

◎热点与综述◎

基于知识图谱的智能问答研究综述

王智悦¹, 于清¹, 王楠¹, 王耀国²

1. 新疆大学 信息科学与工程学院, 乌鲁木齐 830046

2. 新疆维吾尔自治区人民医院 信息中心, 乌鲁木齐 830001

摘要:基于知识图谱的问答是近年来研究热点,从基于模板、语义解析、深度学习、知识图谱嵌入四方面介绍基于知识图谱智能问答实现,归纳了各类方法的优缺点,及尚未解决的关键问题。结合当前人工智能技术发展,重点介绍了基于深度学习的智能问答,有助于更多研究者投身于智能问答研究,根据不同行业需求研发适用于不同领域的问答系统,提高社会智能化信息服务水平。

关键词:智能问答;知识图谱;语义解析;深度学习

文献标志码:A **中图分类号:**TP391 **doi:**10.3778/j.issn.1002-8331.2004-0370

王智悦,于清,王楠,等.基于知识图谱的智能问答研究综述.计算机工程与应用,2020,56(23):1-11.

WANG Zhiyue, YU Qing, WANG Nan, et al. Survey of intelligent question answering research based on knowledge graph. Computer Engineering and Applications, 2020, 56(23):1-11.

Survey of Intelligent Question Answering Research Based on Knowledge Graph

WANG Zhiyue¹, YU Qing¹, WANG Nan¹, WANG Yaoguo²

1. Academy of Information Science and Engineering, Xinjiang University, Urumqi 830046, China

2. Information Center, People's Hospital of Xinjiang Autonomous Region, Urumqi 830001, China

Abstract: The answer selection model based on knowledge graph has become one of the hottest directions at present. This paper introduces the implementation of answer selection model based on knowledge graph from four aspects of template method, semantic parsing, deep learning and knowledge graph embedding, sums up their advantages, disadvantages and unsolved problem. Combined with the development of artificial intelligence technology, this paper introduces intelligent question-answer system based on deep learning. This research is helpful for more researchers to devote themselves to the intelligent question-answer system and develops different kinds of intelligent question-answer system to improve the social intelligent information service.

Key words: intelligent questions and answers; knowledge graph; semantic analysis; deep learning

1 引言

智能问答是自然语言处理中非常重要的研究内容,具体指计算机通过对人类语言的自动分析,回复用户所询问的问题。为了更精准回答用户提问,现通常采用对问句进行深层次语义分析,获取问句丰富内涵信息,再反馈接近用户需要的答案,由此体现计算机智能性。

知识图谱又称科学知识图谱,在图书情报界叫做知识域可视化或知识领域映射地图,用可视化技术描述知识资源及其载体。知识图谱构建需要挖掘、分析、绘制和显示知识及知识之间的相互关系,让原本知识库更具智能。结合知识图谱智能问答,通过知识图谱中实体间存在的直接联系,挖掘推理出潜在关系,与传统搜索引

基金项目:国家自然科学基金新疆联合基金(No.U1603262);国家级大学生创新训练项目(No.20180755075X);自治区大学生创新训练项目(No.S20190755047)。

作者简介:王智悦(1995—),男,硕士,研究领域为自然语言处理、智能问答;于清(1973—),通信作者,女,硕士,副教授,CCF会员,研究领域为自然语言处理、机器翻译、智能问答,E-mail:yuqing0131@126.com。

收稿日期:2020-04-26 **修回日期:**2020-07-23 **文章编号:**1002-8331(2020)23-0001-11

CNKI网络出版:2020-11-02, <https://kns.cnki.net/kcms/detail/11.2127.TP.20201102.0855.002.html>

擎比较,基于知识图谱的信息检索,不再是反馈简单排序的文档结果,而是通过智能语义分析,反馈接近用户需要的问题答案。本文综述近年来基于知识图谱的智能问答技术研究与发展状况,为更多研究者提供信息参考。

1.1 智能问答发展

智能问答最早追溯到人工智能诞生时期,阿兰·图灵^[1]1950年提出通过观察机器是否具备正确回答问题的能力,从而验证机器是否具有智能。不久后,麻省理工学院Weizenbaum在1966年设计了名为ELIZA的聊天机器人^[2],实现了与人简单交流,不过,并不意味着机器对自然语言真正理解。随后又有大量研究成果如:Colby设计的Parry^[3]、ALICE^[4]、Jabberwacky^[5];2011年IBM公司设计研发了超级计算机“沃森”,“沃森”在美国知识竞赛节目《危险边缘》中上演了“人机大战”,战胜两位顶尖人类选手,被视为人工智能发展又一里程碑;华盛顿大学Etzioni教授2011年在Nature上发表文章Search needs a shake-up指出:“以直接而准确的方式回答用户自然语言提问的自动问答系统将构成下一代搜索引擎的基本形态”^[6]。因此,问答系统被看作未来信息智能服务关键性技术之一,是人机交互的重要手段。

1.2 知识图谱发展

知识图谱于2012年由谷歌公司首次提出,具体是将独立的知识转化为三元组形式,形成庞大的知识网络。知识图谱属于人工智能重要研究领域知识工程研究范畴,是利用知识工程建立大规模知识资源的重要应用之一。知识图谱可以理解作为一种语义网络^[7]结构的知识库。在问答系统^[8]中,自然语言通过语义网络来表达和存储十分的方便。21世纪后,语义网络进入新的应用场景即语义Web^[9],这种允许上传图结构(W3C的标准RDF),实现数据之间链接就是知识图谱的雏形^[10],并且知识图谱又不同于早期的语义网络,知识图谱更加强调实体之间的关系或者实体的属性值。知识图谱研究的内容包括知识获取、知识融合、知识计算与应用三大主要步骤^[11]。最具代表性的大规模网络知识图谱包括DBpedia^[12]、Freebase^[13]、KnowItAll^[14]、WikiTaxonomy^[15]和YAGO^[16]、以及BabelNet^[17]、ConceptNet^[18]、DeepDive^[19]、NELL^[20]、Probase^[21]、Wikidata^[22]、XLORE^[23]、Zhishi.me^[24]等。这些知识图谱遵循RDF三元组数据结构,包含数千万级或者亿级规模实体,以及数十亿或百亿事实(即属性值和与其他实体的关系),这些实体被组织在成千上万由语义体现的客观世界概念结构中。图1展示了一个简单的知识图谱。

