.
- [75] WEI J, WANG X Z, SCHUURMANS D, *et al.* Chain-of-thought prompting elicits reasoning in large language models[J]. *Advances in Neural Information Processing Systems*, 2022, **35**: 24824-24837.
- [76] WANG J C, LI W M, LIU F F, *et al.* Hic-KGQA: Improving multi-hop question answering over knowledge graph via hypergraph and inference chain[J]. *Knowledge-Based Systems*, 2023, **277**: 110810. DOI: 10.1016/j.knsys.2023.110810.
- [77] XIONG H B, WANG S T, TANG M R, *et al.* Knowledge graph question answering with semantic oriented fusion model[J]. *Knowledge-Based Systems*, 2021, **221**: 106954. DOI: 10.1016/j.knsys.2021.106954.
- [78] LI X X, ZHAO R C, CHIA Y K, *et al.* Chain of knowledge: A framework for grounding large language models with structured knowledge bases[EB/OL]. 2023: arXiv: 2305.13269. <http://arxiv.org/abs/2305.13269>.
- [79] MEYER L P, STADLER C, FREY J, *et al.* LLM-assisted knowledge graph engineering: Experiments with ChatGPT[EB/OL]. [2023-07-13]. <https://doi.org/10.48550/arXiv.2307.06917>.
- [80] TAFFA T A, USBECK R. Leveraging LLMs in scholarly knowledge graph question answering[EB/OL]. 2023: arXiv: 2311.09841. <http://arxiv.org/abs/2311.09841>.
- [81] LUO H R, TANG Z H, PENG S C, *et al.* ChatKBQA: A generate-then-retrieve framework for knowledge base question answering with fine-tuned large language models[EB/OL]. [2023-10-13]. <https://doi.org/10.4-8550/arXiv.2310.08975>.
- [82] ZHANG Y C, CHEN Z, FANG Y, *et al.* Knowledge-

- able preference alignment for LLMs in domain-specific question answering [EB/OL]. 2023: arXiv: 2311.06503. <http://arxiv.org/abs/2311.06503>.
- [83] AGARWAL D, DAS R, KHOSLA S, *et al.* Bring your own KG: Self-supervised program synthesis for zero-shot KGQA [EB/OL]. [2023-11-14]. <https://doi.org/10.48550/arXiv.2311.07850>.
- [84] RAWTE V, SHETH A, DAS A. A survey of hallucination in large foundation models [EB/OL]. 2023: arXiv: 2309.05922. <http://arxiv.org/abs/2309.05922>.
- [85] SHENG E, CHANG K W, NATARAJAN P, *et al.* The woman worked as a babysitter: On biases in language generation [EB/OL]. 2019: arXiv: 1909.01326. <http://arxiv.org/abs/1909.01326>.
- [86] KALE M, RASTOGI A. Text-to-text pre-training for data-to-text tasks [EB/OL]. [2021-07-09]. <https://doi.org/10.48550/arXiv.2005.10433>.
- [87] YUAN S Y, CHEN J J, SUN C Z, *et al.* ANALOGYKB: Unlocking analogical reasoning of language models with A million-scale knowledge base [EB/OL]. 2023: arXiv: 2305.05994. <http://arxiv.org/abs/2305.05994>.
- [88] DEVLIN J, CHANG M W, LEE K, *et al.* BERT: Pre-training of deep bidirectional transformers for language understanding [EB/OL]. 2018: arXiv: 1810.04805. <http://arxiv.org/abs/1810.04805>.
- [89] LIU W J, ZHOU P, ZHAO Z, *et al.* K-BERT: Enabling language representation with knowledge graph [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, **34** (3): 2901-2908. DOI: 10.1609/aaai.v34i03.5681.
- [90] PETERS M E, NEUMANN M, LOGAN IV R L, *et al.* Knowledge enhanced contextual word representations [EB/OL]. 2019: arXiv: 1909.04164. <http://arxiv.org/abs/1909.04164>.
- [91] ZHANG Z Y, HAN X, LIU Z Y, *et al.* ERNIE: Enhanced language representation with informative entities [EB/OL]. 2019: arXiv: 1905.07129. <http://arxiv.org/abs/1905.07129>.
