

硕士学位论文

面向知识库的知识问答技术研究

**RESEARCH ON QUESTION ANSWERING BASED
ON KNOWLEDGE BASE**

施琦

哈尔滨工业大学

2019 年 6 月

国内图书分类号：TP391.1
国际图书分类号：004.9

学校代码：10213
密级：公开

工程硕士学位论文

面向知识库的知识问答技术研究

硕 士 研 究 生：施琦

导 师：张宇 教授

申 请 学 位：工程硕士

学 科：计算机技术

所 在 单 位：计算机科学与技术学院

答 辩 日 期：2019 年 6 月

授予学位单位：哈尔滨工业大学

Classified Index: TP391.1

U.D.C: 004.9

Dissertation for the Master's Degree in Engineering

**RESEARCH ON QUESTION ANSWERING BASED
ON KNOWLEDGE BASE**

Candidate:	Shi Qi
Supervisor:	Prof. Zhang Yu
Academic Degree Applied for:	Master of Engineering
Speciality:	Computer Technology
Affiliation:	School of Computer Science and Technology
Date of Defence:	June, 2019
Degree-Conferring-Institution:	Harbin Institute of Technology

摘 要

随着人工智能技术的发展,自动问答技术得到了广泛的关注。与传统的搜索引擎不同,问答系统的主要目标是向用户返回正确而简洁的答案,而非大量候选答案的集合。知识库是文本资源的基础,通过利用知识库,可以细致的分析文本资源中所包含的语义知识。在本课题面向知识库的知识问答技术研究中,主要包含以下三个方面的内容:基于单关系信息检索的知识库问答技术、基于多关系信息检索的知识库问答技术和基于语义解析的知识库问答技术。

针对基于单关系信息检索的知识库问答技术,本课题主要探究关系识别技术。先前的工作主要考虑如何建立问题与问句之间的匹配方法,忽略了问题中句法的重要性,所以本课题尝试使用依存句法树,提出了一个基于双向树结构的关系识别模型。实验结果显示本课题提出的方法相对之前的工作有着明显的提升。

针对基于多关系信息检索的知识库问答技术,本课题在基于单关系信息检索知识库问答技术的基础上,定义语义图,添加约束识别技术,构成了端到端的知识库问答系统。实验结果显示本课题构建的知识库问答系统效果显著,在最后,我们对其出现的错误进行了详尽的分析。

针对基于语义解析的知识库问答技术,本课题首先介绍了序列到动作模型、定义了语义图和动作序列,并基于序列到动作模型提出了基于拷贝机制的语义解析模型和执行导向的语义解析模型。实验结果显示两种方法与基于序列到序列语义解析模型的基线方法相比有着显著的提升。在最后,我们对本课题中提到的基于信息检索的模型和基于语义解析的模型进行了详细的对比,比较了二者之间的差异性,阐述了不同因素对两种模型的影响,并对两种模型的优缺点进行了总体的阐述。

关键词: 问答系统; 知识库; 关系识别; 约束识别; 语义解析

Abstract

With the development of artificial intelligence technology, the automatic question answering technology has received extensive attention. Unlike traditional search engines, the main goal of the question answering system is to return correct and concise answers to the user, rather than a collection of a large number of candidate answers. The knowledge base is the basis of text resources. By using the knowledge base, the semantic knowledge contained in text resources can be analyzed in detail. In the research of knowledge quiz technology for knowledge base in this topic, it mainly includes the following three aspects: knowledge base question answering technology based on single-relation information retrieval, knowledge base question answering technology based on multi-relation information retrieval and knowledge base based on semantic parsing.

For the knowledge base question answering technology based on single-relation information retrieval, this topic mainly explores the relationship recognition technology. The previous work mainly considered how to establish the matching method between the question and the question, and ignored the importance of the syntax in the question. Therefore, this paper attempts to use the dependency tree to propose a relationship recognition model based on the bidirectional tree structure. The experimental results show that the method proposed in this paper has a significant improvement over the previous work.

Aiming at the knowledge base question answering technology based on multi-relation information retrieval, this topic adds the constraint recognition technology based on the single-relation information retrieval knowledge base question answering technology, which constitutes an end-to-end knowledge base question answering system. The experimental results show that the knowledge base question answering system constructed by this topic has a significant effect. In the end, we have carried out a detailed analysis of the errors.

Aiming at the knowledge base question answering technology based on semantic parsing, this thesis first introduces the sequence-to-action model, defines the semantic graph and action sequence, and proposes a semantic analysis model based on the copy mechanism and a semantic-analytical model of execution-oriented based on the sequence-to-action model. The experimental results show that the two methods have a significant improvement compared to the baseline method based on the sequence-to-sequence semantic analysis model. In the end, we compare the information retrieval-based model and the semantic parsing-based model mentioned in this topic in detail, compare the differences between both, and explain the influence of different factors on the two models. The advantages and

disadvantages of the two models are generally explained.

Keywords: Question Answering, Knowledge Base, Relation Detection, Constraint Detection, Semantic Parsing

目 录

摘 要	I
ABSTRACT	II
第 1 章 绪 论	1
1.1 课题背景及研究意义	1
1.1.1 课题背景	1
1.1.2 研究意义	2
1.2 国内外研究现状	2
1.2.1 国外研究现状	2
1.2.2 国内研究现状	4
1.2.3 国内外研究现状简析	5
1.3 本文研究内容及章节安排	6
1.3.1 本文研究内容	6
1.3.2 本文章节安排	7
第 2 章 基于句法信息的关系识别技术研究	9
2.1 引言	9
2.2 基于信息检索的关系识别任务	9
2.2.1 任务定义	9
2.2.2 候选项生成方法	10
2.3 数据集介绍	10
2.4 基于残差网络的关系识别模型介绍	11
2.5 基于双向树结构的关系识别模型	13
2.5.1 关系表示模块	14
2.5.2 双向树结构问题表示模块	14
2.5.3 匹配层模块	16
2.5.4 输出层模块	17
2.5.5 训练方式	17
2.6 实验与分析	17
2.6.1 实验设置	18
2.6.2 实验结果与分析	18
2.7 本章小结	19
第 3 章 基于语义图的信息检索方法研究	21

3.1 引言	21
3.2 数据集介绍	23
3.3 基于关系识别与约束识别的知识库问答系统	23
3.3.1 候选项生成方法	23
3.3.2 语义图定义	24
3.3.3 基于双向树结构的关系识别模型	24
3.3.4 约束识别方法	25
3.4 实验与分析	25
3.4.1 实验设置	25
3.4.2 实验结果与分析	26
3.5 本章小结	28
第 4 章 基于序列到动作模型的语义解析方法研究	29
4.1 引言	29
4.2 数据集介绍	29
4.3 动作到序列模型简介	30
4.3.1 动作定义	30
4.3.2 数据预处理	30
4.4 基于序列到动作的语义解析模型	31
4.4.1 基于序列到动作的语义解析模型	31
4.4.2 基于拷贝机制的语义解析模型	32
4.4.3 执行导向的语义解析模型	34
4.5 实验与分析	36
4.5.1 实验设置	36
4.5.2 实验结果与分析	36
4.6 基于信息检索的模型与基于语义解析的模型比较	37
4.7 本章小结	38
结 论	39
参考文献	40
哈尔滨工业大学学位论文原创性声明和使用权限	45
致 谢	46

第 1 章 绪 论

1.1 课题背景及研究意义

1.1.1 课题背景

随着互联网的普及，信息资源越来越丰富，获取信息资源也变得越来越便捷。目前，通过搜索引擎，用户可以方便地获取各类信息。但是搜索引擎存在以下两点不足：一是查询到的结果数量庞大，用户为了获取其真正需要的信息，必须进一步阅读索引返回的信息并加以定位；二是搜索引擎获取信息只是停留在字符串匹配阶段，无法涉及深层次的语义理解。搜索引擎以关键字匹配作为技术基础，其只关注语句的字面含义，没有涉及语义，故检索效果一般。因此近年来自动问答系统得到了广泛的关注。

不同于搜索引擎，问答系统可以返回给用户准确而简洁的答案，而非大量的候选列表。而语义资源的多样性同样使问答系统更加有挑战性。知识库作为质量最高、用语最规范的语义资源之一，受到了很高的关注。在大规模知识库中回答用户提出的问题的技术，也称为知识库 问答技术（Knowledge Base Question Answering, KBQA），成为了研究热点之一。

由于常见问题（Frequently Asked Questions, FAQ）以及社区问答系统（Community Question Answering, CQA）数据存在质量较差，用语不规范，有很多省略词和口语表达方式。近些年，研究人员逐步将研究热点转向知识库和知识图谱，目的是将文本内容转化成以句子中的实体为核心的图结构，图中的边表示实体与实体间的语义关系。知识库是文本资源的基础，通过构建知识库，可以细致分析文本资源中所包含的语义知识。目前，互联网中已经有了一些可以获取的大规模知识库，这些知识库多是以结构为“实体-关系-值”的三元组为基本单元所组成的图结构。由于知识库中的知识大多基于三元组这样的结构化数据，问答系统的首要任务就是分析用户输入的问题，并根据问题的语义信息在知识库中查找、匹配出最佳答案。知识库数据的出现，实际上是对用户可能提问的问题进行了细化，拆分为实体与关系，再通过词向量投射技术以及深度学习技术，完美的解决了 FAQ 数据与 CQA 数据中人机交互性能不佳的问题。

1.1.2 研究意义

随着科学技术的飞速发展，人们对知识获取的准确与便捷性越来越高，传统意义上的搜索引擎已不能满足人们日益增长的信息获取需求。故自动问答系统出现在了人们的视野。知识库问答系统作为其中一种最重要的问答系统之一，受到了广泛的关注。本课题将致力于知识库问答系统中的若干关键技术研究。

目前，得益于深度学习技术的发展，知识库问答技术成为了一个研究热点，相关工作已相对形成一定知识体系，但仍有一些技术问题需要解决。知识库中定义的实体与关系数量是有限的，而用户提出的问题理论上有限种可能。如何将无限种用户问题与有限个实体关系进行匹配，是知识库问答技术的主要挑战之一。本课题将分别从关系链的长度（单关系与多关系）以及方法的类别（基于信息检索的方法与基于语义解析的方法）展开，来解决上述问题。

1.2 国内外研究现状

1.2.1 国外研究现状

自 Freebase^[1]、Wikidata^[2]、Dbpedia^[3]等大规模知识库出现以来，知识库问答技术开始受到广泛关注。目前主流的方法分为基于信息检索的方法与基于语义解析的方法。近年来，随着深度学习技术的迅猛发展，知识库问答技术进入了一个新的发展时期，神经网络被证明对上述两种方法有着很大的提升。本课题将着重阐述结合神经网络的信息检索方法和语义解析方法。

（1）基于信息检索的知识库问答技术

基于信息检索的知识库问答技术主要目标是识别主题实体，构建围绕主题实体的 N 跳子图，并对这些子图进行排序来得到最佳答案。基于信息检索的知识库问答技术主要包括以下几点：实体链接、关系识别和约束识别。

实体链接 实体链接是从问句中识别出实体词并链接到知识库对应实体的过程。识别问句中的实体词并将他们链接到知识库中可以帮理解问句。大部分工作使用知识库来完成，其中包括 Freebase^[1]、Yago^[4]、Dbpedia^[3]等等。Mihalcea 等人^[5]在 2007 年提出 Wikify，将文本连接到维基百科词条的一种方法，其提取了所有维基百科概念的 n-gram 特征作为候选，然后实施基于知识和数据驱动方法的投票方案，以消除候选项的歧义。Cucerzan^[6]在 2007 年使用四种资源生成候选项，即实体页面，重定向页面，消歧页面和列表页面。然后通过计算上下文信息与文档之间的相似度以及维基百科词条中的类别标签来

消除候选项中存在的歧义。Milne 和 Witten^[7]在 2008 年通过收集文本中所有的 n-gram 特征并保留那些概率超过低阈值的候选项作为候选项,然后定义维基百科的超链接机构的共同性和相关性,以消除候选项的歧义。

近年来,关于链接查询中实体的工作已经被广泛研究。Ferrigina 和 Scaiella^[8]于 2010 年提出的 TagME 是关于查询中实体链接的早期工作。它通过搜索维基百科页面标题,锚点和重定向来生成候选。然后根据基于 Milne 和 Witten^[7]的工作启发的相关性度量的投票方案,消除歧义利用维基百科图的结构,并使用 PageRank^[9]对候选实体进行排名。此外,Meij^[10]于 2012 年提出将推文链接到维基百科的两步法。他们首先提取每个 n-gram 的候选概念,然后用监督学习算法对相关概念进行分类。

与围绕查询范围的排序实体进行的工作不同,实体识别和消歧(Carmel 等, 2014^[11])将查询中的实体链接视为查找多个查询解释的问题。获得短文本的 SMAPH 系统(Cornolti 等, 2014^[12])分三个阶段进行:提取,候选实体生成和修剪。首先,他们获取由商业搜索引擎返回的片段。接下来,通过查看搜索片段的粗体部分来解析片段以识别候选实体。最后,他们使用一系列特征来实现二元分类器,比如注释过程的一致性和健壮性,以及片段的排序和组合。他们进一步将 SMAPH-1 扩展到 SMAPH-2(Cornolti 等, 2016^[13])。他们使用注释器 WAT 来注释搜索结果的片段来生成候选项,并联合额外的链接返回步骤以及排名阶段的修剪步骤,从而获得 ERD14 上的最新结果数据集和他们发布的数据集 GERDAQ。还有一项关于 SMAPH 的工作使用了查询日志和锚文本信息(Blanco 等, 2015^[14]),它给出了实体的排名列表,并且通过典型的排名指标进行评估。

关系识别 关系识别即理解问句并识别问句中的关系的过程。现阶段多用表示学习的方法。表示学习,简言之是指通过统计学习自动获取文本的语义表示。在知识库问答领域,常用的方法是利用深度学习技术,将自然语言问句和知识库中的资源映射到同一向量空间中,然后计算余弦相似度找出问句最有可能对应的答案三元组。使用大量问句-答案三元组对来训练。

