



小型微型计算机系统
Journal of Chinese Computer Systems
ISSN 1000-1220, CN 21-1106/TP

《小型微型计算机系统》网络首发论文

题目: 利用知识图谱的多跳可解释问答
作者: 叶蕾, 张宇迪, 杨旭华
收稿日期: 2023-02-28
网络首发日期: 2023-05-19
引用格式: 叶蕾, 张宇迪, 杨旭华. 利用知识图谱的多跳可解释问答[J/OL]. 小型微型计算机系统. <https://kns.cnki.net/kcms/detail/21.1106.TP.20230519.0923.002.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

利用知识图谱的多跳可解释问答

叶 蕾, 张宇迪, 杨旭华

(浙江工业大学计算机科学与技术学院, 杭州 310023)

E-mail: xhyang@zjut.edu.cn

摘 要: 基于知识图谱的多跳问答需要分析和理解自然语言问题并在知识图谱的实体和关系上经过多次推理获取答案, 是自然语言处理的重要研究领域。现有的模型一般通过知识图谱与问题嵌入, 利用神经网络推断答案; 或使用一阶逻辑规则结合概率方法预测答案; 前者缺乏可解释性, 后者在复杂问题中性能欠佳。为解决上述问题, 本文提出一种基于知识图谱的多跳可解释问答方法 (MIQA), 它通过在实体间的多次跳跃推理来获取答案。MIQA 首先使用 BERT 预训练模型获取自然语言问题表征向量以及问题分词后的词向量矩阵, 在每一跳中, 结合问题向量提取问题当前时刻的特征向量, 根据特征向量的分类结果计算下一跳的关系分数和实体分数, 多次跳跃后, 综合分数最高的实体被作为答案, 而获取该答案所对应的路径为推理路径。该方法推理准确率高, 同时具有明显的可解释性。在 MetaQA、WebQuestionsSP、ComplexWebQuestions 三个数据集上, 通过和其他 8 个知名算法相比较, 仿真结果表明 MIQA 性能优异, 达到了当前的 SOTA。

关键词: 知识图谱; 多跳问答; 可解释性; 特征抽取; 注意力机制

中图分类号: TP182

文献标识码 A

Multi-hop Interpretable Question Answering Using Knowledge Graph

YE Lei, ZHANG Yu-Di, YANG Xu-Hua

(College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, PR China)

Abstract: Multi-hop question answering based on knowledge graphs needs to analyze and understand natural language questions and obtain answers through multiple reasoning on the entities and relationships of knowledge graphs, which is an important research field of natural language processing. Existing models generally use knowledge graphs and question embeddings to infer answers using neural networks; or use first-order logic rules combined with probabilistic methods to predict answers; the former lacks interpretability, and the latter performs poorly in complex problems. In order to solve the above problems, this paper proposes a multi-hop interpretable question answering method (MIQA) based on knowledge graph, which obtains answers by reasoning through multiple jumps between entities. MIQA first uses the BERT pre-training model to obtain the natural language question representation vector and the word vector matrix after the question word segmentation. In each hop, it combines the question vector to extract the feature vector at the current moment of the question, and calculates the next hop based on the classification result of the feature vector. The relationship score and entity score, after multiple jumps, the entity with the highest comprehensive score is taken as the answer, and the path corresponding to the answer is the inference path. The method has high inference accuracy and obvious interpretability. On the three data sets of MetaQA, WebQuestionsSP, and ComplexWebQuestions, compared with other 6 well-known algorithms, the simulation results show that MIQA has excellent performance, reaching the current SOTA.

Key words: knowledge graph; multi-hop question answering; interpretability; feature extraction; attention mechanism

1 引言

知识图谱是一种用图模型来建模自然界万物之间关联关系的技术体系, 广泛应用于软件开发^[1]、数据管理^[2]、问答系统^[3]和商品推荐^[4]等领域。得益于 Freebase^[5]和

DBpedia^[6]等大规模知识图谱的构建, 基于知识图谱的问答 (KGQA) 已经成为当前的研究热点。KGQA 是指从给定的自然语言问题出发, 对其进行语义理解与需求解析, 进而从知识图谱中推理得出答案的任务, 它的核心是如何让机器理解自然语言。KGQA 方法大致可分基于嵌入的方法、基于路

径的方法与基于逻辑的方法,根据回答问题的难易程度每种方法又可细分为单跳问答与多跳问答。

单跳问答是指根据问题多次访问单个三元组即可得到全部正确答案的任务。这种任务通常较为简单,应用范围较小。多跳问答是一种通过在知识图谱上多步推理寻找答案的任务,它需要在多个与问题相关联的三元组间跳跃,每次跳跃前一个三元组的尾实体即为后一个三元组的头实体,最后一个三元组的尾实体即为答案。以2跳问题“苹果的首席执行官史蒂夫·乔布斯的父亲在哪出生?”为例,其推理过程如图1。首先根据问题关键词“苹果”、“首席执行官”选择实体“史蒂夫·乔布斯”,利用关键词“父亲”使其沿着关系“父亲是”跳跃到实体“约翰·钱德里”,最后由关键词“出生”可得到答案“叙利亚”。多跳问答能在复杂语境下推理,具有较大的应用前景。

Question : Where was Apple CEO Steve Jobs' father born?

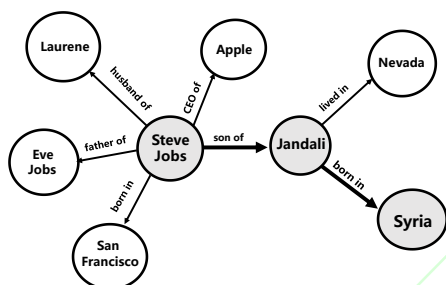


图1 2跳问题的推理过程示意图

Fig 1 Schematic diagram of the reasoning process for the 2-hop problem

基于嵌入的方法通常把词句向量化,通过计算得分函数对实体进行排名来进行推理。例如:模型 EmbedKGQA 模型^[7]将知识图谱嵌入到向量空间,通过链路预测来推断答案。当模型从邻域中选择答案时减少部分选择约束,从而减少了对邻域的依赖,增加了长路径上的推理能力。VRN^[8]在概率框架中加入变分推断,使用逻辑算法优化目标函数下界的同时使用强化学习对模型参数进行调优,依靠问题与答案的组合推理答案。SRN^[9]采用端到端的框架,提出全局语义推理模块,通过多条路径同时推理的方式获得答案。这种方式减少了解码方式对时间的依赖,加快了单向串行数据的传输。Teacher-Student 网络^[10]利用双向推理取得可靠的中间信息来增强实体分布的学习,其中 Student 网络用于寻找答案实体,Teacher 网络用于生成中间信息并指导 Student 网络。基于嵌入的方法精确度比较高,但缺乏解释性。

基于路径的方法首先把自然语言问题的主题实体作为源实体,之后从源实体出发沿着三元组游走寻找答案实体。PRA^[11]将关系路径作为特征训练一个随机游走器,利用它获得关系路径对应的逻辑规则,以便机器发现与利用知识图谱中的隐藏知识。但 PRA 搜索效率不高。DeepPath^[12]提出一种强化学习框架,使用一个基于策略的智能体进行路径搜