- [92] 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/OL]. [2022-07-05]. <https://doi.org/10.48550/arXiv.2107.02137>.
- [93] MOISEEV F, DONG Z, ALFONSECA E, *et al.* SKILL: Structured knowledge infusion for large language models [EB/OL]. 2022: arXiv: 2205.08184. <http://arxiv.org/abs/2205.08184>.
- [94] HE L, ZHENG S C, YANG T, *et al.* Klm0: Knowledge graph enhanced pretrained language model with fine-grained relationships [C]//Findings of the Association for Computational Linguistics: EMNLP 2021. 2021: 4536-4542. DOI: 10.18653/v1/2021.findings-emnlp.384.
- [95] 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. DOI: 10.1162/tacl_a_00360.
- [96] LU Y J, OUYANG S Q, ZHOU K R. Structured knowledge grounding for question answering [EB/OL]. 2022: arXiv: 2209.08284. <http://arxiv.org/abs/2209.08284>.
- [97] ZHANG Q G, DONG J N, CHEN H, *et al.* KnowGPT: Black-box knowledge injection for large language models [EB/OL]. [2024-03-12]. <https://doi.org/10.48550/arXiv.2312.06185>.
- [98] FANG Y, ZHANG Q, ZHANG N Y, *et al.* Knowledge graph-enhanced molecular contrastive learning with functional prompt [J]. *Nature Machine Intelligence*, 2023, **5** (5): 542-553. DOI: 10.1038/s42256-023-00654-0.
- [99] YOUSEFZADEH R, CAO X N. Large language models' understanding of math: Source criticism and extrapolation [EB/OL]. 2023: arXiv: 2311.07618. <http://arxiv.org/abs/2311.07618>.
- [100] YAO S Y, YU D, ZHAO J, *et al.* Tree of thoughts: Deliberate problem solving with large language models [J]. *Advances in Neural Information Processing Systems*, 2024, **36**: 11809-11822.
- [101] BESTA M, BLACH N, KUBICEK A, *et al.* Graph of thoughts: Solving elaborate problems with large language models [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, **38** (16): 17682-17690. DOI: 10.1609/aaai.v38i16.29720.
- [102] GUAN Z H, WU Z H, LIU Z L, *et al.* CohortGPT: An enhanced GPT for participant recruitment in clinical study [EB/OL]. 2023: arXiv: 2307.11346. <http://arxiv.org/abs/2307.11346>.
- [103] TURPIN M, MICHAEL J, PEREZ E, *et al.* Language models don't always say what they think: Unfaithful explanations in chain-of-thought prompting [EB/OL]. 2023: arXiv: 2305.04388. <http://arxiv.org/abs/2305.04388>.
- [104] SUN Y Q, SHI Q, QI L, *et al.* JointLK: Joint reasoning with language models and knowledge graphs for common-sense question answering [EB/OL]. 2021: arXiv: 2112.02732. DOI: 10.48550/arXiv.2112.02732.
- [105] YASUNAGA M, REN H Y, BOSSELU T A, *et al.* QA-GNN: Reasoning with language models and knowledge graphs for question answering [EB/OL]. 2021: arXiv: 2104.06378. <http://arxiv.org/abs/2104.06378>.
- [106] LIU Y H, OTT M, GOYAL N, *et al.* RoBERTa: A

- robustly optimized BERT pretraining approach[EB/OL]. 2019: arXiv: 1907.11692. <http://arxiv.org/abs/1907.11692>.
- [107] YASUNAGA M, BOSSELUT A, REN H Y, *et al.* Deep bidirectional language-knowledge graph pretraining [J]. *Advances in Neural Information Processing Systems*, 2022, **35**: 37309–37323.
- [108] LEWIS P, PEREZ E, PIKTUS A, *et al.* Retrieval-augmented generation for knowledge-intensive NLP tasks [J]. *Advances in Neural Information Processing Systems*, 2020, **33**: 9459–9474.
- [109] BORGEAUD S, MENSCH A, HOFFMANN J, *et al.* Improving language models by retrieving from trillions of tokens[EB/OL]. 2021: arXiv: 2112.04426. <http://arxiv.org/abs/2112.04426>.