1.3 相关工作

近年来,伴随用户对智能应用方面的强大需求,许多公司及机构如Google、百度、维基等对获得的高质量

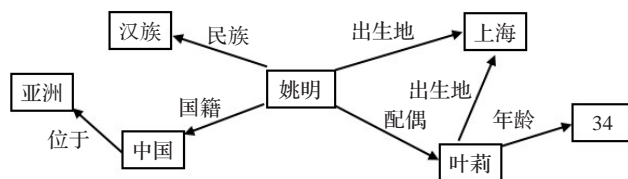


图1 简单的知识图谱

数据,采用自动或半自动化方法设计了一系列完备的知识图谱。例如DBpedia^[12]、Freebase^[13]、YAGO^[16]等,这类图谱由大批量实体、关系及属性构成。同时机器学习与深度学习发展,为智能问答奠定研究基础。例如Baseball^[25]、Lunar^[26]实现了限定域的智能问答,Paralex^[27]、SEMPRE^[28]、ParaSEMPRE^[29]、STAG^[30]尝试更具挑战性的开放域智能问答研究。

大量文献围绕研究智能问答展开,如:郑实福等人^[31]对早期自动问答做了比较全面的综述,另有研究者针对限定域问答进行全面总结^[32],还有针对Web的智能问答,以及研究自然语言处理技术自动问答实现等^[33]。然而,基于知识图谱的智能问答还没有叙述全面的综述类文章,在此对基于知识图谱的智能问答做全面介绍,吸引更多研究者投身该领域研究。

同时,问答系统实现离不开数据集,重点介绍当前较成熟的英文问答数据集有Simplequestion^[34]、Webquestions^[28]、WebquestionsSP^[35]、QALD^[36]、CSQA^[37]等,如表1。Simplequestions是一种大规模问答数据集,这个数据集中的问题可以用一个三元组进行回答,并且数据集包含查询语句。该数据集包含了10万多条问答对,但是简单问题占很大比例,简单问题即由一条三元组就可以回答的问题,也叫single-relation问题,在该数据集上大部分研究者采用向量建模与深度学习相结合的方法并取得了很好的效果。Webquestions是Berant等人借助Google Suggest生成了5 810条问答对,提供了每个答案对应知识库的主题节点。Webquestions数据集更加偏向自然语言,但是同样存在缺点,只是提供了答案而没有对应的查询语句,这就对生成逻辑表达式的模型训练造成了困难,其次是复杂问句较少。Yih等人对Webquestions改进后提出WebquestionsSP, WebquestionsSP是Webquestions的子集,补全了相应的查询语句。QALD是一种开放域问答数据集且表达更加复杂、更加口语化,除了包含问句与答案之外, QALD还为每个问答对配置了关键词以及对应的SPARQL语句。CSQA主要是用于知识图谱序列问答(多轮问答)的数据集,其中每个数据样本均由具有共享上下文的一系列QA对组成。在这些数据集中,尽管序列中的单个问题通常很短,但具有上下文相关性。中文数据集有NLPCC评测(NLPCC2016评测数据包含14 609个问答对的训练集和包含9 870个问答对的测试集。并提供一个知识库,

表1 知识图谱问答数据集

数据集	KG	规模	简单问题	复杂问题
Free917(Cai and Yates,2013)	Freebase	917	YES	YES
WebQuestions(Berant et al.,2013)	Freebase	5 810	NO	YES
SimpleQuestion(Bordes et al.,2015)	Freebase	100 000	YES	NO
WebQuestionsSP(Yih et al.,2016)	Freebase	4 737	YES	YES
ComplexQuestions(Bao et al.,2016)	Freebase	2 100	YES	YES
GraphQuestions(Su et al.,2016)	Freebase	5 166	YES	YES
30M Factoid Questions(Serban et al.,2016)	Freebase	30 000 000	YES	NO
LC-QuAD(Trivedi et al.,2017)	DBpedia	50~500	YES	YES
CSQA(Saha et al.,2018)	Wikidata	1 600 000	YES	YES
QALD-9(Usbeck et al.,2018)	DBpedia	50~500	YES	YES
LC-QuAD 2(Dubey et al.,2019)	DBpedia,Wikidata	30 000	YES	YES

包含 6 502 738 个实体、587 875 个属性以及 43 063 796 个三元组)以及 CCKS 评测(CCKS2018 评测数据集包括 2 298 条问答对,其中有约 1 000 条金融领域问答对,并提供三元组数据)等,随着智能技术发展,研究者们继续尝试构建更高标准的数据集,Liu 等人构建了一个开源的跨语言 OpenQA 数据集 XQA^[38],包含英语训练集以及英语、法语、德语、葡萄牙语、波兰语、中文、俄语、乌克兰语和泰米尔语验证集和测试集。Xiong 等人^[39]收集新闻文本,构建了社交问答数据集。

2 构建模板的问答方法

构建模板的问答方法通过构造一组模板参数,形成查询表达式,对问题文本进行匹配。整个过程不涉及问句分析,通过预设查询模板替代相关实体关系映射。优点:简洁、准确性高,巧妙避开语义解析等难题,适用于简单查询,在实际中应用广泛。

AskJeeves 是国外使用模板库较成功的例子,其中问题理解部分由几百名专家对相关问题构建问题模板手工完成,耗费大量人力,借助模板库返回相关文档链接或者一段文本,回答用户问题,回答方式多样。缺点是后期维护依然需要人工,工作量巨大。

2010 年 TrueKnowledge^[40]模板问答方法被提出,核心理念是:首先使用已知模板成分匹配句子中的内容,例如一些疑问词,其次按照顺序匹配相应模板,一个模板可以覆盖多个问题。例如问题“*What is the capital of France?*”首先匹配疑问词,如 What\Which,这些反映问题意图的疑问词又生成“*what\which a y*”,其中 a 与 y 是问题中内容的映射,确定好这样的待填充结构后,最后继续根据已定义模板将“*is the capital of*”映射到变量 a,将“*France*”映射到变量 y。模板针对不同问题可以复用,但缺点也十分明显,需要大量人工处理形成模板,成本昂贵。同时,生成的模板对数据库依赖性大。

文献[41]构建了基于汽车领域的智能问答系统,针对原有问答模板方法进行了改进,结合汽车领域知识库

提出模板自动生成方法,提前将复杂自然语言处理,预备在模板库构建中取得较好效果。同时,该文献还介绍了模板库自动生成系统结构以及相关开发工具。

Cui 等人^[42]针对简单事实问答,在大规模模板自动化生成方面,提出优化方案。Abujabal 等人^[43]提出 QUINT 模型,通过语料自动学习模板,借助生成的模板将自然语言问句转化为知识库查询。Cocco 等人^[44]提出基于面向对象的问答系统,借助 RDF 形式的 LinkedSpeding 数据集,在现有训练集(相互配对的问答对)上,通过机器学习方法学习 SPARQL 模板。

总而言之,基于模板的问答方法属于比较传统的方法,设计者需要提前设定模板,依据问题相关部分,选取不同模板,生成答案。这种方法优点是:可以获得比较准确的答案,回答响应速度快。缺点:需要耗费大量人力进行模板校对,以及模板库维护。但是,针对问答领域多跳复杂问题,最新的模板方法也能提供解决思路,当前该方法研究重点更侧重于模板自动生成,克服耗时耗力难题。