关于 KBQA 的关系识别也开始于基于特征的方法(Yao 和 Van Durme, 2014^[15]; Bast and Haussmann, 2015^[16]),以适应深度网络的应用(Yih 等, 2015^[17]; Xu 等, 2016^[18]; Dai 等, 2016^[19])和注意力模型(Yin 等, 2016^[20]; Golub 和 He, 2016^[21])。许多上述关系检测研究可以支持大型关系词汇表和开放关系集(如 ParaLex(Fader 等, 2013^[22])),以适应开放领域问题解答的目标。

不同的数据集对上述开放领域的能力有不同的要求。由于知识库的不同,

在知识库上定义的关系种类有着很大的差别,所以识别大型关系集和不可见关系的能力变得更加必要。以上问题有两个主要的解决方案:(1)使用预先训练的关系向量(例如来自 TransE (Bordes 等, 2013^[23])),如 (Dai 等, 2016^[19]); (2)将关系名称分解为序列,并将关系识别分解为序列匹配和排序任务。这种分解方式是有效的,因为关系名称通常包含有意义的单词序列。例如, Yin 等人^[20]于 2016 年将关系拆分为单字序列进行单关系检测。Liang 等人^[24]于 2016 年在 WebQSP 上也取得了良好的表现,在端到端的神经网络模型中使用了词级别关系表示。Yih 等人^[17]于 2015 年使用特征三元组作为问题和相关方面的输入。Golub 和 He^[21]于 2016 年提出了一个生成框架,该框架预测了一个字符级的序列到序列模型的关系。

约束识别 一个只由主题实体和关系或关系链组成的语义图是最简单的。理论上正确答案会包含在由该语义图检索得到的答案集合中,但除正确答案外,仍有一些用户所不关心的冗余的答案,约束识别就是要根据问题对检索到的答案进行约束,来缩小检索到的答案集合。Yih 等人^[17]提出了一种检索式的约束方法来限制答案节点以及复合值类型节点(Compound Value Type, CVT)。

(2) 基于语义解析的知识库问答技术

基于语义解析的知识库问答技术的目标是将自然语言问题转换为一个结构化的逻辑表达式(如 lambda calculus、lambda DCS、SPARQL 等),从而可以直接在知识库中执行来获得答案。Berant 等人^[25]提出了一种通过不依赖逻辑表达式,而使用问题-答案对来训练的方式来进行训练的方法;Reddy 等人^[26]提出了一种通过语义图来表示自然语言的方法,其使用组合范畴语法将句子转换为语义图,然后将其以弱监督的形式引导到 Freebase; Sorokin 等人^[27]使用门机制的图神经网络来建模语义图;Talmor 等人^[28]基于问句分解技术来将复杂问题分解为若干简单问题,并通过这些简单问题预测最终答案的方法。

1.2.2 国内研究现状

在国内,与国外的研究工作相似,同样分为基于语义解析的知识库问答技术和基于信息检索的知识库问答技术。

(1) 基于信息检索的知识库问答技术

Dong 等人^[29]使用 multi-column 的 CNN 来从不同的答案类型中去理解问句; Xu 等人^[30]提出一种混合问答系统,即从结构化的知识库和自由文本中共同获取答案; Yu 等人^[31]提出使用结构化的残差网络来进行关系识别并利用关系识别加强 KBQA 系统的性能; Hao 等人^[32]提出了一种可以根据不同的答案

类型动态表示问题和候选答案的端到端的 KBQA 系统。刘康等人^[33]对现有的基于表示学习的知识库问答技术进行了综述。

(2) 基于语义解析的知识库问答技术

Dong 等人^[34]提出由粗粒度到细粒度的语义解析系统，先生成一系列的模版，然后填充里面缺失的内容；Hu 等人^[35]提出了将复杂自然语言问题转换成语义问句图，用来匹配知识图谱从而找到答案；类似的，Luo 等人^[36]提出将复杂的语义图转换成一个向量来计算与自然语言问句的语义相似度，来找到复杂问句对应的语义图，从而得到答案；Guo 等人^[37]提出了一种在多轮知识库上、弱监督的序列到动作的模型来将问题转化为对应的逻辑表达式，从而得到答案。

1.2.3 国内外研究现状简析

基于知识库的问答系统完美的解决了 FAQ 以及 CQA 中存在的质量较差，用语不规范，有很多省略词和口语表达方式的问题，将自有文本转化为了结构化的知识数据，从而可以细致分析问题中包含哪些语义信息，可以准确快速的回答问题，提升人机交互能力，因此研究知识库问答系统对于获取精准的信息有着重要的应用价值。

现阶段基于知识库的问答技术主要可分为两种方法：基于语义解析（Semantic Parsing）的方法和基于信息检索（Information Retrieval）的方法。基于语义解析的方法目标是将自然语言问句转化成形式化的表示，即关注于构建一个语义解析器将问题转化为一种结构化的逻辑表达式，并可以直接从知识库中获取答案；基于信息检索的方法目标是检索问题中的核心实体，构建围绕该实体的知识子图然后对所有候选答案进行排序从而得到最佳答案。一个问句可以划分成针对三元组中的值进行提问的结构，故在该方法中，理解问句的重点就是识别出问句中的实体以及关系。近年来，由于深度学习技术的发展，两种方法都融合了深度学习技术，并有了很大的提升。

基于信息检索的知识库问答技术主要由三个子问题组成，一是实体链接，即找到问句中的实体和知识库中实体的映射；二是关系识别，即找到问句中自然语言描述和知识库中语义关系的对应。找到实体和关系后，就可以在知识库的三元组唯一确定所需要的问句的答案；三是约束识别，即对前两步得到的简单的语义图根据用户提出的问题进行进一步的约束来缩小得到的答案范围。

实体链接的完整过程包括实体识别、实体消歧。实体识别是指从问句中识别出可能链接到知识库中特定条目的文本片段。由于自然语言中普遍存在一词多义现象，检索到的文本片段并不能唯一确定其所指向的实体。因此，需要借

助实体消歧技术，利用问句中的上下文，确定其所指向的实体。

第一步是识别出文档中的指示实体的文本片段。传统的实体识别所采取的主要技术方法包括：基于规则和词典的方法、基于统计的方法、二者混合的方法等。随着知识库的发展，也出现了利用维基百科进行实体识别的方法。实体链接中采用的实体识别方法主要借鉴基于规则和词典的方法，通过从维基百科中检索实体别名，丰富词典信息，从而提高实体识别的召回率。

实体识别的难点在于对于问句中多个实体，如何识别其中心实体。这里传统的基于统计的方法、基于依存句法的分析很难达到这样的效果，故须采用基于深度学习的方法，这里将其规范为核心实体识别任务。识别问句的核心实体是非常必要的，因为问句的实体是问句的基础，若识别到了错误的实体，将很难得到正确答案。

实体消歧是指根据得到的文本片段的上下文信息，分析其可能的多种含义，判断其在当前上下文中的词义的过程，在问句中特指实体文本片段周围的若干词。即给定实体文本片段、其所在的上下文、候选实体，判断其在当前上下文中所指向的实体的过程。传统的词义消歧采用的方法主要包括基于知识的方法和驱动数据的方法。基于知识的方法依赖从词典中获取的词义信息。而基于数据的方法则构建基于词义标注语料库的机器学习分类器进行词义消歧。

关系识别即理解问句并识别问句中的关系的过程，由于问题的多样性以及知识库中定义的关系的有限性，关系识别任务是非常有挑战性的。关系识别多采用对问题以及候选关系建模并排序的方式，将问题模版与候选关系映射到同一向量空间，并对其进行建模，得到一个语义相似度，值最高的候选答案即为最佳答案。

语义解析过程由于其解决复杂问题的优秀性能，成为了近年来最常用的方法，语义解析模型多采用生成式模型。由于逻辑表达式中字段和实体关系的多样性以及训练数据的稀疏性，语义解析过程存在训练困难的问题。解决语义解析模型中存在的问题是非常困难的，也是非常有意义的。

1.3 本文研究内容及章节安排

1.3.1 本文研究内容

本课题主要从知识库问答技术的关系类型和方法进行分类，即基于单关系信息检索的知识库问答技术、基于多关系信息检索的知识库问答技术、基于语义解析的知识库问答技术。首先对基于句法信息的关系识别技术最后展开研究，

然后对基于语义图的信息检索方法展开研究,最后介绍了基于序列到动作模型的语义解析方法展开研究。由于深度学习以及神经网络的广泛使用以及其良好的特性,本课题所有章节的主要方法围绕基于神经网络的方法展开。

本课题采用两个英文标准数据库进行实验,即 SimpleQuestions^[38]和 WebQSP^[39]。SimpleQuestions 用来实验基于单关系信息检索的知识库问答技术,而 WebQSP 用来实验基于多关系信息检索的知识库问答技术和基于语义解析的知识库问答技术,并对两种方法得出的结果进行对比。

第一部分研究内容是基于信息检索的知识库问答技术,由于大部分问题都是两跳之内的问题,故我们可以很容易的得到主题实体两跳周围的候选关系组成候选答案,从而进行检索排序过程。这样就完成了基本的基于信息检索的知识库问答过程。这是一种非常常规、自然的处理方法。本课题中将基于信息检索的知识库问答技术分为关系识别和约束识别两部分来解决。

第二部分研究内容是基于语义解析的知识库问答技术,基于语义解析的知识库问答技术使用序列到动作模型来生成本课题自定义的动作序列,从而转化为语义图、SPARQL 语言,然后在知识库中执行得到最终答案。训练数据使用广度优先搜索从 SPARQL 语言转化为动作序列。本课题在基线系统的基础上提出两种方法来提升语义解析模型的性能,从而实现将自然语言问题映射为动作序列,从而转化为逻辑表达式的过程。最后对比了两种方式的异同点、优缺点、应用场景以及不同因素对其性能的影响。

1.3.2 本文章节安排

本文主要从基于单关系信息检索的知识库问答技术、基于多关系信息检索的知识库问答技术、基于语义解析的知识库问答技术三个方面阐述我们在知识库问答技术上的研究工作。本文的内容章节安排如下:

第一章为绪论,从知识库问答技术的背景和研究目的出发,主要介绍了本文的研究意义,国内外知识库问答技术的研究现状以及现状分析,并简要总结了本文的主要的研究内容,最后介绍了章节安排。

第二章介绍了基于句法信息的关系识别技术。首先介绍了本章中采用的数据集,然后介绍了一种比较经典的基线模型基于残差网络的关系识别模型,之后介绍了在此基础上提出的基于双向树结构的关系识别模型。最后实验部分在数据集上采用准确率进行评价,对比了本课题使用的方法和基线方法的性能,证明了本章基于句法信息的关系识别技术的有效性。

第三章介绍了基于语义图的信息检索方法。首先介绍了本章中采用的数据

集，然后介绍了基于关系识别与约束识别的知识库问答系统，包括候选项生成方法、语义图等内容。最后实验部分在数据集上采用准确率进行评价，对比了本课题使用的方法和基线方法的性能，证明了本章基于语义图的信息检索方法的有效性，并对模型预测中存在的错误进行了细致的分析。

第四章介绍了基于序列到动作模型的语义解析方法。首先介绍了本章中采用的数据集，然后介绍了序列到动作模型，并定义了语义图和动作序列。之后介绍了本课题使用的基线模型基于序列到序列的语义解析模型，然后改进了该模型，提出了基于拷贝机制的语义解析模型和执行导向的语义解析模型。实验部分同样采用准确率进行评价，对比了两种方法的性能。最后，对基于多关系信息检索的知识库问答模型和基于语义解析的知识库问答模型进行了对比。比较两种方法之间的差异，每种方法的优势，以及未来的研究工作。

第2章 基于句法信息的关系识别技术研究

2.1 引言

在基于信息检索的知识库问答技术中，由于问题的复杂程度不同，回答问题所需要的三元组的个数也会有所不同。根据回答问题所需要的三元组个数，问题可分为单关系问题和多关系问题。单关系问题即通过唯一的一个三元组来指向答案的问题，例如，对于问题“**What is Justin Bieber’s birthday?**”，回答它需要在知识库中找到唯一的一个三元组“**Justin Bieber - date_of_birth - 19940301**”来得到答案，该问题即为单关系问题；多关系问题即需要通过两个或两个以上三元组共同指向答案的问题，例如，对于问题“**What is Justin Bieber’s father’s birthday?**”，回答该问题需要找到三元组“**Justin Bieber - father - Jeremy Bieber**”和“**Jeremy Bieber - date_of_birth - 19750604**”来得到答案，该问题即为多关系问题。按照问题类型，基于信息检索的知识库问答技术可分为基于单关系信息检索的知识库问答技术和基于多关系信息检索的知识库问答技术。本章将介绍基于单关系信息检索的知识库问答技术。首先我们先介绍基于信息检索的关系识别任务，包括其任务定义以及候选项生成方法，然后介绍本章使用的数据集，之后介绍该任务上的一个比较经典且效果优秀的基线方法，然后介绍我们在这章中提出的模型，最后介绍实验部分，分析模型各个模块的结果以及其作用。

本章我们主要介绍基于信息检索知识库问答技术中的关系识别任务。现存的方法大部分仅考虑了问题的序列信息，却没有考虑问题中所包含的句法信息。考虑到句法特征的重要性，我们提出了一种基于依存句法树的双向树结构的关系识别模型，用于关系识别任务。

2.2 基于信息检索的关系识别任务

2.2.1 任务定义

关系识别，简言之，即预测用户对于一个特定实体想要了解的内容。我们可以将关系识别任务看做定义为一个编码-匹配的过程。给定一个自然语言问题 q ，其主题实体 e 和从知识库中检索得到的候选关系集合 R ，关系识别任务旨在识别问题中的关系 r ，即找到可以将问题 q 映射到正确答案的关系 r 。关系识别任务可以看做一个对每个候选项进行打分，然后排序的任务。对在候选集合 R 中的每个候选关系 r ，模型将会预测出一个 q 和 r 之间的语义相似度 $s(q,r)$ ，故

