

知识图谱多跳问答推理研究进展、挑战与展望

杜会芳¹, 王昊奋¹, 史英慧², 王萌³

1. 同济大学设计创意学院, 上海 200092;

2. 东南大学网络空间与安全学院, 江苏 无锡 214100;

3. 东南大学计算机科学与工程学院, 江苏 南京 211189

摘要

近年来,知识图谱问答在医疗、金融、政务等领域被广泛应用。用户不再满足于关于实体属性的单跳问答,而是更多地倾向表达复杂的多跳问答需求。为了应对上述复杂多跳问答,各种不同类型的推理方法被陆续提出。系统地介绍了基于嵌入、路径、逻辑的多跳知识问答推理的最新研究进展以及相关数据集和评测指标,并重点围绕前沿问题进行了讨论。最后总结了现有方法的不足,并展望了未来的研究方向。

关键词

知识图谱;多跳问答;推理

中图分类号:TP311

文献标识码:A

doi: 10.11959/j.issn.2096-0271.2021026

Progress, challenges and research trends of reasoning in multi-hop knowledge graph based question answering

DU Huifang¹, WANG Haofen¹, SHI Yinghui², WANG Meng³

1. College of Design and Innovation, Tongji University, Shanghai 200092, China

2. School of Cyber Science and Engineering, Southeast University, Wuxi 214100, China

3. School of Computer Science and Engineering, Southeast University, Nanjing 211189, China

Abstract

Recently, knowledge graph based question answering has been widely used in many fields such as medical care, finance, and government affairs. Users are no longer satisfied with question answering service of single-hop entity attributes, but want service which can handle complex multi-hop question. In order to accurately and deeply understand multi-hop questions, various types of reasoning methods have been proposed. The latest research methods of multi-hop knowledge graph based question answering were systematically introduced, as well as related datasets and evaluation metrics. These

researches methods were divided into three categories: embedding-based methods, path-based methods, and logic-based methods. In particular, frontier issues discussion were focused. Finally, the shortcomings of existing methods and suggest future research trends were summarized.

Key words

knowledge graph, multi-hop question answering, reasoning

1 引言

1.1 背景介绍

随着大数据时代的来临,帮助用户在海量信息中快速找到想要的信息尤为重要。知识图谱(knowledge graph, KG)以三元组的形式结构化存储海量信息,一个三元组可以表示为 $\langle e_h, r, e_t \rangle$,其中 e_h 表示头实体, e_t 表示尾实体, r 表示头实体与尾实体之间存在的关系。知识图谱三元组之间也存在关联关系,如上述三元组中的头实体 e_h 可能是另外一个三元组的尾实体。知识图谱把海量互联网信息表达成客观世界可认知的语义表示,具有强大的语义表达、存储和表达能力^[1],在工业界和学术界得到了广泛关注和研究应用。智能问答系统^[2]旨在针对用户提出的复杂信息需求,允许用户用自然语言问句的形式提问,并为用户直接返回精准的答案。得益于知识图谱技术的快速发展,基于知识图谱的问答(knowledge graph based question answering, KGQA)技术利用其丰富的结构化语义信息,能够深入理解用户的问题,并给出准确的答案^[3],为用户提供7×24小时的智能问答服务,在医疗、教育、金融等多个领域凸显出重要的应用价值^[4-6]。

传统KGQA以实体、属性等单一具体对象为主,而在实际应用场景中,用户不再满足于单跳的知识问答,如在医疗领域

中的咨询问题“常见的治疗感冒药物有哪些?”。用户更多地倾向表达复杂的多跳问答推理问题,如“请问伴有中耳炎并发症的感冒能用哪种药物治疗?”。而知识图谱多跳问答(以下简称多跳知识问答)^[7-8]即针对包含多跳关系的问题,在知识图谱上进行多步推理,继而推断得到答案的一项任务。

1.2 多跳知识问答推理分类

相比单跳问答,多跳知识问答需要在包含大量知识的知识图谱中找到多个有关联的三元组,并建模多跳长路径,这是更加具有挑战性的任务。同时知识图谱又是不完备的,长路径中任意一个三元组的不完整都会导致找不到正确的答案,这为多跳知识问答带来了更大的挑战。为了解决没有明显答案的复杂多跳查询问题,模型需要基于知识图谱已有知识进行推理。作为多跳知识问答的关键技术支撑,推理为知识图谱智能问答更快落地实际应用发挥了重要作用。

本文总结多跳知识问答推理的最新研究方法,并将这些方法分为基于嵌入的多跳知识问答推理、基于路径的多跳知识问答推理和基于逻辑的多跳知识问答推理3类方法。其中,基于嵌入的多跳知识问答推理方法是将知识图谱元素(如实体、关系等)关联到低维连续向量空间^[9],然后通过定义得分函数或解码器对目标查询对象进行排名来得到答案。而基于路径的多跳知

识问答推理方法首先需要确定问题中的主题实体,然后在知识图谱上随机游走找到答案实体^[10-11],代表性的工作有路径排序算法(path ranking approach, PRA)。关于基于逻辑的多跳知识问答推理方法,本文将重点介绍以一阶逻辑为主的热点研究方法。图1所示为基于知识图谱的多跳问答推理方法分类,接下来综述这些分类中最新方法的研究进展、基本思路和存在的挑战等。

2 基于嵌入的多跳知识问答推理

基于嵌入的多跳知识问答推理方法是将问题和候选答案转化为公共向量空间中的语义向量表示来进行操作的方法,基于嵌入的具体操作方法可以分为基于语义匹配的方法、基于图神经网络的方法和基于记忆网络的方法。

(1) 基于语义匹配的方法

本类方法首先计算问题和候选答案分布式表示之间的语义匹配,然后通过排序候选答案来得到最终答案。Bordes A等人^[12]

将问题和知识图谱三元组用嵌入的方式表示来表达特征的语义。然而,与翻译模型TransE^[13]、TransH^[14]、TransR^[15]等关注嵌入表示的模型类似,这些方法只能回答简单问题^[12,16-17]。为了实现多跳问答推理以及应对多跳知识高效建模的挑战,很多基于语义匹配的方法被提出。Bordes A等人^[18]在原来三元组表示方法^[12]的基础上做了改进,通过对知识图谱中的问答路径和周围子图进行编码得到语义更加丰富的表示来推理得到答案。Dong L等人^[19]提出的多列卷积神经网络(multi-column convolutional neural network, MC-CNN)模型进一步地利用具有更强学习能力的神经网络模型来学习答案路径、答案背景信息以及答案类型的分布式表示,并理解问题,在不使用手动特征及词库等的条件下,在问答数据集WebQuestions^[20]上取得了不错的结果。其中,答案路径是答案节点和被询问实体之间的一组关系;答案上下文指的是连接到答案路径的单跳实体和关系;答案类型是如人名、日期等的类型。Hao Y C等人^[21]认为MC-CNN模型没有充分并合理地考虑候选答案的相关信

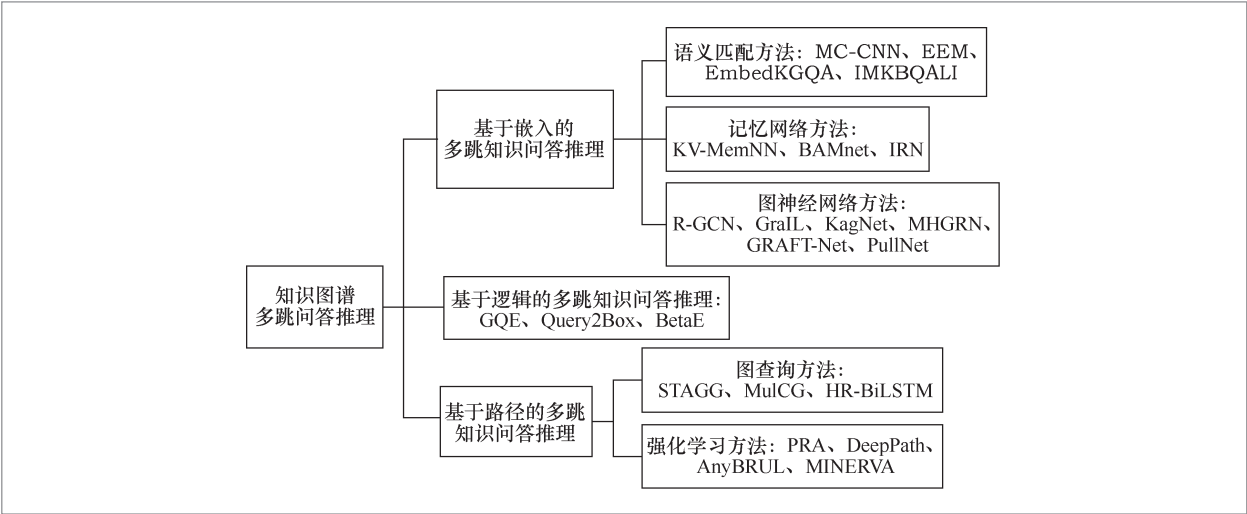


图1 基于知识图谱的多跳问答推理方法分类

息来训练问题嵌入的表示,提出了Cross-Attention机制的神经网络模型来学习知识图谱的全局信息,取得了更好的结果。但是以上几种方法只能完成浅层多跳知识问答推理,对复杂长路径多跳问题的处理能力依然不足。