3 语义解析的问答方法

语义解析与基于模板的方法最明显的区别是逻辑表达式。模板方法需要预设固定表达方式,语义解析方法关键在于对自然语言问句成分进行解析,将查询转化成逻辑表达式,再利用知识图谱的语义信息将逻辑表达式转换成知识图谱查询,最终得到相应结果。逻辑表达式用于面向知识图谱的结构化查询,查找知识库中的实体以及与实体相关的知识,在结构化形式的知识图谱上进行查询,最高效的方法是利用结构化查询语句,类似 SQL、SPARQL 语句等,然而对普通用户来说,设计规范的查询语句存在困难。基于知识图谱的语义解析问答系统实现,需要两个关键步骤:(1)使用语义解析器将问题转换成机器能够理解和运行的语义表示;(2)使用该语义产生结构化查询语言,对知识图谱进行查询,并从返回的实体集合中寻找答案,如图 2 所示。

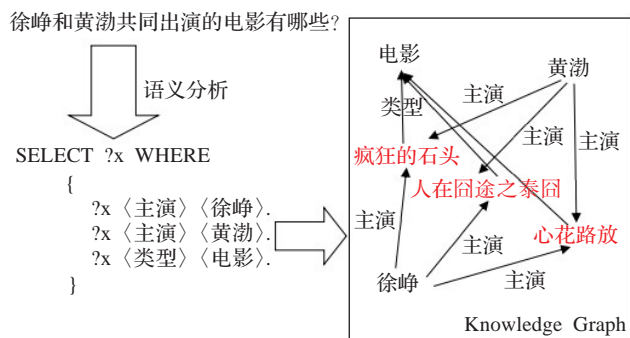


图2 基于语义解析的知识图谱问答过程

常用语义解析方法又分为三类:基于词典-文法的语义解析、基于语义图构建的语义解析、基于神经网络的语义解析。

3.1 基于词典-文法的语义解析方法

该方法由 Berant^[28]提出,它依靠组合法文法,其中组合范畴文法(CCG)^[45]由 ACL 终身成就奖获得者 Steedman 提出,后由 Zettlemoyer 与 Collins^[46]应用于对问句成分进行语义解析。Kwiatkowski 等人提出改进规则方法^[47],有效提升了 CCG 句法分析器鲁棒性;Artzi 等人^[48]使用 AMR (Abstract Meaning Representation) 替换 λ -算子,表示 CCG 语义规则中语义类型部分,同样取得不错效果。语义解析方法还有很多,例如同步上下文无关文法(SCFG)^[49],以及组合语义法(DCS)^[28],DCS 拥有和 λ -算子相似的表示能力,但该方法定义的语义结构更贴近知识图谱的存储结构。

基于词典-文法的语义解析方法可解释性很强,结构清晰,在限定领域问答方面达到很好效果。但是很多重要组成部分(比如 CCG 中的词汇表和规则集)都需人工编写。面对大规模多源异构知识库,该方法存在以下不足:(1)资源(例如词汇表、规则集)标注费时费力,在训练数据有限情况下,性能大打折扣;(2)语义表示与知识库联系不紧密,无法在解析过程中利用知识约束;(3)大规模知识库开放域特性使文本歧义问题严重。

3.2 基于语义图的问句解析方法

Reddy^[50]提出与基于词典-文法的语义解析方法不同的新方法。通过对问句分析,构建相对应的语义图,语义图由节点(实体、变量或者类型)、边(关系)、操作符(count、argmax 等)构成,被看作知识图谱子图,实现将问句映射到知识图谱中,再通过图匹配完成问题回答,该方法即是基于语义图的语义解析方法。较有代表性的是 Reddy 等人^[51]设计的从依存树到语义图的转换文法,将树结构转化为知识图谱的子图结构再结合知识图谱进行匹配;Yih^[52]提出分步骤构建文法,便于理解,可解释性强,提高了语义图构建的准确性;Bast 等人^[53]从另一方面构建了一种基于模板的语义图生成方法。构

建语义图的方法不仅针对简单问题,Hao 等人^[54]还尝试构建复杂语义图解决复杂问题,效果明显,尤其问句中存在多个实体时,更便于语义图构建。

基于语义图的问句解析方法关键技术在于语义图表示,以及语义图构建。语义图结构与自然语言句子结构具有相似性,故使用语义图作为目标语义表示有显著优点,语义解析过程充分利用知识库的知识约束,由组合法文法转换为语义图构建,减少了搜索空间。但是基于语义图的语义解析方法也存在问题,该方法依赖于一些启发式方法构建语义图,导致缺乏通用性。

3.3 基于神经网络的方法

该方法将自然语言及对应的语义看作是两种不同语言,语义分析任务被看做类似于机器翻译任务,利用端到端模型,实现将问句翻译成对应语义的表示序列。

Dong 等^[55]和 Jia 等^[56]用基于 attention 机制的 encoder-decoder 模型(如图3),将自然语言作为输入,得到适用于机器处理的逻辑表示作为输出。Xiao 等人^[57]结合符号先验知识,利用 RNN 模型实现语义解析。Chen 等人^[58]提出 Sequence-to-Action 模型,用语义图表示语义信息。也有研究者在基于神经网络的方法上进行扩展。Krishnamurthy 等人^[59]考虑到目标语言是形式化语言,需要严格条件约束,对比机器翻译的 decoder,语义解析中使用了严格约束条件。Dong 等人^[60]提出用两级 encoder->decoder 改进机器翻译端到端模型,有效解决自然语言与语义表示之间跨度大问题,通过问题分解,提升性能。Chen 等人^[61]不仅考虑从问句中提取相关特征做语义分析,同时结合知识图谱中实体类型、关系路径、上下文关系等,借助注意力机制方法,在 Webquestions 数据集上实验,准确性提高。Lukovnikov 等人^[62]设计了一种端到端基于字符级别的问题编码器,可以处理词库以外的文字问题,捕获文字水平的语义。

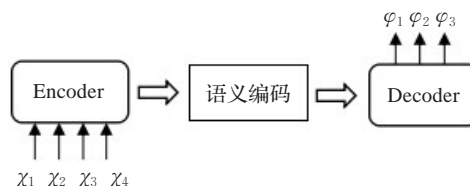


图3 encoder-decoder 模型示意图

基于神经网络方法,相比其他两种方法,模型简单,但是可解释性差,还需要预先准备相应的训练语料,对模型进行训练时,训练过程较长,还需要调整参数。对基于语义解析的三类方法,进一步归纳总结,得到如表2。

综上所述,基于语义解析的知识图谱问答方法,核心任务是将自然语言转化成机器能够理解和执行的语义表示。通常基于符号逻辑的语义表示,缺乏灵活性,在分析问句语义过程中,还容易受到符号间语义鸿沟影

