

## ◆ 机器学习模型 ◆

结合问题-关系注意力和特征视图匹配的关系检测方法<sup>\*</sup>王 栋<sup>1</sup>,王海荣<sup>1,2\*\*</sup>,车 森<sup>1</sup>,孙 崇<sup>1</sup>

(1. 北方民族大学计算机科学与工程学院,宁夏银川 750021;2. 北方民族大学图像图形智能处理国家民委重点实验室,宁夏银川 750021)

**摘要:**问答系统作为信息检索的一种高级形式,已成为人工智能和自然语言处理领域中一个备受关注的研究方向。本文聚焦于知识图谱问答(Knowledge Graph Question Answering,KGQA)研究中的关系检测,针对现有方法中未能较好兼顾全局语义和局部语义信息,以及复杂问题准确率不高的问题,提出了一种结合问题-关系注意力和特征视图匹配的关系检测方法。该方法从问题和知识库中提取多粒度的特征,将提取特征构造成特征对视图作为关系检测模型的输入,视图内部利用双边多视角匹配(Bilateral Multi-Perspective Matching, BiMPM)进行比较匹配,得出关系预测结果。为验证本文提出的方法,在 SimpleQuestions、WebQSP 数据集上,与 6 种主流基线方法进行对比实验,本方法的准确率分别提升 3.42 个和 0.45 个百分点。

**关键词:**关系检测;问答系统;注意力机制;信息检索;特征匹配

中图分类号:TP3 文献标识码:A 文章编号:1005-9164(2023)01-0079-10

DOI:10.13656/j.cnki.gxkx.20230308.009

知识图谱问答(Knowledge Graph Question Answering,KGQA)作为近年来广受关注的自然语言处理任务,目标是在给定知识图谱(知识库)的前提下,对用户提出的问题,在图谱中找到与其答案最匹配的实体或实体集合返回。目前,KGQA 主要有两大类方法:基于语义解析(Semantic Parsing,SP)的方法和基于信息检索(Information Retrieval,IR)的方法。随着深度学习的发展,将深度学习与这两种传统方法相结合的、基于深度学习的方法因适应性更强、灵活

性更好,得到快速发展。基于深度学习的方法一般是使用 n-gram 或序列标注等技术对问题文本进行实体识别、筛选后做实体链接,根据链接后的实体查询知识库得到知识库中的一系列候选事实,用经过训练的神经网络模型计算问题与候选事实间的相关性,最终得到问题的答案。这类方法的核心是设计使用编码器网络对问题和知识库中的知识进行编码,使问题与知识库中对应的知识在编码空间上尽可能相近。通常,KGQA 包括两个子任务:①实体链接,抽取问

收稿日期:2022-09-27

修回日期:2022-10-09

\* 北方民族大学校级科研项目(2021XYZJK06)资助。

【第一作者简介】

王 栋(1998-),男,在读硕士研究生,主要从事知识图谱问答研究。

【\*\*通信作者】

王海荣(1977-),女,博士,副教授,主要从事大数据研究,E-mail:bmdwhr@163.com。

【引用本文】

王栋,王海荣,车森,等. 结合问题-关系注意力和特征视图匹配的关系检测方法[J]. 广西科学,2023,30(1):79-88.

WANG D,WANG H R,CHE M,et al. Relation Detection Method Combining Question-Relation Attention and Feature View Matching [J]. Guangxi Sciences,2023,30(1):79-88.

题中出现的实体,并将其链接到知识库中对应的实体对象;②关系检测,检测候选关系与问题的相关性。

近年来,大量研究集中在提高关系检测任务精确度上。针对问题和候选关系特征提取,Qu等<sup>[1]</sup>从语义级和词级两个粒度对抽象问题和关系的相关性进行建模。张仰森等<sup>[2]</sup>提出一种融合语义信息与问题类型、中心词两类关键问题信息的多阶段注意力模型。Zhang等<sup>[3]</sup>提出基于注意力机制的词级交互模型,采用匹配-聚合架构缓解最大池化操作造成的信息丢失问题。谢金峰等<sup>[4]</sup>提出基于多语义相似性的关系检测方法,通过融合问题和关系的全局语义、局部语义,捕获多样化的特征来提升关系检测性能。Chen等<sup>[5]</sup>提出基于Transformer的深度注意力语义匹配模型(Deep Attention Matching, DAM),该模型应用细粒度的词级注意力提高问题和关系的匹配度。Luo等<sup>[6]</sup>提出一种单关系知识问答(Single-Relation Question Answering, SR-QA)方法,利用构造的基于BERT的关系感知注意力网络以增强候选对象的表示,弥合问题和知识库事实之间的语义鸿沟。

上述研究成果均证明基于注意力机制的循环神经网络(Recurrent Neural Network, RNN)可以显著提高关系检测任务的准确率,但现有方法难以刻画抽象问题与知识库关系间复杂而高级的语义联系<sup>[7-10]</sup>。因此,本文提出一种基于问题-关系注意力机制和特征视图匹配的关系检测模型,结合多层长短期记忆网络(Long Short-Term Memory, LSTM),利用自注意力机制,从多个角度对文本的不同粒度特征进行建模,以提取问题和关系之间更加丰富的语义关系,提高实体关系检测的准确率。

## 1 知识库关系检测

相较于一般的关系检测,KGQA中的关系检测具有以下特点:①关系数量大。一般意义上的关系检测,其关系数量一般小于100,而对于KGQA任务来说,即使是较小的知识库,如Freebase2M<sup>[11]</sup>,其关系数量也高达6000。②KGQA中的关系检测是一种zero-shot学习任务。③像WebQuestions<sup>[12]</sup>这种相对复杂的KGQA任务,预测的不仅是一个单一的关系,而且是一个关系序列。

