

基于层次结构图的多跳知识图谱问答模型

刘昀抒¹, 申彦明¹, 齐恒¹, 尹宝才²

(1.大连理工大学计算机科学与技术学院, 辽宁 大连 116024; 2.北京工业大学人工智能与自动化学院, 北京 100124)

摘要: 知识图谱问答(KBQA)旨在理解用户的自然语言问句,在结构化的知识图谱中通过检索、推理等手段来获取答案实体。近年来,多跳KBQA备受关注,然而,复杂问句中通常存在多个关系意图,已有KBQA方法大多忽视了推理关系链的关系顺序问题。为此,提出一种基于层次结构图的多跳知识图谱问答模型(HSG-KBQA),建模自然语言问句的关系层次顺序,指导模型在每个推理步选择合理的关系意图。设计一种层次结构图,显式地体现问句中关系的层次距离,利用LSTM-BiGCN编码层将词语间的依存信息编码到问句中;提出虚拟节点的概念,利用图池化技术过滤不重要的节点,学习推理过程中知识图谱的状态;设计基于注意力机制和层次权重的解码器来优化指令生成,使推理指令更匹配问句中的关系链顺序。实验结果表明,HSG-KBQA在WebQuestionsSP数据集上取得了71.3%的Hits@1分数,在PathQuestions数据集上取得了97.3%(PQ-2H)和89.7%(PQ-3H)的Hits@1分数,均优于对照基准模型,表明HSG-KBQA模型在KBQA任务中具有更好的性能。

关键词: 知识图谱问答;问答系统;多跳问答;图神经网络;动态推理

中图分类号: TP391

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0066637

Multi-hop Knowledge Base Question Answering Model Based on Hierarchical Structure Graph

LIU Yunshu¹, SHEN Yanming¹, QI Heng¹, YIN Baocai²

(1.School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, Liaoning, China;

2.School of Artificial Intelligence and Automation, Beijing University of Technology, Beijing 100124, China)

[Abstract] Knowledge Base Question Answering (KBQA) aims to interpret natural language questions and obtain answer entities via retrieval and inference in a Knowledge Graph (KG). Recently, considerable attention has been given to multi-hop KBQA, which faces a significant, yet often overlooked challenge: complex questions typically contain multiple relational intentions with an inferential order that is frequently ignored. To address this challenge, this paper proposes a multi-hop KBQA model based on a Hierarchical Structure Graph (HSG-KBQA). This model is designed to recognize and guide the relational hierarchical order of natural language questions, enabling the selection of appropriate relational intents at each inference step. The proposed hierarchical structure graph explicitly represents the relational hierarchical order in questions and encodes the dependency information into the question using an LSTM-BiGCN encoding layer. Furthermore, the concept of a virtual node is introduced to learn the state of the KG during inference utilizing graph pooling techniques. Additionally, an attention mechanism-based decoder with hierarchical weights is presented, aiming to optimize instruction generation. This ensures that each hop instruction aligns more accurately with the relational chain order in the question. Experimental results show that the HSG-KBQA model demonstrates superior performance, achieving Hits@1 scores of 71.3% on the WebQuestionsSP dataset and 97.3% (PQ-2H) and 89.7% (PQ-3H) on the PathQuestions dataset. These results are superior to those of the benchmark model, indicating the enhanced capability of the proposed model in handling the KBQA task.

[Key words] Knowledge Base Question Answering (KBQA); question answering system; multi-hop question answering; graph neural network; dynamic reasoning

0 引言

问答系统是自然语言处理领域的主要研究内容之一,旨在利用计算机自动回答用户所提出的问题以满足用户对知识的需求^[1]。随着知识图谱的发展,基于知识图谱的问答(KBQA)系统开始受到广泛

关注。基于知识图谱的问答系统以知识图谱为数据来源,通过分析用户自然语言问题的语义,在已经构建的结构化知识图谱中通过检索、匹配或推理等手段来获取正确答案。主流的知识图谱问答系统方法主要分为两类,即基于语义解析(SP)的方法和基于信息检索(IR)的方法。

收稿日期:2022-12-28 修回日期:2023-03-10

基金项目:大连市科技创新基金(2022JJ12SN052)。

通信作者 E-mail: Yunshu1994@163.com

知识图谱问答中的问题按照复杂程度可以分为简单问题与复杂问题。简单问题通过单个三元组就可以解答,而复杂问题意味着涉及多个三元组,具有更复杂的查询类型和组合语义。例如在多跳问题中,若一个问题从主语实体出发,经过 n 条三元组到达答案所在的实体,就称该问题为 n 跳问题,对该问题进行回答的过程称为 n 跳问答^[2]。目前,对简单问题的研究已经很充分^[3-5],越来越多的研究人员开始关注复杂问题^[6-8]。基于语义解析的方法需要将问句精确地转化为逻辑表达式或者查询图,复杂的问句提高了精确解析的难度,并且问句所对应的候选逻辑形式更多,大幅增加了计算成本和解析难度。传统的基于信息检索的方法通过计算问句和候选答案实体的语义相似度分数,中间的过程不易解释。由此,出现了对知识图谱多跳推理的研究。本文主要研究基于信息检索方法的多跳推理任务,在该方法中,多跳推理模型模拟图神经网络消息传递的过程,通过在知识图谱中逐步地推理来更新候选实体的状态。因此,多跳推理模型可以追溯到中间的推理过程,得到推理关系链。