表2 语义解析方法比较

方法	核心	优势	劣势
基于词典-组合法的方法	词典 组合法	基于组合语义思想,解析过程清晰可见,可解释性强	需要学习词典和定义组合法,且容易受限于词典的覆盖度
基于语义图的方法	语义图表示 语义图构建	语义图与自然语言句子具有类似的结构;与知识库联系紧密,可充分利用知识库的知识来指导语义图的构建	往往依赖于特定的手段来构建语义图,缺乏一定的通用性
基于神经网络的方法	Decoding 中的 token 形式 Decoding 中的 constraints	端到端,充分利用神经网络模型的表示能力和拟合能力	可解释性差

响,同时从自然语言问句得到结构化语义表示需要许多步操作,多步误差传递对问答准确度造成影响。基于神经网络的方法,需要大量标注数据,然而带有标注的数据集非常有限,加上现有知识图谱覆盖度较低。因此,目前基于语义分析的问答系统在开放域上取得的效果还不近人意。由于基于符号的方法和基于神经网络的方法各具优缺点,所以今后可以考虑两种方法结合,此外,无论是基于符号方法需要构建规则集,还是基于神经网络的方法需要标注语料,都耗费大量人工,因此构建低成本的模型也是研究方向之一。

2006 年 Hinton 等人^[63]在神经网络模型基础上提出深度学习,如今海量数据积累、计算力提升、算法模型改进,促成深度学习迅猛发展。深度学习方法也已应用于问答系统各项任务实现。例如实体识别采用的经典 BILSTM+CRF 方法,如图 4 所示,谷歌提出 BERT^[64]模型;关系分类、意图分类,采用基于字级别的深度学习方法;实体消歧采用深度学习排序方法,判断语义匹配性等。

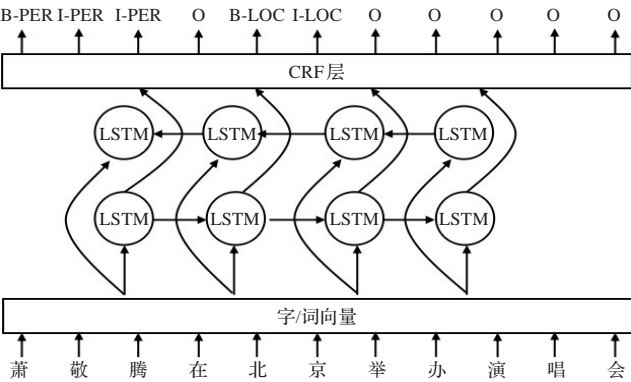


图4 BILSTM+CRF 实体识别模型示意图

当今,在基于知识图谱的问答研究领域,大部分研究者把目光从传统或规则的方法转向深度学习方法。以下章节将重点介绍基于深度学习的知识图谱问答。

4 基于深度学习的答案排序方法

采用深度学习的方法需要将问题以及知识图谱中包含的丰富的语义信息(字、词语、上下文关系,知识图谱中的实体、关系以及属性),投射到一个高维向量空

间,获得字向量或者词向量,通过深度学习模型对向量进行相似度计算,再通过相应打分机制获得候选排序,得出最终问答结果。如图 5 所示。

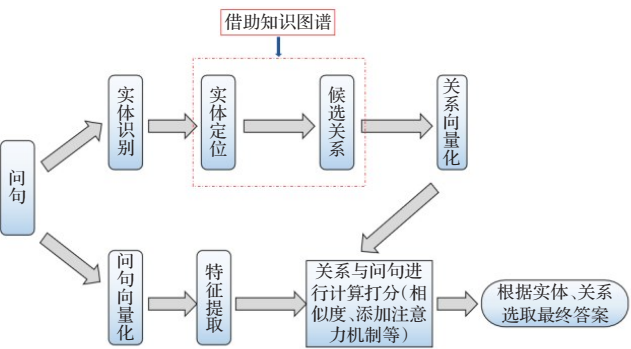


图5 基于深度学习的答案排序方法典型流程

其中,Bordes^[65]在文献[66]基础上提出的方法比较经典,首先对问句中实体准确定位,将问句中的实体抽取出来,再将实体连接到知识库,以该实体为起点,查找与其关系相连的实体作为候选答案。其次计算这些实体关系的组合与问句之间的相似度,通过打分排序,选择相似度最大的候选项返回答案,取得很好效果。Hao 等人^[67]针对前人工作中没有充分考虑候选答案相关信息训练 question representation,提出 Cross-Attention 机制的神经网络模型,训练知识库全局信息,一定程度上减轻了 the Out of Vocabulary (OOV) 问题,在 Web-questions 数据集上取得不错效果。Hao 等人^[68]在问句实体链接方面提出改进方案,同时利用多级编码及多维信息增强效果。Zhang 等人^[69]在原有 APVA 模型上改进,提高关系预测准确性。此外,也有学者尝试将问句实体链接与关系链接合并抽取^[70]。

Bordes 等人^[71]还设计了一种基于记忆网络的问答系统,在多个数据集上验证了系统设计的优越性,具备迁移学习能力。Dong 等人^[72]设计了一种多列卷积神经网络,根据知识图谱特点,定义答案路径、上下文路径、答案类型,作为参考特征,并且每个特征对应一个已经训练的卷积神经网络,用于捕获问句中语义信息,再通过计算问句与答案之间的相似度,对结果打分排序,排名第一的作为最终答案。为更好捕获问句与答案间的交互信息,以及保留更多原始信息,Qu 等人^[73]提出了一

种基于相似矩阵的递归神经网络(AR-SMCNN)模型,利用RNN顺序建模特性捕获语义级相关性,使用注意机制跟踪实体和关系重要部分,并制定了一种准确确定问句中主实体的方法。Lan等人^[74]创新地将“matching-aggregation”框架用于匹配候选项,增强候选实体表示,充分利用问题特有语境关系,提升在公开数据集Webquestions、Simplequestion上的效果。Krishna等人^[75]将问答与实际生活联系起来,将问答过程比作老师与学生之间的关系,学生问题宽泛,同时使用分类方法,由大到小,由粗到细,定位问题意图,最终由老师回答。

以上是国外学者在研究基于知识图谱问答方面,采用深度学习方法取得的成果。国内也有许多研究者,致力于构建领域知识图谱及设计问答系统,将基于知识图谱的单一事实问答拆解成两部分任务,一是实体链接,二是关系识别。罗达等人^[76]按照这一思路提出了一种基于多角度注意力机制的单一事实知识库问答方法,关系识别率达到93.5%。曹明宇等人^[77]构建了一种针对原发性肝癌的知识图谱问答系统。张楚婷^[78]在旅游方面展开知识图谱问答系统研究。张崇宇^[79]设计了关于医疗领域的知识图谱问答系统。史梦飞^[80]设计了一种分布式的问答系统,将问题句进行分类,提高下游任务的准确性,通过构建基于深度学习的End2End问答模型,同时考虑中文问句的复杂性,提出结合语义依存分析的剪枝算法及自动化模板的方法。童培豪^[81]针对复杂问题进行优化,将系统分为三个步骤,实体识别、关系发现、系统整合,并在每一个步骤中实现优化。国内众多评测会议,进一步促进问答系统发展。Lai等人^[82]在NLPPCC2017比赛中设计了一种深度卷积神经网络实体谓词重排序配对方法,取得第一名。Zhou等人^[83]将规则与神经网络相结合,在2019年CCKS评测中,取得第一名。