KGQA中的关系检测任务可形式化定义为给定一个输入问题 $q$ 和一组经过实体检测、实体链接后得

到的候选关系 $R = \{r_1, \dots, r_i\}$ ,其中, $r_i$ 在简单问题中表示单关系,在复杂问题中表示关系路径(可能为单关系也可能为关系链)。关系检测的任务是识别出与问题相关度最高的候选关系 $r_q$ ,因此,关系检测可以转化为通过构造学习评分函数 $S(r, q)$ 优化某些排名损失,这个过程可形式化描述为

$$r_q = \operatorname{argmax}_{r_i \in R} S(r_i, q) \quad (1)$$

作为关系检测任务的关键步骤,自然语言句子匹配(Natural Language Sentence Matching, NLSM)是许多自然语言处理(Natural Language Process, NLP)任务的基础,如释义检测、自然语言推理和答案选择。现有序列相似性计算方法按体系结构可分为两大类:孪生网络(Siamese)体系结构和匹配-聚合体系结构。在Siamese体系结构中,研究了一系列编码器,如LSTM<sup>[13]</sup>、卷积神经网络(Convolutional Neural Network, CNN)<sup>[14]</sup>和基于注意力的模型<sup>[15,16]</sup>,利用这些编码器和模型将两个待匹配序列编码为向量以进行比较。但Siamese体系结构将一系列向量合并为一个向量的池化操作可能会发生信息丢失。匹配-聚合方法<sup>[17-19]</sup>对输入序列在细粒度级别上进行匹配和比较来缓解信息流的瓶颈,通过合并比较结果估计相似性分数。

## 2 方法模型

### 2.1 模型框架

模型通过构造匹配不同的问题关系特征视图实现关系检测,每一个视图都可以看作是一系列标记(词或关系名称)。对于关系 $r$ 的第 $i$ 个视图,有 $r^{(i)} = \{r_1^{(i)}, \dots, r_{M_i}^{(i)}\}$ , $M_i$ 是视图 $r^{(i)}$ 的词序列长度;问题端的定义与关系略有不同,将在第2.2.2节说明。最终得到问题 $q = \{q^{(1)}, \dots, q^{(N_q)}\}$ 和关系 $r = \{r^{(1)}, \dots, r^{(N_r)}\}$ 作为匹配模块的输入, $N_q$ 和 $N_r$ 分别表示 $q$ 和 $r$ 的视图数,它们可能不相等。

本文模型的总体框架见图1。除了实体别名和实体对之外,每对输入都从多个角度进行匹配,然后将实体对的表示和其余对的匹配表示进行聚合以进行最终预测。所提出的关系检测模型包含4个主要组件:输入模块、上下文表示模块、匹配模块和聚合模块。

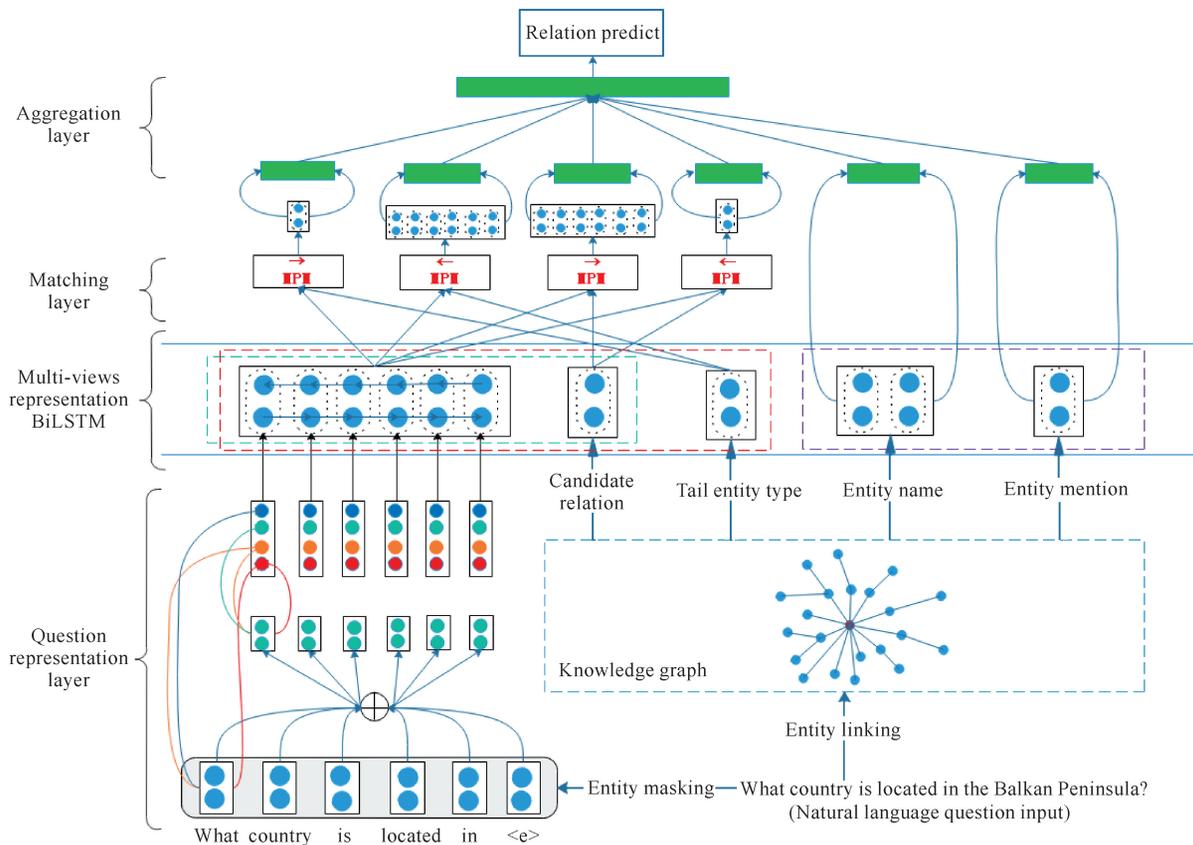


图1 本文模型总体框架图

Fig. 1 Overall frame diagram of the model in this study

## 2.2 模型实现

### 2.2.1 关系检测特征视图

对于输入问题  $q$ , 模型从关系名称及其对应的尾实体(即问题的答案)类型中生成视图, 并在模型中使用以下 3 对输入。

①<实体名, 实体提及>: 从问题和知识库中获取实体信息。

②<关系名, 抽象问题>: 获取输入问题和候选关系之间的交互。按照 Yu 等<sup>[20]</sup>的工作, 将问题中提到的实体替换为一个特殊标记(“Balkan Peninsula”在图 1 中被替换为标记<e>), 这样就得到一个抽象问题, 通过遮蔽问题中的实体, 使模型可以更好地将候选关系名称与问题中实体的上下文相匹配, 减少实体词对这一过程的干扰。