通过排序我们可以得到最佳关系 r_{best} ，如公式(2-1)所示：

$$r_{best} = \operatorname{argmax} s(q, r) \quad (2-1)$$

2.2.2 候选项生成方法

在通过实体链接得到问题中的主题实体后，我们需要通过主题实体在知识库中检索得到若干关系，用于进行关系识别。例如，对于问题“*What city was alex golfis born in?*”，在识别到主题实体为“alex golfis”并链接到知识库中获得其实体 id 后，我们需要根据该实体 id 在知识库中检索满足其实体为该实体 id 的所有三元组，取其关系作为候选项，如“*place_of_birth, date_of_birth*”等等。

2.3 数据集介绍

在本章中，我们使用 SimpleQuestions 作为数据集。SimpleQuestions 由 Facebook 于 2015 年发布，是一个针对于单关系问题的数据集。SimpleQuestions 数据集使用 Freebase 的子集 FB2M 作为知识库，其包含 2150604 个实体，6701 个关系和 14180937 个三元组，数据集由自然语言问题和对应的可以指向正确答案的唯一三元组组成。数据集与知识库具体统计信息如表 2-1 所示。

表 2-1 SimpleQuestions 语料规模

数据集	训练集	验证集	测试集
样本数量	75910	10845	21687

表 2-2 SimpleQuestions 数据样例

问题	三元组
What city was alex golfis born in?	/m/0wzc58l /people/person/place_of_birth /m/0n2z
what film is by the writer phil hay?	/m/0jtw9c /film/writer/film /m/05szq8z
Who created the typeface chicago?	/m/03yctn /base/typefaces/typeface/typeface_creator /m/02f2m7

表 2-2 显示了 SimpleQuestions 数据集中随机选取的几个例子，其中“/m/0wzc58l、/m/0n2z”等代表知识库中定义的实体编号，“/people/person/place_of_birth、/film/writer/film”代表知识库中定义的关系。在本章中，与 Yih 等人^[17]、Yu 等人^[31]的工作相似，我们使用 Yang 和 Chang 等人

[40]的实体链接结果来进行后续的工作。且我们使用 Yu 等人[31]发布的针对关系识别任务的数据集进行实验。

2.4 基于残差网络的关系识别模型介绍

基于残差网络[41]的关系识别模型由 Yu 等人[31]于 2017 年提出，该模型结构简单，在模型参数以及训练速度上都很有优势，且结果表明，该模型达到了当时的最好结果，是很经典的一个简洁有效的模型。下面将详细介绍一下基于残差网络的关系识别模型。图 2-1 展示了该模型的结构图。在关系识别步骤之前，我们使用 Yang 和 Chang 等人[40]的实体链接结果来将问句中的实体片段和知识库中定义的关系进行链接。为了简化模型，在输入关系识别模型之前，我们将问题中的实体进行屏蔽。例如，对于表 2-2 中的第一个例子，我们用“<e>”来代替已识别到的实体，故输入为“What city was <e> born in?”，而模型的期望输出为可以指向正确答案的关系“/people/person/place_of_birth”。我们将这个模型作为本章的基线模型。

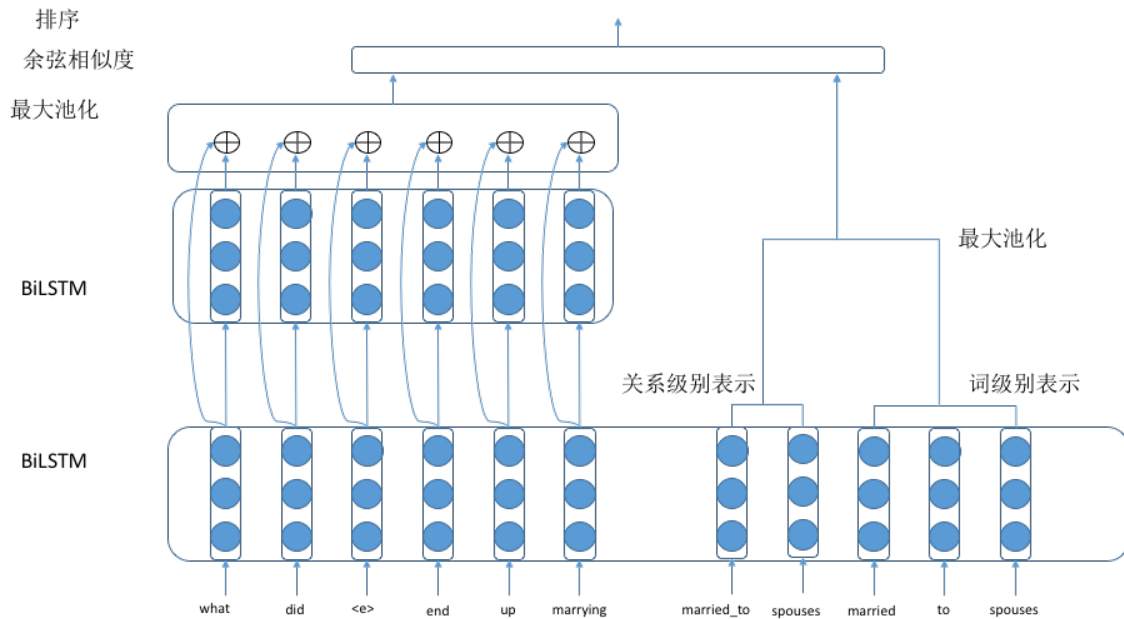


图 2-1 基于残差网络的关系识别模型

基于残差网络的关系识别模型分为输入层、问题表示层、关系表示层、匹配层以及输出层。为了简洁的展示该模型，图中省略了输入层。

(1) 符号约定 为了便于表述，我们使用如下的符号约定：模型的输入为问题-候选关系对， $q = (q_1, q_2, \dots, q_i, \dots, q_n)$ 表示用户输入的问题， q_i 表示问题中的第 i 个词， n 为问题的长度； r 表示某一候选关系，其中 $r \in R$ ，候选关系又分

为关系级别的候选关系和词级别的候选关系，分别表示为 r^{rel} 和 r^{word} ，关系级别即将一个关系词视为一个单独的词，如“place_of_birth”，词级别即将关系词拆分为若干个单词，如“place, of, birth”，我们将二者表示为： $r^{rel} = (r_1^{rel}, r_2^{rel}, \dots, r_j^{rel} \dots r_{l_r}^{rel})$ 和 $r^{word} = (r_1^{word}, r_2^{word}, \dots, r_j^{word} \dots r_{l_w}^{word})$ ，其中， r_j^{rel} 和 r_j^{word} 分别第 j 个关系级别的关系和第 j 个词级别的关系， l_r 和 l_w 分别表示关系级别关系的长度和词级别关系的长度。自然地，训练数据的表示形式为 $\{q, r^{rel}, r^{word}\}$ 的标注对。

(2) **输入模块** 对于一组问题-候选关系对的输入，我们使用预训练词向量 Glove^[43]将其表示成向量形式，问题、关系级别的关系以及词级别的关系分别表示为： e^q, e^{rel}, e^{word} 。

(3) **关系表示模块** 在关系表示模块中，分别使用循环神经网络表示关系级别的关系和词级别的关系，具体方法如公式(2-2)和公式(2-3)所示：

$$r_t^{rel} = \text{BiLSTM}(r_{t-1}^{rel}, e_t^{rel}) \quad (2-2)$$

$$r_t^{word} = \text{BiLSTM}(r_{t-1}^{word}, e_t^{word}) \quad (2-3)$$

然后，我们在两种表示上应用最大池化方法来得到最终的关系表示 r^{final} ，如公式(2-4)所示：

$$r^{final} = \text{MaxPooling}([r^{rel}, r^{word}]) \quad (2-4)$$

这种表示方法的主要优点为：关系级别的表示更关注于全局信息，但会有数据稀疏的问题；而词级别的表示更关注于局部信息，但会缺失原始关系词中的一些全局信息。故两者结合可以结合二者的优点。

(4) **问题表示模块** 在问题表示模块中，使用残差网络的方法来对问题进行建模。与常规的多层 LSTM 使用最后一层来进行预测不同，该模型中，以两层 LSTM 为例，作者希望两层问题表示可以相互补充，并且两者都应该与关系表示进行比较，这种方式称为层次化的匹配。这在该任务中很重要，因为由于句法的多样性，每个关系词可以对应不同长度的短语。

所以可以通过分别计算问题和关系每一层之间的相似性并且在两个分数之间做加权和来执行层次化匹配。然而，这种方式并未带来性能上的提高，因为这种方法受训练困难的影响，该模型的收敛训练损失远高于单层基线模型。这主要是因为深度 BiLSTM 不保证两层的问题隐藏层表示是可比较的，训练通常属于局部最优，其中一层具有良好的匹配分数，而另一层的权重总是接近 0，而且层数更高的模型本身训练更加困难。

所以，为了解决以上问题，该模型通过在两个 BiLSTM 层之间添加连接线来使用残差网络结构。具体操作如公式(2-5)~公式(2.7)所示：

$$q_t^{(1)} = \text{BiLSTM}(q_{t-1}^{(1)}, e_t^Q) \quad (2-5)$$

$$q_t^{(2)} = \text{BiLSTM}(q_{t-1}^{(2)}, q_t^{(1)}) \quad (2-6)$$

$$q_t^{final} = q_t^{(1)} + q_t^{(2)} \quad (2-7)$$

最后，我们得到问题的最终表示 $q^{final} = \{q_t^{final}\}_{t=1}^N$ 。

(5) **输出模块** 在输出模块中，如公式(2-8)所示，我们通过余弦相似度来作为问题和关系的语义相似度 $s(r, q)$ ：

$$s(r, q) = \cos(r^{final}, q^{final}) \quad (2-8)$$

(6) **训练方式** 该模型使用 Hinge Loss 来进行训练，通过一个排序损失来最大化正确关系和候选关系集合中的其他关系之间的界限，具体做法如公式(2-9)所示：

$$loss = \max\{0, \gamma - s(r^+, q) + s(r^-, q)\} \quad (2-9)$$

其中 γ 为一个介于 0 到 1 之间的常数， r^+ 指正确关系， r^- 指候选关系集合 R 中的其他关系。

2.5 基于双向树结构的关系识别模型

考虑到在现有方法没有考虑句法信息，所以我们提出了一种基于双向树结构的关系识别模型来将问题的句法信息融合进模型中。模型共分为关系表示模块，双向树结构问题表示模块，匹配模块，以及输出模块，模型结构图如图 2-2 所示。下面将对以上几个模块进行详细阐述。

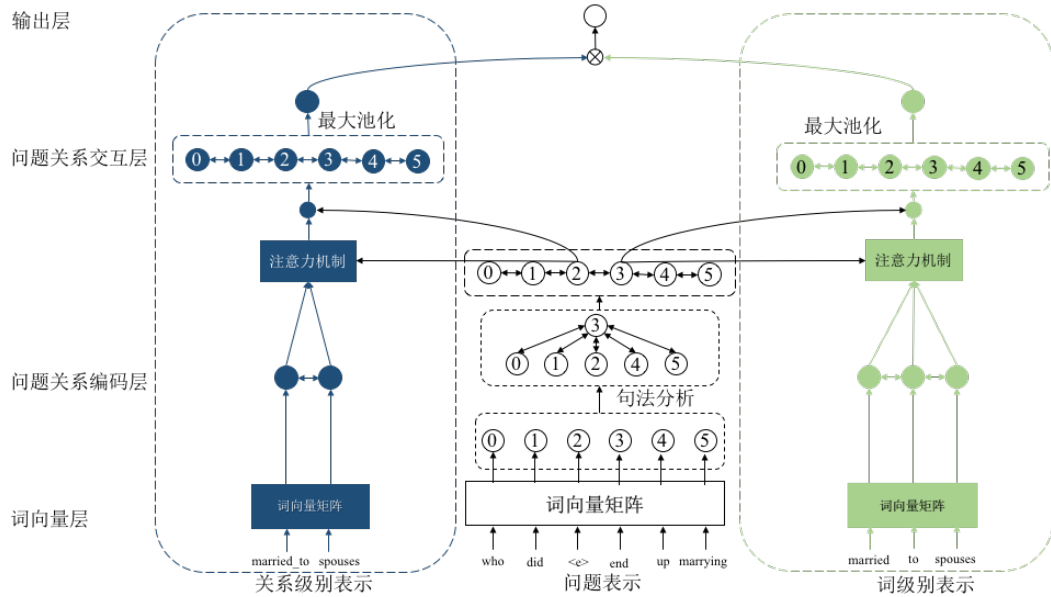


图 2-2 基于双向树结构的关系识别模型

2.5.1 关系表示模块

在关系表示模块中，我们同样分别采用关系级别的关系表示方法和词级别的关系表示方法来结合二者的优点，如公式(2-10)~公式(2-11)所示：

$$r_t^{rel} = \text{BiLSTM}(r_{t-1}^{rel}, e_t^{rel}) \quad (2-10)$$

$$r_t^{word} = \text{BiLSTM}(r_{t-1}^{word}, e_t^{word}) \quad (2-11)$$

通过这种方式，我们获得了关系级别的关系表示 r^{rel} 和词级别的关系表示 r^{word} 。

2.5.2 双向树结构问题表示模块

问题表示模块的目标是获得问题中每个词的表示。为了利用问题中的句法信息，我们提出使用树结构的循环神经网络，问题表示模块结构图如图 2-3 所示。

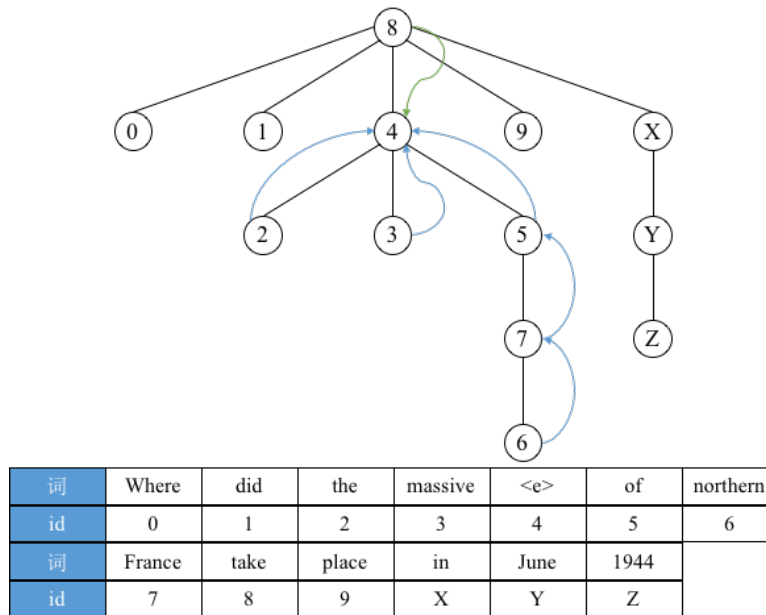


图 2-3 问题表示模块结构示意图

在问题表示模块中，我们将 Tree-LSTM 结构^[42]制定为一个双向的过程。给定一个问题，我们首先将其转化为词向量。然后分别自底向上和自顶向下的构建循环神经网络。对于一个依存句法树来说，对于任意的节点 j ， $C(j)$ 代表节点 j 的孩子节点集合。然后我们按照公式(2-12)~公式(2-18)来构建自底向上的循环神经网络：