在基于深度学习答案排序的方法中,计算输入问题和候选答案实体之间的相关性是核心任务。当前采用问题与答案对直接训练的问答模型取得较好效果。Costa等人^[84]针对当前多领域多知识图谱发展,提出针对开放域的基于知识图谱的问答框架,尝试设计一个高效且高质量的问答系统,实现在任何领域对任何类型问题都有效的问答系统。

5 基于知识图谱的嵌入学习的问答方法

知识图谱嵌入(Knowledge Graph Embedding, KGE)学习是针对知识图谱中的三元组做word embedding操作,在一般的自然语言处理任务中,将文本的表述转化为word2vec这种产物,KGE不同于普通的词向量表示,知识图谱嵌入学习面向知识库中的实体和关系进行表

示学习,从而更为直接地构造实体与关系之间的语义相关性。相比于普通的词向量,KGE更能反映相近的实体以及关系间的相似程度,KGE的目的是为了保存知识图谱中原有的信息,并将知识图谱用连续向量空间中的低维密集向量或者矩阵来表示,可以有效解决数据稀疏问题,使知识获取、融合、推理的性能得到显著提升。现有研究已经提出了许多的KGE,如线性映射方法DistMult^[85]、基于平移的TransE^[86]、基于张量因子化的RESCAL^[87]、神经张量网络NTN^[88]等,并被证明在知识图谱补全、问答系统和关系抽取等许多应用中都是有效的。

Wang等人^[89]借助知识图谱的嵌入学习,提出了一种解决通过SPARQL查询没有答案的方法。Wang针对一些在知识图谱中可以找到答案,但是通过SPARQL匹配时,要求每个指定的查询项都需要匹配,从而无法找到正确答案的问题,通过专门为SPARQL查询匹配设计的保留实体上下文的转化模型,将知识图谱三元组映射到连续的向量空间中,使得语义相似的实体在向量空间中接近。这样的模型框架可以方便地生成高质量的近似答案。

Huang等人^[90]设计了一种基于知识图谱嵌入的问答系统,针对谓词在问题中的不同表达和实体被识别后的消歧等问题,Huang等人结合知识图谱嵌入学习在问答系统的优点来解决,知识图谱嵌入学习表示图谱中的每个谓词/实体为低维向量,给定一个简单问题,目标是找到谓词嵌入空间的一个点作为谓词的表示向量,并在实体嵌入空间中找到一个点作为头实体的表示向量。对于所有知识图谱可以解答的问题,它们的谓词向量必然在谓词嵌入空间中。因此,作者设计谓词与头实体学习模型,将问题作为输入,并返回尽可能接近于问题的谓词/实体嵌入表示的向量,确定谓词与头实体,最终找到尾实体,完成问答。

Hamilton等人^[91]认为知识图谱是一张有实体和关系组成的一张图,通过学习实体低纬度嵌入表示,可以预测潜在或者缺失的边。目前知识图谱查询的难点在于处理更为复杂的逻辑查询,因为这涉及多个未发现的边、实体和属性。针对这一问题,Hamilton等人设计了一种框架,实现在不完整知识图谱上有效地对连接逻辑查询进行预测,在低维空间中对图谱节点embedding操作,并在这个embedding space中将逻辑运算符表示为学习过的几何运算(例如平移、旋转)。通过在低维embedding space中执行逻辑运算,可以预测图谱中的关系。

知识图谱问答中一种有效的方法就是将自然语言问句转化为图结构的查询,最为重要的过程就是将实体/关系映射到知识图谱的点与边,从而构造结构查询。

Wang 等人^[92]考虑到语言的灵活性以及模糊性,映射的过程存在挑战性,提出了一种新的问答框架 KemQA,尝试去解决自然语言问题的映射问题和构造结构查询,通过大量的实验,证明框架的有效性。

知识图谱的嵌入学习方法是将知识图谱中的每个对象编码到连续向量空间,这样的做法能够反映出关系密切的实体以及关系的相似程度,保存了知识图谱中的重要信息。这样的做法为问答的实现做了很好的铺垫。知识图谱嵌入式学习的问答方法具有很好的可行性以及鲁棒性。如何能让知识图谱嵌入学习学到更多更细致的知识图谱信息是这种方法的关键。

6 多跳推理的知识图谱问答

近年来,有论文指出 KGQA 任务中的简单问题已经基本被解决^[93]。从知识图谱的发展来看,当前学术界更关注解决包含多跳、组合或者需要借助推理解决的复杂问题,由于问题和知识的多样性和复杂性,知识图谱问答仍然是一项具有挑战的任务,尤其是在多跳的问答中,许多研究者在这方面进行了许多尝试。

多跳问题主要分为两类,一类是路径问题,另一类是联合问题。路径问题指的是问题中只包含一个主题实体,但是含有多个关系,这类问题的解决需要沿着图谱中的某些线路遍历才能找到问题的答案(借助一些关系和中间实体);联合问题则是包含多个主题实体,这类问题的答案可能是多个路径问题结果的交集。语义解析方法、深度学习答案排序以及知识图谱嵌入的方法不能有效地解决多条复杂问题,如图6。

Q: NBA 联赛中被称为“黑曼巴”的球员的小女儿的生日是?
A: 2006 年 5 月 2 日

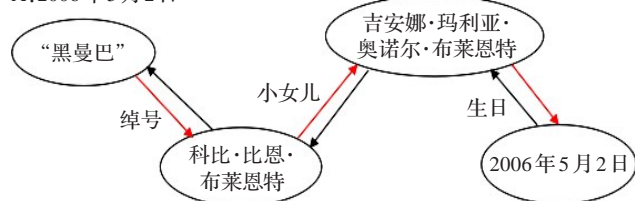


图6 多跳问题

Zhang 等人^[94]的方法是将多跳问答分为两个步骤:一是通过概率模型来识别问句中的实体(得到图谱中每个实体是问句中实体的概率);二是设计了一种逻辑推理模型,提出了一种嵌入推理图的体系结构,其中包含了所有推理规则及其复杂性,在知识图谱上做逻辑推理。Zhou 等人^[95]针对多跳提出了一种新的可解释推理网络模型,该模型采用可解释的逐跳推理过程来回答问题。该模型动态地决定输入问题的某部分应该对应图谱中的某一跳进行分析;预测与当前解析结果相对应的关系;利用预测的关系更新问题表示和推理过程的状态;然后驱动下一跳推理。Vakulenko 等人^[96]提出了一