索,按最有希望的关系扩展路径,通过控制路径属性来减小搜索空间。DIVA^[13]结合先验算法、后验算法以及推理算法对自然语言问题进行变分推理,将寻找路径与答案推理两个任务融合在一起,用 AEVB^[14]方法对模型进行求解,使模型能更好地处理含有噪音的场景。基于路径的方法对关系型知识处理能力较强,但对于属性型、事实型知识时通常需要复杂的建模并且比较费时。

基于逻辑的方法通常以使用一阶逻辑为主,例如:KB^[15]利用马尔可夫逻辑网络结合概率图形的方法进行推理,它为每个图形域指定一个简单的马尔可夫网络,通过优化损失函数的最大似然估计学习有效权重,依此推理答案。VP^[16]根据给定谓词优化条件似然函数,在求解器中使用加权的方法解析实体并进行链路预测,但这种方法很难处理大型知识图谱复杂的三元组关系。PGAT^[17]利用概率逻辑注意网络推理自然语言问题,通过结合图注意网络,能有效提取邻域节点信息,同时优化了所有可能的三元组的联合分布,从而更准确地预测答案。基于逻辑的方法通常在多关系多实体的复杂语境下表现欠佳。

为解决上述问题,本文提出了多跳问答模型 MIQA。它通过解析自然语言问题来计算推理中下一跳的关系分数和实体分数,经多次推理后,选择综合分数最高的实体作为答案,相应的推理路径为多跳过程中所产生最高分数的路径,因而具有明显的可解释性。

本文的主要贡献如下。

(1) 提出了 MIQA 框架,通过分析自然语言问题特征向量计算推理每一步所涉及的实体和关系分数来获取最合理的推理路径,在实现高效推理的同时具有可解释性;

(2) 设计了一种自然语言问题特征向量提取模式,在推理下一跳过程中,使之和词向量矩阵相结合生成新的问题向量,使模型既可以记忆已有的推理步骤,又可以展开新的下一跳推理;

(3) 设计了一种跳数注意力机制,它为每一跳的实体分数列表分配权重,通过计算推理路径中实体分数和权重得到候选答案综合分数,选取综合分数最高的作为答案。

论文的其余部分如下:第二节为相关工作,第三节描述本文提出的模型,第四节为数值仿真和结果分析,最后对本文工作作出总结并展望未来的研究方向。

2 相关工作

基于知识图谱的多跳可解释问答,本文提出的算法与目前已有算法存在明显不同。本文提出提取自然语言问题特征向量和跳数注意力机制来计算推理路径。以下为相关研究基础。

2.1 知识图谱问答中的语义匹配

语义匹配可以判断两个文本相似程度。目前 KGQA 中的语义匹配方法一般是通过知识图谱与自然语言问题嵌入来计算问题表征与候选答案表征之间的语义匹配分数,将排

名前 n 的实体作为候选答案。根据其使用的嵌入方法可分为基于表示的方法与基于交互的方法。

基于表示的方法通常使用一个 Encoder 模块提取问句的特征信息并编码, 利用得到的语义向量参与接下来的运算。例如: MC-CNN 模型^[18]利用多列卷积神经网络来学习答案路径、上下文信息和答案实体的分布, 但没有充分利用候选答案的相关信息。Cross-Attention 模型^[19]学习知识图谱的全局信息, 使用交叉注意力机制动态更新候选答案的表征, 依此计算实体分数并推测答案, 此模型在短路径预测时有较好表现, 但面对复杂的长路径预测时性能欠佳。 λ -DCS^[20]模型使用 lambda 表达式解析自然语言问题中的语义与逻辑, 可以消除非必要变量以提高模型的专注度, 从而提高准确率。Signature-Based^[21]模型可以寻找从主题实体出发到任一点的最短路径, 利用前 n 个候选路径构造查询树, 时间性能较佳。

基于交互的方法首先对问句分词并提取每个部分的表征, 之后在这些表征中进行交互匹配, 从而更好地把握语义焦点。例如: EmbedKGQA^[7]模型通过链路预测来增加长路径上的推理能力, 使用 ComplEx^[22]模型与 RoBERTa^[23]模型将实体与关系嵌入复数空间进行求解。PST^[24]模型使用期望最大化算法将语义分解到多个模块, 每个模块根据语义匹配结果奖励正样本同时惩罚负样本, 然后给出预测答案。3DSSM^[25]模型使用立体特征网络结构, 将一维的自然语言问题向量嵌入到多维空间, 利用神经网络提取问题向量的局部特征并生成匹配表示, 然后进行下游推理。Teacher-Student^[10]模型使用 Student 网络获取答案实体, Teacher 网络学习推理过程中的信号并指导 Student 网络进行参数修正, 从而提高模型的推理能力。ReifKB^[26]模型利用稀疏矩阵重构符号知识库, 重构任务可以分布到多个 GPU, 具有较高的效率。AMPA 模型^[27]提出一种融合注意力机制和指针标注的实体关系联合抽取方法, 利用多头注意力机制获取问题的语义特征信息, 之后使用指针标注抽取问题的主语与宾语, 通过指针标注处理问题。

2.2 多跳推理

多跳推理的核心是如何依据已有信息找到多个数据结构的内在联系, 并通过推理得到答案。多跳推理可分为基于逻辑规则的方法、基于问题分解的方法和基于图神经网络等方法。基于逻辑规则的方法通过定义一套逻辑规则来预测答案, 例如: 概率逻辑图注意力网络 PGAT^[17]结合一阶逻辑与神经网络, 使用变分算法对所有三元组的分布进行优化, 为多跳推理领域开辟了一个新方向。BetaE 模型^[28]进行实体与关系嵌入, 同时定义了一阶逻辑的关系投影、交集与取负三个操作, 从可以在不同的实体之间实现推理。

基于问题分解的方法将自然语言问题拆分为多个相关子问题, 求解每个子问题获得相应的答案域, 最后从所有答案域的交集中选择答案。例如: DECOMPRC^[29]模型将复杂

自然语言问题分解为多个简单阅读理解问题, 每个简单阅读理解问题都被重构为跨度预测问题, 对多个简单阅读理解问题的答案评分, 选择分数最高的实体作为结果。MAC Network^[30]模型将问题分解为基于注意力的多个推理步骤, 每个步骤都包含新的循环记忆、注意力与合成单元, 通过施加结构约束来调节它们的相互作用, 采用端到端的方法直接从数据中推断出执行过程以获得答案。KVMemNN^[31]提出了一种新的键值记忆网络, 它通过在内存读取操作的寻址和输出阶段使用不同的编码, 使读取文档更加可行。