$$\widetilde{h}_j^\uparrow = \sum_{k \in C(j)} h_k^\uparrow, \quad (2-12)$$

$$i_j^\uparrow = \sigma(W^{(i)}x_j + U^{(i)}\widetilde{h}_j^\uparrow + b^{(i)}), \quad (2-13)$$

$$f_{jk}^\uparrow = \sigma(W^{(f)}x_j + U^{(f)}h_k^\uparrow + b^{(f)}), \quad (2-14)$$

$$o_j^\uparrow = \sigma(W^{(o)}x_j + U^{(o)}\widetilde{h}_j^\uparrow + b^{(o)}), \quad (2-15)$$

$$u_j^\uparrow = \tanh(W^{(u)}x_j + U^{(u)}\widetilde{h}_j^\uparrow + b^{(u)}), \quad (2-16)$$

$$c_j^\uparrow = i_j^\uparrow \odot u_j^\uparrow + \sum_{k \in C(j)} f_{jk}^\uparrow \odot c_k^\uparrow, \quad (2-17)$$

$$h_j^\uparrow = o_j^\uparrow \odot \tanh(c_j^\uparrow), \quad (2-18)$$

其中, x_j 是节点 j 的输入, h_k^\uparrow 是 j 节点的每个孩子的隐层状态, $W^{(i)}$ 、 $W^{(f)}$ 、 $W^{(o)}$ 、 $U^{(i)}$ 、 $U^{(f)}$ 、 $U^{(o)}$ 、 $b^{(i)}$ 、 $b^{(f)}$ 、 $b^{(o)}$ 为需要学习的参数。 \uparrow 符号代表自底向上的过程。在这个过程中, 我们对依存句法树的每个叶子节点按公式(2-19)进行初始化:

$$h_{leaf}^\uparrow = 0 \quad (2-19)$$

自顶向下的构建过程如公式(2-20)~公式(2-25)所示:

$$i_j^\downarrow = \sigma(W^{(i)}x_j + U^{(i)}h_j^\downarrow + b^{(i)}), \quad (2-20)$$

$$f_j^\downarrow = \sigma(W^{(f)}x_j + U^{(f)}h_j^\downarrow + b^{(f)}), \quad (2-21)$$

$$o_j^\downarrow = \sigma(W^{(o)}x_j + U^{(o)}h_j^\downarrow + b^{(o)}), \quad (2-22)$$

$$u_j^\downarrow = \tanh(W^{(u)}x_j + U^{(u)}h_j^\downarrow + b^{(u)}), \quad (2-23)$$

$$c_j^\downarrow = i_j^\downarrow \odot u_j^\downarrow + f_j^\downarrow \odot c_j^\downarrow, \quad (2-24)$$

$$h_j^\downarrow = o_j^\downarrow \odot \tanh(c_j^\downarrow), \quad (2-25)$$

与自底向上的过程相似, x_j 是节点 j 的输入, h_j^\downarrow 是 j 节点的每个孩子的隐层状态, $W^{(i)}$ 、 $W^{(f)}$ 、 $W^{(o)}$ 、 $U^{(i)}$ 、 $U^{(f)}$ 、 $U^{(o)}$ 、 $b^{(i)}$ 、 $b^{(f)}$ 、 $b^{(o)}$ 为需要学习的参数。 \downarrow 符号代表自顶向下的过程。在这个过程中, 为了在一定程度上结合整个依存句法树的信息, 对于依存句法树的根节点, 我们利用自底向上过程中根节点的状态作为自底向下过程的初始状态, 如公式(2-26)所示:

$$h_{root}^\downarrow = h_{root}^\uparrow \quad (2-26)$$

以依存句法树中的某一节点 j 为例, 自底向上的过程结合了节点 j 的子孙节点信息, 但缺少祖先节点信息; 而自顶向下的过程结合了节点 j 祖先节点的信息, 而缺少子孙节点的信息, 所以我们将两者结合得到节点 j 的最终表示, 如

公式(2-27)所示：

$$q_j = [h_j^{\uparrow}; h_j^{\downarrow}] \quad (2-27)$$

由此，我们得到了整个问题的最终表示 $\{q_t\}_{t=1}^n$ ， t 为原始问题的 id。与 Yu 的模型相似，我们将主题实体用“<e>”代替，这样可以给模型位置信息，并且可以消除主题实体对关系识别的影响。此外，我们注意到，主题实体通常是一个名词或名词短语，所以屏蔽主题实体对构建依存句法树几乎没有影响。此外，在我们的工作中，由于问题长度不是特别长，所以我们没有设置窗口大小。

2.5.3 匹配层模块

在之前的步骤中，我们得到了问题的语义信息表示。在匹配层模块中，我们结合问题和关系的信息并将序列信息与句法信息融合。

首先，输入我们在之前步骤得到的问题表示 $\{q_t\}_{t=1}^n$ 、关系级别的关系表示 $\{r_t^{rel}\}_{t=1}^{m_1}$ 和词级别的关系表示 $\{r_t^{word}\}_{t=1}^{m_2}$ ，与 Wang 和 Jiang 的工作类似，我们将 q_t 和 \tilde{q}_t 连接作为循环神经网络的输入，如公式(2-28)所示：

$$g_q = \text{BiLSTM}(g_{q-1}, [q_t; \tilde{q}_t]) \quad (2-28)$$

其中， \tilde{q}_t 是由如下的注意力机制的池化方法得到，如公式(2-29)~(2-31)所示：

$$s_j^t = v^T \tanh(W_r r_j + W_q q_t + W_g g_{t-1}) \quad (2-29)$$

$$a_i^t = \frac{e^{s_i^t}}{\sum_{j=1}^n e^{s_j^t}} \quad (2-30)$$

$$\tilde{q}_t = \sum_{i=1}^n a_i^t r_i \quad (2-31)$$

其中， v ， W_r ， W_q ， W_g 是需要学习的参数， r_i 代表 r_i^{rel} 或者 r_i^{word} 。对于关系级别的关系表示和词级别的关系表示这两种方式来说，我们使用相同的模型结构。在这一步骤之后，我们得到了 $G_Q^{rel} = \{g_q^{rel}\}_{t=1}^n$ 和 $G_Q^{word} = \{g_q^{word}\}_{t=1}^n$ ，分别结合了两种级别的关系表示。每个问题表示动态的结合了来自关系的信息。

这个模块实际上是一个混合匹配过程。通过这一模块后，我们将语义信息和序列信息融合在了一起。这个方法已应用于多个任务，尤其是阅读理解任务。我们的工作与阅读理解模型不同的是：传统的阅读理解任务通常作为一个序列匹配任务，而我们的模型则是一个混合匹配任务，可以同时捕获序列和句法信

息。

2.5.4 输出层模块

在这个步骤中，我们首先在 G_Q^{rel} 和 G_Q^{word} 上应用最大池化操作得到了向量表示，如公式(2-32)~(2-33)所示：

$$\widetilde{g_Q^{rel}} = \text{MaxPooling}(G_Q^{rel}) \quad (2-32)$$

$$\widetilde{g_Q^{word}} = \text{MaxPooling}(G_Q^{word}) \quad (2-33)$$

然后，我们通过构建特征向量来得到最终的向量表示 g_q ，如公式(2-34)所示：

$$g_q = [\widetilde{g_Q^{rel}}; \widetilde{g_Q^{word}}] \quad (2-34)$$

最终，我们用一个线性层来计算问题-关系对的语义相似度，如公式(2-35)所示：

$$S_{q,r} = \text{Sigmoid}(l \cdot (W \cdot g_q + b)) \quad (2-35)$$

其中， W ， b 是需要学习的参数。

2.5.5 训练方式

与我们之前介绍的基于残差网络的关系识别模型相同，我们使用 Hinge Loss 来作为训练损失函数来扩大正负例样本之间的差别。在训练过程中，候选关系集合 R 可以划分为两个子集，即正确关系集 R_{pos} 和其他关系集 R_{neg} ，对于每个 $r_{pos} \in R_{pos}$ 和 $r_{neg} \in R_{neg}$ ，我们使用公式(2-36)来计算训练损失：

$$l = \max\{0, \gamma - s_{q,r_{pos}} + s_{q,r_{neg}}\} \quad (2-36)$$

其中 γ 为一个介于 0 到 1 之间的常数，用来扩大正负例之间的差距，这样，我们可以保证正例的分数比负例的分数高。

在测试阶段，我们计算问题 q 与每个候选关系 r 之间的语义相似度 $s(q, r)$ ，然后选择分数最高的关系作为最佳关系，具体操作如公式(2-37)所示：

$$r_{best} = \text{argmax}_{r \in R} s(q, r) \quad (2-37)$$

2.6 实验与分析

本小节介绍基于句法信息的关系识别技术的实验，分为实验设置及结果和分析。为了验证我们采用的模型的性能，我们详尽的对比了两个模型，结果显示我们提出的模型比基线模型有着明显的提高。

2.6.1 实验设置

本章使用 SimpleQuestions 数据集进行实验。我们使用 Pytorch 构建关系识别模型,使用 300 维 Glove 词向量^[43]作为预训练词向量。由于关系词为复合词,缺少预训练词向量,我们随机初始化关系级别的关系词,并将所有未登录词初始化为 0。我们使用 Stanford Parser^[44]来解析每个自然语言问题,从而获得依存句法树结构。在模型训练的过程中,我们使用 Adagrad 算法^[45]来最小化训练损失,学习率设为 0.00005,隐层维度设置为 400,此外,神经网络层与层之间的 dropout 设置为 0.3,损失函数中的常数设为 0.5。

2.6.2 实验结果与分析

表 2-3 显示了我们的模型与基线系统在关系识别任务上结果的比较。AMPCNN 是 Yin 等人^[20]的工作,其应用了注意力最大池化机制的卷积神经网络结构;BiCNN 是 Yih 等人^[17]的工作,在他们的工作中,问题和关系通过字符级别卷积神经网络来建模表示,其中窗口大小为 3;BiLSTM 代表两个 BiLSTM 模型用来分别表示问题和关系,词级别的关系表示和关系级别的关系表示带来了不同的效果;HR-BiLSTM 为 Yu 等人^[31]的工作,即我们在 2.4 节介绍的基于残差网络的关系识别模型。

我们的模型超过了基于残差网络的关系识别模型 0.4 个百分点,这意味着句法信息一定程度上在问题理解中有着一定作用。

表 2-3 关系识别模型在 SimpleQuestions 数据集上的实验结果

模型	准确率 (%)
AMPCNN (Yin 等)	91.3
BiCNN (Yih 等)	90.0
BiLSTM (只用词级别关系表示)	91.2
BiLSTM (只用关系级别关系表示)	88.9
HR-BiLSTM (Yu 等)	93.3
基于双向树结构的关系识别模型 (本章工作)	93.7

模型消融实验如表 2-4 所示。我们每次移除一个部分,保持另外几个部分不变来分析模型中每个部分对于整个模型的贡献。我们可以看出,双向树结构的问题表示和混合注意力匹配机制贡献了最多的性能提升。

首先,移除混合注意力匹配机制带来了 0.51 个百分点的下降,这说明结合句法信息和序列信息起到了一个重要的作用,其中,注意力机制和匹配算法都

对最后的结果起到了一定的作用；然后，我们移除了双向树结构的问题表示结构，并替换为常规的 BiLSTM，这个操作带来了 0.28 个百分点的下降。

从表格中可以观察到，自底向上的结构化问题表示比自顶向下的结构化问题表示起到了更大的作用。因为对于依存句法树中的大部分节点来说，子孙节点的数量要比祖先节点的数量多，这意味着自底向上的结构化问题表示比自顶向下的结构化问题表示可以捕获更多信息。此外，仅仅使用自底向上的结构化问题表示或自顶向下的结构化问题表示并不会比使用 BiLSTM 效果好，只有二者结合才能带来更好的效果。

表 2-4 消融实验

模型	准确率 (%)
基于双向树结构的关系识别模型（本章工作）	93.65
-将双向树结构的问题表示替换为 BiLSTM	93.37
-自底向上的问题表示	93.27
-自顶向下的问题表示	93.51
-混合注意力匹配	93.14
-匹配（双向树结构问题表示+注意力机制）	93.37
-注意力机制	93.44
单层 BiLSTM	92.80

为了清晰的论证在匹配层中混合注意力匹配机制的效果，我们选择采用注意力机制热力图的方式来可视化每个问题-关系对的权重。如图 2-4 所示，我们使用一个例子来展示二者之间的注意力权重关系。由于在单关系任务中，关系级别的关系只有一个词，所以我们只展示词级别的关系表示的热力图。可以看到，语义相近的两个词获得了更高的语义匹配权重。

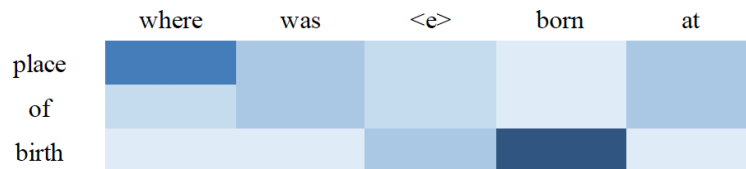


图 2-4 单关系识别注意力机制热力图

2.7 本章小结

本章介绍了基于句法信息的关系识别技术，首先介绍了关系识别任务，包括其任务定义以及候选项生成方法，然后介绍了本章使用的数据集，之后介绍

了该任务上的一个比较经典且效果优秀的基线方法，然后介绍了我们在这章中提出的模型，最后介绍了实验部分，分析了模型各个模块的结果以及其作用，实验显示我们提出的基于双向树结构的关系识别模型相比于之前的工作有着明显提升，这说明句法信息与序列信息结合的方式有着一定的作用。