Saxena A等人^[22]提出的EmbedKGQA模型通过基于知识图谱嵌入模型进行链接预测来缓解多跳问答面临的数据不完整问题,使其具有可以在复杂长路径上的多跳推理能力。EmbedKGQA模型使用ComplEx模型^[23]将知识图谱中的实体和关系嵌入复数向量空间,同时采用ComplEx的打分函数 ϕ 预测答案。具体而言,对于一个给定的问题 q ,首先使用RoBERTa^[24]模型编码初始向量,然后通过一个前馈神经网络将该向量表示投射到复数嵌入向量空间。问题 q 和其主题实体 h 以及知识图谱中的任一实体 a 可以构成三元组,其嵌入向量分别表示为 q 、 h 和 a ,如果 a 是 q 的目标答案实体,则将 (h,q,a) 视为正样本,并且使 $\phi(h,q,a)>0$,否则将 (h,q,a) 视为负样本,并使 $\phi(h,q,a)<0$,负样本可通过将正样本中的答案实体替换为知识图谱中其他非答案实体来获得。EmbedKGQA使用大量的正负样本训练数据学习问题和实体的嵌入向量表示,在推理阶段则在嵌入空间中通过打分函数选择得分最高的实体作为可能的目标答案。

He G L等人^[25]认为多跳知识问答推理算法只接收最终答案的反馈会使学习不稳定或无效,学习推理过程中的监督信号也非常重要,同时也能提升模型的可解释性。由此,He G L等人提出了一种创新的Teacher-Student模型。Teacher-Student框架最早由Hinton G^[26]等人提出,用来做知识蒸馏,其中复杂的Teacher模型的预测被视为“软标签”,一个轻量级Student模型被用于训练拟合软标签。后

来,一些Teacher-Student框架的研究逐渐被应用到问答任务中^[27]来加快模型的推理速度。在He G L等人^[25]提出的模型中,Student网络的目标是找到问题的正确答案,而Teacher网络试图学习预测过程中的监督信号,以提高学生网络的推理能力。Teacher网络利用了正向和逆向双向推理产生可靠的中间监督信号来增强中间实体分布表示学习。在3个公开的数据集上证明了该模型的有效性。

基于语义匹配的方法多采用弱监督学习,减少了人工标注的工作量,同时该方法能够反映出用户问题和知识图谱知识的相似语义,为多跳推理问答的实现做了很好的基础。如何让基于语义匹配的方法学到更丰富和全面的知识图谱信息是这种方法的关键。另外,增强表示方法的可解释性也是重要的考虑因素。

(2) 基于图神经网络的方法

图卷积神经网络(graph convolutional network, GCN)^[28]通过聚合知识图谱中每个实体的邻居实体来传递消息。GCN因为其在图结构上的有效性和可伸缩性而受到广泛利用,它是多跳推理方法采用的复杂图神经网络的基础,如图2所示。Schlichtkrull M等人^[29]提出关系图卷积神经网络(relational graph convolutional network, R-GCN),最早将图卷积神经网络用于知识图谱链接预测的研究任务,通过聚合特定关系来扩展GCN,使其适用于编码多关系图来预测多跳问题的答案。R-GCN模型整体是一个encoder-decoder架构,encoder通过图卷积神经网络将实体的邻域信息进行聚合,更新实体的表示;decoder基于更新后的实体表示,采用打分函数来预测边。然而,该模型缺乏透明度,无法为关系选择提供可解释的依据。相比R-GCN,Teru K等人^[30]提出的GraIL框架采用一种基于注意力机制的多

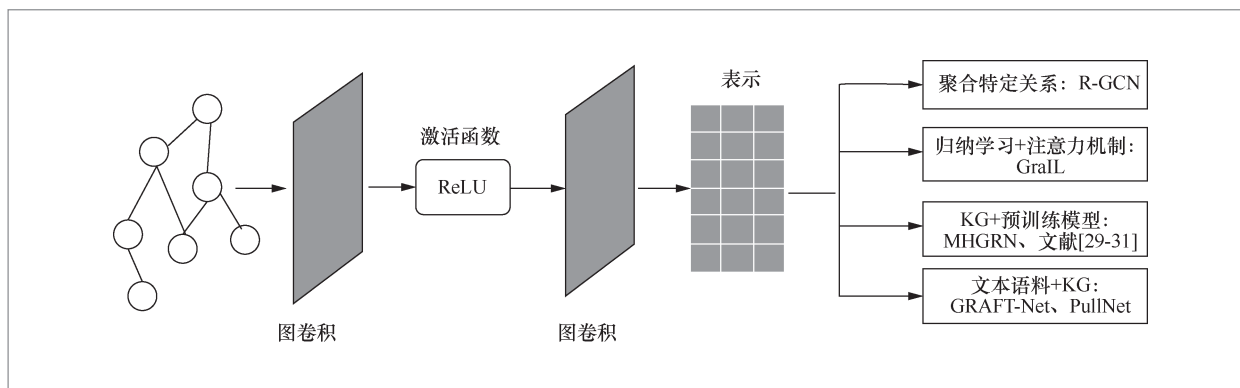


图2 图卷积神经网络

关系图神经网络模型来解决在知识图谱上进行归纳式关系预测的问题。该注意力机制不仅和两个相邻实体以及它们之间的关系有关,也和需要被预测的目标关系有关。最终利用两个目标实体的表示、子图的表示以及预测关系的表示进行打分,将最高得分作为预测结果。归纳学习方法GraIL显式编码知识图谱中的规则以及利用的注意力机制增强了模型的可解释性。

近年来, BERT (bidirectional encoder representations from transformers)^[31]、RoBERTa等预训练语言模型在智能问答等自然语言处理 (natural language processing, NLP) 任务中取得了巨大的成功。但是预训练模型的知识是隐式学习的,无法明确表示出来,因此无法提供可解释的预测。为了利用预训练模型学习到的丰富知识,并且结合知识图谱中的显式知识提高可解释性,许多预训练模型结合知识图谱的问答模型^[32-34]被提出, Feng Y L等人^[33]提出的多跳图关系网络 (multi-hop graph relation network, MHGRN) 模型就是其中一种。MHGRN模型是一种新颖的多跳图关系网络模型,它结合了基于路径的模型,具有可解释性和基于GNN模型扩展性强的优点,利用图神经网络通过在实体之

间传递消息来编码结构化信息,同时为了进一步使模型具有显式建模关系路径的能力,将图分解为路径,并类似Lin B Y等人^[32]提出的知识感知型图神经网络 (knowledge-aware graph network, KagNet) 采用长短期记忆 (long short-term memory, LSTM) 网络^[35]对限定连接长度的问题实体和答案实体的所有路径进行编码,然后通过注意力机制聚合所有路径嵌入来预测结果。

传统智能问答解决方法是单独从知识图谱或者文本中推断出答案。文本虽然是非结构化的知识,但是可以提供丰富的上下文信息,将文本与知识图谱结构化的优势结合是最新的研究热点。

GRAFT-Net (graphs of relations among facts and text network)^[36]和 PullNet (pull network)^[37]将外部文本语料库和知识图谱结合起来完成多跳问答任务。GRAFT-Net将Wikipedia语料库中的文档和知识图谱中的实体建模作为节点,将实体和文档的链接关系 (句子中包含此实体名词) 以及知识图谱中实体之间存在的关系建模为边,构建问题子图,以执行多跳推理。具体而言, GRAFT-Net从问题出发,首先由问题主题实体链接一些种子实体,再以种子实体为起点,通过个性化页面