然而,以往的基于信息检索的方法忽略了推理关系链的关系顺序问题。当用户输入的问题为复杂问题时,问句中可能存在多个推理关系和多个隐藏的中间实体。对于多跳推理模型来说,每个推理步所偏向的关系都是至关重要的,错误的关系顺序将导致错误的推理结果,在特殊情况下,即使得到了正确的答案,也无法得到正确的推理路径。一些模型通过抽取答案关系路径来关注这个问题,如 Rce-KGQA^[9],然而从主题实体到答案实体的最短路径不一定是正确的推理路径。例如,问题“*What organization did the wife of [Bill Gates] found?*”的主题实体为 Bill Gates,答案实体为 Bill & Melinda Gates Foundation,在知识图谱中从主题实体出发有2条路径可以到达答案实体节点,第1条路径为“*Bill Gates*→*founder*→*Bill & Melinda Gates Foundation*”,第2条路径为“*Bill Gates*→*spouse*→*Melinda Gates*→*founder*→*Bill&Melinda Gates Foundation*”。显然,第2条路径是正确的推理过程。有模型^[10-12]通过循环神经网络和注意力机制关注到问句中不同的部分,但是这种机制所派生出的一系列推理指令可能是以不正确的顺序来匹配知识图谱中的关系。例如,问题“*Who acted in the movies directed by the director [Mel Gibson]?*”,该问题中正确的关系链为“*Mel Gibson*→*directed_by*→?*→starred_actors*→?”,如果“*directed_by*”和“*starred_actors*”的关系顺序错误,将会导致错误的回答,如推理路径“*Mel Gibson*→*starred_actors*→*The Expendables*→*directed_by*→*Patrick Hughes*”。

针对上述问题,本文提出一种基于层次结构图的知识图谱多跳问答模型(HSG-KBQA),该模型分为问题嵌入模块、图池化模块、推理模块这3个模

块。为了缓解关系顺序问题,本文设计问题嵌入模块,为问句构造层次结构图,层次结构图是一种由依存句法树构造的图,图的结构信息能够帮助理解复杂的语义,捕获问句中词语与主题实体间的层次距离,进而显式地体现推理关系链的关系顺序,充分利用语义信息逐跳推理,加强信息检索方法的可解释性。同时,提出虚拟节点的概念,使用图池化技术获取当前推理步图谱的虚拟节点,用于生成下一步推理指令,使推理指令既包含问句语义信息,又与知识图谱中的推理状态紧密相关。最后,利用基于图神经网络的推理模型在知识图谱上进行多跳推理。为了评估所提模型的性能,在大规模数据集上进行实验,验证模型在多跳问答任务中的有效性。本文主要工作如下:

1) 提出一种基于层次结构图的知识图谱多跳问答模型 HSG-KBQA,该模型通过构造层次结构图来关注推理过程中的关系顺序问题,从而加强多跳推理的可解释性。

2) 使用双向图神经网络将问句的语义和结构信息融入问句嵌入,多角度理解复杂语义。本文模型提出虚拟节点的概念,使得推理指令能够融合知识图谱的当前状态,对下一步推理进行正确指导,并且在实验中验证了虚拟节点的可行性。

3) 在 WebQuestionsSP 和 PathQuestions 数据集上验证基于层次结构图的知识图谱多跳问答模型的有效性,表明该模型通过关注每个推理步应偏向的关系而提升了问答的准确率。

1 相关工作

1.1 基于语义解析的方法

基于语义解析的方法^[13-16]首先对问句进行语义解析,将问题转化为某种逻辑形式,然后将逻辑形式转化为能够在知识图谱上运行的查询语句,如 SPARQL,最后通过执行查询语句来获取答案。

传统的语义解析方法主要依赖规则模板对自然语言问句进行解析。随后,文献[13]提出了基于查询图的知识图谱问答框架 STAGG,该框架通过定义一系列的状态和行为,将语义解析过程转换为查询图的生成过程,通过 LambdaRank 算法对候选查询图进行排序,进而选出最佳答案。后续很多工作都在此之上对查询图的生成过程进行改进。基于语义解析的方法在回答多跳问题时,着重解决理解复杂语义的问题,利用句法标注、定义更多的运算符来生成结构更复杂的逻辑形式。文献[15]提出一种新型的骨架句法分析模型,通过不断地拆分问句,抽取复杂问句的主干,进而更准确地理解长且难的问句。文献[16]引入抽象语义表示(AMR),更好地实现了问题与逻辑表达式中元素(如实体、关系、属性和类型)的对齐。

基于语义解析的方法能够反映问题语义结构,

紧密结合知识图谱信息,但是,由于语义解析需要复杂的推理过程以及大量的手工规则特征,因此基于语义解析的方法实现起来很有难度^[17]。

1.2 基于信息检索的方法

基于信息检索的方法^[18-19]通常抽取特定于问题的子图作为候选答案实体集合,并假设子图包含正确答案实体。然后,通过基于向量的计算对候选实体进行评分和排序,选择得分最高的候选实体作为答案。为了能够在不完整的知识图谱上问答,文献[18]提出的 GraftNet 模型引入外部文本语料来丰富知识图谱,文献[19]提出的 EmbedKGQA 模型则是引入知识图谱嵌入模型 ComplEx^[20],通过预训练的 ComplEx 模型学习实体和关系的向量表示以缓解知识图谱的稀疏性问题。

为了能够追溯更具解释性的推理过程,越来越多的研究人员开始关注基于知识图谱的多跳推理^[21-24]。文献[21]提出 KVMem 模型,该模型把知识图谱中的三元组以键值对的形式存储在记忆网络中,每次输出与问题相关性高的三元组,从而模拟多跳推理的过程。但是,这对体量庞大的图谱来说具有一定的挑战,因为该模型需要存储大量的键值对。文献[22]提出一个可解释的推理网络 (IRN),该网络在推理过程中动态地关注问题,并通过上一跳预测的关系来更新问题嵌入。但是,大多数的知识图谱问答数据仅包含问答对,推理的中间实体和关系路径没有标注,而 IRN 在弱监督条件下表现一般,需要中间结果的指导才能获取较优的结果。为了缓解这个问题,文献[23]提出一种基于强化学习的逐步推理网络模型 (SRN),将多关系问题的推理任务视为一个顺序的决策过程,并且使用基于潜力的奖励策略激励模型进行推理。文献[24]受到图双向搜索算法的启发,设计一种基于双向推理的教师网络框架,该教师网络通过双向推理学习中间监督信号,从而指导学生网络在弱监督条件下进行基于图神经网络的推理。基于图神经网络的知识推理可以较好地