③<尾实体类型, 抽象问题>: 用来计算与关系连接的尾实体类型和问题的匹配程度。受之前工作的启发, 模型也使用了尾实体类型, 表明关系的后半部分在逻辑上可以暗示尾实体。一般情况下, 知识库关系只有特定类型的尾实体。例如, 对于示例问题 “What country is located in the Balkan Peninsula?”,

Freebase 中的对应关系是 contains, 尾实体只能是 location 类型。这种关系和其他关系, 如 adjoins, street\_address, nearby\_airport, people\_born\_here 中, 只能将 location 作为尾实体, 但这些关系没有明确包含预期作为答案的实体类型的单词, 换言之, 这种对应只反映在语义上而非字面上。因此, 利用尾实体类型信息可能有助于提高关系检测性能。关系实例中的尾实体可能与多种类型关联, 给定三元组<The\_Audacity\_of\_Hope, author, Barack\_Obama>, Barack\_Obama 的类型从一个人到更具体的 writer、politician、us\_president 等。所以, 当给定关系 author, 删减不相关的实体类型 (politician、us\_president) 并保留相关的实体类型 (person、writer) 就显得至关重要。为了实现这一步, 从 Freebase 为每个唯一关系获取最多 600 个实例并查询获得每个尾实体的类型。最后, 只保留至少 98% 的尾实体类型。如果这种方法找不到关系的任何尾实体类型, 则使用默认的特殊标记。一旦获得了特定关系的尾实体类型, 通过连接每个尾部类型中的单词来形成一个字符串。

所有候选关系的第一对输入均保持不变, 即不需

要像其他输入对那样通过多角度匹配进行彻底比较, 以用于区分候选关系。对于第二视图和第三视图的输入, 模型生成两个匹配特征向量, 每个匹配方向一个(即对于视图对 $\langle a, b \rangle$ , 方向是 $a \rightarrow b$ 和 $a \leftarrow b$ )。最后, 将这两对交互信息结合起来, 形成一个高级的联合视图, 可以帮助检测最有可能正确的关系, 得出给定问题、候选关系名称以及相应的尾实体类型的匹配程度。

### 2.2.2 输入模块

所有视图的输入都是词序列, 除问题词序列外, 模型分两步对每个序列进行编码。首先, 为每个词构造一个 $d$ 维向量, 该向量由两部分组成: 词嵌入和字符嵌入。词嵌入是一种固定的、经过预训练的向量, 如 Glove、word2vec。字符嵌入是通过将单词中的每个字符(也由向量表示)馈入 LSTM 计算。

对于输入问题, 若问题中所有标记总数为 $|Q|$ , 则问题词序列可表示为 $q = \{\omega_1, \omega_2, \dots, \omega_{|Q|}\}$ 。问题序列通过预训练的固定向量进行词嵌入, 得到词嵌入向量 $E_q = \{E_{q_1}, E_{q_2}, \dots, E_{q_{|Q|}}\}$ 。

对词嵌入向量应用注意力机制, 使模型在处理长序列时可以获得更好的性能, 对 $q = \{\omega_1, \omega_2, \dots, \omega_{|Q|}\}$ 和 $r = \{r_1, r_2, \dots, r_{|R|}\}$ 序列,  $q$ 中的每个词 $q_i$ 的注意力权重

$$a_{i,j} = q_i^T r_j. \quad (2)$$

使用 Softmax 公式计算 $q$ 的问题-关系注意力加权问题向量 $\bar{q} = (\bar{q}_1, \dots, \bar{q}_{|Q|})$ :

$$\bar{q}_i = \sum_{j=1}^{|R|} \frac{\exp(a_{i,j})}{\sum_{k=1}^{|R|} \exp(a_{i,k})} q_i. \quad (3)$$

最后, 将加权后的问题嵌入馈入双向长短期记忆(Bi-directional Long Short-Term Memory, BiLSTM)网络层得到问题的语义向量

$$v_q = \text{BiLSTM}(m_q), \quad (4)$$

$$m_{q_i} = [q_i, \bar{q}_i, (q_i - \bar{q}_i), (q_i \odot \bar{q}_i)], \quad (5)$$

式中, “-”表示逐元素减法, “ $\odot$ ”表示逐元素乘法,  $v_q \in \mathbb{R}^{|Q| \times d_q}$ ;  $m_{q_i} \in \mathbb{R}^{4 \times d_e}$ , “[, ]”表示连接操作。至此, 得到了问题的语义向量。

### 2.2.3 上下文表示模块

使用相同的 BiLSTM 对输入的视图进行编码, 在匹配模块中使用每个 BiLSTM 输出上下文向量。将问题的上下文向量输入多个匹配模块, 以匹配关系和尾实体类型。本模块的目的是将上下文信息整合到每个输入序列的每个时间步的表示中。利用 BiL-

STM 对输入序列的每个时间步进行上下文嵌入编码, 并获得每个单词位置的隐状态。实际操作中, 为这两种编码器提供一个单独的共享编码器效果更好。

作为输入, 单词由行向量 $x \in \mathbb{R}^n$ 表示,  $x$ 可以是单词嵌入和单词特征的串联。第 $t$ 个单词的单词向量是 $x_t$ 。使用具有 LSTM 的 RNN 编码器处理单词序列, 它被证明在许多类型的 NLP 任务中有效, 包括机器阅读理解<sup>[21-23]</sup>和神经机器翻译<sup>[24-26]</sup>任务。对于每个位置 $t$ , LSTM 使用输入 $x_t$ 和前一个隐状态 $h_{t-1}$ 计算 $h_t$ :

$$i_t = \sigma(W_i x_t + U_i h_{t-1}), \quad (6)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1}), \quad (7)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1}), \quad (8)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1}), \quad (9)$$