排名 (personalized pagerank, PPR)^[38]算法从它们的邻居实体中取出PPR分数最高的几个实体及相关联的边,并将它们都加入问题子图中。同时从文本语料库中检索出5个与问题高度相关的句子,并将可以链接到这些句子的实体一并加入问题子图中。最终问题子图由句子节点、实体节点以及句子和实体之间的链接关系、实体和实体之间的关系构成。GRAFT-Net将实体节点随机初始化为一个固定长度的向量,然后基于图卷积神经网络学习和更新问题子图中节点的向量表示,最后对问题子图中的实体节点进行二分类来预测出哪些是答案实体,以完成推理过程。GRAFT-Net使用启发式算法构建的问题子图规模过大,并且很多时候可能不包含答案。PullNet同样基于图卷积网络,但不同于GRAFT-Net, PullNet不要求将整个子图一步构建完成,而是迭代式学习构建的过程。PullNet迭代地构建问题子图,初始子图只包括问题及其中的实体,每一次迭代时,首先使用图卷积网络计算将子图中的实体节点加入下一个迭代过程的概率,确定所有输出概率大于设定阈值的实体,然后对每一个被选择的实体,从文本语料库检索出相关的句子集合,从知识图谱检索出相关的三元组集合,将新检索到的句子、三元组、句子中的实体以及三元组的头实体和尾实体都视为新的节点,将新节点之间的联系视为新边,用新节点和新边更新问题子图。子图构建完成后, PullNet使用与GRAFT-Net相同的方法对节点进行分类,找出最可能的答案实体。

基于图神经网络的方法在处理知识图谱数据时具有更强的表达力,能够充分考虑问题中主题实体在图谱中的相关实体和关系信息,但是依然存在实体邻居信息稀疏时语义表示效果较差的挑战。另外,随着实体和关系不断加入子图中,计算成本

和内存消耗会迅速增加。尽管基于注意力机制的图神经网络能够自适应地学习邻居的重要性权重,但是依然不能完全解决此类问题。

(3) 基于记忆网络的方法

传统的循环神经网络 (recurrent neural network, RNN)、LSTM网络等深度学习模型使用隐藏层状态作为其记忆模块,但是这种方法产生的记忆力太短程,无法精确记住被转化为稠密向量的长路径知识。Weston J等人^[39]提出了一种可读写的外部记忆模块,联合记忆模块保存场景信息,以实现长期记忆的目标。该方法中的记忆网络包括I (input feature map)、G (generalization)、O (output feature map)、R (response) 4个组件。I用来将输入转化为内部特征向量表示;G用来更新记忆,并插入记忆槽中;O根据新的输入和当前的记忆状态输出特征映射表示;R把组件O的结果转化为想要的输出形式,如文本回答。该模型是本文接下来要介绍的复杂记忆网络多跳知识问答推理方法的基础,如图3所示。

基于Weston J等人^[39]提出的记忆网络模型, Miller A等人^[40]提出了键值记忆网络 (key-value memory network, KV-MemNN) 模型,通过迭代更新存储键值对 (key-value) 的记忆槽来完成多跳推理任务。此处的key指三元组中头实体和关系的组合, value指尾实体。与Weston J等人^[39]提出的记忆网络相比,该模型能够处理结构复杂的知识图谱知识。该模型通过以下过程完成一次迭代更新:首先使用向量表示用户问题和知识图谱三元组,通过用户问题检索知识图谱得到相关子集,并作为键值记忆槽,然后计算每一个key和问题的相关程度评分,根据评分对每个value加权求和,最后用得到的value向量更新问题向量。经过以上 N 步迭代后,对

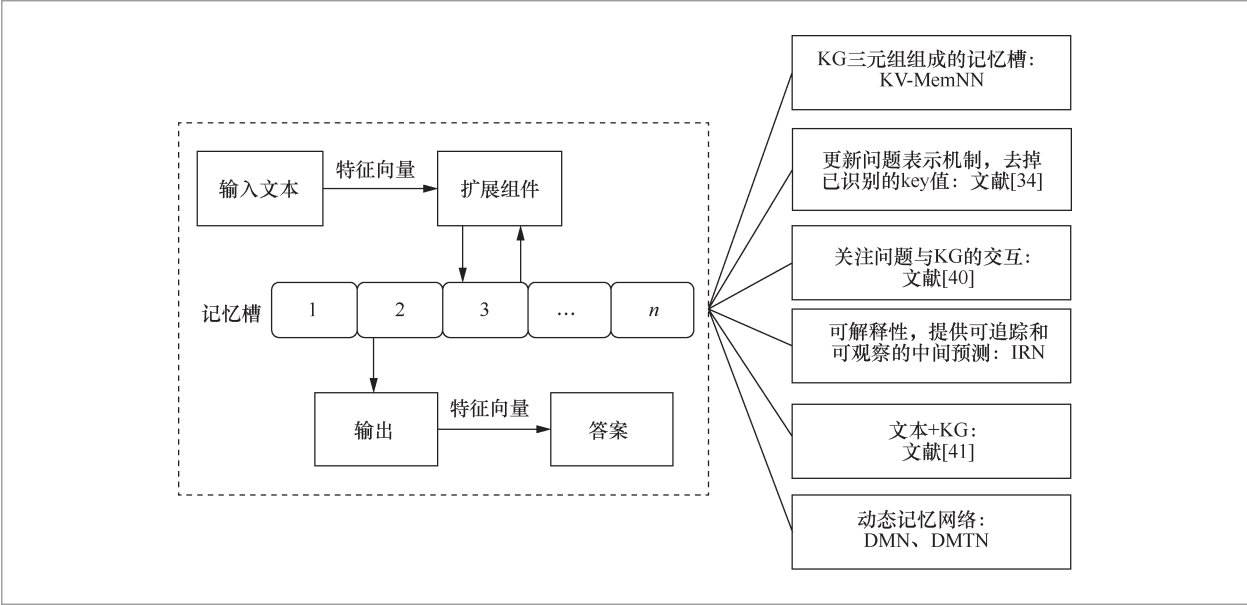


图3 记忆网络模型框架

问题向量进行分类，最终得到多跳问题的答案。KV-MemNN模型结构简单，通用性强，通过向量的迭代更新来进行隐式推理。然而，复杂问题需要进行多次迭代表示，同时需要对应构造记忆槽，容易造成内存不足以及训练时间慢等问题。基于KV-MemNN模型，Xu K等人^[34]提出了一个新的用户问题表示更新机制，即在更新时不考虑问题中已经定位到的key值。该模型将问题分解为针对记忆的查询序列，基于上述更新机制，可以增强多跳复杂问题的推理能力。但是上述方法对问题和知识图谱三元组分开进行编码，忽略了两者的交互作用。因此，Chen Y等人^[41]提出了双向注意记忆网络模型，该模型利用注意机制捕捉问题与知识图谱信息之间的相关性，并利用此相关性增强问题的表征来提高推理答案的能力。

Das R等人^[42]在2017年提出采用记忆网络并基于通用模式在知识图谱和文本上进行多跳知识问答推理，该框架通过将结构化知识图谱和非结构化文本在一个公

共嵌入空间中对齐，相比单独使用知识图谱或文本取得了更好的效果。动态记忆网络(dynamic memory network, DMN)模型和动态记忆张量网络(dynamic memory tensor network, DMTN)模型^[43-44]采用动态记忆网络，允许模型将注意力集中在之前迭代的输入和结果上，形成情景记忆，然后在一个层次递归序列模型中推理得到答案。两者不同的是，DMN采用门函数实现注意力机制，而DMTN采用的是神经张量网络，以实现更好的推理效果。

以上记忆网络模型都取得了不错的效果，但是模型依然具有“黑盒”特性，可解释性差。为了使多跳知识问答推理过程更加可信，一些方法在保证模型准确率的同时也尝试增强模型的可解释性。Zhou M T等人^[45]提出的解释推理网络(interpretable reasoning network, IRN)是一种新颖的具有可解释性的记忆网络推理模型，它采用可解释的逐跳推理过程来回答问题。该模型可以动态地决定输入问题的哪一部分应该在哪一跳进行分