能够捕捉知识图谱的结构信息,在该领域已逐渐成为当下主流的研究方法^[25]。

2 问题定义

给定一个实体集合 E 和关系集合 R ,知识图谱 G 是一组三元组集合,表示为 $G = \{(h, r, t)\} \subseteq E \times R \times E$,其中, $r \in R$ 表示头实体 $h \in E$ 和尾实体 $t \in E$ 之间的关系。本文使用 $e_i \in \mathbb{R}^d$ 和 $r_j \in \mathbb{R}^d$ 分别表示实体 i 和关系 j 的嵌入,使用 $E \in \mathbb{R}^{d \times |E|}$ 和 $R \in \mathbb{R}^{d \times |R|}$ 分别表示知识图谱中的实体和关系的嵌入矩阵,分别由 e 和 r 组成。知识图谱问答任务定义为给定一个自然语言问句 Q 与主题实体,通过对问句 Q 的语义理解,在知识图谱 G 上从主题实体经过 T 步推理得到答案实体集合 $A \subseteq E$ 。

3 本文模型设计与实现

本文提出基于层次结构图的知识图谱多跳问答模型 HSG-KBQA,其整体结构如图1所示。模型输入为问句和知识图谱,本文遵从文献[24]的工作,抽取特定于该问句的两跳内的剪枝子图,用于后续的推理工作。模型主要包括3个部分:

1)问题嵌入模块。问题嵌入模块利用依存句法树构造层次结构图和层次权重,并提出基于长短期记忆(LSTM)网络和双向图卷积神经网络(BiGCN)的编码器,从多维度编码问句,用于为每个推理步生成推理指令。

2)图池化模块。图池化模块提出虚拟节点的概念,利用基于自注意力机制的双层图池化模型得到子图的虚拟节点,虚拟表示中间实体节点,用于后续的推理工作。

3)推理模块。推理模块执行基于图神经网络的推理过程,该模块负责生成推理指令,然后通过基于向量的计算,推理指令引导模型选择匹配度高的关系路径,每次推理后更新候选实体和子图的虚拟节点状态,重复多个推理步后,选择分数最高的候选答案实体作为预测结果。

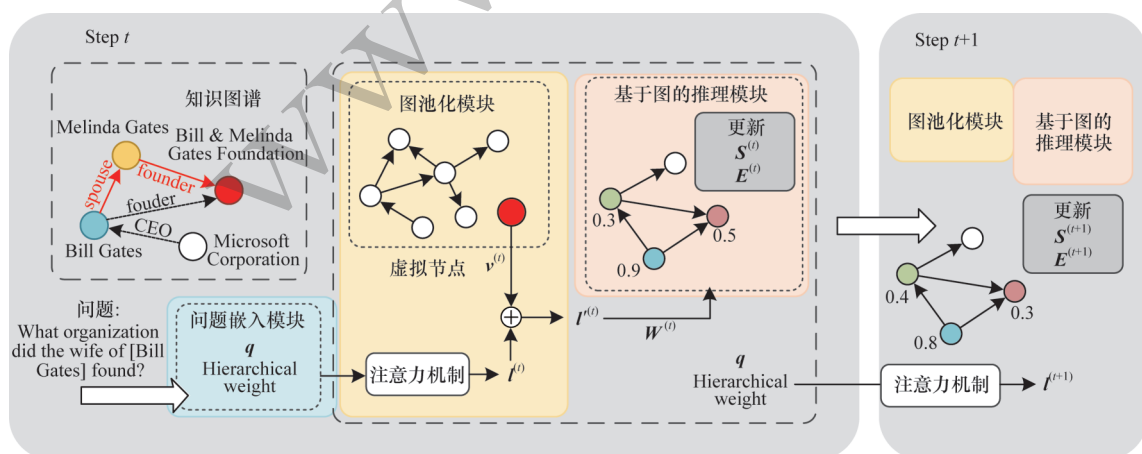


图1 HSG-KBQA 模型整体框架

Fig.1 Overall framework of HSG-KBQA model

3.1 问题嵌入模块

3.1.1 LSTM-BiGCN 编码器

以往模型常采用 LSTM 来获取问句嵌入, LSTM 是一种用于挖掘序列数据时序信息的循环神经网络, 能够通过门控机制记忆长距离的序列信息。LSTM 擅长捕捉句子中词语从前到后的位置顺序特征, 但无法捕获词语与主题实体间的层次距离特征。为了捕获这种层次特征, 本文借助依存句法树。依存句法树是用于描述词语之间依存关系的一种依存句法分析结果。从图 2(a) 可以看出, 自然语言问句中名词、谓词之间的依存关系与图谱中的实体、关系是相互对应的。利用这一特点, 本文提出将依存句法树转换为层次结构图的方法, 使其能够体现问句中待推理的关系与主题实体的层次距离, 并利用 LSTM-BiGCN 编码器编码层次结构图。

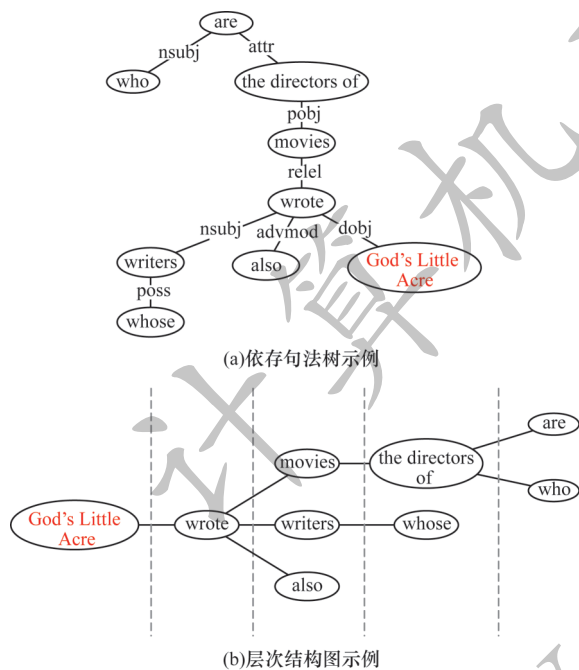


图 2 依存句法树和层次结构图

Fig.2 The dependency tree and hierarchical structure graph

在图 2 中, 以“who are the directors of movies whose writers also wrote [God's Little Acre]?”为例, 首先构建该问句的依存句法树[图 2(a)], 然后以主题实体“God's Little Acre”作为第 1 层节点, 转换树的结构, 形成层次结构图[图 2(b)], 这种图结构可以反映出句子中其他词语与主题实体在依存关系上的距离。