$$c_t = f_t \cdot \tilde{c}_{t-1} + i_t \cdot \tilde{c}_t, \quad (10)$$

$$h_t = o_t \cdot \tanh(c_t), \quad (11)$$

式中,  $i_t, f_t, o_t, c_t$ 和 $h_t \in \mathbb{R}^d$ 分别是输入门、遗忘门、输出门、存储单元和 $d$ 维隐状态;  $\tilde{c}_t$ 是单元状态的更新值;  $W_{\{i,f,o,c\}}, U_{\{i,f,o,c\}} \in \mathbb{R}^{n \times d}$ 和 $U_{\{i,f,o,c\}}, U \in \mathbb{R}^{d \times d}$ 是 LSTM 的参数;  $\sigma$ 是 sigmoid 函数。对于 $t$ 中的一个词, 使用来自前向 LSTM 的隐藏状态 $\vec{h}_t$ 作为前一下文文的表示, 以及来自反向 LSTM、对文本进行反向编码作为后一下文文的 $\overleftarrow{h}_t$ , 让 $h_t = [\vec{h}_t; \overleftarrow{h}_t]$ , 形成 $x_t$ 的双向上下文编码, 如式(12)所示。为区分不同来源的隐状态, 将 $P$ 中第 $j$ 个单词的 $h_j$ 和 $Q$ 中第 $k$ 个单词的 $h_k$ 分别表示为 $h_j^p$ 和 $h_k^q$ 。

$$\begin{aligned} \vec{h}_i &= \overrightarrow{\text{LSTM}}(\vec{h}_{i-1}, p_i), i = 1, \dots, M, \\ \overleftarrow{h}_i &= \overleftarrow{\text{LSTM}}(\overleftarrow{h}_{i+1}, p_i), i = M, \dots, 1. \end{aligned} \quad (12)$$

### 2.2.4 匹配模块

关系检测的核心任务是计算关系与给定问题之间的信息相关性, 为此设计了带有注意力机制的匹配模块, 以更好地匹配关系的每个视图与给定问题。注意力机制之所以重要, 是因为关系的不同视图通常对应问题的不同部分, 如疑问词通常更能代表关系类型。

本模块使用经过修改的双边多视角匹配(BiMPPM)模型<sup>[27]</sup>, 其结构如图2所示。BiMPPM 在多个文本匹配任务中表现出良好的性能。假设 BiMPPM 也可以有效地用于关系检测, 因为问题的唯一视图可能需要与关系或尾实体类型匹配, 并且匹配方法应该在多粒度和多角度上匹配与问题的关系。

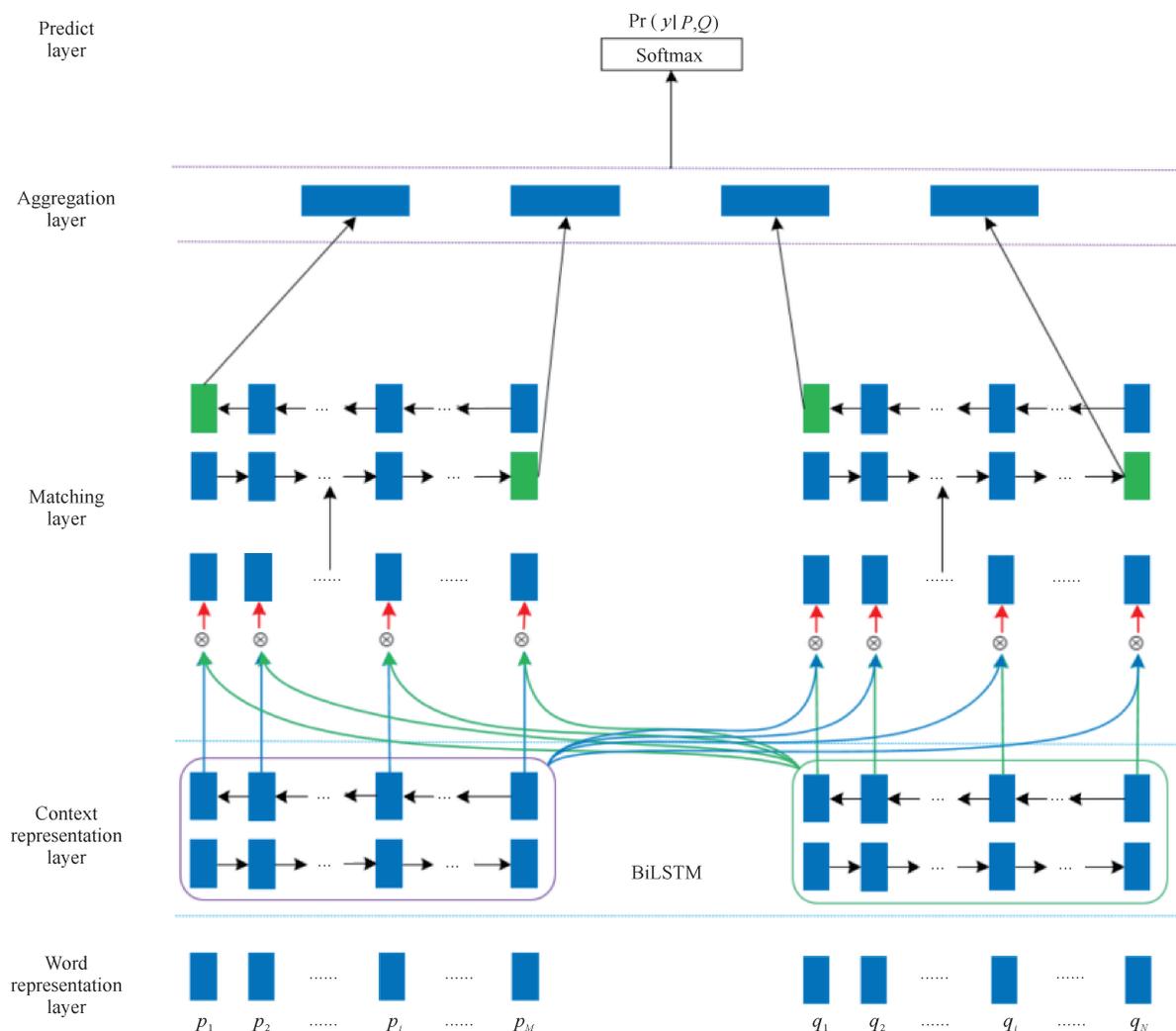


图2 双边多视角匹配模型框架图

Fig.2 Framework diagram of BiMPM model

在图1中,匹配层的每个框是一个单向多视角匹配(MPM)块,两个MPM块一起构成一个BiMPM模块。在实验中,所有视图上的模块都具有相同的参数。每个MPM模块将两个序列(一个起始序列和一个目标序列)作为输入,并将起始序列的每个上下文向量与目标序列的所有上下文向量进行匹配。图1中MPM框内的箭头表示匹配的方向,即起始→目标,问题序列和关系序列被分别视为起始序列和目标序列,反之也成立。