析,预测与当前解析结果相对应的关系,并利用预测的关系更新问题表示和推理过程的状态,然后驱动下一跳推理。该模型可以为推理分析和故障诊断提供可追踪和可观察的中间预测,从而允许人工操作来预测最终答案,这个过程提高了模型的透明度和可信赖度。

3 基于路径的多跳知识问答推理

为了更好地建模多跳知识,基于路径的方法受到了广泛关注。知识图谱中的结构化知识以三元组的形式存在,基于路径的多跳知识问答推理以用户问题中的主题实体为源实体,沿着知识图谱多个三元组的头实体、关系、尾实体的逐跳路径进行搜索,找到答案实体或者关系,通常包括3个阶段:处理输入问题、对知识图谱进行推理、预测答案。以下将基于路径的多跳知识问答推理方法分为基于强化学习的方法、基于查询图的方法。

(1) 基于强化学习的方法

路径排序算法(path-ranking algorithm, PRA)^[46]是一种有效的大规模知识图谱推理路径学习方法。基于路径排序算法的多跳知识问答推理方法的主要思想是利用实体之间的复杂路径特征来学习随机游走器,进而推断出答案。

与PRA基于随机游走的路径查找模型不同,Xiong W H等人^[47]提出的DeepPath是一种创新的可控多跳推理方法。该方法将路径搜索转化为强化学习过程,该过程允许通过控制所找到的路径的属性来减小搜索空间。模型如图4所示,该强化学习方法包括两部分:第一部分是建模为马尔可夫决策过程的外部环境,指定智能体和知识图谱之间的动态交互;第二部分是策略网络智能体,将状态向量映射

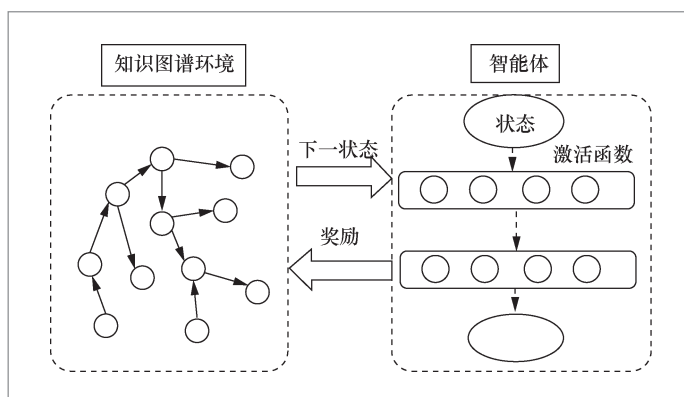


图4 DeepPath模型

到随机策略中。在每一步中,通过与环境交互,智能体学习选择一个关系链接来扩展推理路径。为了更好地指导强化学习智能体学习到目标关系路径,与之前的研究相比,该方法更多地考虑了奖励的设置,在准确性、多样性和精确度3个方面采用奖励机制,监督每一跳的行动,具体方式如下。

在环境设置中,强化学习智能体的决策数量包括正确决策的数量以及错误决策的数量,总体数量很大。其中,错误决策序列的数量会随着路径的长度呈指数增长。鉴于此挑战,DeepPath模型添加的第一个奖励函数定义如下:

$$r_{\text{global}} = \begin{cases} +1, & \text{如果到达目标实体} \\ -1, & \text{其他} \end{cases} \quad (1)$$

相比长路径,短路径通常能提供更多可靠的推理证据。短路径链条通过限制智能体与环境交互的长度,提升推理的效率。因此,DeepPath定义高效奖励如下:

$$r_{\text{efficiency}} = \frac{1}{\text{length}(p)} \quad (2)$$

其中, p 为一个序列关系,如 $r_1 \rightarrow r_2 \rightarrow \dots \rightarrow r_n$ 。

知识图谱中存在很多语义相似的实体,这导致智能体有可能寻找到具有相似语法和语义的路径,这些路径通常包含冗余信息。为了鼓励智能体找到多样化的路径,

DeepPath使用当前路径和已有路径之间的余弦相似度定义一个多样化奖励函数:

$$r_{\text{diversity}} = -\frac{1}{|F|} \sum_{i=1}^{|F|} \cos(\mathbf{p}, \mathbf{p}_i) \quad (3)$$

其中, $\mathbf{p} = \sum_1^n \mathbf{r}_i$, 表示 $r_1 \rightarrow r_2 \rightarrow \dots \rightarrow r_n$ 的嵌入向量, F 表示已经存在的关系路径个数。

Meilicke C等人^[48]提出的AnyBRUL (reinforced anytime bottom up rule learning) 方法同样利用强化学习对路径进行采样, 基于采样路径构造基本规则, 并将其推广为抽象规则。DeepPath和AnyBRUL都要求首先对头实体和尾实体之间的所有路径进行取样, 然后利用它们来评估尾实体是否为正确的答案, 因此, 它无法应对找不到尾实体的情况。与DeepPath和AnyBRUL需要预先计算路径的方式不同, 有些方法通过给定的头实体和查询关系利用训练模型来获得正确的答案实体。在这些模型中, Das R等人^[49]在2018年提出的MINERVA (meandering in networks of entities to reach verisimilar answer) 是一个代表性模型。与DeepPath不同的是, MINERVA的状态由查询关系和部分路径的嵌入组成, 在抽样过程中不需要嵌入答案实体。

强化学习模型通常采用一个0/1的硬奖励来监督抽样过程, 指示抽样实体是否为正确的答案实体。Lin X V等人^[50]提出的RS (reward shaping) 在最后一步采用基于正确答案实体和采样实体之间的软奖励, 而不是使用0/1的硬奖励。此外, 受dropout技术的启发, RS模型在训练过程中为了避免选择大量的重复路径, 缓解过拟合, 采用了dropout技术。Shen Y L等人^[51]提出的M-Walk引入了一种基于值的强化学习方法, 并使用蒙特卡洛树搜索来克服稀疏正奖励的挑战。Chen W H等人^[52]提出的DIVA将推理任务当作一个由寻找

路径和答案推理组成的统一模型, 其中路径建模为隐变量, 采用AEVB (auto-encoding variational Bayes)^[53]对模型进行求解。

基于强化学习的方法把多跳知识问答推理任务当作序列决策过程, 只需一些目标即可得到泛化能力较强的模型, 无须大量的标注语料。但是在面临复杂的多跳知识推理场景时, 智能体的决策能力往往不足。

(2) 基于查询图的方法

基于路径排序的方法容易造成搜索空间爆炸, 而基于查询图的方法是基于路径方法的扩展, 与基于路径的方法生成从主题实体到答案的路径方式不同, 基于查询图的方法使用图来表示问题, 并围绕主题实体扩展出一个子图, 基于查询图的方法比基于路径的方法更具表示能力。

早期的基于语义解析的方法将自然语言问题转换为结构化的查询 (如SPARQL查询), 在知识图谱上执行查询可以得到问题的答案。Reddy S等人^[54]等人充分使用了组合范畴语法 (combinatory categorial grammar, CCG) 的表示能力, 并提出Graph Parser模型来解析问题, 受此启发, Yih W T等人^[20]在2015年定义了查询图的概念, 并提出了一个分阶段的查询图生成 (staged query graph generation, STAGG) 模型来处理知识图谱问答, 查询图可以直接匹配为问题的逻辑形式, 进而翻译成查询, 因此语义解析问题可归结为查询图生成问题。STAGG定义了3个阶段来生成查询图: 首先, 使用现有的实体链接工具获取候选主题实体, 并对其评分; 然后, STAGG探索主题实体和答案节点之间的所有关系路径, 为了限制搜索空间, 它仅在下一跳的节点是一个复合类型 (compound value type, CVT) 节点时, 探索长度为2的路径, 否则只考虑

长度为1的路径,使用深度卷积神经网络对所有关系路径进行打分,以判断当前选择的关系与问题的匹配程度;最后,根据启发式规则将约束节点附加到关系路径上。在这3个阶段的每个阶段,都使用对数线性模型对当前的部分查询图进行评分,并输出最佳的最终查询图来查询知识图谱。STAGG有效地使用了知识图谱中的信息来裁剪语义解析空间,从而简化了任务难度。