为了将层次结构图转换为计算机能够处理的形式, 本文在传统 LSTM 编码器的基础上增加双向 GCN 模型来学习层次结构图的嵌入表示, 获取层次结构图中分词节点局部的依存特征。最初的 GCN 是为无向图设计的, 但考虑到层次结构存在前后方向, 本文使用 BiGCN^[27] 编码层次结构图, 分别学习

出边和入边邻接节点的特征, 将两者的特征进行拼接作为最终的隐藏状态表示, 计算公式如下:

$$h_{g,u}^{l+1} = \text{ReLU} \left(\sum_{v \in \tilde{\mathcal{N}}(u)} \left(\vec{W}^l h_{g,v}^l + \vec{b}^l \right) \right) \quad (1)$$

$$h_{g,u}^{l+1} = \text{ReLU} \left(\sum_{v \in \tilde{\mathcal{N}}(u)} \left(\tilde{W}^l h_{g,v}^l + \tilde{b}^l \right) \right) \quad (2)$$

$$h_{g,u}^{l+1} = h_{g,u}^{l+1} \oplus h_{g,u}^{l+1} \quad (3)$$

其中: ReLU 为激活函数; $h_{g,u}^l$ 表示词语 u 在第 l 层的隐藏向量; $\tilde{\mathcal{N}}(u)$ 和 $\tilde{\mathcal{N}}(u)$ 分别表示词语节点 u 的入边邻域和出边邻域; \vec{W} 、 \vec{b} 和 \tilde{W} 、 \tilde{b} 分别表示出边和入边的权重矩阵、偏置项; \oplus 表示向量的连接操作。然后, 将节点特征输入 LSTM, 得到问句嵌入 q 和一组新的隐藏状态 $\{h_j\}_{j=1}^{|q|}$, $h_j \in \mathbb{R}^d$, 计算公式表示为:

$$\left[q, (h_1, h_2, \dots, h_{|q|}) \right] = \text{LSTM} \left(\{h_{g,j}\}_{j=1}^{|q|} \right) \quad (4)$$

3.1.2 层次权重

以往模型获取到问句嵌入之后, 使用基于注意力机制的解码器生成推理指令, 希望在学习不同推理步的推理指令时能够关注到问句的特定部分, 这种方式无法充分地挖掘出问句中的关系链顺序。为了能够充分利用问句中显式的关系链, 本节提出层次权重的概念, 用于学习能够体现推理顺序的推理指令, 使其在每个推理步关注正确的关系。

如图 2(b) 所示, 层次结构图能够体现待推理关系的层次特点, 即该图按照从左到右的顺序, 符合本文所预期的正确的推理关系顺序, 距离主题实体较近的关系应在推理前期被关注, 而距离较远的关系则在推理后期更重要。因此, 本文为每个层次的词语引入层次权重, 计算公式如下:

$$h_{hw,i}^{(0)} = \text{Softmax} \left(1 - \frac{i}{n+1} \right) \quad (5)$$

其中: i 表示层次结构图中的层数; n 为总层数, 规定主题实体与其后继节点属于第 1 层。为了避免重复推理某一个关系, 需要将已经推理过的关系的层次权重降低, 具体来说, 类比循环队列数据结构, 在下一步推理中, 把已被利用的关系放在最后一层, 后续层的关系向前提升一层。

3.2 图池化模块

对于自然语言问句“What organization did the wife of [Bill Gates] found?”来说, 它所包含的信息包括主题实体和一系列关系, 即“Bill Gates”“the wife of”“found organization”, 而推理的中间节点“Melinda Gates”是隐藏的。为了能够结合图谱状态来更好地推理, 本节提出虚拟节点的概念来虚拟表示中间实体节点。

如图 3 所示, 受到 SAGPool 模型^[28] 的启示, 本文采用图池化技术实现虚拟节点。

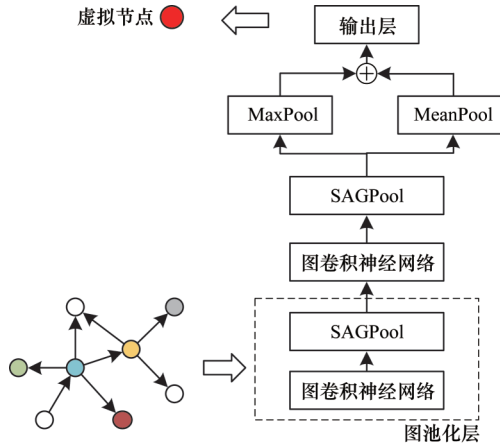


图3 图池化模块结构

Fig.3 Structure of graph pooling module

SAGPool是一种通过节点特征和图拓扑结构进行分层表示学习的技术,这种分层池化技术利用自注意力机制从图中学习节点的重要性,然后利用topK机制丢弃不重要的节点。本文利用SAGPool的思想,将子图输入基于自注意力机制的图池化层,过滤不重要的实体节点,利用2次图池化操作,经过输出层得到虚拟节点的嵌入。基于自注意力机制的图池化模块不仅能够学习当前推理步的图谱状态,还有助于减少子图中噪声实体的影响。根据SAGPool的思想,使用Kipf^[29]的图卷积公式得到自注意力分数,计算公式如下:

$$Z = \sigma \left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{X} \Theta \right) \quad (6)$$

其中: σ 是激活函数; $\tilde{\mathbf{A}} \in \mathbb{R}^{N \times N}$ 是包含自连接的邻接矩阵, N 表示子图中实体节点数目; $\tilde{\mathbf{D}} \in \mathbb{R}^{N \times N}$ 是度矩阵; $\mathbf{X} \in \mathbb{R}^{N \times d}$ 表示节点特征,即实体节点的嵌入表示; Θ 表示滤波器的参数矩阵。然后依据分数 Z 和池化比率 $k \in (0, 1]$,通过topK机制保留前 kN 个分数最高的节点,由以下公式建模:

$$\mathbf{E}_{\text{pool}}^{(t)} = \text{topK}(Z, kN) \quad (7)$$

其中: k 是超参数,决定了筛选节点的数量; $\mathbf{E}_{\text{pool}}^{(t)}$ 表示经过筛选后子图中的实体节点集合在第 t 个推理步中的嵌入矩阵。