在匹配过程中,通过合成目标的所有上下文向量,为起始序列的每个上下文向量计算匹配向量。然后,该模型使用多角度余弦相似性函数从多个角度计算起始序列的上下文向量和匹配向量之间的相似性。使用比较两个向量的多角度余弦匹配函数 $f_m$ 计算匹配向量:

$$m = f_m(v_1, v_2; \mathbf{W}), \quad (13)$$

式中, $v_1$ 和 $v_2$ 是两个 $d$ 维向量; $\mathbf{W} \in R^{l \times d}$ 是一个形为 $l \times d$ 的可训练参数, $l$ 是视图数;返回值 $m$ 是一个 $l$ 维向量。每个元素 $m_k \in m$ 是第 $k$ 个角度的匹配值,由两个加权向量之间的余弦相似度计算得到,即

$$m_k = \text{cosine}(\mathbf{W}_k \odot v_1, \mathbf{W}_k \odot v_2), \quad (14)$$

式中, $\odot$ 是元素相乘, $\mathbf{W}_k$ 是 $\mathbf{W}$ 的第 $k$ 行,它控制第 $k$ 个视图,并为 $d$ 维空间的不同维度分配不同的权重。此处MPM模块使用了以下4种匹配策略。

①全匹配策略。该策略将起始序列的每个上下文向量与代表整个目标序列的最后一个上下文向量进行比较,匹配向量 $\vec{m}_i^{\text{full}}$ 和 $\overleftarrow{m}_i^{\text{full}}$ 按下式计算:

$$\begin{aligned} \vec{m}_i^{\text{full}} &= f_m(\vec{h}_i^p, \vec{h}_N^q; \mathbf{W}^1), \\ \overleftarrow{m}_i^{\text{full}} &= f_m(\overleftarrow{h}_i^q, \overleftarrow{h}_1^p; \mathbf{W}^2). \end{aligned} \quad (15)$$

②最大池匹配策略。该策略利用多视角余弦相似函数将起始序列的每个上下文向量与目标的每个上下文向量进行比较, 只保留每个维度的最大值, 匹配向量  $\vec{m}_i^{\max}$  和  $\vec{h}_i^{\max}$  的计算公式为

$$\begin{aligned}\vec{m}_i^{\max} &= \max_{j \in \{1, \dots, N\}} f_m(\vec{h}_i^p, \vec{h}_j^p; \mathbf{W}^3), \\ \vec{h}_i^{\max} &= \max_{j \in \{1, \dots, N\}} f_m(\vec{h}_i^p, \vec{h}_j^q; \mathbf{W}^4).\end{aligned}\quad (16)$$

③注意力匹配策略。该策略计算两个序列中所有上下文向量对之间的余弦相似度, 通过获取目标所有上下文向量的加权和来计算匹配向量, 其中注意力权重是最开始计算的余弦相似度。注意力匹配策略的匹配向量  $\vec{m}_i^{\text{att}}$  和  $\vec{h}_i^{\text{att}}$  按下式计算:

$$\begin{aligned}\vec{m}_i^{\text{att}} &= f_m(\vec{h}_i^p, \vec{h}_i^{\text{mean}}; \mathbf{W}^5), \\ \vec{h}_i^{\text{att}} &= f_m(\vec{h}_i^p, \vec{h}_i^{\text{mean}}; \mathbf{W}^6),\end{aligned}\quad (17)$$

$$\begin{aligned}\vec{h}_i^{\text{mean}} &= \frac{\sum_{j=1}^N \alpha_{i,j} \cdot \vec{h}_j^p}{\sum_{j=1}^N \alpha_{i,j}}, \\ \vec{h}_i^{\text{mean}} &= \frac{\sum_{j=1}^N \alpha_{i,j} \cdot \vec{h}_j^q}{\sum_{j=1}^N \alpha_{i,j}},\end{aligned}\quad (18)$$

$$\begin{aligned}\alpha_{i,j} &= \text{cosine}(\vec{h}_i^p, \vec{h}_j^q), j = 1, \dots, N, \\ \bar{\alpha}_{i,j} &= \text{cosine}(\vec{h}_i^p, \vec{h}_j^q), j = 1, \dots, N.\end{aligned}\quad (19)$$

④最大注意力匹配策略。不同匹配策略见图3。从图3可以看出, 最大注意力匹配策略与注意力匹配类似, 不同之处在于它不把所有上下文向量的加权和作为匹配向量, 而是从目标中选取具有最大余弦相似度的上下文向量。