针对STAGG无法回答涵盖复杂约束的问题,Bao J W等人^[55]扩展了约束类型和运算符,新增了类型约束以及显式和隐式时间约束等,将多重约束问题转化为多重约束查询图(multi-constraint query graph, MulCG)来实现推理,MulCG仍然遵循STAGG的框架,但提供了更多规则,以应对复杂问题。STAGG和MulCG等方法都要求首先将问题中的候选主题实体链接到知识图谱的实体上,实体链接的质量将影响后续的推理效果。Yu M等人^[56]将知识图谱问答分为实体链接和关系检测两个关键的子任务,并提出了一个残差学习增强的分级的双向LSTM(hierarchical residual bidirectional LSTM, HR-BiLSTM)用于关系检测。对于实体链接,他们观察到,在SimpleQuestions数据集上主题实体识别的Top-1准确率只有72.7%,因此在由实体链接器产生初始的候选主题实体后,又将问题文本输入HR-BiLSTM中,HR-BiLSTM对问题以及在知识图谱中与候选主题实体相关联的关系进行不同抽象级别的编码,并计算两者的相似度得分,仅保留与那些得分较高的关系相关联的候选主题实体。在确定了新的主题实体后,同样使用HR-BiLSTM选择新的关系,以逐步地生成查询。在处理问题约束时,HR-BiLSTM也遵循STAGG中的惯例。

为了限制搜索空间,STAGG只能探索

2跳之内的路径,无法回答涉及更多跳推理的问题。Lan Y S等人^[57]改进了STAGG方法,使其可以应对更长的关系路径,即在扩展路径的同时加入约束,而不是只在建立关系路径之后再添加约束,如此可以有效地缩减搜索空间,改进的分级查询图生成方法可以同时处理包含约束的问答和涉及多跳的问答,该方法在WebQuestionsSP(WebQSP)^[58]系列数据集上取得了优异的表现。

基于路径的多跳知识问答推理方法具有较好的可解释性。但只能处理关系型知识,在处理属性型、事实型知识时,建模复杂。

4 基于逻辑的多跳知识问答推理

基于符号逻辑规则的方法因其准确率高、可解释性强等优点,在多跳知识问答推理研究中被广泛探讨。其中,作为逻辑规则的主要方法,或结合概率逻辑方法,或结合知识嵌入方法,一阶逻辑近几年已成为研究热点。本节重点介绍基于一阶逻辑的多跳知识推理方法的思路。

结合一阶逻辑的自然性和概率逻辑模型的不确定性优点,马尔可夫逻辑网络方法已被证明在知识图推理上的有效性^[59-60]。然而,在大规模知识图谱上由于三元组之间的复杂结构,以上方法推理过程困难,效率较低。基于注意力机制的图神经网络擅长处理高度复杂的图问题,Vardhan V H等人^[61]提出的概率逻辑图注意力网络(probabilistic logic graph attention network, pGAT)用变分EM算法优化了由马尔可夫逻辑网络定义的所有可能三元组的联合分布。这有助于模型有效地结合一阶逻辑和图注意力网络。该推理模型的提出为多跳知识问答提供了有利的参考。

近年来,为了解决知识图谱存在的规模庞大和不完整问题,逻辑规则与知识嵌入相结合的研究^[61-66]受到很多关注。图查询嵌入(graph query embedding, GQE)^[62]、Query2Box^[67]以及BetaE^[63]等方法将查询表示成有向无环的计算图,计算图指明了在知识图谱上进行多跳推理以获得目标答案的步骤。这些方法将一阶逻辑运算符看作可通过训练习得的几何操作,它们都以查询包含的主题实体的嵌入向量为起点,迭代地使用几何操作生成查询的嵌入向量,然后通过向在向量空间中计算实体嵌入与查询嵌入的距离来预测答案。

GQE主要关注一阶逻辑查询的一个子集,即只涉及存在量词和合取运算的查询。GQE提出了两个核心的几何运算符 P (projection)和 I (intersection),其中 P 负责根据下一跳的关系 r 将当前生成的查询嵌入 q 投射为新的查询嵌入向量, I 负责聚集计算图中出边指向同一个节点的所有顶点的嵌入向量表示,以模拟逻辑合取运算。 P 和 I 由式(4)实现:

$$\begin{aligned} P(q, r) &= R_r q \\ I(\{q_1, \dots, q_n\}) &= W_r \Psi(NN_k(q_i), \forall i = \{1, \dots, n\}) \end{aligned} \quad (4)$$

其中, R_r 和 W_r 是关于关系 r 的可训练的参数矩阵, NN_k 是一个 k 层的前馈神经网络, Ψ 是一个对称的向量函数。

GQE将查询嵌入为向量空间中的一个点。Query2Box认为复杂的一阶逻辑查询对应一个答案实体集合,但目前一个点不能有效地建模一个集合,并且在向量空间中对两个点做逻辑运算也不自然,故Query2Box将查询嵌入为向量空间中的box。box由中心点和偏移两部分组成,答案实体嵌入为一个点应包含在对应的box中,关系嵌入也表示成中心点和偏移项两部分。查询包含的实体节点的嵌入被初始化为偏移项为0的box,对于当前输入的box嵌入 p 和下一

跳关系嵌入 r ,关系嵌入也表示成中心点和偏移项两部分, P 操作通过在向量空间中将 p 和 r 相加产生新的box, I 操作取多个box的交集作为结果,按照计算图中的运算逐步推理,即可生成查询最终的box嵌入。Query2Box还在GQE的基础上增加了析取 V 运算,扩充了逻辑查询的功能。GQE和Query2Box等方法将查询嵌入为闭区域,比如欧几里得空间中的一个点或一个box,而因为对闭区域取补并不能得到闭区域,所以这些方法不能支持逻辑取负运算,并且不能对不确定性进行建模。有别于Query2Box等,BetaE将实体和查询嵌入为 $[0,1]$ 区间上的 β 分布,并为 β 向量定义了概率关系投射、概率交和概率取负3个概率逻辑运算符,以支持任何形式的一阶逻辑查询。对 β 分布的参数取倒数即实现了取负运算,即将高概率密度区域转换为低概率密度区域,反之亦然。对于概率投射,BetaE为每一种关系类型学习一个多层感知器,以将一个 β 向量通过下一跳的关系 r 匹配到另一个 β 向量,BetaE将概率交定义为多个 β 向量的带权积。同样,按照计算图中的推理次序对起始的 β 向量做变换,即可生成查询对应的 β 分布,然后在嵌入空间选择与查询嵌入概率最相近的实体集作为答案。

基于一阶逻辑查询的多跳知识问答推理方法具有严密性、自然性,且准确率高。为了模拟不确定性,一些研究将逻辑运算和概率逻辑结合,或者将逻辑运算转换为嵌入向量空间的几何操作,理论上可以进行任意跳数的推理。但这些模型都不能很好地处理复杂知识查询。

5 实验评测

对多跳知识问答推理模型在现有的数据集上进行评估可以验证模型的性能表现。

目前常用的自然语言问答数据集有MetaQA、WebQuestionsSP和ComplexWebQuestions (CWQ) 等, 另外, 链接预测领域常用的数据集 (如FB15k和FB15k-237等) 也被用来构造查询形式的问答, 实现在知识图谱上的推理。本节将介绍这些常用的评测数据集及评测指标, 并对前文所述方法在以上数据集上的表现进行了对比分析。

5.1 评测数据集

(1) MetaQA数据集

MetaQA数据集^[68]是基于WikiMovies^[40]数据集构建的多跳问答数据集, 它包含超过40万个电影领域的多跳问题, 这些问题有Vanilla、NTM和Audio 3个版本。Vanilla版本的MetaQA常被用于多跳知识问答推理任务, 它除了包含1跳、2跳和3跳3种类型的问答数据, 还包含一个知识图谱, 其有约135 000个三元组、43 000个实体以及9种关系。

(2) WebQuestionsSP数据集

Yih W T等人^[58]对WebQuestions数据集改进后提出了WebQuestionsSP数据集。WebQuestions是为了解决真实问题而构造的数据集, 其问题来源于谷歌建议应用程序接口 (Google suggest API), 答案由众包平台Amazon Mechanic Turk通过人工标注生成。WebQuestions只包含了问题的答案, 没有提供问题对应的查询语句, 针对此缺点, WebQSP为每个问题构造了其对应的SPARQL查询表达, 并删除了一些表达有歧义以及无清晰意图或答案的问题。WebQSP包含4 737个1跳或2跳问题, 问题的答案可以在Freebase^[69]知识库中找到。