经过2层图卷积神经网络和SAGPool的筛选操作,将需要保留的重要节点通过输出层输出。输出层使用平均池化和最大池化,之后经过一个全连接层,得到第 t 个推理步中子图的虚拟节点表示 $\mathbf{v}^{(t)}$,计算公式如下:

$$\mathbf{v}^{(t)} = \text{ReLU}(\mathbf{W}^{(t)} \mathbf{v}^{(t)} + \mathbf{b}^{(t)}) \quad (8)$$

$$\mathbf{v}^{(t)} = \text{MaxPool}(\mathbf{E}_{\text{pool}}^{(t)}) \oplus \text{MeanPool}(\mathbf{E}_{\text{pool}}^{(t)}) \quad (9)$$

其中:最大池化计算公式为 $\max_{i=1}^{N'} e_i^{(t)}$;平均池化计算公式为 $\frac{1}{N'} \sum_{i=1}^{N'} e_i^{(t)}$, N' 表示输出时子图包含的实体数量。

3.3 推理模块

如图1所示,经过以上模块获得了问句嵌入 \mathbf{q} 、层次权重 h_{hw} 和虚拟节点嵌入 $\mathbf{v}^{(t)}$,基于图的推理模块的主要任务是利用这些嵌入向量计算推理指令,指导子图中的候选实体选择匹配度高的关系,然后通过相邻的关系传播信息,更新推理后候选实体的分数,最终选出分数最高的候选实体作为答案。

为了获取推理指令,本文设计融合注意力机制与层次权重的解码器。以往模型采用基于注意力机制的解码器为每个推理步生成一条特定的推理指令向量。注意力机制可以使神经网络专注于输入的某一部分,其核心是为输入序列中的元素学习一组注意力权重来衡量元素的重要程度,即关注问句中的重要部分。本文在注意力权重的基础上加入层次权重,生成一系列推理指令 $\mathbf{l}^{(t)}$,目的是利用层次权重所体现的关系层次顺序信息,使推理指令在不同推理步能够选择合理的关系。然后,将图池化模块得到的虚拟节点嵌入 $\mathbf{v}^{(t)}$ 与推理指令向量连接,得到最终的推理指令 $\mathbf{l}^{(t)}$,这种连接形成一种(虚拟节点,推理关系,?)的形式,与知识图谱中的(头实体,关系,尾实体)相对应,意味着该推理指令融合了知识图谱和问句语义信息。需要注意的是,问句中包含了主题实体的信息,因此,在第1步推理时不需要使用虚拟节点。

$$\mathbf{l}^{(t)} = \sum_{i=1}^{|q|} \left((1-\lambda) a_i^{(t)} + \lambda h_{\text{hw},i}^{(t)} \right) \mathbf{h}_i \quad (10)$$

$$\mathbf{l}^{(t)} = \mathbf{W}_l^{(t)} (\mathbf{v}^{(t)} \oplus \mathbf{l}^{(t)}) + \mathbf{b}_l^{(t)} \quad (11)$$

$$a_i^{(t)} = \text{Softmax}(\mathbf{W}_a f_s(\mathbf{h}_i, \mathbf{q}) + \mathbf{b}_a) \quad (12)$$

其中: $a_i^{(t)}$ 是第 i 个语义单元在第 t 个推理步中的注意力权重; $h_{\text{hw},i}^{(t)}$ 是第 i 个语义单元在第 t 个推理步中的层次权重; $f_s(\mathbf{h}_i, \mathbf{q})$ 是注意力函数,本文使用 $f_s(\mathbf{h}_i, \mathbf{q}) = \mathbf{h}_i^\top \mathbf{q}$ 。

在获取推理指令之后,需要初始化实体的嵌入和分数以及关系的嵌入。首先,子图中的候选实体分数定义为 $s^{(0)} \in [0, 1]$,主题实体的初始值设置为1。实体间的关系嵌入初始化为 $\mathbf{r}^{(0)}$,通过聚合实体相邻的关系嵌入初始化实体嵌入 $\mathbf{e}^{(0)}$ 。对于与实体 u 相连的关系 r ,计算其关系权重 $\mathbf{g}_{(u,r,u)}^{(t)}$,公式如下:

$$\mathbf{g}_{(u,r,u)}^{(t)} = \sigma(\mathbf{l}^{(t)} \cdot (\mathbf{W}_{rq} \mathbf{r})) \quad (13)$$

其中: $\mathbf{W}_{rq} \in \mathbb{R}^{d \times d}$ 是可学习的权重系数矩阵。

受到R-GCN^[30]的启示,本文通过聚合相邻关系和相邻实体的信息,以更新实体在该推理步的嵌入,计算公式如下:

$$\mathbf{e}_u^{(t)} = \sigma \left(\mathbf{W}_0^{(t-1)} \mathbf{e}_u^{(t-1)} + \sum_{r \in \mathcal{R}_{u' \in \mathcal{N}_r(u)}} \sum_{c_{u,r}} \frac{1}{c_{u,r}} \mathbf{W}_r^{(t-1)} \tilde{\mathbf{e}}_{u'}^{(t-1)} \right) \quad (14)$$

$$\tilde{\mathbf{e}}_{u'}^{(t-1)} = s_{u'}^{(t-1)} \mathbf{g}_{(u',r,u)}^{(t)} \quad (15)$$

其中: $\vec{\mathcal{N}}_r(u)$ 表示与实体 u 具有关系 r 的相邻实体集合; $\mathbf{W}_r^{(t-1)} \in \mathbb{R}^{d \times d}$ 表示关系 r 的权重矩阵; $c_{u,r}$ 表示正则化系数, 取 $|\vec{\mathcal{N}}_r(u)|$ 。接着, 通过实体的嵌入对实体进行评分, 计算公式为:

$$s_u^{(t)} = \text{Softmax}\left(\left(e_u^{(t)}\right)^\top \mathbf{w}\right) \quad (16)$$