第 3 章 基于语义图的信息检索方法研究

3.1 引言

在基于信息检索的知识库问答技术中，除单关系问题外，还有多关系问题。由于问题的复杂性和关系的多样性，多关系问题比单关系问题更具挑战性。本章将介绍基于多关系信息检索的知识库问答技术，并研究如何在此基础上构建一个完整的知识库问答系统。基于信息检索的知识库问答系统流程图如图 3-1 所示。

与单关系问题不同，多关系问题往往需要多跳推理过程，直接应用关系识别模型不足以回答完整的多关系问题。此外，只有一条从主题实体指向答案的关系链并不足以完全限制答案的范围。故在本章，我们研究如何解决多关系问题。多关系问题可以分为两个子任务来解决，即关系识别和约束识别，关系识别用来识别问题的核心关系推导链，而约束识别用来向核心关系推导链上添加约束。本章首先将介绍数据集，然后介绍基于关系识别与约束识别的知识库问答系统，最后将会介绍实验部分。

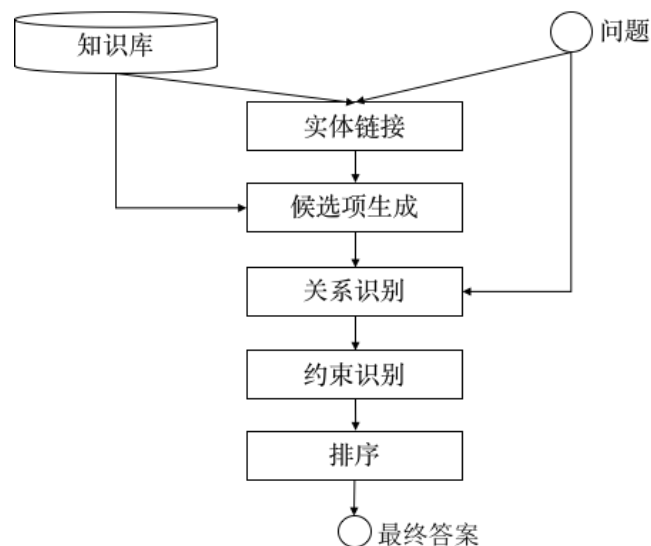


图 3-1 基于信息检索的知识库问答系统流程图

表 3-1 WebQSP 语料规模

数据集	训练集	测试集
样本数量	3778	1639

表 3-2 WebQSP 数据样例

问题	SPARQL
Who was richard nixon married to?	<pre> PREFIX ns: http://rdf.freebase.com/ns/ SELECT DISTINCT ?x WHERE { FILTER (?x != ns:m.06c97) FILTER (!isLiteral(?x) OR lang(?x) = " OR langMatches(lang(?x), 'en')) ns:m.06c97 ns:people.person.spouse_s ?y . ?y ns:people.marriage.spouse ?x . ?y ns:people.marriage.type_of_union ns:m.04ztj . } </pre>
what year did president william henry harrison take office?	<pre> PREFIX ns: http://rdf.freebase.com/ns/ SELECT DISTINCT ?x WHERE { FILTER (?x != ns:m.0835q) FILTER (!isLiteral(?x) OR lang(?x) = " OR langMatches(lang(?x), 'en')) ns:m.0835q ns:government.politician.government_positions_held ?y . ?y ns:government.government_position_held.from ?x . ?y ns:government.government_position_held.office_position_or_title ns:m.060d2 . } </pre>
who will play mr gray in the film?	<pre> PREFIX ns: http://rdf.freebase.com/ns/ SELECT DISTINCT ?x WHERE { FILTER (?x != ns:m.0nfjvrm) FILTER (!isLiteral(?x) OR lang(?x) = " OR langMatches(lang(?x), 'en')) ns:m.0nfjvrm ns:film.film_character.portrayed_in_films ?y . ?y ns:film.performance.actor ?x . } </pre>

3.2 数据集介绍

在本章中，我们使用 WebQSP 作为数据集。WebQSP 由 Yih 等人^[39]于 2016 年发布，是一个针对于多关系问题的数据集，在 WebQuestions 数据集的基础上添加了 SPARQL 表达式的标注，以用来解决语义解析问题。该数据集使用 Freebase 全集作为知识库，数据集样本由自然语言问题和对应的可以指向正确答案的 SPARQL 逻辑表达式组成。数据集与知识库具体统计信息如表 3-1 所示。

表 3-2 显示了 WebQSP 数据集中随机选取的几个例子，其中，“ns:m.06c97、ns:m.04ztj”等代表知识库中定义的实体编号，“ns:people.person.spouse_s、ns:people.marriage.spouse、ns:people.marriage.type_of_union”等词代表知识库中定义的关系。以关系“ns:people.person.spouse_s”为例，“people”代表关系的域（domain），“person”代表关系的类型（type），“spouse_s”代表关系的属性（property），观察可知，属性与自然语言中的问题关系最紧密，其语义信息也最可能在自然语言问题中出现，故为了简化模型，本课题中只使用关系的属性进行关系表示，不使用域和类型。在本章中，与 Yih 等人^[17]、Yu 等人^[31]的工作相似，我们使用 Yang 和 Chang 等人^[40]的实体链接结果来进行后续工作。由于我们使用信息检索的方法，故任务可分为关系识别与约束识别两个子任务，我们同样使用 Yu 等人^[31]发布的针对于信息检索方法版本的数据集进行实验。

3.3 基于关系识别与约束识别的知识库问答系统

3.3.1 候选项生成方法

与单关系问题不同，多关系的问题标签不是一个唯一的三元组，而是一个由很多三元组构成的 SPARQL 语句。根据信息检索方法的特性，我们将整个任务分为关系识别与约束识别两个步骤。故首先我们需要在 SPARQL 语句中提取从主题实体指向答案的关系链作为核心关系推导链来进行关系识别，其余约束用来进行约束识别。以表 3-2 的第二个例子为例，自然语言问题为“What year did president William henry Harrison take office?”其对应的 SPARQL 语句为

```

PREFIX ns: http://rdf.freebase.com/ns/\nSELECT DISTINCT ?x\nWHERE {\nFILTER (?x != ns:m.0835q)\nFILTER (!isLiteral(?x) OR lang(?x) = " OR langMatches(lang(?x), 'en'))\nnns:m.0835q ns:government.politician.government_position_s_held ?y .\n?y ns:government.government_position_held.from ?x .\n?y ns:government.government_position_held.office_position_or_title ns:m.060d2 .\n} c 的 SPARQL 语句 “PREFIX ns: http://rdf.freebase.com/ns/\nSELECT DISTINCT ?x\

```

`nWHERE { \nFILTER (?x != ns:m.0835q)\nFILTER (!isLiteral(?x) OR lang(?x) = " OR langMatches(lang(?x), 'en'))\nns:m.0835q ns:government.politician.government_positions_held ?y .\n?y ns:government.government_position_held.from ?x .\n?y ns:government.government_position_held.office_position_or_title ns:m.060d2 .\n}`”来说,“government_positions_held-from”为核心关系推导链,而“office_position_or_title”为约束。本课题中,我们采用如上方法从 SPARQL 语句中提取得到核心关系推导链和约束。

3.3.2 语义图定义

为了形式化的表示 SPARQL,我们定义一种语义图结构来刻画 SPARQL 中实体与关系的连接方式。图 3-2 为语义图的实例,以该图为例,语义图中共有三种节点,其中,矩形节点表示实体节点,圆形节点表示变量节点,即连接若干关系的 CVT 节点,菱形节点表示聚合约束节点;语义图中共有两种边,其中,r1~r4 表示关系边,MaxAtN 表示聚合函数组成的边。在语义图结构中,核心关系推导链和约束均为均可视为语义图中的边,其中 r1-r2 为核心关系推导链,r3、r4、MaxAtN 为约束。

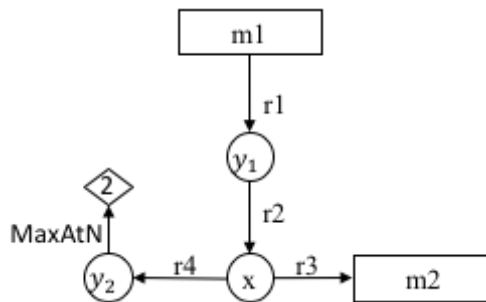


图 3-2 语义图实例

3.3.3 基于双向树结构的关系识别模型

由第二章介绍的关系识别模型可知,关系识别模型的输入为自然语言问题和候选关系(在本章中为核心关系推导链),输出为二者之间的语义相似度,然后对所有的语义相似度进行排序,得到最终的核心关系推导链。这与我们上一章提出的关系识别模型类似。我们在上一章模型的基础上完成本章工作,为了得到与上一章的模型相同的结构,首先我们将核心关系推导链做如下处理,以“government_positions_held-from”为例,关系级别的表示中,我们将其以“government_positions_held”、“from”的形式作为循环神经网络的输入,词级

别的表示中，我们将其以“government”、“positions”、“held”、“from”的形式作为神经网络的输入。我们通过观察可以得到，越靠近答案节点的关系，其重要程度可能越高，所以这种表示方式在一定程度上是合理的。于是，我们可以使用第二章中提出的模型来进行关系识别操作。

3.3.4 约束识别方法

核心关系推导链只能构成最简单的语义图，虽然可以通过该语义图得到包含正确答案的答案集合，但该集合中仍包含许多不满足问题中约束条件的不正确的实体答案，故我们需要通过添加约束来进一步缩小答案集合。与 Yih 等人^[17]和 Yu 等人^[31]的工作相似，我们使用一种约束识别方法来识别问题中可能存在的约束条件，然后将这些约束条件添加进语义图中来得到更准确的答案。

为了进一步限制答案集合，我们通过以下两种方式为只包含核心关系推导链的语义图添加约束：即通过添加约束关系或添加聚合函数。下面将分别介绍两种方法：

添加约束关系即对语义图中的答案节点和 CVT 节点来说，寻找该节点周围一跳之内的所有三元组，然后将关系作为边，三元组的尾实体作为节点连接在语义图中，组成新的语义图。

添加聚合函数即同样对语义图中的答案节点和 CVT 节点来说，有一些约束无法用三元组的形式进行表达，对于这种情况我们需要为特定节点添加聚合函数，聚合函数由“argmin”、“argmax”等一系列聚合操作组成。

在得到了上述两种方法构建的语义图集合后，我们需要使用一种匹配方法来判断我们需要保留哪条约束。这里我们计算输入问题的 n-gram 与约束三元组的关系的余弦相似度（对于聚合函数，计算输入问题的 n-gram 与聚合函数节点值的余弦相似度），保留相似度大于 θ 的约束， θ 为介于 0 与 1 之间的常数。

3.4 实验与分析

3.4.1 实验设置

本章使用 WebQSP 数据集进行实验，我们使用 Pytorch 构建关系识别模型，使用 300 维 Glove 词向量^[43]作为预训练词向量。由于关系词为复合词，缺少预训练词向量，我们随机初始化关系级别的关系词，并将所有未登录词初始化为 0。我们使用 Stanford Parser^[44]来解析每个自然语言问题，从而获得依存句法树结构。在模型训练的过程中，我们使用 Adagrad 算法^[45]来最小化训练损失，学

习率设为 0.00001，隐层维度设置为 400，此外，神经网络层与层之间的 dropout 设置为 0.3，损失函数中的常数设为 0.5。

3.4.2 实验结果与分析

首先我们来验证关系识别模型对于多关系问题带来的提升。表 3-3 显示了我们的模型与基线系统在多关系问题的关系识别任务上结果的比较。BiCNN 是 Yih 等人的工作，在他们的工作中，问题和关系通过字符级别卷积神经网络来建模表示，其中窗口大小为 3；BiLSTM 代表两个 BiLSTM 模型用来分别表示问题和关系，词级别的关系表示和关系级别的关系表示带来了不同的效果；HR-BiLSTM 为 Yu 等人的工作，即我们在 2.4 节介绍的基于残差网络的关系识别模型。我们的模型在 WebQSP 上超过了基于残差的关系识别模型 2.61 个百分点，相比于第二章在 SimpleQuestions 上的实验，有着更加明显的作用。这是因为由于 WebQSP 中的问题要比 SimpleQuestions 中的问题要复杂，所以 WebQSP 中问题的句法特征要比 SimpleQuestions 中的问题要明显得多，所以在 WebQSP 上的效果比 SimpleQuestions 上更加显著。

表 3-3 关系识别模型在 WebQSP 数据集上的实验结果

模型	准确率 (%)
BiCNN (Yih 等)	77.74
BiLSTM (只用词级别关系表示)	79.32
BiLSTM (只用关系级别关系表示)	78.96
HR-BiLSTM (Yu 等)	82.53
基于双向树结构的关系识别模型 (本章工作)	85.14

我们在关系识别的基础上，加上实体链接与约束识别模块，构成完整的问答系统，其结果如表 3-4 所示。可以看出，我们的系统比两个基线系统搞出了 0.6 个百分点。

表 3-4 基于多关系信息检索的知识库问答系统结果

模型	准确率 (%)
STAGG (Yih 等)	63.9
HR-BiLSTM (Yu 等)	63.9
基于双向树结构的关系识别模型+约束识别 (本章工作)	64.5

模型消融实验如表 3-5 所示。我们可以看出，相比于 SimpleQuestions 上，每个位置都有了更加显著的提升。同样地，双向树结构的问题表示和混合注意

力匹配机制贡献了最多的性能提升。

表 3-5 消融实验

模型	准确率 (%)
基于双向树结构的关系识别模型（本章工作）	85.14
-将双向树结构的问题表示替换为 BiLSTM	83.93
-自底向上的问题表示	83.51
-自顶向下的问题表示	84.29
-混合注意力匹配	82.35
-匹配（双向树结构问题表示+注意力机制）	83.57
-注意力机制	83.32
单层 BiLSTM	78.41