(3) ComplexWebQuestions数据集

Talmor A等人^[70]基于WebQSP构建了ComplexWebQuestions数据集。首先从WebQSP中采样问题及其SPARQL查询,

并自动地构造更复杂的包含组合、连接、比较级以及最高级等形式的SPARQL查询, 最后由Amazon Mechanic Turk众包平台将这些SPARQL查询重组为自然语言问题, 问题的答案通过在Freebase中执行SPARQL查询获得。CWQ共包含34 689个问题及其对应的答案和SPARQL查询。

(4) 其他数据集

FB15k^[14]是Freebase的一个子集, 也是知识图谱补全领域的一个基准数据集, 由三元组构成。为了解决FB15k测试集数据泄露的问题, Toutanova K等人^[71]在FB15k的基础上构建了FB15k-237, 并且移除了FB15k中反向的关系。NELL995^[47]数据集由NELL系统构建而来。

5.2 评测指标

Hits@1、F1分数、平均倒数排名 (mean reciprocal rank, MRR) 和平均精度均值 (mean average precision, MAP) 是常用来评估多跳知识问答模型效果的指标。把问答视为问题答案检索任务, Hits@1是指正确答案是最终排序的第一个的占比。F1分数是指检索到的答案的精确率和召回率的调和平均, 取在所有问题上的F1分数的平均作为模型最终的F1结果。MRR指所有正确答案实体排名的倒数的均值。MAP指对所有查询的平均正确率 (average precision, AP) 取均值的结果。

5.3 实验结果及分析

表1列出了部分多跳知识推理方法在相同实验数据集 (MetaQA、WebQSP和CWQ) 上Hits@1的结果^[25], “-”表示在原文献中对应指标的结果未给出, 1-hop、2-hop和3-hop表示由源实体到目标答案实体推理时分别需要经过1跳、2跳和3跳。

表1 模型在 MetaQA、WebQSP、CWQ 上的 Hits@1

模型	MetaQA			WebQSP	CWQ
	1-hop	2-hop	3-hop		
KV-MemNN	96.2%	82.7%	48.9%	46.7%	21.1%
GRAFT-Net	97.0%	94.8%	77.7%	66.4%	32.8%
PullNet	97.0%	99.9%	91.4%	68.1%	45.9%
EmbedKGQA	97.5%	98.8%	94.8%	66.6%	-
NSM _{th}	97.2%	99.9%	98.9%	74.3%	48.8%

由表1可以看出,在MetaQA数据集上,对于单跳问题,几个模型的Hits@1结果相近,而随着推理跳数增加,KV-MemNN和GRAFT-Net两个模型的性能有了明显的下降。相较于KV-MemNN和GRAFT-Net,PullNet在3跳问题上仍然取得了不错的表现,且PullNet在3个数据集上都取得了显著的结果,充分显示了其提出的迭代构建问题子图方法的有效性以及图神经网络在基于知识图谱的多跳推理问题上的适用性。相比KV-MemNN、GRAFT-Net和PullNet,EmbedKGQA在MetaQA数据集的平均Hits@1值更高,在WebQSP数据集上也表现突出,证明了EmbedKGQA基于语义匹配的方法在解决多跳问答问题上的有效性。NSM_{th}^[25]总体上优于表1中的其他所有模型,表明了使用Teacher-Student网络模型学习中间监督信号可以很好地提升推理能力。

表2列出了基于逻辑的模型(GQE、Q2B(Query2Box)、BetaE)在FB15k、FB15k-237和NELL-995 3个数据集上多跳推理的结果^[63]。

对比它们的结果可以看出,对于表2中所有数据集,Q2B在1-hop、2-hop、3-hop问题上的表现均优于GQE,证明了在理论上将查询(问题)嵌入为向量空间中的一个box而不是一个点的思想更合理,而且在实际中也能取得更好的结果。3个模

型中BetaE在多跳推理任务上的平均MRR最大,显示了概率嵌入在逻辑查询领域相对于传统方法的有效性。

表3列出了部分基于强化学习的多跳知识问答方法^[48-51]在FB15k、FB15k-237和NELL-995 3个数据集上的结果,采用的评测指标有Hits@1和MAP,“-”表示在原文献中对应指标的结果未给出。由表3可以看出,在FB15k-237数据集上,RewardShaping^[50]的Hits@1结果比MINERVA高出了11%,这很可能因为MINERVA中的智能体在训练时会受到假负样本的影响,导致低质量的奖励,并且由于缺乏正确的动作序列用于训练,智能体可能被假搜索轨迹误导却偶然地被引向了正确答案。RewardShaping针对以上问题做了改进,使智能体探索更广泛的路径集合,以抵消对虚假路径的敏感度。但在NELL-995数据集上,RewardShaping模型的表现并不强于MINERVA,因为RewardShaping适合处理一对多的关系类型,而NELL-995数据集中的关系类型大多是一对一的。从MAP结果来看,DIVA比DeepPath效果要好,这是因为DeepPath关注于找到实体对间的路径,缺少了对实体对是正样本还是负样本的判断,导致实验结果对噪声和对抗样本敏感,而DIVA更擅长处理噪声和应对更复杂的推理场景。M-Walk使用RNN编码状态,并将蒙特卡罗搜索树与神经策略结合来应对稀疏奖励的挑战,与DIVA、DeepPath相比,取得了更好的结果。

6 结束语

近年来,国内外涌现了很多基于知识图谱的多跳问答推理方法,这些方法都取得了一定的进展,但仍然存在一些

挑战。

随着KGQA技术的不断发展,智能问答越来越多地深入人们的日常生活中,为人们的生活带来便利。而在复杂的实际场景中,人们更加倾向表达更加复杂的多跳问题,这使得多跳知识智能问答成为必需,知识图谱多跳问答推理方法也得到越来越多的关注。本文基于3个主流分类,总结了近年来多跳知识问答推理的研究方法。本文中很多方法是多种方式结合的形式,而分类方法是按照具体研究方法的重点创新和贡献来划分的。比如基于逻辑的分类中,BetaE方法虽然将知识图谱实体表示为嵌入的方式,但是该方法表达的是一阶逻辑求解答案的思想,是神经网络增强符号逻辑的一种方法体现,故本文将该方法归类为基于逻辑的方法。本文也阐述了常用的公开数据集、评价指标,以及不同方法在每个数据集上的最新表现和分析。

本文阐述的方法在实际应用中依然存在瓶颈,主要面临以下3个挑战。

- 多跳知识问答建模问题:难以充分理解用户复杂多跳问题以及建模知识图谱中的多跳关系。
- 小样本学习问题:在知识资源和训练数据有限的条件下,难以有效表征知识和推理答案。
- 可解释性:深度学习在多跳知识问答任务中有卓越的表现,但是可解释性差,导致模型缺少可信度。

表2 GQE、Q2B、BetaE 的 MRR 结果

数据集	模型	1-hop	2-hop	3-hop
FB15k	BetaE	65.1%	25.7%	24.7%
	Q2B	68.0%	21.0%	14.2%
	GQE	54.6%	15.3%	10.8%
FB15k-237	BetaE	39.0%	10.9%	10.0%
	Q2B	40.6%	9.4%	6.8%
	GQE	35.0%	7.2%	5.3%
NELL-995	BetaE	53.0%	13.0%	11.4%
	Q2B	42.2%	14.0%	11.2%
	GQE	32.8%	11.9%	9.6%

表3 基于强化学习的模型的结果

指标	方法	FB15k	FB15k-237	NELL-995
Hits@1	MINERVA	—	21.7%	66.3%
	RewardShaping	—	32.7%	65.6%
MAP	DeepPath	57.2%	—	79.6%
	DIVA	59.8%	—	88.6%
	M-Walk	—	—	89.9%

各类多跳知识问答推理方法的优点和不足见表4。

未来,将从以下两方面进行更深入的研究。

- 随着BERT等预训练模型的出现,多项自然语言处理任务(如智能问答的性能)被刷新纪录。随着各类基于Transformer注意力机制的预训练模型层出不穷,很多研究者看好预训练模型和知识图谱结合的方法。通过两者结合,知识图谱中的知识可以显式指导要学习的知识,进而增强BERT对背景知识或常识信