种新的方法使用无监督消息传递(Message passing)来解决复杂问题,首先识别实体、关系和类名,并将其映射到图中的对应项,从而定位答案实体。最后,根据确定的问题类型对这些问题进行聚合。这种方法可以借助一系列稀疏矩阵乘法来模拟小型局部子图上的连接实现。

多跳推理的问答是当前的研究难点与热点问题,还有许多的工作需要去做。

7 基于知识图谱问答系统研究机遇与挑战

问答系统早就存在,从2011年Siri,到Google Now、语音助手、智能音箱等本质上都是问答系统。问答系统在日常生活中帮人们导航、搜索问题,偶尔与用户交互聊天,让生活越来越方便。智能时代职业场所及公共场所中人们每天面对大量数据、海量信息及知识更新,“懂行”的“专家助手”成为刚需。知识图谱用于描述真实世界中存在的实体和概念,以及实体和概念之间的关系,可通过人工方式构建或定义行业领域的知识体系,基于此开发各种高级应用,智能问答便是应用之一。

基于知识图谱的问答系统,研究方法很多,其中基于深度学习的答案排序法略显优势。但是依然存在不足:缺乏泛化性、可解释性,对时序性问题不能很好回答,例如:“现在谁统治丹麦?”。本章结合智能问答系统研究中遇到的问题,总结尚需解决的关键问题。

7.1 回答复杂问题

现有的基于知识图谱的问答技术,在单一问题上已取得非常好的效果,如:BERT模型实验。然而,在实际问答场景下,用户的问题往往复杂,现有的基于深度学习的知识图谱问答方法受到挑战,复杂问题研究仍有很大提升空间。一旦问题中存在多种关系与实体,或者存在潜在关系,甚至在开放性领域,如CCKS评测中的一个问题“我现在饿了,数据库里有什么吃的?”,现有的方法的效果不尽人意。

7.2 深度推理机制

现有的知识图谱规模日趋庞大,能覆盖多个领域,但是仍面临信息缺失问题,知识图谱构建是项大工程。面向问答的深度推理显示出作用,传统的基于符号的推理需要严格符号匹配,推理规则有限,以致领域适应性差,无法实现大规模推理。但是,深度学习作用于分布式语义表示,可以利用语义空间中的相似度计算,弥补符号推理需要严格规则的缺陷。所以,如何利用深度学习的大规模、可学习优点、融入传统逻辑推理规则,构建精准的大规模知识推理引擎是自动问答系统迫切需要解决的又一个热点问题。

7.3 多轮交互式问答

传统的问答系统大多采用一问一答的形式,但是现

实应用场景中,用户与机器之间需要多轮交互,由此反馈给用户的答案不只是单一的实体、概念、关系等形式,需要考虑到先前的对话对即将发生的对话的影响或者当前的对话可能对后续对话造成的影响。除此之外,还需要采用用户理解的自然语言形式回复问题。在这一过程中,如何结合知识库,将知识库问答的答案加入自然语言回复中,是亟待解决的问题。

7.4 长尾问句挑战

问句的长尾(long tail)的问题也是近年来研究的热点问题,长尾问句是指在问答数据上,少量表示相同意思,出现频率高的问句在整个数据集中占有大部分的比例。将问句以柱状图的形式展示就会有“很长的尾巴”,尾巴上都是一些出现频率不高的问句。例如在数据集中与“ehr系统包含什么功能”意思相近的问句有“ehr系统有哪些功能”“ehr系统有什么功能”“ehr系统能实现什么功能”。如果数据集大部分都是相似的问句,这就会导致数据集的特征比较单一,学习的模型泛化能力不强。这就可能需要使用一些其他的算法来增强模型的泛化能力。

8 结束语

智能时代来临,海量信息充斥在社会每个角落,用户对自动问答需求越来越强烈,现有的问答系统还处在起步阶段,仅具备最基本的智能与推理能力。人工智能专家在智能问答领域的研究从未中断过,基于知识图谱的问答系统作为自动问答研究的重要方向,其技术发展趋势从限定领域向开放领域发展,从单一数据源向多元数据源发展,从浅层语义分析向深度推理发展。本文关于知识图谱智能问答研究综述期望相关技术早突破,开拓不同行业智能问答应用需求。

参考文献:

- [1] Turing A M. Computing machinery and intelligence[J]. Mind, 1950, 59(236): 433-460.
- [2] Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine[J]. Communications of The ACM, 1966, 9(1): 36-45.
- [3] Colby K M, Weber S, Hilf F D. Artificial paranoia[J]. Artificial Intelligence, 1971, 2(1): 1-25.
- [4] Wallace R S. The anatomy of ALICE[C]// Parsing the Turing test. Berlin: Springer, 2009: 181-210.
- [5] Carpenter R. Jabbervacky[EB/OL]. [2019-05-16]. <http://www.jabbervacky.com>.
- [6] Etzioni O. Search needs a shake-up[J]. Nature, 2011, 476(7358): 25-26.
- [7] Sowa J F. Principles of semantic networks: exploration in the representation of knowledge[J]. Frame Problem in Artificial Intelligence, 1991(2/3): 135-157.
- [8] Simmons R F. Natural language question-answering systems: 1969[J]. Communications of the ACM, 1970, 13(1): 15-30.
- [9] Berners-Lee T, Hendler J, Lassila O. The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities[J]. Scientific American, 2001, 284(5): 34-43.
- [10] Auer S, Barnaghi P. Linked data—the story so far[J]. International Journal on Semantic Web and Information Systems, 2009, 5(3): 1-22.
- [11] 漆桂林, 高桓, 吴天星. 知识图谱研究进展[J]. 情报工程, 2017, 3(1): 4-25.
- [12] Bizer C, Lehmann J, Kobilarov G, et al. DBpedia—a crystallization point for the web of data[J]. Journal of Web Semantics, 2009, 7(3): 154-165.
- [13] 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, 2008: 1247-1250.
- [14] Etzioni O, Cafarella M, Downey D, et al. Web-scale information extraction in knowitall (preliminary results)[C]// the Web Conference, 2004: 100-110.
- [15] Ponzetto S P, Strube M. WikiTaxonomy: a large scale knowledge resource[C]// European Conference on Artificial Intelligence, 2008: 751-752.
- [16] Suchanek F M, Kasneci G, Weikum G. YAGO: a core of semantic knowledge[C]// Proceedings of the 16th International Conference on World Wide Web, 2007: 697-706.
- [17] Navigli R, Ponzetto S P. BabelNet: building a very large multilingual semantic network[C]// Annual Meeting of the Association for Computational Linguistics, 2010: 216-225.
- [18] Speer R, Havasi C. Representing general relational knowledge in ConceptNet 5[C]// International Conference on Language Resources and Evaluation, 2012: 3679-3686.
- [19] Niu F, Zhang C, Re C, et al. DeepDive: web-scale knowledge-base construction using statistical learning and inference[C]// VLDS, 2012: 25-28.
- [20] Carlson A, Betteridge J, Kisiel B, et al. Toward an architecture for never-ending language learning[C]// International Conference on Artificial Intelligence, 2010: 1306-1313.
- [21] Wu W, Li H, Wang H, et al. Probase: a probabilistic taxonomy for text understanding[C]// International Conference on Management of Data, 2012: 481-492.
- [22] Vrandečić D, Krotzsch M. Wikidata: a free collaborative knowledgebase[J]. Communications of the ACM, 2014,