其中: $\mathbf{w} \in \mathbb{R}^d$ 表示可学习参数。

3.4 模型训练

模型使用 KL 散度作为损失函数, KL 散度是 2 个概率分布间差异的非对称性度量^[31]。对于训练集中的样本, 正确答案实体集合为 $A_{\text{groundtruth}} = \{e_1, e_2, \dots, e_{|A_{\text{groundtruth}}|}\}$, 则实体的概率分布表示为:

$$y_i = \begin{cases} \frac{1}{|A_{\text{groundtruth}}|}, & e_i \in A_{\text{groundtruth}} \\ 0, & e_i \notin A_{\text{groundtruth}} \end{cases} \quad (17)$$

将预测的分数作为预测概率分布, 则损失函数计算为:

$$\mathcal{L} = \text{KL}(s \parallel y) = \sum s_x \lg \frac{s_x}{y_x} \quad (18)$$

损失函数越小, 预测结果的概率分布越接近正确结果。

4 实验验证

4.1 实验数据与参数设置

WebQuestionsSP^[32]是以 Freebase 图谱为基础而构建的数据集, 是目前最常用的自然语言问答数据集之一, 该数据集包含了简单问题和复杂问题、多实体问题和单实体问题, 共有 4 737 条问题。本文实验采用 GraftNet 的划分结果, 训练数据集 2 849 条, 测试数据集 1 639 条, 验证数据集 249 条。Freebase 是一个超大型的知识图谱, 拥有超过 5 800 万个实体、32 亿个三元组以及近 4 万个属性, 是规模最大的知识图谱之一。因此, WebQuestionsSP 数据集的训练集规模相较于其所依赖的大规模复杂知识图谱来说较小, 使得实验具有一定的挑战性。

PathQuestions 是 IRN 文章引用 Freebase 的子集作为知识图谱而构建的多跳问答数据集, 具有 2 215 个实体和 14 个关系。该数据集分别包含了 1 908 个两跳数据 PQ-2H 和 5 198 个三跳数据 PQ-3H, PQ-M 是指混合数据。本文实验采用 IRN 的划分方法, 其中, 两跳数据的训练集为 1 526 条, 验证集为 191 条, 测试集为 191 条, 三跳数据的训练集为 4 158 条, 验证集为 520 条, 测试集为 520 条。

实验参数设置情况为: 采用 PyTorch 深度学习框架, 使用 Adam 优化器, 推理步数设置为 3, 式(10)中的参数 λ 设置为 0.8, 池化比率 k 设置为 0.5。在 WebQuestionsSP 中学习率设置为 0.000 5, 实体、问句的词向量维度为 100, 批大小设置为 30, 子图包含 2 000 个实体。在 PathQuestions 中学习率设置为 0.000 5, 实体、问句的词向量维度为 150, 批大小设置

为 50, 子图包含 400 个实体。

4.2 对比模型与评价指标

为了验证本文模型的有效性, 采用基于信息检索方法的主流模型作为对照基准, 具体如下:

1) KVMem 是较早实现多跳知识图谱问答推理的模型之一, 将 Memory Network 引入 KBQA 中, 使用哈希方法选取三元组以键值对的形式存储, 键存储头实体和关系的组合, 值存储尾实体, 通过迭代地检索键值对操作完成多跳推理任务。

2) Rce-KGQA 在基于知识图谱嵌入的问答模型上引入关系链推理, 抽取主题实体到候选实体之间的关系链, 使用孪生网络模型学习关系链和问题之间的语义相似度。

3) NSM^[10] 是用于视觉问答的神经状态机模型, 在多跳推理的过程中不断更新实体状态, 以得到最终的答案实体。

4) SRN^[23] 是一种基于强化学习的逐步推理模型, 该模型将多关系问题转化为顺序决策的过程, 并使用波束搜索减少候选实体的数量。

5) DCRN^[33] 是为了解决在多跳问答中候选实体集过大问题而提出的深度认知推理网络。DCRN 由两阶段推理组成, 设计一个基于 RNN 的路径解码器, 在第 1 阶段利用该解码器和候选实体的语义信息来缩小第 2 阶段的候选实体集。

6) sMoCo^[34] 用于解决多跳问答中知识图谱缺少相邻实体和关系的问题。模型利用图谱中的关系和实体生成新的简单问答对, 提出 sMoCo 对比学习算法以最大化正问答对之间的相似性, 最后利用整体识别模块推断得到最佳答案。

7) IRN^[22] 模型是一个具有可解释性的记忆推理网络, 该网络在推理过程中动态地关注问题, 通过上一跳预测的关系来更新问题嵌入, 并构建一个状态空间对历史推理信息进行编码。

8) MINERVA^[35] 模型考虑到知识图谱的不完整性以及缺失关键关系边的问题, 提出一种基于强化学习的路径搜索模型。该模型通过在知识图谱上走来查询答案, 并在抵达答案节点后停止。

9) HRPE^[36] 设计一种具有领域自适应性的前提假设对推理模型, 由候选实体到主题实体的三元组转化的若干条短问句组成前提集合, 由候选实体替换原始问句疑问词得到的语句作为假设集合, 两者交互预测答案集合。

本文选取 Hits@1 作为评价指标, Hits@1 指排序第 1 个的实体为正确答案的占比。

4.3 结果分析

本文模型以及对照基准模型在 WebQuestionsSP 数据集上的实验结果如表 1 所示, 最优结果加粗标注, 可以看出本文模型的性能相对于基准模型有明显提升。与考虑关系顺序和方向的最优框架 Rce-KGQA 相比, 本文模型的 Hits@1 结果高出 0.9 个百

分点;与多跳推理基础框架NSM相比,本文模型的Hits@1结果高出2.8个百分点。所有模型在PathQuestions数据集上的实验结果如表2所示,可以看出本文模型在PathQuestions数据集上的整体结果均高于基线模型,其中,在两跳数据集上的结果有明显提升,说明本文模型提出的层次结构图和层次权重能够更有效地理解推理意图,联合虚拟节点在每一个推理步选择合理的关系,可以提升多跳推理的准确性。