基于图神经网络的方法通过提取知识图谱中的邻域信息来传递消息, 由此获取实体与关系的向量表征并依此推测答案。例如: R-GCN^[32]模型使用 GCN^[33]构建关系网络, 同时提出权值共享与系数约束的方法, 使模型能完成关系复杂的推理任务。GraftNet 模型^[34]用无结构的文本数据对结构化的知识图谱进行补充, 可以处理文本和知识图谱的混合数据。Pullnet 模型^[35]可以识别子图是否需要扩展, 若需要则从语料库中向已有子图中注入信息将其动态地拓展, 利用 GNN^[36]动态获取实体的表征并对其评分, 从中选择分数最高的实体作为答案。TTQA^[37]模型通过语法分析方法构建查询图, 将其与图匹配网络、预训练模型和注意力机制结合得到更加丰富的结构匹配表示, 增强了模型的可解释性。

3 基于知识图谱的多跳可解释问答

本文提出基于自然语言问题特征向量和跳数注意力机制的推理方法, 使问答系统具有长距离的推理能力和良好的可解释性。

3.1 问题描述

给定一个知识图谱 $G = (E, R)$, 其中 E 表示实体集合, R 表示关系集合。对于一个包含主题实体 e_s 的自然语言问题 q , 若存在实体 e_a 满足 q 的要求, 且存在 $e_s \xrightarrow{r_{s1}} e_1 \xrightarrow{r_{12}} e_2 \dots e_n \xrightarrow{r_{na}} e_a$ 的路径 (n 为有限正整数), 则 e_a 被认为是 q 的一个答案。本文的目标是通过分析自然语言问题 q , 在知识图谱上通过多跳可解释推理找到所有通往 e_a 的路径, 并将所有 e_a 作为答案。

3.2 算法框架

模型主要包括预处理模块, 特征提取模块, 分类模块与推理模块, 模型架构如图 2 所示。

预处理模块: 使用 BERT 预训练模型获取自然语言问题表征向量以及问题分词后的词向量矩阵。

特征提取模块: 利用 CNN 网络提取问题特征, 之后计算问题中每个词对问题语义的贡献程度, 将其归一化得到每个词在问句中的权重分布。

分类模块: 提取当前时刻自然语言问题的特征向量, 利用多层感知机计算知识图谱中下一跳推理中可能的每个关系的分数以及每一跳的分数。

推理模块: 利用头实体分数、关系分数计算尾实体的分数, 计算知识图谱中可能的推理路径中每个实体的综合分

数, 分数最高的实体将被作为答案。

3.3 预处理模块

MIQA 首先统计知识图谱中的实体总数与关系总数, 并为每个实体与关系分配对应的索引。本文将知识图谱中所有的三元组的头实体、尾实体与关系替换为对应的索引并纵向拼接, 得到三元组索引矩阵 TM , 具体而言, $TM(0)$ 表示头实体索引向量, $TM(1)$ 表示关系索引向量, $TM(2)$ 表示尾实体索引向量。对一个自然语言问题而言, 本文首先使用 BERT 预训练模型对问题编码得到问题表征向量 QV 与词向量矩阵 WVM , 其次利用分词器将自然语言问题分词后找出其中的所有实体(即主题实体), 之后从知识图谱中找到主题实体, 从主题实体出发开始推理。

MIQA 创建一个长度为实体种数的集合 S_E , 其中每个主题实体索引指向的位置填入初始分数 1, 其余位置填入初始分数 0。假设知识图谱 G 上一共有 K 个实体, 初始化的 S_E 表示为

$$S_E^0 = [S_{E_0}, S_{E_1}, \dots, S_{E_K}], S_{E_i} \in \{0, 1\} \quad (1)$$

其中 S_{E_i} 表示实体 E_i 的分数。本文利用 TM 更新 S_E 来实现三元组间的跳跃, 每次更新均创建新的 S_E 来储存新的分数, 其上标为已跳跃的次数。同时本文设置了最大跳数 Max_Hop 。

3.4 特征提取模块

本文使用 CNN 处理 QV 以得到特征更为明显的过渡向量

$$V_q = \text{Maxpool}(\text{ReLU}(\text{Conv}(QV))) \quad (2)$$

其中 $V_q \in \mathbb{R}^d$, Conv 代表卷积操作, 用于初步提取模型感兴趣的特征信息; ReLU 为激活函数; Maxpool 代表最大池化操作。

为使模型具有记忆能力, 本文整合过渡向量与前一时刻的问题特征向量来获得当前时刻带有记忆的问题向量

$$V_Q^T = \text{Tanh}(\text{MLP}(V_q | V_F^{T-1})) \quad (3)$$

其中 Tanh 是激活函数, 符号 $|$ 代表向量的拼接。 V_F^{T-1} 代表前一时刻的问题特征向量, MLP 代表多层感知机。二者将在分类模块详细介绍。本文利用 V_Q^T 与 WVM 计算 T 时刻问句中每个词的权重分布

$$WD^T = \frac{dt^T}{\sum dt^T} \quad (4)$$

其中

$$dt^T = \text{Softmax}(\sum V_Q^T \times WVM) \times MV \quad (5)$$

其中 $WD^T \in \mathbb{R}^l$, $dt^T \in \mathbb{R}^l$ 表示 T 时刻问句中的每个词与问句的相似度, 符号 \times 代表先将 WVM 按行分为 l 个向量, 之后将 V_Q 与每个向量进行点乘运算。此操作以问题向量为标准, 将每个词向量与之进行 \times 运算, 以此作为两者的相似度。

Softmax 函数将相似度归一化。 $MV \in \mathbb{R}^l$ 代表屏蔽向量。假设一个自然语言问题分词后长度为 n , 填充后长度为 l , 其对应 MV 从左往右将前 n 个元素设置为 1, 其余元素设置为 0。 MV 的作用是屏蔽由于填充长度产生的额外分数, 避免了分数稀释。式(4)将 dt^T 向量中剩余的非 0 元素归一化, 由此获取第 T 步时每个词的权重分布 WD^T 。

3.5 分类模块

本文使用多层感知机对特征向量分类, 每个类别的值为对应索引的关系或跳数的分数, 由此得到关系分数集合 S_R^T 与跳数分数集合 S_H^T 。当前时刻的特征向量为