首先，移除混合注意力匹配机制带来了 2.78 个百分点的下降，这说明结合句法信息和序列信息起到了一个非常重要的作用；然后，移除双向树结构的问题表示结构，并替换为常规的 BiLSTM，这个操作带来了 1.21 个百分点的下降。与 SimpleQuestions 相同，自底向上的结构化问题表示比自顶向下的结构化问题表示起到了更大的作用。同样，结合自底向上的结构化问题表示和自顶向下的结构化问题表示才能带来更好的效果。

为了清晰的论证在匹配层中混合注意力匹配机制的效果，我们同样选择采用注意力机制热力图的方式来可视化每个问题-关系对的权重。如图 3-3 和图 3-4 所示，我们使用两个例子来展示二者之间的注意力权重关系。可以看到，在关系级别的关系表示和词级别的关系表示中，语义相近的两个词都获得了更高的语义匹配权重。

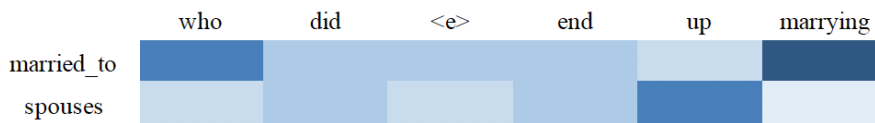


图 3-3 多关系识别注意力机制热力图（关系级别关系表示）

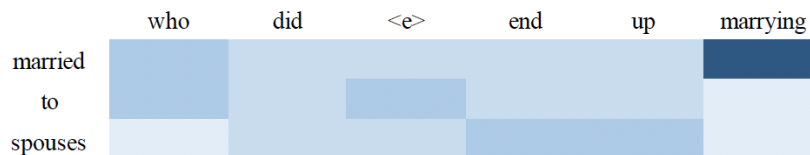


图 3-4 多关系识别注意力机制热力图（词级别关系表示）

与此同时，我们对模型预测的错误进行了分析。我们将错误的预测划分成

以下三类：

(1) **十分相似的关系** 我们预测的关系与其正确的关系非常相似。例如，对于问题 “What is the name of a game developed by free radical design”，模型预测得到的结果为 “game_versions_developed”，而不是 “games_developed”。这种错误的类型占了最大的比例。

(2) **具有不同表达方式的相似关系** 由于表达方式的多样性，将自然语言表述映射到知识库定义的关系中是非常困难的。例如，对于问题 “What does David Wiesner do for a living?”，模型预测的关系为 “notable_types”。然而，正确的关系为 “profession”。错误的原因是虽然 “do for a living” 和 “profession” 是一组同义词，但字面上差距很大，机器很难理解。

(3) **与实体相关的关系** 在本课题中，为了得到主题实体位置信息以及消除主题实体信息对于关系识别的影响，我们屏蔽了主题实体信息，但也会在一定程度上造成一些问题。例如，对于问题 “What is the name of an art gallery?”，我们将主题实体 “an art gallery” 用 “<e>” 来代替，即输入模型的问题为 “What is the name of <e>”。因此，由于主题实体信息的缺失，我们的模型未能预测得到其正确的关系 “museums”。

3.5 本章小结

本章介绍了基于多关系信息检索的知识库问答技术，首先介绍了本章使用的数据集，然后介绍了一套经典且效果优秀的知识库问答系统，然后完善了第二章提出的模型，将其组成一套端到端的问答系统，并介绍了关系识别方法和约束识别方法，最后介绍了实验部分，分析了模型各个模块的结果以及其作用，并对模型预测的错误进行了分析，实验显示我们的关系识别模型相比于之前的工作有着明显提升。此外，在此基础上，我们构建的知识库问答系统相比于之前的工作也有着明显的提升。

第 4 章 基于序列到动作模型的语义解析方法研究

4.1 引言

基于信息检索的知识库问答系统已经广泛的用于开放域知识问答系统、在线客服系统等多个场景，是一个近乎成熟的方法，但仍存在一些问题，比如模型难以预测问题中的隐含问题（“do for a living”与“profession”）。针对以上问题，我们采用基于语义解析的方法来一定程度上避免以上情况。语义解析的目标是将自然语言问题转换成可以直接查询知识库的逻辑表达式，从而得到正确答案。本章中，我们使用 SPARQL 作为目标逻辑表达式，为了形式化的表示 SPARQL，我们定义一种语义图结构来刻画 SPARQL 中实体与关系的连接方式。基于语义解析的知识库问答系统的流程图如图 4-1 所示。我们首先将自然语言问题转换为一系列的动作序列，然后将动作序列转换为语义图，之后将语义图转换为 SPARQL 逻辑表达式，最后查询知识库得到答案。与基于信息检索的知识库问答系统类似，我们同样使用 Yang 和 Chang 等人^[40]的实体链接结果，以便于对比两种方法，为了简洁，我们在图 4-1 中省略了实体链接步骤。由于从动作序列到语义图到 SPARQL 都是基于一系列的规则进行转换，故本章的重点在如何将问题转换为动作序列。

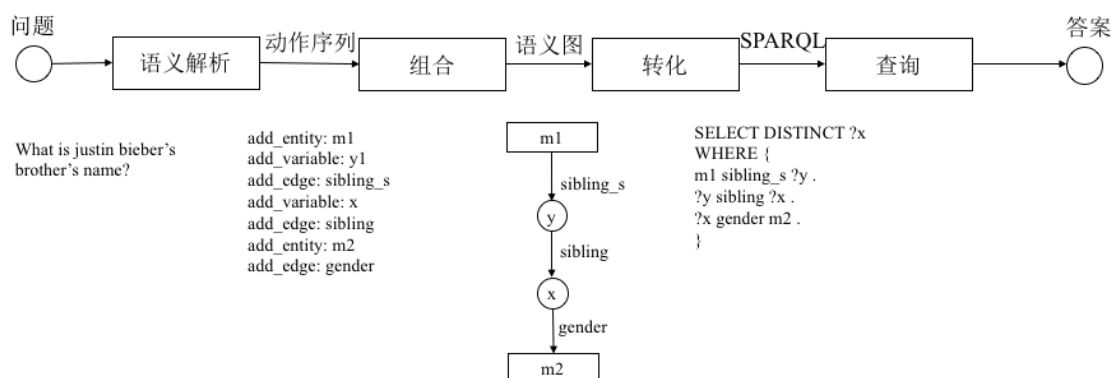


图 4-1 基于语义解析的知识库问答系统流程图

4.2 数据集介绍

在本章中，我们仍使用 WebQSP 数据集，与第三章不同的是，我们不在将语义解析任务拆分为关系识别和约束识别，而是将其所有的核心关系推导链和约束都视为语义图中的边，而主题实体和约束实体均视为语义图中的节点。故

我们使用原始数据集进行实验。此外，我们采用与 3.3.2 节相同的语义图定义。

4.3 动作到序列模型简介

4.3.1 动作定义

为了生成可以表示语义图的动作序列，我们定义了如表 4-1 所示的一系列动作。其中，动作语义表示动作所包含的语义信息，参数 1 和参数 2 表示动作所包含的位置信息。“add_entity”表示向语义图中添加一个实体节点，其动作语义代表该节点所具有的语义信息，同理，“add_variable”表示向语义图中添加一个变量，“add_answer_variable”同样表示向语义图中添加一个节点，但该节点指向答案；“add_general_edge”表示向语义图中添加一条关系边，动作语义代表关系所具有的语义信息，而参数 1 和参数 2 分别代表与关系相连的头实体与尾实体；“add_oper_edge”表示向语义图中添加一条聚合函数边，动作语义代表据韩函数名称（如 MaxAtN, argmax 等等），参数 1 和参数 2 代表需要约束的节点和具体约束条件（可为空，当为空时用占位符替代）。

表 4-1 动作定义

动作名称	动作语义	参数 1	参数 2	注释
add_entity	<e>	-	-	添加实体节点
add_variable	y1	-	-	添加变量节点
add_answer_variable	x	-	-	添加答案变量节点
add_general_edge	sibling_s	<e>	y1	添加关系边
add_oper_edge	MaxAtN	x	2	添加聚合边

4.3.2 数据预处理

为了进行序列到动作模型的实验，我们需要对数据集进行预处理。本章使用的数据集为 WebQSP 数据集，数据集中提供自然语言问题与其对应的 SPARQL 语言。为了进行有监督训练，我们针对原始 SPARQL 语言进行如下处理：首先，提取 WHERE 从句中的所有信息，我们通过观察可以发现，WHERE 从句由若干条关联的三元组组成，通过分析三元组的组成成分，我们可以轻易的得到由这些三元组构成的语义图；在得到语义图后，我们从语义图中主题实体出发，利用广度优先搜索得到对应的动作序列。这样，我们就可以使用问题与动作序列来训练模型。

此外，我们从实体链接步骤得到主题实体后，我们利用主题实体在知识库

中检索该主题实体周围两跳的所有关系，作为候选关系集合，用于后续模型中。

4.4 基于序列到动作的语义解析模型

本节将介绍本课题中使用的从序列到动作的三种语义解析模型，分别为基于序列到序列的语义解析模型、基于拷贝机制^[46]的语义解析模型和执行导向的语义解析模型。

4.4.1 基于序列到动作的语义解析模型

序列到序列模型是一个非常经典的模型，由编码器和解码器组成，具体模型如图 4-2 所示。问题经过问题编码器编码成向量，然后使用其隐层状态来进行解码。此外，我们使用了注意力机制来计算解码器中每个时间戳与输入问题之间的注意力分布。模型分为编码器，解码器以及注意力机制。下面我们将分别介绍几个模块。

(1) **符号约定** 为了便于表述，我们使用如下的符号约定：模型的输入为自然语言问题， $q = (q_1, q_2, \dots, q_i, \dots, q_n)$ 表示自然语言问题， q_i 表示问题中的第 i 个词， n 为问题的长度；模型期望的输出为一组动作序列， $a = (a_1, a_2, \dots, a_i, \dots, a_m)$ 代表若干动作序列，每个动作序列又由动作名称、动作语义和若干参数组成，由于模型为序列模型，我们将其展开成序列结构，所有元素总数量为 m 。一组训练数据为一个问题和一个动作序列构成，训练数据的表示形式为 $\{q, a\}$ 的标注对。

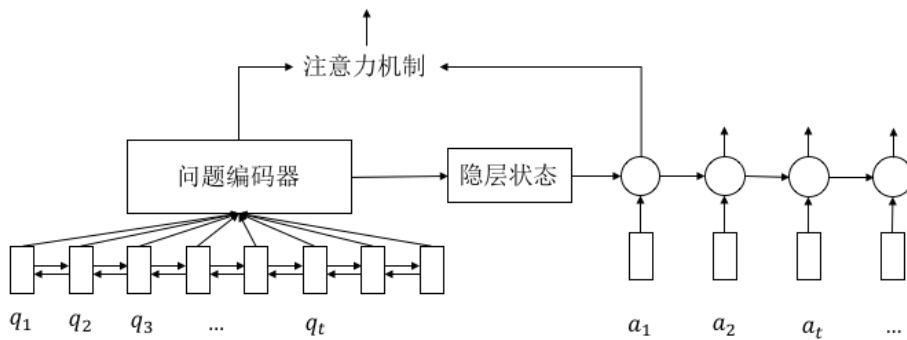


图 4-2 基于序列到序列的语义解析模型

(2) **输入模块** 对于一组问题-动作序列的输入，我们使用预训练词向量 Glove^[43] 将问题表示成向量形式，记为： $e^Q = \{e_1^Q, e_2^Q, \dots, e_t^Q, \dots, e_n^Q\}$ ，由于动作序列不存在预训练词向量，故我们随机初始化动作序列向量，记为 $e^A = \{e_1^A, e_2^A, \dots, e_t^A, \dots, e_m^A\}$ 。

(3) **编码器模块** 对于一个自然语言问题，我们使用公式(4-1)对其进行编码：

$$h_t^e = \text{BiGRU}(h_{t-1}^e, e_t^Q) \quad (4-1)$$

最后我们得到问题隐层表示 $q_{hidden} = q_N$ ，问题表示 $q = \{h_t^e\}_{t=1}^N$ 。

(4) **解码器模块** 我们在得到问题隐层表示后，我们使用公式(4-2)~(4-3)进行解码：

$$h_t^d = \text{GRUCell}(h_{t-1}^d, e_t^A) \quad (4-2)$$

$$h_0^d = q_{hidden} \quad (4-3)$$

(5) **注意力机制** 我们对解码器每个时间戳编码器的隐层表示进行注意力机制的操作，具体操作如公式(4-4)~(4-6)所示，其中 o_t 是应用注意力机制后解码器第 t 个时间戳的输出， p_t 为词表概率分布：

$$\alpha_t = \text{softmax}(w_1 h_t^d + w_2 q) \quad (4-4)$$

$$o_t = \alpha_t q \quad (4-5)$$

$$p_t = \text{log_softmax}(o_t) \quad (4-6)$$

(6) **训练方式** 模型采用交叉熵损失函数进行训练，具体操作如公式(4-7)~(4-8)所示：

$$\text{loss}_t = - \sum_{x \in a_t} p_t[x] \quad (4-7)$$

$$\text{loss} = \frac{1}{T} \sum_{t=0}^T \text{loss}_t \quad (4-8)$$

4.4.2 基于拷贝机制的语义解析模型

在经典的序列到序列模型的基础上，我们通过引入拷贝机制^[46]来改进我们的模型。我们的基于拷贝机制的语义解析模型在 See 等人的 pointer-generator 网络的基础上进行改进。Pointer-generator 网络使用拷贝机制将输入端的词拷贝至输出端来提升解码器的性能。受这一思路的启发，我们设计了如图 4-3 所示的模型。考虑到动作序列中几乎不会出现输入端的词，同时考虑到实体周围的关系对解码的重要性，我们设计了一种方法，将实体周围两跳的关系利用拷贝网络拷贝至解码器，得到拷贝概率和生成概率后，计算得到最终概率分布。模型分为问题编码器、候选关系编码器、解码器、注意力机制以及拷贝模块，下面我们将详细介绍这几个模块。