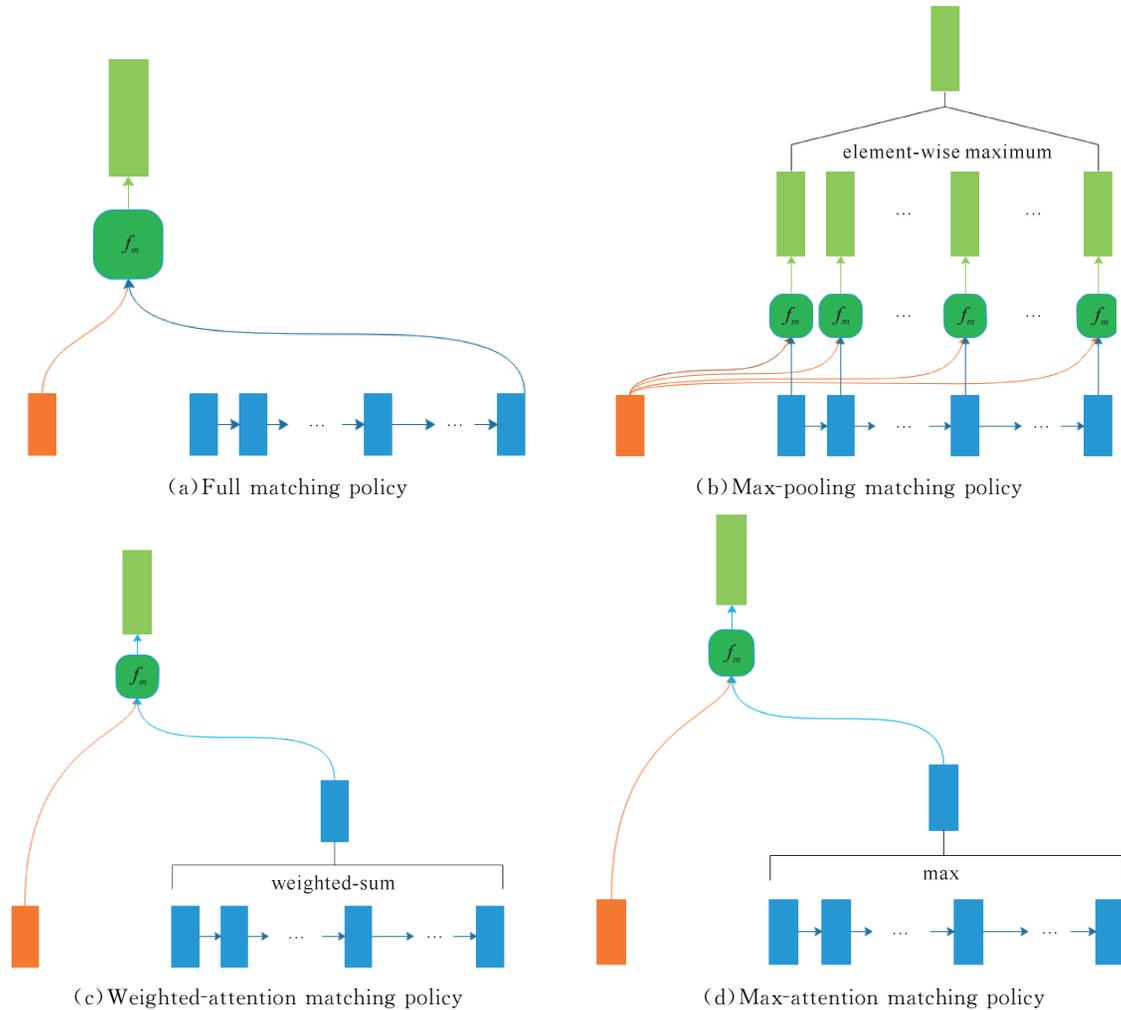


图3 不同匹配策略示意图

Fig. 3 Schematic diagram of different matching strategies

### 2.2.5 聚合模块

本模块分别对两个匹配向量序列应用另一个 BiLSTM,然后通过连接来自 BiLSTM 模型最后一个时间步的隐状态向量,构造一个定长的匹配向量作为一个视图的整体匹配表示。为了将输入对和实体对的不同视图匹配结果进行组合,设置了一个聚合层,该聚合层从不同视图提取匹配分数,并提取实体对的特征表示,然后构造一个用于关系预测的特征向量。在本文中,只使用了匹配模块从所有视图生成的不同匹配表示的串联,所有多个视图的组合表示通过多个感知层转化为最终预测。

## 3 验证实验

为验证本文提出的方法,在 SimpleQuestions (SQ)<sup>[28]</sup> 和 WebQSP(WQ)<sup>[12]</sup> 公开数据集上进行实验。两个数据集中的每个问题都标有头实体和关系信息。SQ 只有单关系问题,即每个问题对应一个  $\langle \text{head}, \text{relation}, \text{tail} \rangle$  三元组。相比而言,WQ 既有单关系问题,也有多关系问题。对于多关系问题,在连接头实体和尾实体的路径上存在多个关系,采用与 Yu 等<sup>[29]</sup> 相同的方法来创建正负实例,使用罗达等<sup>[30]</sup> 提出的结合 ELMO 语言模型的动态词向量以

表 1 各模型在 SimpleQuestions 和 WebQSP 公开数据集上关系检测任务的准确率

Table 1 The accuracy of each model in relation detection task on SimpleQuestions and WebQSP

序号 ID	模型 Model	关系输入 Relation input level	SimpleQuestions (%)	WebQSP (%)
1	BiCNN	char-3-gram	90.00	77.74
2	AMPCNN	word level	91.30	-
3	BiLSTM (w/words)	word level	91.20	79.32
4	BiLSTM (w/relation names)	relation level	88.90	78.96
5	HR-CNN	word level + relation level	92.90	79.08
6	HR-BiLSTM	word level + relation level	93.30	82.53
7	Multi-View Matching (word embedding + char embedding)	entity pair + abstract question + relation name + tail entity type	93.69	<b>86.19</b>
8	Multi-View Matching (word embedding)	entity pair + abstract question + relation name + tail entity type	<b>93.79</b>	85.41

Note: "-" indicates that the model performance has not been experimentally verified on this data set, the bold font represents the best performance of the compared models on this data set

从表 1 可以看出,本文的模型在 WQ (第 7 行) 和 SQ (第 8 行) 公开数据集上的关系检测任务准确率都得到了提升,分别比之前的模型高出 3.66 和 0.49 个百分点。正如表中准确率所示,WQ 由于加入了复杂问题,问题的整体难度要高于 SQ,有以下 3

及词性特征的实体链接方法。

本文实验分别使用包含约 210 万个实体的 Freebase 子集 FB2M<sup>[11]</sup> 和包含约 490 万个实体的 FB5M 作为知识库。训练集有 571 000 个实例,平均每个问题有 8 个候选关系。通过 WebQSP 数据集上的关系检测任务来评估关系检测模型,该训练集有 215 000 个实例,平均每个问题有 71 个候选关系。

### 3.1 实验设置

实验参数:关系候选词、实体候选词以及问题的最大输入长度在单词级别均设置为 30,关系候选词在关系级别的输入长度设置为 10,单词的词嵌入维度  $d$  为 300,词向量经过 BiLSTM 层的输出维度  $d$  为 300,maxpooling 层的窗口大小设置为 600,drop-out 均设置为 0.5,epoch 设置为 30,学习率设置为  $1e-4$ ,使用预训练词向量 Glove 进行随机初始化。