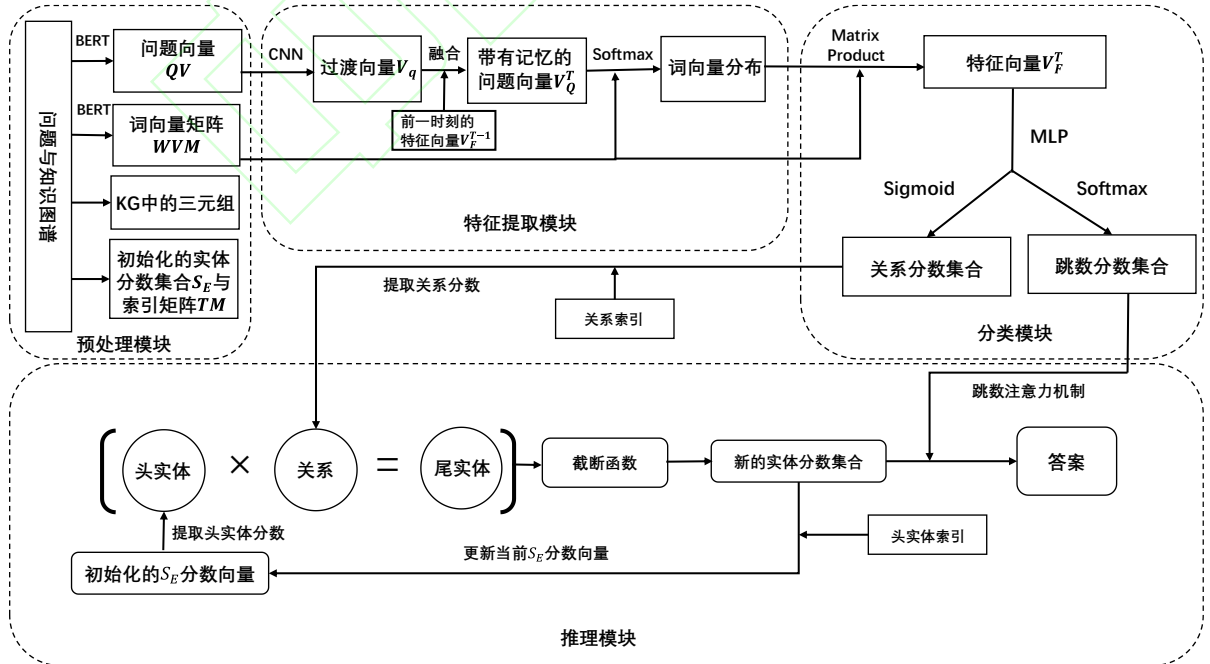


图 2 MIQA 模型的整体框架

Fig 2 Overall framework of MIQA model

$$V_F^T = \text{sum}_{\text{col}}(WD^T \otimes WVM) \quad (6)$$

其中 $V_F^T \in \mathbb{R}^d$, 符号 \otimes 表示将 WD^T 中的每个元素与 WVM 中对应的全部元素相乘, 函数 $\text{sum}_{\text{col}}(x)$ 的作用是将矩阵 x 每列元素相加以降低向量的维度。此操作为句子中的每个元素分配百分比的权重的同时降低了上层计算的复杂度。基于 V_F^T , 本文计算每个关系的分数

$$S_R^T = \text{Sigmoid}(MLP(V_F^T)) \quad (7)$$

与每一跳的分数

$$S_H^T = \text{Softmax}(MLP(V_F^T)) \quad (8)$$

3.6 推理模块

本文从预训练模块获得的三元组索引矩阵 TM 中提取每个三元组头实体、关系与尾实体的索引向量, 分别记为 Sub 、 Rel 和 Obj , 具体公式如下:

$$Sub, Rel, Obj = TM(0), TM(1), TM(2) \quad (9)$$

其中函数 $TM(x)$ 表示提取 TM 矩阵的第 x 列数据。本文利用式 (7) 与当前时刻实体分数集合 S_E^T 提取每个实体与关系的分数。

$$SubS, RelS = S_E^T(Sub), S_R^T(Rel) \quad (10)$$

其中 $SubS, RelS$ 分别表示头实体与关系的分数集合。 $S^T(x)$ 函数表示从 T 时刻的分数集合 S^T (包括 S_E^T 、 S_R^T 和 S_H^T) 中提取以向量 x 中每个值作为索引的分数。 T 时刻的尾实体分数集合

$$ObjS = \left(\sum (SubS \times RelS)_{\text{same } Obj} \right)_{Obj \text{ Index}} \quad (11)$$

本文将同个实体获得的分数累加, 并以向量 Obj 中的元素为索引将结果放入集合 $ObjS$ 中, 流程示例如图 3 所示。假设图 3 知识图谱中一共存在五个实体, 第一次跳跃结束时实体 e_1 的分数通过三元组 (e_2, r_2, e_1) 内部的乘法运算得到, 实体 e_2 的分数通过三元组 (e_3, r_3, e_2) 与三元组 (e_5, r_5, e_2) 分别进行乘法运算后将结果相加得到, 同理可计算剩余实体的分数。若没有三元组指向某个实体 (如 e_3), 本文将将其分数设置为 0。本文计算所有三元组的尾实体分数, 并将结果按实体索引更新 S_E^T 。本文将分数最大的实体作为答案。需要注意的是分数最高的实体有多个时, 本文将所有分数最高的实体都作为答案。基于式 (11), MIQA 模型实现了聚合头实体与关系的信息。在多次跳跃后一个实体将包含多个其他实体传递来

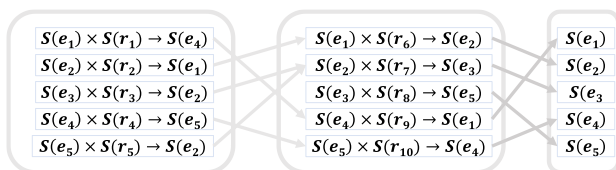


图 3 基于分数的跳跃方法示意图

Fig 3 Schematic diagram of score-based jumping method

的信息, 实现了信息聚合效果, 并且具有可解释性。

由式 (11) 可知若存在某个非答案实体 U 在某一跳的预测分数远大于其他实体, 那么即使之后有答案实体取得了较高的分数, 模型也会因为 U 在某次跳跃中分数过高而做出错误的选择 (具体细节将在式 (15) 中介绍)。这将导致模型很难离开局部最优值, 因此需要使用截断函数模块对最终分数进行修剪。

本文首先设计了一个获取阈值的算法:

$$\alpha = \frac{MMR}{|S_{SE}^T|} \times \sum_{i=0}^{|S_{SE}^T|} \max(S_{SE}^T) \quad (12)$$

其中 α 表示阈值, MMR 是最大 - 平均比率 (Max - Mean Rate), 用于适应不同的使用场景, $|S_{SE}^T|$ 表示 T 时刻一个批次中分数向量的个数。本文希望模型不被某些极值误导, 同时充分利用每一组数据, 因此考虑在模型中使用动态值的方法, 以避免出现模型不收敛的情况。基于此需求本文提出了 max_mean 算法, 其核心思想是提取一个批次中每个分数向量的最大值, 将它们相加后取平均值。之后本文通过 $\text{get}()$ 函数获取数值超过 α 的元素位置矩阵

$$MX = \text{get}(S_E^T, \alpha) \quad (13)$$

其中 $\text{get}(M, x)$ 表示创建一个与矩阵 M 形状相同的新矩阵 MX , 若 M 中存在值大于 x 的元素, 则在 MX 中将相应位置填入 1, 其余元素填入 0。接下来本文使用一种可导的方式来获得截断后的分数

$$S_E^{T+1} = \frac{S_E^T}{\frac{MX \times S_E^T}{\alpha} + (1 - MX)} \quad (14)$$