- 57(10):78-85.
- [23] Wang Z, Li J, Wang Z, et al. XLoRe: a large-scale English-Chinese bilingual knowledge graph[C]//International Semantic Web Conference, 2013:121-124.
- [24] Niu X, Sun X, Wang H, et al. Zhishi.me: weaving Chinese linking open data[C]//International Semantic Web Conference, 2011:205-220.
- [25] Green B F, Alice K W, Chomsky C, et al. Baseball: an automatic question answerer[C]//Proceedings of the Western Joint Computer Conference, 1961:219-224.
- [26] Woods W A. Progress in natural language understanding: an application to lunar geology[C]//National Computer Conference, 1973:441-450.
- [27] Fader A, Zettlemoyer L, Etzioni O, et al. Paraphrase-driven learning for open question answering[C]//Meeting of the Association for Computational Linguistics, 2013:1608-1618.
- [28] Berant J, Chou A K, Frostig R, et al. Semantic parsing on freebase from question-answer pairs[C]//Conference on Empirical Methods in Natural Language Processing, 2013:1533-1544.
- [29] Berant J, Liang P. Semantic parsing via paraphrasing[C]//Meeting of the Association for Computational Linguistics, 2014:1415-1425.
- [30] Yih W, Chang M, He X, et al. Semantic parsing via staged query graph generation: question answering with knowledge base[C]//International Joint Conference on Natural Language Processing, 2015:1321-1331.
- [31] 郑实福, 刘挺, 秦兵, 等. 自动问答综述[J]. 中文信息学报, 2002, 16(6):47-53.
- [32] 王东升, 王卫民, 王石, 等. 面向限定领域问答系统的自然语言理解方法综述[J]. 计算机科学, 2017, 44(8):1-8.
- [33] 郭天翼, 彭敏, 伊穆兰, 等. 自然语言处理领域中的自动问答研究进展[J]. 武汉大学学报(理学版), 2019, 65(5):417-426.
- [34] Bordes A, Usunier N, Chopra S, et al. Large-scale simple question answering with memory networks[J]. arXiv: 1506.02075, 2015.
- [35] Yih W, Richardson M, Meek C, et al. The value of semantic parse labeling for knowledge base question answering[C]//Meeting of the Association for Computational Linguistics, 2016:201-206.
- [36] Usbeck R, Gusmita R H, Saleem M, et al. 9th challenge on Question Answering over Linked Data (QALD-9)[C]//Semdeep/NLIWoD@ISWC, 2018:58-64.
- [37] Saha A, Pahuja V, Khapra M M, et al. Complex sequential question answering: towards learning to converse over linked question answer pairs with a knowledge graph[C]//Thirty-Second AAAI Conference on Artificial Intelligence, 2018:705-713.
- [38] Liu J, Lin Y, Liu Z, et al. XQA: a cross-lingual open-domain question answering dataset[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019:2358-2368.
- [39] Xiong W, Wu J, Wang H, et al. TWEETQA: a social media focused question answering dataset[J]. arXiv: 1907.06292, 2019.
- [40] Tunstall-Pedoe W. True knowledge: open-domain question answering using structured knowledge and inference[J]. AI Magazine, 2010, 31(3):80-92.
- [41] 丁斌. 汽车领域智能问答系统中模板库自动生成方法的研究[D]. 上海: 上海交通大学, 2017.
- [42] Cui W, Xiao Y, Wang H, et al. KBQA: learning question answering over QA corpora and knowledge bases[J]. Proceedings of the VLDB Endowment, 2017, 10(5):565-576.
- [43] Abujabal A, Yahya M, Riedewald M, et al. Automated template generation for question answering over knowledge graphs[C]//the Web Conference, 2017:1191-1200.
- [44] Cocco R, Atzori M, Zaniolo C, et al. Machine learning of SPARQL templates for question answering over LinkedSpending[C]//Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, 2019:156-161.
- [45] Steedman M. Surface structure and interpretation[J]. Linguistic Inquiry Monographs, 1996, 43(3):271-276.
- [46] Zettlemoyer L S, Collins M. Learning to map sentences to logical form: structured classification with probabilistic categorial grammars[J]. arXiv: 1207.1420, 2012.
- [47] Kwiatkowski T, Zettlemoyer L, Goldwater S, et al. Lexical generalization in CCG grammar induction for semantic parsing[C]//Conference on Empirical Methods in Natural Language Processing, 2011:1512-1523.
- [48] Artzi Y, Lee K, Zettlemoyer L, et al. Broad-coverage CCG semantic parsing with AMR[C]//Conference on Empirical Methods in Natural Language Processing, 2015:1699-1710.
- [49] Li J, Zhu M, Lu W, et al. Improving semantic parsing with enriched synchronous context-free grammar[C]//Conference on Empirical Methods in Natural Language Processing, 2015:1455-1465.
- [50] Reddy S, Lapata M, Steedman M, et al. Large-scale semantic parsing without question-answer pairs[J]. Transactions of the Association for Computational Linguistics, 2014, 2(1):377-392.
- [51] Reddy S, Tackstrom O, Collins M, et al. Transforming dependency structures to logical forms for semantic parsing[J]. Transactions of the Association for Computational Linguistics, 2016, 4(1):127-140.