### 3.2 结果与分析

#### 3.2.1 关系检测

本文所提出的模型与主流基线模型在 SimpleQuestions 和 WebQSP 公开数据集上关系检测任务的准确率见表 1。表中“关系输入”一栏标明了这个模型所输入的关系粒度。

个原因可能导致这种趋势:首先,由于存在多跳关系,WQ 中每个问题的平均候选关系数更多。其次,WQ 的问题更多,更接近人类提出的真实问题。相比较而言,SQ 中的问题本质上是合成的,因为它们是由以知识库为单位的真实答案组成的。再次,WQ 需要更

复杂的知识库推理,因为从头实体到答案的路径通常由多跳实体节点组成。因此, SQ 的准确率在 90% 左右, 而 WQ 仍有较大提升空间。

### 3.2.2 消融实验

本文模型在 SimpleQuestions 和 WebQSP 公开

表 2 Multi Feature View Matching 模型进行消融实验结果

Table 2 Results of ablation experiment with Multi Feature View Matching model

序号 ID	模型 Model	w/o 字符嵌入 w/o character embedding	SimpleQuestions (%)	WebQSP (%)
1	( Q' ,Relation)	Y	93.15	75.26
2	( Q' ,Relation)	N	93.27	75.63
3	( Q' ,Relation + Type)	Y	93.32	76.41
4	( Q' ,Relation + Type)	N	93.46	75.95
5	( Q ,Relation)( Q ,Type)	Y	93.16	83.71
6	( Q' ,Relation)( Q' ,Type)	Y	93.42	84.74
7	( Q' ,Relation)( Q' ,Type)	N	93.55	84.86
8	[(Entity Pair)( Q' ,Relation)( Q' ,Type)],BiGRU	Y	93.68	86.05
9	[(Entity Pair)( Q' ,Relation)( Q' ,Type)],BiGRU	N	93.71	86.15
10	[(Entity Pair)( Q' ,Relation)( Q' ,Type)],BiLSTM	Y	93.73	<b>86.19</b>
11	[(Entity Pair)( Q' ,Relation)( Q' ,Type)],BiLSTM	N	<b>93.79</b>	85.91

Note: the bold font represents the best performance of the compared ablation experimental models on this data set

根据表 2 第 5-7 行,使用关系和尾部类型作为两个视图单独输入优于将它们作为单个输入(第 3-4 行)。比较第 5 行和第 6 行可以看出,使用替换实体提及的问题文本比使用原始问题文本更有助于模型关注问题中的重要信息。另外,在单词嵌入的基础上使用字符嵌入并没有显著影响。分析其原因,可能是由于 KB 关系和尾部类型的数量太少,尽管 Freebase 中有数千个这样的词汇,但与从大型文本语料库中获取的词汇相比,它们的数量仍然少得多。第 8-9 行是使用 BiGRU 替代 BiLSTM 作为视图编码器的模型性能,可以看到,BiLSTM 作为编码器仍然具有一定优势。从最后两行可以看到,本文所提出的模型在 WebQSP 和 SimpleQuestions 数据集上都达到了先进性能。虽然替换实体提及产生了改进,但会导致模型无法在关系检测过程中使用实体信息,所以考虑分别从实体对输入中提取特征用于最终预测,实验结果表明该操作可行。使用多视角匹配取代带有实体提及的问题,有助于模型关注所需的信息,也减少了关于实体的信息,对最终预测有一定帮助,实验结果也证实了这个结论。从表 2 也可以清楚地看到添加关系类型是有益的。

数据集上进行消融实验的关系检测准确率见表 2。分别对单视图、两视图、三视图作为输入时的模型表现进行评估,并在调整试图数量的同时对特征视图中的问题表示和关系粒度进行分别组合,以此来分析模型各部分对模型最终表现所起到的作用。

## 4 结论

关系检测是 KGQA 中的关键步骤,针对现有方法中未能较好兼顾全局语义和局部语义信息,以及复杂问题准确率不高的问题,本文提出了一种新的知识库关系检测模型,模型在构造的多个问题-知识库关系视图之间进行双边多视角匹配。实验结果表明,本文的模型在知识库问答关系检测任务上较以前的方法准确率有所提高,并且有望通过提高关系检测精度来使现有的 KGQA 系统获得更好表现。下一步,将考虑在问句表示中融入知识库信息以及联合实体选择与关系检测任务来进一步优化关系检测过程。

### 参考文献

- [1] QU Y, LIU J, KANG L, et al. Question answering over freebase via attentive RNN with similarity matrix based CNN [Z/OL]. (2018-04-10)[2022-09-27]. <https://arxiv.org/abs/1804.03317v2>.
- [2] 张仰森,王胜,魏文杰,等.融合语义信息与问题关键信息的多阶段注意力答案选取模型[J].计算机学报,2021,44(3):491-507.
- [3] ZHANG H, XU G, LIANG X, et al. An attention-based word-level interaction model: relation detection for