其中 S_E^{T+1} 表示 S_E^T 经过更新后的结果。经过计算后, S_E^{T+1} 中大于 α 的数都变为 α , 小于 α 的数不变。截断函数模块在完成截断工作的情况下增加了梯度的概念, 使损失函数收敛更快, 提高了模型的性能。之后 S_E^{T+1} 参与迭代直到到达最大跳数 Max_Hop 。由此可知模型在计算过程中难免会进行不必要的跳跃, 因此本文使用每跳分数

$$S_{\text{last}} = \sum_{i=0}^{\text{Max_Hop}} S_H^i \otimes S_E^i \quad (15)$$

来修正偏差, 其中 S_{last} 表示所有实体最终的分数。本文将同个实体在每个时刻的分数与每一跳分数加权后相加, 把最终结果作为实体的最终分数。本文采用 $L2$ 范数作为损失函数

$$\text{Loss} = \frac{|E| \times W \times \sum (S_{\text{last}} - \text{Answer})^2}{\sum (|E| \times W)} \quad (16)$$

本文将 $|E|$ 与 W 的组合式称为权重模块, 其中 $|E|$ 表示所有实体的数量, 它的作用是扩大损失函数对模型的影响力, 使模型能较大范围地更新权重矩阵, 有效防止了过拟合现象的发生。 W 代表预设的惩罚权重。由于 $Answer$ 采用的是 $One - Hot$ 编码, 因此

$$W = Answer \times PW + 1 \quad (17)$$

其中 PW 是可调参数, 用于适应不同的使用场景。本文将元素为零的位置的惩罚力度设置为1, 而成倍增大元素非零位置的惩罚力度。

3.7 模型实现步骤

MIQA 的过程如算法 1 所示, 输入为知识图谱与自然语言问题, 输出为分数最高的答案实体。

算法 1: MIQA 的过程

输入: 知识图谱 G , 自然语言问题 q , 最大跳数

Max_Hop , 最大-平均比率 MMR

输出: 分数最高的实体

1. $BERT(q) \rightarrow WVM, QV$

2. $S_E^0 = if \ index \rightarrow e: score = 1, else: score = 0$

3. $V_Q = CNN(QV)$

4. **for** $i < H$:

5. $dt^T = softmax(V_Q \text{ 与 } WVM \text{ 在时间 } T \text{ 的相似度})$

6. $WD^T = dt^T$ 中归一化后的元素分布

7. $V_F^T =$ 通过元素分布从 WVM 提取的特征向量

8. $S_R^T, S_H^T = V_F^T$ 分类得到的分数

9. $SubP, RelP = S_E^T(Sub), S_R^T(Rel)$

10. $SubP \times RelP = ObjP$, 依据 $ObjP$ 更新 S_E^T

11. $\alpha = MMR * sum(max(S_{P_E^T}))/size(S_{P_E^T})$

12. $S^{T+1} = S^T$ 中大于 α 的部分替换为 α

13. $S_{last} = \sum_{i=0}^{Max_Hop} S_H^i \otimes S_E^i$

14. **return** $max(S_{last})$

4 实验与结果分析

在三个公开数据集上, 本文将 MIQA 与 8 个知名算法做了比较, 从而验证 MIQA 模型的有效性。

4.1 数据集

本文在如下三个 KBQA 任务中常用的数据集上进行实验。

- (1) **MetaQA**^[7]: 此数据集中所有的自然语言问题都是从 MovieQA 的 Wikidata 知识库中生成, 它包含超过 40 万个问题, 135000 个三元组, 43000 个实体以及 9 种关系。数据集提供 Vanilla、NTM 和 Audio 3 个版本, 并将基础知识库放置于 kb 文件中, 该文件的每一行都是一个以 $\langle Subject, Relation, Object \rangle$ 形式储存的三元组。本文中采用了 Vanilla 版本的数据集, 它移除了原数据

集中含有歧义的实体, 并将剩余部分分为 1-hop、2-hop 和 3-hop 三个组件, 每个组件均由十种文本模板产生并提供了 train、dev 和 test 三个数据集划分相同的文件。

- (2) **WebQuestionsSP(WebQSP)**^[38]: 此数据集是 WebQuestions 数据集经过改进后诞生的, 它包含数千个基于 Freebase 的自然语言问题, 数百万个实体和三元组。原始数据集 WebQuestions 是为了解决真实问题而构造的, 问题的答案是由人工标注的。但它仅仅提供了答案而没有提供与问题对应的查询语句。WebQuestionsSP 在它的基础上为每个问题构造了 SPARQL 查询表示, 并且删除了有歧义或意图不清晰的答案与问题。
- (3) **ComplexWebQuestions(CWQ)**^[39]: 此数据集是基于 WebQuestionsSP 生成的。首先从 WebQuestionsSP 中提取自然语言问题与对应的 SPARQL 查询表示, 然后自动生成更复杂的组合。其中包含组合 (44.7%)、连接 (43.6%)、比较级 (6.2%) 和最高级 (5.4%) 形式, 最后重新组合为问题。问题的答案通过在 Freebase 上经过 SPARQL 查询得到。

三个数据集的统计信息如表 1 所示:

表 1 MetaQA、WebQSP 和 CWQ 的统计信息

Table 1 Statistics for MetaQA, WebQSP and CWQ

Document	MetaQA			WebQSP	CWQ
	1-hop	2-hop	3-hop		
Train	96106	118980	114196	2998	27734
Dev	9992	14872	14274	100	3480
Test	9947	14872	14274	1639	3475
Max_Hop	3			2	4

4.2 基准算法

本文将 MIQA 与如下八种知名模型进行比较。

KVMemNN^[31]: 构造键值记忆网络, 将三元组向量化后用 $key - value$ 方式存储与输出信息, 从原始文本中选取答案。

SRN^[9]: 将多关系问题表示为顺序决策问题, 使用注意力机制与神经网络增强问题特征, 使用路径搜索的方法获得答案。

GraftNet^[34]: 使用文本数据对知识图谱进行补充, 依据问题的主题实体提取相应子图, 使用 GNN 从子图中提取答案。

PullNet^[35]: 首先从自然语言问题中提取主题实体, 从主题实体开始不断加入其他与自然语言问题相关的实体, 以此形成知识图谱, 最后使用图 CNN 网络提取答案。

ReifKB^[26]: 将标签形式的数据采用稀疏矩阵储存, 使用

基于概率的方法在知识图谱上跳跃,根据实体综合概率选择答案。

EmbedKGQA^[7]: 基于知识图谱与自然语言问题嵌入,通过链路预测的方法取评分最高的实体作为答案。

TransferNet^[40]: 每次跳跃关注问题的不同部分并计算关系激活分数,然后以可微的方式沿着激活的关系传递先前的实体分数。

NSM^[41]: 使用学生网络选择正确答案,教师网络学习中间的监督信号,从而提高学生网络推理能力。

本文使用的八种对比模型均为 SOTA 方法,其中 KVMemNN 与 NSM 是基于嵌入的方法,SRN、EmbedKGQA 与 TransferNet 是基于路径的方法,GraftNet 与 PullNet 是基于子图的方法,ReifKB 是基于概率的方法。这些模型在同类方法中具有优秀的表现,因此被作为基线广泛应用于 KGQA 领域^[10,38,41]。