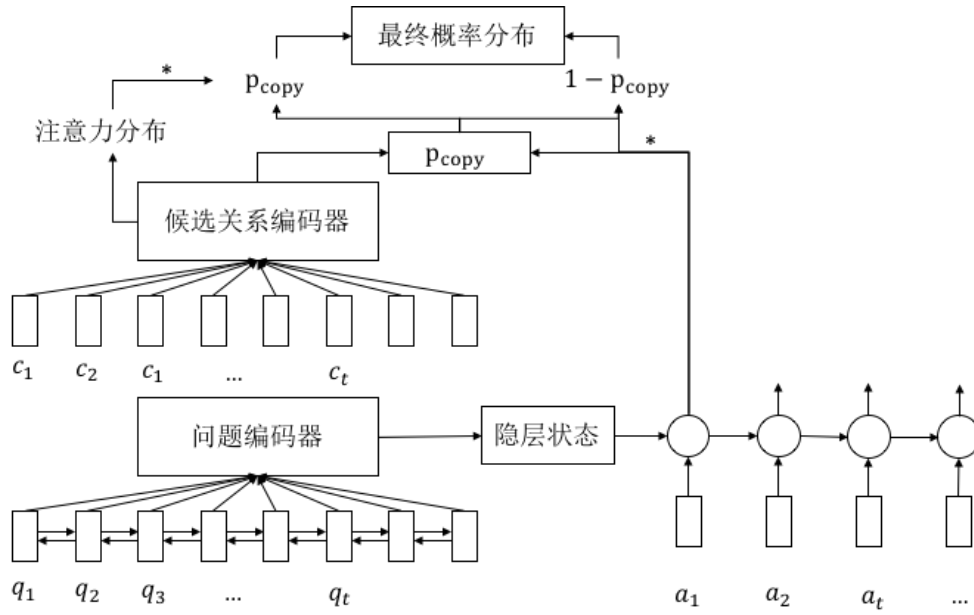


图 4-3 基于拷贝机制的语义解析模型

(1) **符号约定** 为了便于表述，我们使用如下的符号约定：模型的输入为自然语言问题和主题实体周围两条以内的候选关系，其中 $q = (q_1, q_2, \dots, q_i, \dots, q_n)$ 表示自然语言问题， q_i 表示问题中的第 i 个词， n 为问题的长度； $c = (c_1, c_2, \dots, c_i, \dots, c_k)$ 表示主题实体周围两跳的所有候选关系， c_i 表示第 i 个关系， k 为候选关系的数量；模型期望的输出为一组动作序列， $a = (a_1, a_2, \dots, a_i, \dots, a_m)$ 代表若干动作序列，每个动作序列又由动作名称、动作语义和若干参数组成，由于模型为序列模型，我们将其展开成序列结构，所有元素总数量为 m 。一组训练数据为一个问题、一个候选关系集合以及一个动作序列构成，训练数据的表示形式为 $\{q, c, a\}$ 的标注对。

(2) **输入模块** 对于一组 $\{q, c, a\}$ 训练数据的输入，我们使用预训练词向量 Glove^[43] 将问题表示成向量形式，记为： $e^q = \{e_1^q, e_2^q, \dots, e_t^q, \dots, e_n^q\}$ ，由于候选关系和动作序列不存在预训练词向量，故我们随机初始化候选关系向量和动作序列向量，分别记为： $e^c = \{e_1^c, e_2^c, \dots, e_t^c, \dots, e_m^c\}$ 和 $e^a = \{e_1^a, e_2^a, \dots, e_t^a, \dots, e_m^a\}$ 。

(3) **问题编码器模块** 对于一个自然语言问题，我们使用公式(4-9)对其进行编码：

$$h_t^{eq} = \text{BiGRU}(h_{t-1}^{eq}, e_t^q) \quad (4-9)$$

(4) **候选关系编码器模块** 对于一个候选关系集合，由于其不存在序列关系，我们使用线性编码器进行编码，最终得到候选关系的表示 c ，如公式(4-10)所示：

$$c = we^c + b \quad (4-10)$$

(5) **解码器模块** 我们在得到问题隐层表示后，我们以公式(4-11)~(4-12)进行解码：

$$h_t^d = \text{GRUCell}(h_{t-1}^d, e_t^A) \quad (4-11)$$

$$h_0^d = q_{\text{hidden}} \quad (4-12)$$

(6) **注意力机制** 我们对解码器每个时间戳编码器的隐层表示对输入的候选关系进行注意力机制的操作，具体操作如公式(4-13)~(4-15)所示，其中 s_t 是应用注意力机制后解码器第 t 个时间戳的输出， p_{vocab} 为词表概率分布：

$$\alpha_t = \text{softmax}(w_1 h_t^d + w_2 c) \quad (4-13)$$

$$s_t = \alpha_t c \quad (4-14)$$

$$p_{\text{vocab}} = \text{softmax}(s_t) \quad (4-15)$$

(7) **拷贝模块** 在拷贝模块中，由于输入端的自然语言问题不会出现在动作序列中，故我们设计一种可以将候选关系拷贝至解码端的方法来提了解码性能，具体过程如公式(4-16)~(4-18)所示：

$$p_{\text{copy}} = \sigma(w_1 s_t + w_2 h_t^d + w_3 e_t^A) \quad (4-16)$$

$$p_{\text{final}} = p_{\text{copy}} \sum_{i:w_i=w} \alpha_i^t + (1 - p_{\text{copy}}) p_{\text{vocab}}(w) \quad (4-17)$$

$$p_t = \log(p_{\text{final}}) \quad (4-18)$$

其中， p_t 为结合拷贝机制的最终概率分布， p_{copy} 为需要从候选答案中拷贝的概率。与 Pointer-Generator 网络相同，如果 w 是一个不存在于词表的词(out-of-vocabulary, OOV)，则 p_{vocab} 为0；如果 w 在候选关系中未出现，则 $\sum_{i:w_i=w} \alpha_i^t$ 为0。

(8) **训练方式** 模型采用交叉熵损失函数进行训练，具体操作如公式(4-19)~(4-20)所示：

$$\text{loss}_t = - \sum_{x \in a_t} p_t[x] \quad (4-19)$$

$$\text{loss} = \frac{1}{T} \sum_{t=0}^T \text{loss}_t \quad (4-20)$$

4.4.3 执行导向的语义解析模型

在基于拷贝机制的语义解析模型的基础上，我们发现引入候选关系后，效果有了显著的提升，但仍有一些不足。在使用候选关系集合之前，模型预测的

关系序列中会出现很多不在候选关系集合中的词，在使用候选关系集合后，发现这一情况有了很大的改善，但预测得到的很多关系词虽然在两跳范围内，但仍与正确答案相差较大，在此基础上，我们使用执行导向的语义解析模型来解决这一问题，执行导向的语义解析模型结构如图 4-4 所示。模型分为问题编码器、候选关系编码器、解码器、注意力机制以及拷贝模块以及执行导向机制。除执行导向机制外，其余与基于拷贝机制的语义解析模型相同，为了表述简洁，这里只介绍执行导向模块。

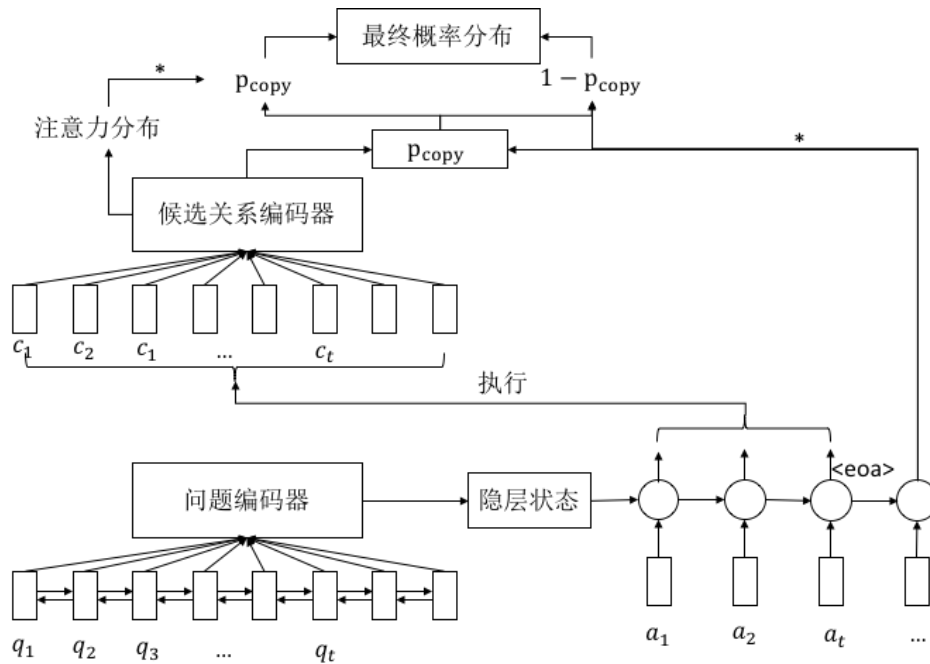


图 4-4 执行导向的语义解析模型

执行导向模块 在之前的模型中，我们发现引入候选关系集合的同时，会加强解码性能，但由于主题实体周围两跳的候选关系数量巨大，引入信息的同时也引入了部分噪声。我们提出一种执行导向的语义解析模型，模型将解码范围控制在解码过程中当前节点（实体节点与 CVT 节点）一跳范围内，从而解决引入大量候选关系所带来的数据噪声问题。模型描述如下：

针对生成的动作序列，在一个可以执行的动作解码结束后，我们执行生成的部分序列，得到实体集合作为中间结果，根据该实体集合，我们检索其一跳以内的所有关系作为动态词表，用于之后的解码过程。通过这种方式可以降低噪声信息的影响。例如，对于问题“**What is the name of <e> brother?**”其对应的正确的动作序列为“**add_entity <e> add_variable y1 add_general_edge ns:people.person.sibling_s <e> y1 <coa> add_answer_variable x add_general_edge ns:**

people.sibling_relationship.sibling y1 x <ea> add_entity <ge> add_general_edge ns:people.person.gender x <ge> <ea>”。其中“<e>”为主题实体，由实体链接识别得到，在语义解析过程中已被屏蔽；“<ge>”为问题中的其他实体，由于我们可以从主题实体出发，在知识库中根据生成的关系检索到其他实体，所以这里我们同样选择将其他实体屏蔽，用“<ge>”代替，这样模型将会专注于生成关系。由于添加边的动作结束后可得到一个可执行的三元组，故我们在动作序列的“add_general_edge”动作结束后添加标签“<ea>”来表示一个可执行动作的结束，当模型预测得到“<ea>”时，我们执行这一阶段得到的三元组并进行后续操作。

此外，我们在预测过程中使用波束搜索（beam search）来缓解错误级联问题。即在生成动作序列时，始终保留 top-k 的最佳预测序列（k 为波束大小，beam size），这在一定程度上可以缓解当解码出错误的关系时，对解码后面的元素所产生的影响。

4.5 实验与分析

4.5.1 实验设置

本章使用 WebQSP 数据集进行实验，我们使用 Pytorch 构建关系识别模型，使用 300 维 Glove 词向量^[43]作为与训练词向量。由于关系词为复合词，缺少与训练词向量，我们随机初始化关系级别的关系词，并将所有未登录词初始化为 0。在模型训练的过程中，我们使用 Adam 算法来最小化训练损失，学习率设为 0.0002，隐层维度设置为 1000，此外，神经网络层与层之间的 dropout 设置为 0.3。

4.5.2 实验结果与分析

本章实验结果如表 4-2 所示。我们分别给出了基于序列到序列的语义解析模型、基于拷贝机制的语义解析模型、执行导向的语义解析模型的结果。我们将基于序列到序列的语义解析模型作为基线模型。结果证明，每次模型的改进都带来了大幅度的效果的提升。

首先是基于序列到序列的语义解析模型，在这一经典的模型中，我们达到了 54.8% 的准确率；在基于拷贝机制的语义解析模型中，我们的准确率达到 57.5%，比基线系统提升了 2.7 个百分点，这证明，主题实体的候选关系信息对解码过程中生成关系有着很大的帮助，同时也证明拷贝机制对于序列到动作模型有很大的作用；在执行导向的语义解析模型中，我们的准确率达到 60.9%，

比基线系统提高了 6.1 个百分点，这证明，采用边执行边解码的策略来限制解码范围对于序列到动作模型产生了很大的作用。

表 4-2 语义解析模型结果

模型	准确率 (%)
基于序列到序列的语义解析模型	54.8
基于拷贝机制的语义解析模型	57.5
执行导向的语义解析模型	60.9

4.6 基于信息检索的模型与基于语义解析的模型的比较

为了阐述本课题不同章节之间的联系，我们在这一节对基于信息检索的知识库问答技术与基于语义解析的知识库问答技术进行详尽的比较与分析。首先我们给出在第三章和第四章中所提到的最好实验结果以及本课题所介绍的基线方法，如表 4-3 所示。

在表中，我们可以看到，基于语义解析的方法效果不佳，尽管我们已经极大程度的提升了语义解析方法的性能，但其仍比基于信息检索的方法低 3.6 个百分点。在这里我们根据实验数据分析得到了以下原因：

(1) **基于信息检索的模型为检索排序过程，而基于语义解析的模型为生成过程。**基于信息检索的模型实际上是一个检索排序过程，我们首先检索得到若干候选答案，然后对其进行打分排序，在候选答案中一定包含正确答案，模型要做的是如何通过表示学习的方法找到正确答案；而语义解析模型实际上是一个生成过程，我们使用的基于序列到动作的模型目标是生成我们自己定义的动作序列，从而得到最佳答案，分析数据得知，模型优秀的学习到了我们定义的动作序列生成规则，故未预测正确的动作序列中，占比最大的是因为模型生成了错误的关系，虽然我们使用执行导向的方法和波束搜索极大限度的避免了生成不可执行动作的可能性，但仍无法完全避免错误级联问题，即当前面时间戳的解码得到的关系错误时，后面的解码过程会因此受到影响，而基于信息检索的方法则不会存在该问题，因为所有的候选答案均为检索得到的。