- [52] Yih W, Chang M, He X, et al. Semantic parsing via staged query graph generation: question answering with knowledge base[C]//International Joint Conference on Natural Language Processing, 2015: 1321-1331.
- [53] Bast H, Haussmann E. More accurate question answering on freebase[C]//Conference on Information and Knowledge Management, 2015: 1431-1440.
- [54] Hao Z, Wu B, Wen W, et al. A subgraph-representation-based method for answering complex questions over knowledge bases[J]. Neural Networks, 2019, 119: 57-65.
- [55] Dong L, Lapata M. Language to logical form with neural attention[C]//Meeting of the Association for Computational Linguistics, 2016: 33-43.
- [56] Jia R, Liang P. Data recombination for neural semantic parsing[C]//Meeting of the Association for Computational Linguistics, 2016: 12-22.
- [57] Xiao C, Dymetman M, Gardent C, et al. Symbolic priors for RNN-based semantic parsing[C]//International Joint Conference on Artificial Intelligence, 2017: 4186-4192.
- [58] Chen B, Sun L, Han X, et al. Sequence-to-action: end-to-end semantic graph generation for semantic parsing[C]//Meeting of the Association for Computational Linguistics, 2018: 766-777.
- [59] Krishnamurthy J, Dasigi P, Gardner M, et al. Neural semantic parsing with type constraints for semi-structured tables[C]//Conference on Empirical Methods in Natural Language Processing, 2017: 1516-1526.
- [60] Dong L, Lapata M. Coarse-to-fine decoding for neural semantic parsing[C]//Meeting of the Association for Computational Linguistics, 2018: 731-742.
- [61] Chen Y, Wu L, Zaki M J, et al. Bidirectional attentive memory networks for question answering over knowledge bases[J]. arXiv: 1903.02188, 2019.
- [62] Lukovnikov D, Fischer A, Lehmann J, et al. Neural network-based question answering over knowledge graphs on word and character level[C]//the Web Conference, 2017: 1211-1220.
- [63] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [64] Devlin J, Chang M W, Lee K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[J]. arXiv: 1810.04805, 2018.
- [65] Bordes A, Chopra S, Weston J. Question answering with subgraph embeddings[J]. arXiv: 1406.3676, 2014.
- [66] Bordes A, Weston J, Usunier N, et al. Open question answering with weakly supervised embedding models[C]//European Conference on Machine Learning, 2014: 165-180.
- [67] 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]//Meeting of the Association for Computational Linguistics, 2017: 221-231.
- [68] Hao Y, Liu H, He S, et al. Pattern-revising enhanced simple question answering over knowledge bases[C]//International Conference on Computational Linguistics, 2018: 3272-3282.
- [69] Zhang R, Wang Y, Mao Y, et al. Question answering in knowledge bases: a verification assisted model with iterative training[J]. ACM Transactions on Information Systems, 2019, 37(4): 1-26.
- [70] Dubey M, Banerjee D, Chaudhuri D, et al. EARL: joint entity and relation linking for question answering over knowledge graphs[C]//International Semantic Web Conference, 2018: 108-126.
- [71] Weston J, Chopra S, Bordes A. Memory networks[J]. arXiv: 1410.3916, 2014.
- [72] Dong L, Wei F, Zhou M, et al. Question answering over freebase with multi-column convolutional neural networks[C]//International Joint Conference on Natural Language Processing, 2015: 260-269.
- [73] Qu Y, Liu J, Kang L, et al. Question answering over freebase via attentive RNN with similarity matrix based CNN[J]. arXiv: 1804.03317, 2018.
- [74] Lan Y, Wang S, Jiang J, et al. Knowledge base question answering with a matching-aggregation model and question-specific contextual relations[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2019, 27(10): 1629-1638.
- [75] Krishna K, Iyyer M. Generating question-answer hierarchies[J]. arXiv: 1906.02622, 2019.
- [76] 罗达, 苏锦钊, 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.
- [77] 曹明宇, 李青青, 杨志豪, 等. 基于知识图谱的原发性肝癌知识问答系统[J]. 中文信息学报, 2019, 33(6): 88-93.
- [78] 张楚婷. 基于知识图谱的旅游问答系统研究与实现[D]. 广西桂林: 桂林电子科技大学, 2019.
- [79] 张崇宇. 基于知识图谱的自动问答系统的应用研究与实现[D]. 北京: 北京邮电大学, 2019.
- [80] 史梦飞. 基于中文知识库的问答系统研究与实现[D]. 上海: 华东师范大学, 2018.
- [81] 童培豪. 针对复杂问题优化的知识图谱问答系统的研究和实现[D]. 上海: 华东师范大学, 2019.
- [82] Lai Y, Jia Y, Lin Y, et al. A Chinese question answering system for single-relation factoid questions[C]//National CCF Conference on Natural Language Processing and Chinese Computing. Cham: Springer, 2017: 124-135.
- [83] Zhou W C, Lu D. Combining neural network models with

- rules for chinese knowledge base question answering[C]//China Conference on Knowledge Graph and Semantic Computing,2019:312-324.
- [84] Costa J O,Kulkarni A.Leveraging knowledge graph for open-domain question answering[C]//2018 IEEE/WIC/ACM International Conference on Web Intelligence,2018:389-394.
- [85] Yang B,Yih W,He X,et al.Embedding entities and relations for learning and inference in knowledge bases[J].arXiv:1412.6575,2014.
- [86] Bordes A,Usunier N,Garciaduran A,et al.Translating embeddings for modeling multi-relational data[C]//Advances in Neural Information Processing Systems,2013:2787-2795.
- [87] Nickel M,Tresp V,Kriegel H,et al.A three-way model for collective learning on multi-relational data[C]//International Conference on Machine Learning,2011:809-816.
- [88] Socher R,Chen D,Manning C D,et al.Reasoning with neural tensor networks for knowledge base completion[C]//Advances in Neural Information Processing Systems,2013:926-934.
- [89] Wang Meng,Wang Ruijie,Liu Jun,et al.Towards empty answers in SPARQL:approximating querying with RDF embedding[C]//Proceedings of the 17th International Semantic Web Conference,2018:513-529.
- [90] Huang Xiao,Zhang Jingyuan,Li Dingcheng,et al.Knowledge graph embedding based question answering[C]//Proceedings of the 12th ACM International Conference on Web Search and Data Mining,2019:105-113.
- [91] Hamilton W,Bajaj P,Zitnik M,et al.Embedding logical queries on knowledge graphs[C]//Proceedings of the 32nd Annual Conference on Neural Information Processing Systems,2018:2026-2037.
- [92] Wang Ruijie,Wang Meng,Liu Jun,et al.Leveraging knowledge graph embeddings for natural language question answering[C]//Proceedings of the 24th International Conference of Database Systems for Advanced Applications,2019:659-675.
- [93] Petrochuk M,Zettlemoyer L.SimpleQuestions nearly solved: a new upperbound and baseline approach[J].arXiv:1804.08798,2018.
- [94] Zhang Y,Dai H,Kozareva Z,et al.Variational reasoning for question answering with knowledge graph[C]//Proceedings of the 32nd AAAI Conference on Artificial Intelligence,New Orleans,Louisiana,USA,2018:6069-6076.
- [95] Zhou M,Huang M,Zhu X,et al.An interpretable reasoning network for multi-relation question answering[C]//International Conference on Computational Linguistics,2018:2010-2022.
- [96] Vakulenko S,Garcia J D,Polleres A,et al.Message passing for complex question answering over knowledge graphs[C]//Conference on Information and Knowledge Management,2019:1431-1440.