表1 模型在WebQuestionsSP数据集上的实验结果

模型	Hits@1	%
NSM	68.5	
SRN	68.5	
DCRN	67.8	
sMoCo	69.2	
Rce-KGQA	70.4	
HSG-KBQA	71.3	

表2 模型在PathQuestions数据集上的实验结果

模型	PathQuestions			%
	PQ-2H	PQ-3H	PQ-M	
KVMem	89.50	79.20	85.2	
MINERVA	75.90	71.20	73.1	
IRN	91.90	83.30	85.8	
SRN	96.30	89.20	89.3	
HRPE	56.54	88.96	—	
HSG-KBQA	97.30	89.70	90.2	

本文对WebQuestionsSP数据集上的训练集损失和验证集损失进行分析,结果如图4所示。从图4可以看出,在100轮次后,训练集损失曲线和验证集损失曲线均趋于收敛,验证集损失略小于训练集损失,原因可能是训练集样本数较小,但数值相差不大。图4结果说明该模型得到了良好的拟合效果。

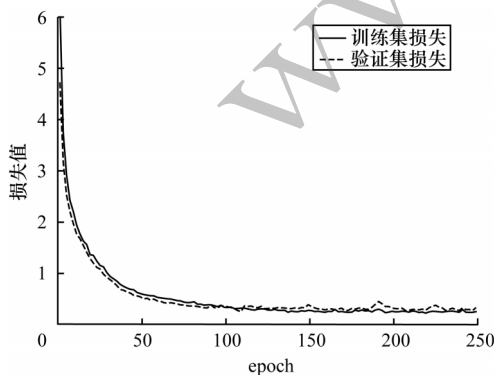


图4 损失函数曲线

Fig.4 Loss function curves

此外,本文还在WebQuestionsSP数据集上对参数 λ 的取值进行了分析,实验结果如表3所示。参数 λ 控制了层次权重在推理指令中的占比,参数 λ 越大,层次权重占比越大。从表3可以看出,随着 λ 的增加,Hits@1提升,验证了层次权重的有效性。

表3 参数 λ 的影响分析

λ 取值	Hits@1/%
0.0	68.7
0.3	69.4
0.8	71.3

4.4 消融实验

为了验证模型各个模块的有效性,本文在WebQuestionsSP数据集上进行消融实验,去掉不同模块得到的实验结果如表4所示。

表4 消融实验结果

模型	Hits@1	%
HSG-KBQA	71.3	
w/o GraphPool	66.0	
w/o Hierarchical weight	68.7	
w/o BiGCN	70.8	

从表4可以看出:

1) w/o GraphPool表示删除虚拟节点,其Hits@1相对HSG-KBQA下降了5.3个百分点,说明仅使用问句生成推理指令导致推理与所依赖的知识图谱分离,这对模型的影响较大。在推理指令中加入虚拟节点可以有效地融合子图的推理状态,从而更好地进行决策。

2) w/o Hierarchical weight表示去掉层次权重,将式(10)中的系数 λ 设置为0。实验结果显示,w/o Hierarchical weight的Hits@1相对HSG-KBQA下降了2.6个百分点,说明使用层次权重引导选择正确的关系顺序是有必要的。

3) w/o BiGCN表示去掉问句嵌入中的层次结构图编码模块,仅使用LSTM编码问句。实验结果显示,w/o BiGCN的Hits@1下降了0.5个百分点,说明层次结构图有利于模型理解问句语义,但模型效果提升不高,可能是因为层次结构图所包含的节点和边较少,并且仅使用一层图神经网络也可能导致效果不佳。

5 结束语

针对多跳知识图谱问答中推理关系链的关系顺序问题,本文提出一种基于层次结构图的多跳知识图谱问答模型。该模型首先通过层次结构图和层次权重显式地提炼问句的关系链,使其能够在每个推理步关注到问句中合理的推理意图;其次利用虚拟

节点表示每个推理步的知识图谱状态,联合问句的语义信息共同生成推理指令,进一步提升模型的答案准确率;最后在 WebQuestionsSP 和 PathQuestions 数据集上进行实验,结果表明,该模型表现出了较好的预测性能,消融实验结果也验证了各模块的有效性。但是,本文模型需要借助工具生成问句的依存句法树,而依存句法树的合理性将会对模型准确率产生较大影响。因此,下一步的研究将分为3个方向:1)优化层次结构图的编码模块;2)优化层次权重的计算公式,尝试设计为可学习的层次权重;3)优化依存句法树的生成工具,提升其处理长难句的准确率。