(2) **训练信息检索模型比训练语义解析模型所需要的数据少很多。**众所周知，训练检索排序模型所需要的数据量比训练序列到序列模型所需要的数据量少很多，这在一定程度上也会体现在模型的性能上。WebQSP 数据集中训练数据的数量为 3778，虽然我们已极大程度的提升了语义解析模型的准确率，但数据的数量仍会对模型学习的效果产生一定的影响。而更大量数据集则意味着更大量更复杂的逻辑表达式标注，SPARQL 语言语法结构复杂，更加复杂的语法

结构导致了我们所构建的动作序列难以囊括所有的动作。

(3) **语义解析模型受未登录词的影响比信息检索模型更加严重。**当测试集中出现未登录词时,在信息检索模型中,我们一般只需定义一个特殊的词“<unk>”来表示该词,然后跟随模型一起训练即可,此外,少量未登录词一般情况下不会过大的影响正确答案的得分;而在语义解析模型中,尽管我们使用了一些技巧,但仍无法完全避免未登录词对语义解析模型所带来的影响,此外,由于动作序列的逻辑严谨性,在序列到动作模型中,未登录词对预测正确动作序列所带来的影响是毁灭性的。故未登录词会对语义解析模型产生很大的影响。

故我们可以得出结论,在知识库问答系统中,基于信息检索的方法更倾向于得到更好的效果,而语义解析模型虽在小规模数据上表现不佳,但其可以反应数据本身的相似度,在推理任务中仍然行之有效。

表 4-3 基于信息检索的方法与基于语义解析的方法对比结果 (WebQSP 数据集)

模型	准确率 (%)
STAGG (Yih 等)	63.9
HR-BiLSTM (Yu 等)	63.9
基于双向树结构的关系识别模型+约束识别 (信息检索方法)	64.5
执行导向的语义解析模型 (语义解析方法)	60.9

4.7 本章小结

本章介绍了基于语义解析的知识库问答技术。首先介绍了本章使用的数据集;然后介绍了序列到动作模型的基本情况,包括本章使用的语义图的定义、动作定义和数据预处理过程;之后介绍了我们本章使用的三个模型,分别是基于序列到序列的语义解析模型,基于拷贝机制的语义解析模型和基于执行导向的语义解析模型;最后介绍了实验部分,分析了各个模型的结果及其优势,对比了三个模型,对每一个优化的部分进行了分析,实验显示我们的语义解析模型相比于基线语义解析模型有着很大的提升。此外,我们对本章的不同章节所提到的知识库问答技术进行了详细的对比与分析,我们详尽的对比了模型类别、数据数量、未登录词等因素对语义解析模型和信息检索模型所产生的影响,并对两种模型的优缺点进行了总体的概述。

结 论

针对问答系统的不同方法以及问题的不同复杂程度,本课题分为三个部分,基于单关系信息检索的知识库问答技术、基于多关系信息检索的知识库问答技术和基于语义解析的知识库问答技术,三种方法均取得了良好的效果,对知识库问答系统的性能有着显著的提升。

在基于单关系信息检索的知识库问答技术方面,本课题提出了一种基于双向依存句法树的关系识别模型,其考虑了问题中所包含的句法信息。实验结果证明句法信息对于关系识别任务有着很大的作用。

在基于多关系信息检索的知识库问答技术方面,本课题在基于单关系信息检索知识库问答技术的基础上,使用了一种约束识别方法,构成了端到端的知识库问答系统。实验结果显示信息检索模型对于多关系的复杂问题仍然表现良好,约束识别方法对约束复杂问题时不可缺少的。

在基于语义解析的知识库问答技术方面,本课题首先介绍了序列到动作模型、定义了语义图和动作序列,并在基于序列到序列模型的基础上提出了拷贝机制的语义解析模型和执行导向的语义解析模型。实验结果显示两种方法与基线方法相比有着显著的提升,这证明候选关系对于语义解析模型有着非常重要的作用,消除数据噪声的影响对语义解析模型也有着至关重要的作用。

此外,我们对比了基于信息检索的模型和基于语义解析的模型,比较了两者的优缺点以及不同因素对两种模型的影响。我们得出结论,基于信息检索的模型注重于学习关系之间的差异性,以至于当新的关系输入时,我们可以根据其差异性预测其与问题之间的语义相似度;基于语义解析的模型注重于刻画数据本身的相似度,在推理任务中有着很大的发挥空间。

在未来的工作中,我们可以尝试使用更加有效的语义解析模型,以提高生成动作序列效果,并使用规模更大的数据集来减少数据数量对模型学习程度的影响,并发掘语义解析模型在推理任务中的作用。

参考文献

- [1] 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. AcM, 2008: 1247-1250.
- [2] Vrandečić D, Krötzsch M. Wikidata: a free collaborative knowledge base[J]. 2014.
- [3] Auer S, Bizer C, Kobilarov G, et al. Dbpedia: A nucleus for a web of open data[J]. The semantic web, 2007: 722-735.
- [4] Suchanek F M, Kasneci G, Weikum G. Yago: a core of semantic knowledge[C]//Proceedings of the 16th international conference on World Wide Web. ACM, 2007: 697-706.
- [5] Mihalcea R, Csomai A. Wikify!: linking documents to encyclopedic knowledge[C]//Proceedings of the sixteenth ACM conference on Conference on information and knowledge management. ACM, 2007: 233-242.
- [6] Cucerzan S. Large-scale named entity disambiguation based on Wikipedia data[J]. 2007.
- [7] Milne D, Witten I H. Learning to link with wikipedia[C]//Proceedings of the 17th ACM conference on Information and knowledge management. ACM, 2008: 509-518.
- [8] Ferragina P, Scaiella U. Tagme: on-the-fly annotation of short text fragments (by wikipedia entities)[C]//Proceedings of the 19th ACM international conference on Information and knowledge management. ACM, 2010: 1625-1628.
- [9] Page L, Brin S, Motwani R, et al. The PageRank citation ranking: Bringing order to the web[R]. Stanford InfoLab, 1999.
- [10] Meij E, Weerkamp W, De Rijke M. Adding semantics to microblog posts[C]//Proceedings of the fifth ACM international conference on Web search and data mining. ACM, 2012: 563-572.
- [11] Carmel D, Chang M W, Gabrilovich E, et al. ERD'14: Entity recognition and disambiguation challenge[C]//ACM SIGIR Forum. ACM, 2014, 48(2): 63-77.
- [12] Cornolti M, Ferragina P, Ciaramita M, et al. The SMAPH system for query entity recognition and disambiguation[C]//Proceedings of the first international workshop on Entity recognition & disambiguation. ACM, 2014: 25-30.
- [13] Cornolti M, Ferragina P, Ciaramita M, et al. A piggyback system for joint entity

- mention detection and linking in web queries[C]//Proceedings of the 25th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2016: 567-578.
- [14] Blanco R, Ottaviano G, Meij E. Fast and space-efficient entity linking for queries[C]//Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. ACM, 2015: 179-188.
- [15] Yao X, Van Durme B. Information Extraction over Structured Data: Question Answering with Freebase[C]//ACL (1). 2014: 956-966.
- [16] Bast H, Haussmann E. More accurate question answering on freebase[C]//Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. ACM, 2015: 1431-1440.
- [17] Yih S W, Chang M W, He X, et al. Semantic parsing via staged query graph generation: Question answering with knowledge base[J]. 2015.
- [18] Xu K, Reddy S, Feng Y, et al. Question Answering on Freebase via Relation Extraction and Textual Evidence[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016, 1: 2326-2336.
- [19] Dai Z, Li L, Xu W. CFO: Conditional Focused Neural Question Answering with Large-scale Knowledge Bases[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016, 1: 800-810.
- [20] Yin W, Yu M, Xiang B, et al. Simple Question Answering by Attentive Convolutional Neural Network[C]//Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. 2016: 1746-1756.
- [21] He X, Golub D. Character-Level Question Answering with Attention[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016: 1598-1607.
- [22] Fader A, Zettlemoyer L, Etzioni O. Paraphrase-driven learning for open question answering[C]//Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2013, 1: 1608-1618.
- [23] Bordes A, Usunier N, Garcia-Duran A, et al. Translating embeddings for modeling multi-relational data[C]//Advances in neural information processing systems. 2013: 2787-2795.
- [24] Liang C, Berant J, Le Q, et al. Neural symbolic machines: Learning semantic parsers on freebase with weak supervision[C]//55th Annual Meeting of the Association for Computational Linguistics, ACL 2017. Association for

- Computational Linguistics (ACL), 2017: 23-33.
- [25] 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. 2013: 1533-1544.
- [26] 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.
- [27] Sorokin D, Gurevych I. Modeling Semantics with Gated Graph Neural Networks for Knowledge Base Question Answering[C]//Proceedings of the 27th International Conference on Computational Linguistics. 2018: 3306-3317.
- [28] Talmor A, Berant J. The Web as a Knowledge-Base for Answering Complex Questions[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). 2018: 641-651.
- [29] Dong L, Lapata M. Language to Logical Form with Neural Attention[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016, 1: 33-43.
- [30] Xu K, Reddy S, Feng Y, et al. Question Answering on Freebase via Relation Extraction and Textual Evidence[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016, 1: 2326-2336.
- [31] Yu M, Yin W, 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 (Volume 1: Long Papers). 2017: 571-581.
- [32] Hao Y, Zhang Y, 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 (Volume 1: Long Papers). 2017: 221-231.
- [33] 刘康,张元哲,纪国良,等.基于表示学习的知识库问答研究进展与展望[J].自动化学报,2016,42(6):807-818. DOI:10.16383/j.aas.2016.c150674.
- [34] Dong L, Lapata M. Coarse-to-Fine Decoding for Neural Semantic Parsing[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2018: 731-742.
- [35] Hu S, Zou L, Zhang X. A State-transition Framework to Answer Complex Questions over Knowledge Base[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018: 2098-2108.

- [36] Luo K, Lin F, Luo X, et al. Knowledge Base Question Answering via Encoding of Complex Query Graphs[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018: 2185-2194.
- [37] Guo D, Tang D, Duan N, et al. Dialog-to-Action: Conversational Question Answering Over a Large-Scale Knowledge Base[C]//Advances in Neural Information Processing Systems. 2018: 2942-2951.
- [38] Bordes A, Usunier N, Chopra S, et al. Large-scale simple question answering with memory networks[J]. arXiv preprint arXiv:1506.02075, 2015.
- [39] 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 (Volume 2: Short Papers). 2016, 2: 201-206.
- [40] Yang Y, Chang M W. S-MART: Novel Tree-based Structured Learning Algorithms Applied to Tweet Entity Linking[C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 2015, 1: 504-513.
- [41] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [42] Tai K S, Socher R, Manning C D. Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks[C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 2015, 1: 1556-1566.
- [43] Pennington J, Socher R, Manning C. Glove: Global vectors for word representation[C]//Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). 2014: 1532-1543.
- [44] Manning C, Surdeanu M, Bauer J, et al. The Stanford CoreNLP natural language processing toolkit[C]//Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations. 2014: 55-60.
- [45] Duchi J, Hazan E, Singer Y. Adaptive subgradient methods for online learning and stochastic optimization[J]. Journal of Machine Learning Research, 2011, 12(Jul): 2121-2159.
- [46] See A, Liu P J, Manning C D. Get To The Point: Summarization with Pointer-Generator Networks[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017:

1073-1083.

哈尔滨工业大学学位论文原创性声明和使用权限

学位论文原创性声明

本人郑重声明：此处所提交的学位论文《面向知识库的知识问答技术研究》，是本人在导师指导下，在哈尔滨工业大学攻读学位期间独立进行研究工作所取得的成果，且学位论文中除已标注引用文献的部分外不包含他人完成或已发表的研究成果。对本学位论文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。

作者签名：施琦

日期：2019年6月25日

学位论文使用权限

学位论文是研究生在哈尔滨工业大学攻读学位期间完成的成果，知识产权归属哈尔滨工业大学。学位论文的使用权限如下：

(1) 学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文，并向国家图书馆报送学位论文；(2) 学校可以将学位论文部分或全部内容编入有关数据库进行检索和提供相应阅览服务；(3) 研究生毕业后发表与此学位论文研究成果相关的学术论文和其他成果时，应征得导师同意，且第一署名单位为哈尔滨工业大学。

保密论文在保密期内遵守有关保密规定，解密后适用于此使用权限规定。

本人知悉学位论文的使用权限，并将遵守有关规定。

作者签名：施琦

日期：2019年6月25日

导师签名：张

日期：2019年6月25日

致 谢

时光荏苒，转眼在哈尔滨工业大学社会计算与信息检索研究中心度过了美好的两年时光。在这两年中，我学到了很多知识，认识了很多老师，交到了很多朋友，也从一个本科毕业生成长成为了一名准博士生。在这里，我要感谢所有的老师、同学，感谢这宝贵的两年中出现在我生活中的朋友们，让我有了一段难忘的经历。

感谢我的导师张宇教授。您学识渊博、视野广阔。这两年多的时间内，张老师在学习和生活上给予了我无微不至的帮助，在此衷心的感谢张老师。

感谢刘挺教授带我进入了赛尔大家庭，让我有机会认识很多如此优秀的师兄师姐师妹们，并同意让我在硕士阶段后读了您的博士，让我在人生的道路上迈出了至关重要的一步。

感谢秦老师在两年内对我的帮助，虽然不在您的组里，但仍感受到了您各方面的关怀，在此衷心的感谢您。感谢车万翔老师、赵妍妍老师、张伟男老师、丁效老师、刘铭老师、冯骁骋老师，感谢各位老师对实验室的付出，让我们可以安心的学习、科研，不断进步！

感谢 QA 组的所有同学们，两年来大家一起学习、一起工作，共同进步。大家在一起的时间充实且快乐。尤其感谢尹庆宇师兄在两年多的时间内对我的帮助，带我从小白慢慢的成长，如今师兄也即将毕业，在这里祝庆宇师兄在今后的道路上万事如意，越来越顺利。感谢我的室友们，两年时间以来对我学术和生活上的帮助。感谢女朋友孙卓的陪伴，让我在繁忙的科研生活中仍享受到了温馨的时光。

最后，要感谢我的父母，感谢父母这么多年来对我在精神上和经济上的支持。这么多年来，有父母在身后，我才可以追随自己的想法，毫无顾忌的求学、做自己想做的事情，愿父母身体健康、万事如意！