4.3 实验设置

本文使用的硬件配置如下:GPU 采用 NVIDIA GeForce GTX 3090,内存为 24GB。模型建立在 Pytorch 搭建的框架上,首先使用三层 CNN 网络提取自然语言问题的特征,并且在关系图中添加了反关系与闭环检测。本文使用 Adamw 作为优化器,它采用自适应机制自动调整梯度下降的学习率,在反向传播的公式中加入了正则项的梯度,提高了计算效率。初始的学习率设置为 0.001,权重衰减系数设置为 $1e-5$ 。本模型使用学习率预热算法,预热比例设置为 0.1。这种算法先将学习率从 0 线性增加到优化器中的预设值,之后从预设值线性下降为 0。采用学习率预热算法有助于缓解模型前期的过拟合现象,使模型的收敛更迅速且有更好的稳定性。

对于 MetaQA 数据集的实验,本文将 Max_Hop 设置为 3,使用 BiGRU 网络对自然语言问题进行编码,其中隐藏层维度设置为 1024,层数设置为 1,随机丢弃概率设置为 0.2。超参数 PW 设置为 9, MMR 设置为 1。

对于 WebQSP 数据集的实验,本文使用 BERT 预训练模型中的 bert-base-uncased 模型进行编码。它使用 12 层编码器,在小写的标记中训练,训练后生成维度为 768 的向量。本文将 Max_Hop 设置为 2,超参数 PW 设置为 70, MRR 设置为 0.9。

对于 CWQ 数据集实验,本文将 Max_Hop 设置为 3,使用 bert-base-uncased 模型进行编码。超参数 PW 设置为 13, MMR 设置为 1。

4.4 评估指标

本文使用

$$Hits@1 = \frac{1}{|Tr|} \sum_{i=1}^{|Tr|} \sigma(rank_i = 1) \quad (18)$$

与

$$MRR = \frac{1}{|Tr|} \sum_{i=1}^{|Tr|} \frac{1}{rank_i} \quad (19)$$

作为评分标准, $Hits@1$ 指标是指链接预测中排名第 1 的三元组的平均占比,其中 $|Tr|$ 表示三元组的数量, $\sigma(x)$ 是判断函数,若 x 为真,函数值为 1; 否则函数值为 0。 $rank$ 是指三元组的排名。 MRR 指标是指链接预测给出的所有答案中正确答案所在排名的倒数之和。

4.5 实验与结果分析

本文利用每个数据集的 train 部分进行训练,每隔一定的轮数用 dev 文件进行验证,最后使用 test 文件对模型的性能进行评估。首先本文利用编码器对自然语言问题与给定的知识图谱进行嵌入处理,依据注意力机制提取问题向量在本步骤的特征向量,据此对关系进行评分。之后将上一轮实体分数矩阵与关系分数矩阵相乘得到本轮的实体分数矩阵,循环以上步骤直到达到 Max_Hop 。最后将实体在每一跳的分数与每跳的权重相乘得到该实体的最终分数。本文取得分最高的实体作为答案。

本文的实验结果如表 2 所示。对于 MetaQA 数据集, MIQA 在 1 跳数据集上取得了 97.6% 的准确率,超越了基线模型,并在 2 跳与 3 跳数据集上取得了接近 100% 的准确率。本文寻找了造成 1 跳数据集上不能完全回答正确的原因,具体解释如图 4 所示。对于部分自然语言问题,假设本文从主题实体出发有且仅有两个符合条件的三元组,它们的头实体和关系均相同,但尾实体不同。MIQA 模型准确地找到了全部的答案,但 1 跳测试数据集中记录的答案可能仅为全部正确答案中的一个,因而导致了不可避免的错误。

对于 WebQSP 数据集与 CWQ 数据集, MIQA 分别做到了 70% 与 48.1% 的准确率。这些数据集比 MetaQA 更大更复杂,包含更多的实体与三元组,但训练样本却更少。相比于六个知名模型, MIQA 取得了最好的效果,证明本文的方法是有效且可行的。

综上所述, MIQA 模型在所有的数据集上,和八种知名算法相比较,都获得了最佳性能。

4.6 消融实验

本文对 MIQA 进行了消融实验,分别检测了删除 CNN 特征提取模块 MIQA-CNN、截断函数模块 MIQA-TF、损失权重模块 MIQA-W 以及影响力扩展参数 MIQA-|E| 对模型的影响,结果如表 3 所示。结果显示删除截断函数模块对 MetaQA 数据集的影响最大,其中 3-hop 部分两个指标的评分下降最多,达到 10.7% 与 5.9%; 删除损失权重模块对 WebQSP 数据集影响最大,分数下降了 27.9% 与 27.1%; 删除影响力扩展参数对 CWQ 数据集影响最大,分数下降了 19% 与 14.7%。删除 CNN 特征提取模块对分数的影响相对

不明显,但依旧显著降低了模型的总体水平。

表 2 不同模型在不同数据集上的 Hits@1/MRR 分数对比

Table 2 Comparison of Hits@1/MRR scores of different models on different data sets

Model	MetaQA			WebQSP	CWQ
	1-hop	2-hop	3-hop		
	Hits@1/MRR	Hits@1/MRR	Hits@1/MRR	Hits@1/MRR	Hits@1/MRR
KVMemNN	95.6%/96.1%	24.7%/28.1%	10.2%/15.5%	46.9%/50.8%	21.0%/25.9%
SRN	97.0%/97.7%	95.6%/97.8%	75.2%/81.3%	—	—
GraftNet	96.9%/97.8%	94.7%/96.3%	78.1%/82.5%	65.7%/71.8%	32.3%/36.6%
PullNet	97.2%/97.3%	99.8%/99.9%	91.3%/94.6%	68.1%/74.3%	47.3%/48.3%
ReifKB	96.5%/96.4%	81.7%/86.1%	72.4%/76.0%	52.5%/56.7%	—
EmbedKGQA	97.4%/97.1%	98.9%/99.2%	94.7%/98.1%	66.7%/70.2%	—
TransferNet	97.4%/98.3%	99.8%/99.9%	99.7%/99.9%	68.7%/72.4%	45.6%/48.6%
NSM	97.1%/97.9%	99.8%/99.9%	98.9%/99.7%	68.7%/72.8%	47.6%/49.5%
MIQA(ours)	97.6%/98.9%	99.9%/99.9%	99.8%/99.9%	70.0%/75.4%	48.1%/50.3%

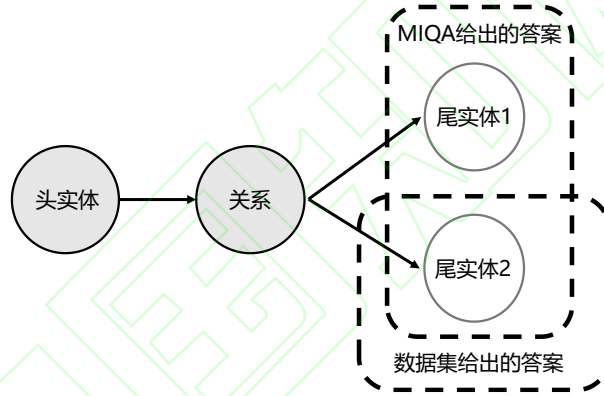


图 4 1跳推理不能完全回答正确的原因示意图 该问题的答案为尾实体 1 和尾实体 2, 下方框代表模型给出的答案, 上方框表示测试集里的正确答案。虽然模型找到了所有正确答案, 但是测试数据集中可能答案缺失, 因此 MIQA 模型给出的答案被判定为错误。

Fig 4 Schematic diagram of why 1-hop reasoning cannot fully answer the correct answer. The answers to this question are Tail Entity 1 and Tail Entity 2. The lower box represents the answer given by the model, and the upper box represents the correct answer in the test set. Although the model found all the correct answers, there may be missing answers in the test dataset, so the answers given by the MIQA model are judged to be wrong.