表4 基于知识图谱的多跳问答推理方法优缺点

多跳知识问答推理方法	优点	不足
基于嵌入的方法	减少对手工定义模板或规则的依赖;嵌入具有不确定语义	可解释性方面存在欠缺;知识图谱中实体邻居信息稀疏时,语义表示效果差
基于路径的方法	具有较好的可解释性	处理属性型、事实型知识时,建模复杂
基于逻辑的方法	具有严密性、自然性、准确率高等特点	不能模拟不确定性;将逻辑运算转换为嵌入向量空间的几何操作不能很好地处理复杂知识查询

息的编码能力,同时文本上下文和知识上下文的融合能够增强多跳推理中长路径的特征表示,从而提升多跳知识问答推理任务的表现性能。

- 越来越多的互联网数据以及行业数据呈现出多模态的特点,表现为对同一事物采用图谱知识、文本描述、图像、信息图表以及视频等多种模态进行表达。不久的将来将有更多关于多模态知识图谱多跳问答的研究,满足人们日益增长的理解和认知服务需求。多模态知识图谱能够提升知识的相关性和多样性,从而更完整地表达对应的图谱中的知识。通过多模态丰富的知识表示,能够进一步保障多模态知识图谱多跳问答推理的效果。

参考文献:

- [1] 王昊奋, 丁军, 胡芳槐, 等. 大规模企业级知识图谱实践综述[J]. 计算机工程, 2020, 46(7): 1-13.
WANG H F, DING J, HU F H, et al. Survey on large scale enterprise-level knowledge graph practices[J]. Computer Engineering, 2020, 46(7): 1-13.
- [2] 邱楠, 王昊奋, 邵浩. 从聊天机器人到虚拟生命-人工智能技术的新机遇[J]. 中国人工智能学会通讯, 2017, 11(7): 32-40.
QIU N, WANG H F, SHAO H. From chatbots to virtual life - a new opportunity for artificial intelligence technology[J]. Communications of the CAAI, 2017, 11(7): 32-40.
- [3] 陈成, 陈跃国, 刘宸, 等. 意图知识图谱的构建与应用[J]. 大数据, 2020, 6(2): 57-68.
CHEN C, CHEN Y G, LIU C, et al. Constructing and analyzing intention knowledge graphs[J]. Big Data Research, 2020, 6(2): 57-68.
- [4] WU T X, WANG H F, LI C, et al. Knowledge graph construction from multiple online encyclopedias[J]. World Wide Web, 2020, 23(5): 2671-2698.
- [5] WANG M, WANG H F, QI G L, et al. Richpedia: a large-scale, comprehensive multi-modal knowledge graph[J]. Big Data Research, 2020, 22: 100159.
- [6] 邹艳珍, 王敏, 谢冰, 等. 基于大数据的软件项目知识图谱构造及问答方法[J]. 大数据, 2021, 7(1): 22-36.
ZOU Y Z, WANG M, XIE B, et al. Software knowledge graph construction and Q&A technology based on big data[J]. Big Data Research, 2021, 7(1): 22-36.
- [7] 官赛萍, 靳小龙, 贾岩涛, 等. 面向知识图谱的知识推理研究进展[J]. 软件学报, 2018, 29(10): 2966-2994.
GUAN S P, JIN X L, JIA Y T, et al. Knowledge reasoning over knowledge graph: a survey[J]. Journal of Software, 2018, 29(10): 2966-2994.
- [8] 王昊奋, 奚宁, 周扬. 企业计算中的机器人智能问答[J]. 中国计算机学会通讯, 2019, 15(5): 28-37.
WANG H F, XI N, ZHOU Y. Intelligent question answering in enterprise computing[J]. Communications of the CCF, 2019, 15(5): 28-37.
- [9] WANG M, WANG R J, LIU J, et al. Towards empty answers in SPARQL: approximating querying with RDF embedding[C]//Proceedings of the 17th International Semantic Web Conference. Cham: Springer, 2018: 513-529.
- [10] ZHU C H, REN K, LIU X, et al. A graph traversal based approach to answer non-aggregation questions over DBpedia[C]//Proceedings of the 5th Joint International Semantic Technology Conference. Cham: Springer, 2015: 219-234.
- [11] WANG H F, LIU Q L, PENIN T, et al. Semplore: a scalable IR approach to search the Web of data[J]. Web Semantics: Science, Services and Agents on the World Wide Web, 2009, 7(3): 177-188.
- [12] BORDES A, WESTON J, USUNIER N.

- Open question answering with weakly supervised embedding models[J]. Lecture Notes in Computer Science, 2014, 8724(PART 1): 165–180.
- [13] BORDES A, USUNIER N, GARCIA-DURAN A, et al. Translating embeddings for modeling multi-relational data[C]// Proceedings of the 26th International Conference on Neural Information Processing Systems. New York: ACM Press, 2013: 2787–2795.
- [14] WANG Z, ZHANG J W, FENG J L, et al. Knowledge graph embedding by translating on hyperplanes[C]// Proceedings of the 28th AAAI Conference on Artificial Intelligence. [S.l.:s.n.], 2014: 1112–1119.
- [15] LIN Y K, LIU Z Y, SUN M S, et al. Learning entity and relation embeddings for knowledge graph completion[C]// Proceedings of the 29th AAAI Conference on Artificial Intelligence. [S.l.:s.n.], 2015.
- [16] LI D C, ZHANG J Y, LI P. Representation learning for question classification via topic sparse autoencoder and entity embedding[C]// Proceedings of the 2018 IEEE International Conference on Big Data. Piscataway: IEEE Press, 2019: 126–133.
- [17] HUANG X, ZHANG J Y, LI D C, et al. Knowledge graph embedding based question answering[C]// Proceedings of the 12th ACM International Conference on Web Search and Data Mining. [S.l.:s.n.], 2019: 105–113.
- [18] BORDES A, CHOPRA S, WESTON J. Question answering with subgraph embeddings[C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. [S.l.:s.n.], 2014: 615–620.
- [19] DONG L, WEI F R, ZHOU M, et al. Question answering over freebase with multi-column convolutional neural networks[C]// Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. [S.l.:s.n.], 2015: 260–269.
- [20] YIH W T, CHANG M W, HE X D, et al. Semantic parsing via staged query graph generation: question answering with knowledge base[C]// Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. [S.l.:s.n.], 2015: 1321–1331.
- [21] HAO Y C, ZHANG Y Z, LIU K, et al. An end-to-end model for question answering over knowledge base with Cross-Attention combining global knowledge[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. [S.l.:s.n.], 2017: 221–231.
- [22] SAXENA A, TRIPATHI A, TALUKDAR P. Improving multi-hop question answering over knowledge graphs using knowledge base embeddings[C]// Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. [S.l.:s.n.], 2020: 4498–4507.
- [23] TROUILLON T, WELBL J, RIEDEL S, et al. Complex embeddings for simple link prediction[C]// Proceedings of the International Conference on Machine Learning. [S.l.:s.n.], 2016: 2071–2080.
- [24] LIU Y H, OTT M, GOYAL N, et al. RoBERTa: a robustly optimized bert pretraining approach[J]. arXiv preprint, 2019, arXiv:1907.11692.
- [25] HE G L, LAN Y S, JIANG J, et al. Improving multi-hop knowledge base question answering by learning intermediate supervision signals[J]. arXiv preprint, 2021, arXiv:2101.03737v1.
- [26] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[J]. arXiv preprint, 2015, arXiv:1503.02531.