- knowledge base question answering [Z/OL]. (2018-04-10)[2022-09-27]. <https://arxiv.org/abs/1801.09893>.
- [4] 谢金峰,王羽,葛唯益,等. 基于多语义相似性的关系检测方法[J]. 西北工业大学学报,2021,39(6):1387-1394.
- [5] CHEN Y,LI H. DAM:transformer-based relation detection for question answering over knowledge base [J]. Knowledge-Based Systems,2020,201:106077.
- [6] LUO D,SU J,YU S. A BERT-based approach with relation-aware attention for knowledge base question answering [C]//2020 International Joint Conference on Neural Networks (IJCNN). [S. l.]: IEEE, 2020: 1-8. DOI:10.1109/IJCNN48605.2020.9207186.
- [7] 赵丹丹,黄德根,孟佳娜,等. 基于 BERT-GRU-ATT 模型的中文实体关系分类[J]. 计算机科学,2022,49(6):319-325.
- [8] 潘理虎,陈亭亭,闫慧敏,等. 基于滑动窗口注意力网络的关系分类模型[J]. 计算机技术发展,2022,32(6):21-27,33.
- [9] YAO X,VAN DURME B. Information extraction over structured data:question answering with freebase [C]// Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, Maryland, United States: Association for Computational Linguistics,2014:956-966.
- [10] BAST H,HAUSSMANN E. More accurate question answering on freebase [C]//Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. New York, United States: Association for Computing Machinery, 2015: 1431-1440.
- [11] 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, United States: Association for Computing Machinery, 2008. DOI: 10.1145/1376616.1376746.
- [12] BERANT J,CHOU A,FROSTIG R,et al. Semantic parsing on freebase from question-answer pairs [C]// Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. Seattle, Washington, United States: IEEE,2013:1533-1544.
- [13] TAN M,DOS SANTOS C,XIANG B,et al. Lstm-based deep learning models for non-factoid answer selection [Z/OL]. (2015-11-12)[2022-09-27]. <https://arxiv.org/abs/1511.04108>.
- [14] SEVERYN A,MOSCHITTI A. Learning to rank short text pairs with convolutional deep neural networks [C]//Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, United States: Association for Computing Machinery, 2015: 373-382. DOI: 10.1145/2766462.2767738.
- [15] YIN W,SCHÜTZE H,XIANG B,et al. ABCNN:attention-based convolutional neural network for modeling sentence pairs [J]. Transactions of the Association for Computational Linguistics,2016,4:259-272.
- [16] ROCKTÄSCHEL T,GREFENSTETTE E,HERMANN K M,et al. Reasoning about entailment with neural attention [Z/OL]. (2016-04-01)[2022-09-27]. <https://arxiv.org/pdf/1509.06664.pdf>.
- [17] WANG S,JIANG J. A compare-aggregate model for matching text sequences [Z/OL]. (2016-11-06)[2022-09-27]. <https://arxiv.org/abs/1611.01747>.
- [18] WANG Z,HAMZA W,FLORIAN R. Bilateral multi-perspective matching for natural language sentences [Z/OL]. (2017-07-14)[2022-09-27]. <https://arxiv.org/abs/1702.03814>.
- [19] GONG Y,LUO H,ZHANG J. Natural language inference over interaction space [Z/OL]. (2018-05-26)[2022-09-27]. <https://arxiv.org/abs/1709.04348>.
- [20] YU Y,HASAN K S,YU M,et al. Knowledge base relation detection via multi-view matching [Z/OL]. (2018-04-09)[2022-09-27]. <https://arxiv.org/abs/1803.00612>.
- [21] SHI T,LI X,LIU Z,et al. Research on Bi-LSTM machine reading comprehension algorithm based on attention mechanism [C]//Journal of Physics: Conference Series. Ireland, UK: IOP Publishing, 2022, 2258(1): 012007.
- [22] YANG Z,SUN Y,KUANG Q. Question answering model based on machine reading comprehension with knowledge enhancement and answer verification [J]. Concurrency and Computation: Practice and Experience,2022,34(12):e5828.
- [23] XU L,ZHANG X,ZONG B,et al. Zero-shot cross-lingual machine reading comprehension via inter-sentence dependency graph [Z/OL]. (2021-12-01)[2022-09-27]. <https://arxiv.org/abs/2112.00503>.
- [24] VYAWAHARE A,TANGSALI R,MANDKE A,et al. PICT@ DravidianLangTech-ACL2022: neural machine translation on dravidian languages [Z/OL]. (2022-04-19)[2022-09-27]. <https://arxiv.org/abs/2204.09098v1>.

- [25] LI Y, HE W. English-Chinese machine translation model based on bidirectional neural network with attention mechanism [J]. *Journal of Sensors*, 2022; 5199248.
- [26] VELPURI S Y, KARANWAL S, ANITA R. Neural machine translation using attention mechanism [C]// *Proceedings of International Conference on Recent Trends in Computing*. Singapore: Springer, 2022: 717-729.
- [27] WANG Z, HAMZA W, FLORIAN R. Bilateral multi-perspective matching for natural language sentences [Z/OL]. (2017-07-14) [2022-09-27]. <https://arxiv.org/abs/1702.03814>.
- [28] BORDES A, USUNIER N, CHOPRA S, et al. Large-scale simple question answering with memory networks [Z/OL]. (2016-06-05) [2022-09-27]. <https://arxiv.org/abs/1506.02075>.
- [29] YU M, YIN W P, HASAN K S, et al. Improved neural relation detection for knowledge base question answering [Z/OL]. (2017-04-20) [2022-09-27]. <https://arxiv.org/abs/1704.06194>.
- [30] 罗达, 苏锦钿, 李鹏飞. 基于多角度注意力机制的单一事实知识库问答方法[J]. *计算机科学*, 2019, 46(10): 215-221.

---

## Relation Detection Method Combining Question-Relation Attention and Feature View Matching

WANG Dong<sup>1</sup>, WANG Hairong<sup>1,2\*\*</sup>, CHE Miao<sup>1</sup>, SUN Chong<sup>1</sup>

(1. College of Computer Science and Engineering, North Minzu University, Yinchuan, Ningxia, 750021, China; 2. The Key Laboratory of Images & Graphics Intelligent Processing of State Ethnic Affairs Commission, North Minzu University, Yinchuan, Ningxia, 750021, China)

**Abstract:** As an advanced form of information retrieval, question answering system has become a research direction that has received much attention in the field of artificial intelligence and natural language processing. This article focuses on relation detection in Knowledge Graph Question Answering (KGQA) research. In view of the problems that the existing methods fail to take into account global semantic and local semantic information and the low accuracy of complex problems, a relation detection method combining question-relation attention and feature view matching is proposed. This method extracts multi-granularity features from questions and knowledge base, constructs the extracted features into feature pair view as the input of the relation detection model. Bilateral Multi-Perspective Matching (BiMPM) is used to compare and match within the view to obtain the relation prediction results. In order to verify the method proposed in this article, the accuracy of this method is improved by 3.42 and 0.45 percentage points respectively compared with the six mainstream baseline methods on the SimpleQuestions and WebQSP datasets.

**Key words:** relation detection; question answering system; attention mechanism; information retrieval; feature matching

责任编辑: 陆媛峰

---