4.7 可解释性分析

针对本文第一部分所述自然语言问题“苹果的首席执行官史蒂夫·乔布斯的父亲在哪出生?”, 本文对其进行可视化的可解释性分析, 如图 5 所示。假设模型将“史蒂夫·乔布斯”作为主题实体, 并将其初始分数设为 1, 本文沿着将其作为头实体的三元组, 分别计算与之相连接的两个关系的分数。之后本文将头实体分数与关系分数相乘的结果作为尾实体的分数, 并计算 1-hop 的权重, 结果如左侧方框内的 1 跳分数所示。将已有实体的分数进行更新, 并重复上述步骤直到到达预设的 *Max_Hop* (右侧方框内的 2 跳分数), 最终每个实体每次跳跃后的分数与每跳分数相乘后相加作为最终结果。例如对于实体“叙利亚”, 由于第一跳没有参与,

因此第一跳分数为 0。它的总分计算如下: $S = 0 \times 0.4 + 0.6 \times 0.56 = 0.336$ 。计算图中全部的实体后它的得分是最高的, 因此选择它为最终答案。值得注意的是两个“美国”实体, 如果没有到达阈值, 它们将在后一步合并到一起。在大型知识图谱中这种现象可能造成某个实体的分数经过累加远远超过其他实体的情况, 说明了使用截断函数模块的必要性。

4.8 参数敏感性分析

本文绘制了最大平均比率 *MMR* 分别和 *Hits@1*、*MRR* 分数的关系图, 即 *MMR-Score* 分数曲线, 如图 6 所示。其中 *Hits@1* 与 *MRR* 分数是多次计算获取的平均值。该结果表明 CWQ^[39]数据集的分数相对变化最小, 在 *MMR* 为 1 时准确

率最高，WebQSP^[38]数据集的分数变化最大，在MMR为0.9时准确率最高。MetaQA^[7]数据集变化幅度与CWQ数据集相似但比CWQ更大，在MMR为1时准确率最高。本文认为导致这些现象的原因是CWQ的数据集中训练集数据、验证集的数据数量之和与测试集数据数量的比值约为10:1，WebQSP的数据集中该比值约为2:1，因此同样回答一个问

题CWQ的可用样本数量明显多于WebQSP，对应的分数曲线较为稳定。而MetaQA数据集的对应平均比值约为9:1，因此曲线波动幅度与CWQ相似但比它更大。本文在训练模型时使用了每个数据集表现最优时的MMR数值。

表3 消融实验：移除不同模块的MIQA模型与原模型的Hits@1/MRR分数对比

Table 3 Ablation experiment: Remove the MIQA model of different modules to compare with the Hits@1/MRR score of the original model

	MetaQA			WebQSP	CWQ
	1-hop	2-hop	3-hop		
	Hits@1/MRR	Hits@1/MRR	Hits@1/MRR	Hits@1/MRR	Hits@1/MRR
Original	97.6%/98.9%	99.9%/99.9%	99.8%/99.9%	70.0%/75.4%	48.1%/50.3%
MIQA - CNN	97.3%/98.4%	99.0%/99.7%	97.1%/98.7%	63.3%/67.5%	45.9%/48.1%
MIQA - TF	97.4%/98.3%	99.0%/99.6%	89.1%/94.0%	54.1%/59.1%	44.4%/47.2%
MIQA - W	97.2%/98.0%	99.2%/99.8%	96.4%/98.1%	42.1%/48.3%	36.9%/43.2%
MIQA - E	97.4%/98.2%	99.6%/99.9%	99.2%/99.7%	56.0%/59.8%	29.1%/35.6%

问题：苹果的首席执行官史蒂夫·乔布斯的父亲在哪出生？

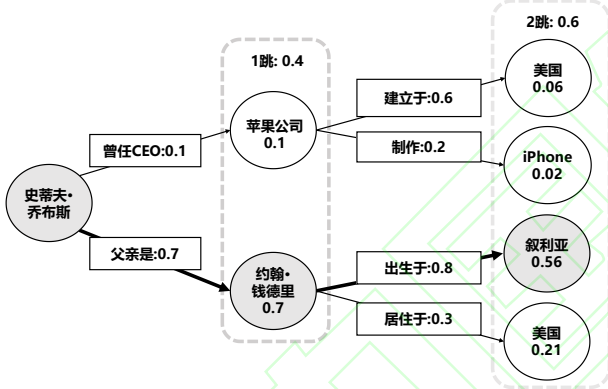


图5 可解释的问答过程

Fig 5 Explainable Question Answering Process

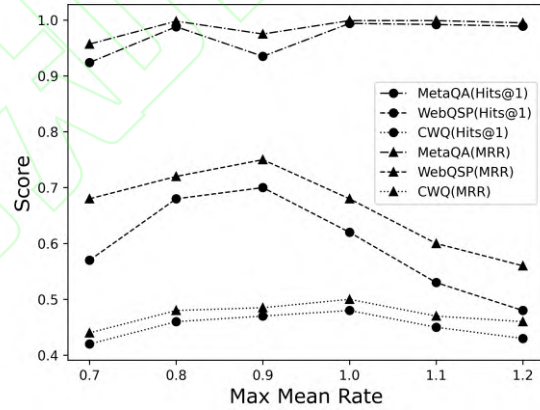


图6 MMR - Hits@1变化曲线

Fig 6 MMR - Hits@1 change curve

5 总结

本文提出了一种基于知识图谱的多跳可解释问答网络，它通过更新实体分数矩阵实现实体间跳跃并获取答案，其决策过程有很好的可解释性。其中使用的基于自然语言问题特征向量和跳数注意力机制的推理方法使问答系统具有长距离的推理能力和良好的可解释性。将来，本文计划在特征向量处理和跳数自适应方面进行扩展研究，使之具有更强的问答性能，在路径预测、可解释问答所涉及的金融、医疗和教育等领域具有广阔的应用前景。

References:

[1] Wang Fei, Liu Jing-ping, Liu Bin, et al. Research on code knowledge graph construction and intelligent software development methods[J]. Journal of Software, 2020, 31(1): 47-66.