- [110] LI X Z, CAO Y X, PAN L M, *et al.* Towards verifiable generation: A benchmark for knowledge-aware language model attribution [EB/OL]. 2023: arXiv: 2310.05634. <http://arxiv.org/abs/2310.05634>.
- [111] LIU L L, LI X, HE R D, *et al.* Enhancing multilingual language model with massive multilingual knowledge triples [EB/OL]. 2021: arXiv: 2111.10962. <http://arxiv.org/abs/2111.10962>.
- [112] CHEN Z Y, SINGH A K, SRA M. LMExplainer: A knowledge-enhanced explainer for language models [EB/OL]. [2023-08-03]. <https://doi.org/10.48550/arXiv.2303.16537>.
- [113] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, *et al.* Graph attention networks [EB/OL]. 2017: arXiv: 1710.10903. <http://arxiv.org/abs/1710.10903>.
- [114] CHEN Z, CHEN J, GAIDHANI M, *et al.* XplainLLM: A QA Explanation Dataset for Understanding LLM Decision-Making [EB/OL]. [2023-11-15]. <https://doi.org/10.48550/arXiv.2311.08614>.
- [115] ARSENYAN V, BUGHDARYAN S, SHAYA F D, *et al.* Large language models for biomedical knowledge graph construction: Information extraction from EMR notes [EB/OL]. [2023-12-09]. <https://doi.org/10.48550-arXiv.2301.12473>.
- [116] GAO Y J, LI R Z, CASKEY J, *et al.* Leveraging a medical knowledge graph into large language models for diagnosis prediction [EB/OL]. 2023: arXiv: 2308.14321. <http://arxiv.org/abs/2308.14321>.
- [117] JIANG P C, XIAO C, CROSS A, *et al.* GraphCare: Enhancing healthcare predictions with personalized knowledge graphs [EB/OL]. 2023: arXiv: 2305.12788. <http://arxiv.org/abs/2305.12788>.
- [118] WAN Z, CHENG F, MAO Z Y, *et al.* GPT-RE: In-context learning for relation extraction using large language models [EB/OL]. 2023: arXiv: 2305.02105. <http://arxiv.org/abs/2305.02105>.
- [119] ZHU Y H, REN C Y, XIE S Y, *et al.* REALM: RAG-driven enhancement of multimodal electronic health records analysis via large language models [EB/OL]. 2024: arXiv: 2402.07016. <http://arxiv.org/abs/2402.07016>.
- [120] GAO J, ZHAO H, YU C L, *et al.* Exploring the feasibility of ChatGPT for event extraction [EB/OL]. 2023: arXiv: 2303.03836. <http://arxiv.org/abs/2303.03836>.
- [121] CHAN C, CHENG J Y, WANG W Q, *et al.* ChatGPT evaluation on sentence level relations: A focus on temporal, causal, and discourse relations [EB/OL]. 2023: arXiv: 2304.14827. <http://arxiv.org/abs/2304.14827>.
- [122] YUAN C H, XIE Q Q, ANANIADOU S. Zero-shot temporal relation extraction with ChatGPT [EB/OL]. 2023: arXiv: 2304.05454. <http://arxiv.org/abs/2304.05454>.
- [123] LEE D H, AHRABIAN K, JIN W, *et al.* Temporal knowledge graph forecasting without knowledge using in-context learning [EB/OL]. [2023-10-20]. <https://doi.org/10.48550/arXiv.2305.10613>.
- [124] LIAO R T, JIA X, MA Y P, *et al.* GenTKG: Generative forecasting on temporal knowledge graph [EB/OL]. 2023: arXiv: 2310.07793. <http://arxiv.org/abs/2310.07793>.
- [125] TAN Q Y, NG H T, BING L D. Towards benchmarking and improving the temporal reasoning capability of large language models [EB/OL]. 2023: arXiv: 2306.08952. <http://arxiv.org/abs/2306.08952>.
- [126] YUAN C H, XIE Q Q, HUANG J M, *et al.* Back to the future: Towards explainable temporal reasoning with large language models [EB/OL]. 2023: arXiv: 2310.01074. <http://arxiv.org/abs/2310.01074>.

□