参考文献

- [1] 吴天波,刘露平,罗晓东,等. 基于弱依赖信息的知识库问答方法[J]. 计算机工程, 2021, 47(6): 76-82.
WU T B, LIU L P, LUO X D, et al. Knowledge base question answering method based on weak dependency information[J]. Computer Engineering, 2021, 47(6): 76-82. (in Chinese)
- [2] 吴天波,周欣,程军军,等. 基于位置和注意力联合表示的知识图谱问答[J]. 计算机工程, 2022, 48(8): 98-104, 112.
WU T B, ZHOU X, CHENG J J, et al. Knowledge graph question-answering based on joint location and attention representation[J]. Computer Engineering, 2022, 48(8): 98-104, 112. (in Chinese)
- [3] BORDES A, USUNIER N, CHOPRA S, et al. Large-scale simple question answering with memory networks[EB/OL]. [2022-11-05]. <https://arxiv.org/abs/1506.02075>. pdf.
- [4] 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.]: Association for Computational Linguistics, 2015: 260-269.
- [5] LAN Y S, WANG S H, JIANG J. Knowledge base question answering with topic units[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. Washington D. C., USA: IEEE Press, 2019: 5046-5052.
- [6] MAHESHWARI G, TRIVEDI P, LUKOVNIKOV D, et al. Learning to rank query graphs for complex question answering over knowledge graphs[EB/OL]. [2022-11-05]. <https://arxiv.org/abs/1811.01118>.
- [7] QIU Y Q, ZHANG K, WANG Y Z, et al. Hierarchical query graph generation for complex question answering over knowledge graph[C]//Proceedings of the 29th ACM International Conference on Information & Knowledge Management. New York, USA: ACM Press, 2020: 1285-1294.
- [8] LUO K Q, LIN F L, LUO X S, et al. Knowledge base question answering via encoding of complex query graphs[C]//Proceedings of 2018 Conference on Empirical Methods in Natural Language Processing. [S. l.]: Association for Computational Linguistics, 2018: 2185-2194.
- [9] JIN W, ZHAO B, YU H, et al. Improving embedded knowledge graph multi-hop question answering by introducing relational chain reasoning[EB/OL]. [2022-11-05]. <https://arxiv.org/abs/2110.12679>. pdf.
- [10] DREW A H, CHRISTOPHER D. Learning by abstraction: the neural state machine[C]//Proceedings of 2019 Conference on Neural Information Processing Systems. New York, USA: ACM Press, 2019: 5901-5914.
- [11] HAN J L, CHENG B, WANG X. Two-phase hypergraph based reasoning with dynamic relations for multi-hop KBQA[C]//Proceedings of the 29th International Joint Conference on Artificial Intelligence. New York, USA: ACM Press, 2020: 3615-3621.
- [12] HAN J L, CHENG B, WANG X. Open domain question answering based on text enhanced knowledge graph with hyperedge infusion[EB/OL]. [2022-11-05]. <https://aclanthology.org/2020.findings-emnlp.133/>.
- [13] 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.]: Association for Computational Linguistics, 2015: 1321-1331.
- [14] 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.]: Association for Computational Linguistics, 2020: 969-974.
- [15] SUN Y W, ZHANG L L, CHENG G, et al. SPARQA: skeleton-based semantic parsing for complex questions over knowledge bases[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(5): 8952-8959.
- [16] KAPANIPATHI P, ABDELAZIZ I, RAVISHANKAR S. Question answering over knowledge bases by leveraging semantic parsing and neuro-symbolic reasoning[EB/OL]. [2022-11-05]. <https://arxiv.org/abs/2012.01707>.
- [17] 张鹏举,贾永辉,陈文亮. 基于多特征实体消歧的中文知识图谱问答[J]. 计算机工程, 2022, 48(2): 47-54.
ZHANG P J, JIA Y H, CHEN W L. Chinese knowledge based question answering based on multi-feature entity disambiguation[J]. Computer Engineering, 2022, 48(2): 47-54. (in Chinese)
- [18] SUN H T, DHINGRA B, ZAHEER M, et al. Open domain question answering using early fusion of knowledge bases and text[C]//Proceedings of 2018 Conference on Empirical Methods in Natural Language Processing. [S. l.]: Association for Computational Linguistics, 2018: 4231-4242.
- [19] 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.]: Association for Computational Linguistics, 2020: 4498-4507.
- [20] TROUILLON T, WELBL J, RIEDEL S, et al. Complex embeddings for simple link prediction[C]//Proceedings of the 33rd International Conference on Machine Learning. New York, USA: ACM Press, 2016: 2071-2080.
- [21] MILLER A, FISCH A, DODGE J, et al. Key-value memory networks for directly reading documents[C]//Proceedings of 2016 Conference on Empirical Methods in Natural

- Language Processing. [S. l.]: Association for Computational Linguistics, 2016; 1400-1409.
- [22] 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.]: Association for Computational Linguistics, 2018; 2010-2022.
- [23] QIU Y Q, WANG Y Z, JIN X L, et al. Stepwise reasoning for multi-relation question answering over knowledge graph with weak supervision [C]// Proceedings of the 13th International Conference on Web Search and Data Mining. New York, USA: ACM Press, 2020; 474-482.
- [24] HE G L, LAN Y S, JIANG J, et al. Improving multi-hop knowledge base question answering by learning intermediate supervision signals[C]// Proceedings of the 14th ACM International Conference on Web Search and Data Mining. New York, USA: ACM Press, 2021; 553-561.
- [25] 孙水发, 李小龙, 李伟生, 等. 图神经网络应用于知识图谱推理的研究综述[J]. 计算机科学与探索, 2023, 17(1): 27-52.
SUN S F, LI X L, LI W S, et al. Review of graph neural networks applied to knowledge graph reasoning[J]. Journal of Frontiers of Computer Science & Technology, 2023, 17(1): 27-52. (in Chinese)
- [26] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [27] FU T J, LI P H, MA W Y. GraphRel: modeling text as relational graphs for joint entity and relation extraction[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. [S. l.]: Association for Computational Linguistics, 2019; 1409-1418.
- [28] LEE J, LEE I, KANG J. Self-attention graph pooling[C]// Proceedings of the 36th International Conference on Machine Learning. New York, USA: ACM Press, 2019; 3734-3743.
- [29] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[C]// Proceedings of the 5th International Conference on Learning Representations. New York, USA: ACM Press, 2017; 126-135.
- [30] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks [EB/OL]. [2022-11-05]. <https://arxiv.org/abs/1703.06103>.
- [31] KULLBACK S, LEIBLER R A. On information and sufficiency[J]. The Annals of Mathematical Statistics, 1951, 22(1): 79-86.
- [32] 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. [S. l.]: Association for Computational Linguistics, 2016; 201-206.
- [33] CAI J Y, ZHANG Z Q, WU F. Deep cognitive reasoning network for multi-hop question answering over knowledge graph [C]// Proceedings of the 59th Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing. [S. l.]: Association for Computational Linguistics, 2021; 219-229.
- [34] CN J E, MEI S J, HU X R, et al. Seeing the wood for the trees; a contrastive regularization method for the low-resource knowledge base question answering [EB/OL]. [2022-11-05]. <https://aclanthology.org/2022.findings-naacl.82.pdf>.
- [35] 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 [EB/OL]. [2022-11-05]. <https://arxiv.org/abs/1711.05851>. pdf.
- [36] GUPTA A, ANNERVAZ K M, DUKKIPATI A, et al. An inference approach to question answering over knowledge graphs [EB/OL]. [2022-11-05]. <https://arxiv.org/abs/2112.11070>. pdf.

编辑 吴云芳