[2] Wang Xin, Zou Lei, Wang Chao-kun, et al. A review of knowledge graph data management research[J]. Journal of Software, 2019, 30(7): 2139-2174.

[3] Saxena A, Chakrabarti S, Talukdar P. Question answering over temporal knowledge graphs[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2021: 6663-6676.

[4] Cheng Zhang-tao, Zhong Ting, Zhang Sheng-ming, et al. Survey of recommender systems based on graph learning[J]. Computer Science, 2022, 49(9): 1-13.

[5] Bollacker K, Evans C, Paritosh P, et al. Freebase: a collaboratively created graph database for structuring human knowledge[C]// Proceedings of the 2008 Association for Computing

- Machinery Special Interest Group on Management of Data International Conference on Management of Data, 2008:1247-1250.
- [6] Auer S, Bizer C, Kobilarov G, et al. Dbpedia: a nucleus for a web of open data[C]//Semantic Web, Springer Berlin Heidelberg, 2007:722-735.
- [7] 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, 2020:4498-4507.
- [8] Zhang Y, Dai H, Kozareva Z. Variational reasoning for question answering with knowledge graph[C]//32nd Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence, 2018, 32(1):.
- [9] Qiu Yun-qi, Wang Yuan-zhuo, Jin Xiao-long, et al. Stepwise reasoning for multi-relation question answering over knowledge graph with weak supervision[C]//International Conference on Web Search and Data Mining, 2020:474-482.
- [10] Yang Z, Shou L, Gong M. Model compression with two-stage multi-teacher knowledge distillation for web question answering system[C]//Proceedings of the 13th International Conference on Web Search and Data Mining, 2020:690-698.
- [11] Gardner M, Talukdar P, Kisiel B. 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, 2013:833-838.
- [12] Xiong W, Hoang T, Wang W Y. DeepPath: a reinforcement learning method for knowledge graph reasoning[C]//Conference on Empirical Methods in Natural Language Processing, 2017:564-573.
- [13] Chen W H, Xiong W H, Yan X F. Variational knowledge graph reasoning[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2018:1823-1832.
- [14] Kingma D P, Welling M. Auto-encoding variational Bayes[C]//Proceedings of the 2nd International Conference on Learning Representations, 2014:1-14.
- [15] Richardson M, Domingos P. Markov logic networks[J]. Machine Learning, 2006, 62(1):107-136.
- [16] Singla P, Domingos P. Discriminative training of Markov logic networks[C]//Proceedings of the Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence, 2005:868-873.
- [17] Harsha Vardhan L V, Jia G, Kok S. Probabilistic logic graph attention networks for reasoning[C]//Companion Proceedings of the Web Conference, 2020:669-673.
- [18] Dong L, Wei F, 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, 2015:260-269.
- [19] Hao Y C, Zhang Y Z, Liu K. 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, 2017:221-231.
- [20] Liang P. Lambda dependency-based compositional semantics[J]. arXiv preprint arXiv:1309.4408, 2013.
- [21] Yuan Y, Lian X, Chen L, et al. Keyword search over distributed graphs with compressed signature[J]. IEEE Transactions on Knowledge and Data Engineering, 2017, 29(6):1212-1225.
- [22] Trouillon T, Welbl J, Riedel S. Complex embeddings for simple link prediction[C]//International Conference on Machine Learning, Proceedings of Machine Learning Research, 2016:2071-2080.
- [23] Liu Y, Ott M, Goyal N, et al. Roberta: a robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692, 2019.
- [24] Yang B, Wan F, Liu C, et al. Part-based semantic transform for few-shot semantic segmentation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(12):7141-7152.
- [25] Lu W, Yu R, Wang S. Sentence semantic matching based on 3D CNN for human-robot language interaction[J]. Association for Computing Machinery Transactions on Internet Technology, 2021, 21(4):1-24.
- [26] William W Cohen, Haitian Sun, R Alex Hofer, et al. Scalable neural methods for reasoning with a symbolic knowledge base[J]. arXiv preprint arXiv:2002.06115, 2020.
- [27] Tang Nan-nan, Chen Ji, Hou Lei, et al. Entity-relationship joint extraction method combining attention mechanism and pointer annotation [J]. Journal of Chinese Computer Systems, 2023, 44(2):256-262.
- [28] Ren H, Leskovec J. Beta embeddings for multi-hop logical reasoning in knowledge graphs[J]. Advances in Neural Information Processing Systems, 2020, 33:19716-19726.
- [29] Min S, Zhong V, Zettlemoyer L, et al. Multi-hop reading comprehension through question decomposition and rescoring[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019:6097-6109.
- [30] Hudson D A, Manning C D. Compositional attention networks for machine reasoning[C]//International Conference on Learning Representations, 2018.
- [31] 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, 2016:1400-1409.
- [32] Schlichtkrull M, Kipf T N, Bloem P, et al. Modeling relational data with graph convolutional networks[C]//European Semantic Web Conference, Springer, Cham, 2018:593-607.
- [33] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[J]. arXiv preprint arXiv:1609.02907, 2016.

- [34] Sun H,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,2018:4231-4242.
- [35] Sun H,Bedrax-Weiss T,Cohen 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 and the 9th International Joint Conference on Natural Language Processing,2019:2380-2390.
- [36] Scarselli F,Tsoi A C,Gori M,et al.Graphical-based learning environments for pattern recognition[C]//Joint International Association for Pattern Recognition International Workshops on Statistical Techniques in Pattern Recognition and Structural and Syntactic Pattern Recognition,Springer,Berlin,Heidelberg,2004:42-56.
- [37] Sun Ya-wei,Cheng Gong,Li Xiao,et al.Interpretable knowledge graph complex question answering method based on graph matching network[J].Computer Research and Development,2021,58(12):2673-2683.
- [38] Yih W,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,2016,2:201-206.
- [39] 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,2018:641-651.
- [40] Shi J,Cao S,Hou L,et al.TransferNet:an effective and transparent framework for multi-hop question answering over relation graph[C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing,2021:4149-4158.
- [41] He G,Lan Y,Jiang J,et al.Improving multi-hop knowledge base question answering by learning intermediate supervision signals[C]//Proceedings of the 14th Association for Computing Machinery International Conference on Web Search and Data Mining,2021:553-561.
- [37] 孙亚伟,程 龚,厉 肖,等.基于图匹配网络的可解释知识图谱复杂问答方法[J].计算机研究与发展,2021,58(12):2673-2683.

附中文参考文献:

- [1] 王 飞,刘井平,刘 斌,等.代码知识图谱构建及智能化软件开发方法研究.软件学报,2020,31(1):47-66.
- [2] 王 鑫,邹 磊,王朝坤,等.知识图谱数据管理研究综述.软件学报,2019,30(7): 2139-2174.
- [4] 程章桃,钟 婷,张晟铭,等.基于图学习的推荐系统研究综述[J].计算机科学,2022,49(9):1-13.
- [27] 唐楠楠,陈 吉,侯 磊,等.融合注意力机制和指针标注的实体关系联合抽取方法[J].小型微型计算机系统,2023,44(2):256-262.