- [27] YANG Z, SHOU L J, GONG M, et al. Model compression with two-stage multi-teacher knowledge distillation for web question answering system[C]//Proceedings of the ACM 13th International Conference on Web Search and Data Mining. New York: ACM Press, 2020: 690–698.
- [28] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[C]//Proceedings of 5th International Conference on Learning Representations. [S.l.:s.n.], 2017.
- [29] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks[C]//The Semantic Web: 15th International Conference. [S.l.:s.n.], 2018: 593–607.
- [30] TERU K, DENIS E, HAMILTON W. Inductive relation prediction by subgraph reasoning[C]//Proceedings of the 37th International Conference on Machine Learning. [S.l.:s.n.], 2020: 9448–9457.
- [31] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. [S.l.:s.n.], 2019: 4171–4186.
- [32] LIN B Y, CHEN X Y, CHEN J, et al. KagNet: knowledge-aware graph networks for commonsense reasoning[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. [S.l.:s.n.], 2019: 2829–2839.
- [33] FENG Y L, CHEN X Y, LIN B Y, et al. Scalable multi-hop relational reasoning for knowledge-aware question answering[C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing. [S.l.:s.n.], 2020: 1295–1309.
- [34] XU K, LAI Y X, FENG Y S, et al. Enhancing key-value memory neural networks for knowledge based question answering[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. [S.l.:s.n.], 2019: 2937–2947.
- [35] GREFF K, R K SRIVASTAVA, KOUTNÍK J, et al. LSTM: a search space odyssey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 28(10): 2222–2232.
- [36] SUN H T, DHINGRA B, ZAHEER M, et al. Open domain question answering using early fusion of knowledge bases and text[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL Press, 2018: 4321–4242.
- [37] SUN H T, BEDRAX-WEISS T, COHEN W W. PullNet: open domain question answering with iterative retrieval on knowledge bases and text[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing. [S.l.:s.n.], 2019: 2380–2390.
- [38] HAVELIWALA T H. Topic-Sensitive Pagerank: a context-sensitive ranking algorithm for web search[J]. IEEE Transactions on Knowledge and Data Engineering, 2003, 15(4): 784–796.
- [39] WESTON J, CHOPRA S, BORDES A. Memory networks[C]//Proceedings of the 3rd International Conference on Learning Representations. [S.l.:s.n.], 2015: 1–15.
- [40] MILLER A, FISCH A, DODGE J, et al. Key-value memory networks for directly reading documents[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. [S.l.:s.n.], 2016: 1400–1409.
- [41] CHEN Y, WU L F, ZAKI M J. Bidirectional attentive memory networks for question

- answering over knowledge bases[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. [S.l.:s.n.], 2019: 2913–2923.
- [42] DAS R, ZAHEER M, REDDY S, et al. Question answering on knowledge bases and text using universal schema and memory networks[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. [S.l.:s.n.], 2017: 358–365.
- [43] KUMAR A, IRSOY O, ONDRUSKA P, et al. Ask me anything: dynamic memory networks for natural language processing[C]//Proceedings of the 33rd International Conference on Machine Learning. New York: ACM Press, 2016: 2068–2078.
- [44] RAMACHANDRAN G S, SOHMSHETTY A. Ask me even more: dynamic memory tensor networks (extended model)[J]. arXiv preprint, 2017, arXiv:1703.03939v1.
- [45] ZHOU M T, HUANG M L, ZHU X Y. An interpretable reasoning network for multi-relation question answering[C]//Proceedings of the 27th International Conference on Computational Linguistics. [S.l.:s.n.], 2018: 2010–2022.
- [46] GARDNER M, TALUKDAR P, KISIEL B, et al. Improving learning and inference in a large knowledge-base using latent syntactic cues[C]//Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. [S.l.:s.n.], 2013: 833–838.
- [47] XIONG W H, HOANG T, WANG W Y. DeepPath: a reinforcement learning method for knowledge graph reasoning[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. [S.l.:s.n.], 2017: 564–573.
- [48] MEILICKE C, CHEKOL M W, FINK M, et al. Reinforced anytime bottom up rule learning for knowledge graph completion[J]. arXiv preprint, 2020, arXiv:2004.04412.
- [49] DAS R, DHULIAWALA S, ZAHEER M, et al. Go for a walk and arrive at the answer: reasoning over paths in knowledge bases using reinforcement learning[C]//Proceedings of the 6th International Conference on Learning Representations. [S.l.:s.n.], 2018.
- [50] LIN X V, SOCHER R, XIONG C M. Multi-hop knowledge graph reasoning with reward shaping[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. [S.l.:s.n.], 2018: 3243–3253.
- [51] SHEN Y L, CHEN J S, HUANG P S, et al. M-Walk: learning to walk over graphs using monte carlo tree search[C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. [S.l.:s.n.], 2018: 6786–6797.
- [52] CHEN W H, XIONG W H, YAN X F, et al. Variational knowledge graph reasoning[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. [S.l.:s.n.], 2018: 1823–1832.
- [53] KINGMA D P, WELING M. Auto-encoding variational Bayes[C]//Proceedings of the 2nd International Conference on Learning Representations. [S.l.:s.n.], 2014: 1–14.
- [54] REDDY S, LAPATA M, STEEDMAN M. Large-scale semantic parsing without question-answer pairs[J]. Transactions of the Association for Computational Linguistics, 2014, 2: 377–392.
- [55] BAO J W, DUAN N, YAN Z, et al. Constraint-based question answering with knowledge graph[C]//Proceedings

- of the 26th International Conference on Computational Linguistics. [S.l.:s.n.], 2016: 2503–2514.
- [56] YU M, YIN W P, HASAN K S, et al. Improved neural relation detection for knowledge base question answering[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. [S.l.:s.n.], 2017: 571–581.
- [57] LAN Y S, JIANG J. Query graph generation for answering multi-hop complex questions from knowledge bases[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. [S.l.:s.n.], 2020: 969–974.
- [58] YIH W T, RICHARDSON M, MEEK C, et al. The value of semantic parse labeling for knowledge base question answering[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL Press, 2016: 201–206.
- [59] RICHARDSON M, DOMINGOS P. Markov logic networks[J]. Machine Learning, 2006, 62(1–2): 107–136.
- [60] SINGLA P, DOMINGOS P. Discriminative training of Markov logic networks[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2005: 868–873.
- [61] VARDHAN L V H, JIA G, KOK S. Probabilistic logic graph attention networks for reasoning[C]//Proceedings of the Web Conference 2020. [S.l.:s.n.], 2020: 669–673.
- [62] HAMILTON W, BAJAJ P, ZITNIK M, et al. Embedding logical queries on knowledge graphs[C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. [S.l.:s.n.], 2018: 2030–2041.
- [63] REN H Y, LESKOVEC J. Beta embeddings for multi-hop logical reasoning in knowledge graphs[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. [S.l.:s.n.], 2020: 19716–19726.
- [64] DING B Y, WANG Q, WANG B, et al. Improving knowledge graph embedding using simple constraints[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. [S.l.:s.n.], 2018: 110–121.
- [65] GUO S, LI L, HUI Z, et al. Knowledge graph embedding preserving soft logical regularity[C]//Proceedings of the 29th ACM International Conference on Information & Knowledge Management. New York: ACM Press, 2020: 425–434.
- [66] QU M, TANG J. Probabilistic logic neural networks for reasoning[C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems. [S.l.:s.n.], 2019: 7710–7720.
- [67] REN H Y, HU W H, LESKOVEC J. Query2Box: reasoning over knowledge graphs in vector space using box embeddings[C]//Proceedings of the 8th International Conference on Learning Representations. [S.l.:s.n.], 2020.
- [68] ZHANG Y Y, DAI H J, KOZAREVA Z, et al. Variational reasoning for question answering with knowledge graph[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2018.
- [69] BOLLACKER K, EVANS C, PARITOSH P, et al. Freebase: a collaboratively created graph database for structuring human knowledge[C]//Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data. New York: ACM Press, 2008: 1247–1250.
- [70] TALMOR A, BERANT J. The web as a knowledge-base for answering complex questions[C]//Proceedings of the 16th Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. [S.l.:s.n.], 2018: 641–651.

[71] TOUTANOVA K, CHEN D Q. Observed versus latent features for knowledge base and text inference[C]//Proceedings of

the 3rd Workshop on Continuous Vector Space Models and Their Compositionality. [S.l.:s.n.], 2015: 57–66.

作者简介



杜会芳 (1991–), 女, 同济大学设计创意学院博士生, 主要研究方向为知识图谱、智能问答。



王昊奋 (1982–), 男, 同济大学设计创意学院特聘研究员, 中国计算机学会 (CCF) 理事、计算机术语审定委员会副主任、CCF TF SIGKG主席, OpenKG联合创始人, 主要研究方向为知识图谱、自然语言处理、问答对话、智能内容生成。



史英慧 (1998–), 女, 东南大学网络空间与安全学院硕士生, 主要研究方向为知识图谱、多模态数据。



王萌 (1989–), 男, 博士, 东南大学计算机科学与工程学院讲师, CCF会员, 东南大学“至善青年学者”支持计划获得者, 主要研究方向为知识图谱、多模态数据、自然语言处理。

收稿日期: 2021-03-29

通信作者: 王昊奋, carter.whfcarter@gmail.com

基金项目: 中央高校基本科研业务专项资金资助项目 (No.22120210109)

Foundation Item: Fundamental Research Funds for the Central Universities(No.22120210109)