

## 面向知识图谱的知识推理研究进展<sup>\*</sup>

官赛萍<sup>1,2</sup>, 靳小龙<sup>1,2</sup>, 贾岩涛<sup>1,2</sup>, 王元卓<sup>1,2</sup>, 程学旗<sup>1,2</sup>



<sup>1</sup>(网络数据科学与技术重点实验室(中国科学院 计算技术研究所), 北京 100190)

<sup>2</sup>(中国科学院大学 计算机与控制学院, 北京 100049)

通讯作者: 靳小龙, E-mail: jinxiaolong@ict.ac.cn

**摘要:** 近年来,随着互联网技术和应用模式的迅猛发展,引发了互联网数据规模的爆炸式增长,其中包含大量有价值的知识.如何组织和表达这些知识,并对其进行深入计算和分析备受关注.知识图谱作为丰富直观的知识表达方式应运而生.面向知识图谱的知识推理是知识图谱的研究热点之一,已在垂直搜索、智能问答等应用领域发挥了重要作用.面向知识图谱的知识推理旨在根据已有的知识推理出新的知识或识别错误的知识.不同于传统的知识推理,由于知识图谱中知识表达形式的简洁直观、灵活丰富,面向知识图谱的知识推理方法也更加多样化.将从知识推理的基本概念出发,介绍近年来面向知识图谱知识推理方法的最新研究进展.具体地,根据推理类型划分,将面向知识图谱的知识推理分为单步推理和多步推理,根据方法的不同,每类又包括基于规则的推理、基于分布式表示的推理、基于神经网络的推理以及混合推理.详细总结这些方法,并探讨和展望面向知识图谱知识推理的未来研究方向和前景.

**关键词:** 知识推理;知识图谱;规则推理;分布式表示;神经网络

**中图法分类号:** TP18

中文引用格式: 官赛萍,靳小龙,贾岩涛,王元卓,程学旗.面向知识图谱的知识推理研究进展.软件学报,2018,29(10):2966–2994.  
http://www.jos.org.cn/1000-9825/5551.htm

英文引用格式: Guan SP, Jin XL, Jia YT, Wang YZ, Cheng XQ. Knowledge reasoning over knowledge graph: A survey. Ruan Jian Xue Bao/Journal of Software, 2018, 29(10): 2966–2994 (in Chinese). http://www.jos.org.cn/1000-9825/5551.htm

## Knowledge Reasoning Over Knowledge Graph: A Survey

GUAN Sai-Ping<sup>1,2</sup>, JIN Xiao-Long<sup>1,2</sup>, JIA Yan-Tao<sup>1,2</sup>, WANG Yuan-Zhuo<sup>1,2</sup>, CHENG Xue-Qi<sup>1,2</sup>

<sup>1</sup>(CAS Key Laboratory of Network Data Science and Technology (Institute of Computing Technology, The Chinese Academy of Sciences), Beijing 100190, China)

<sup>2</sup>(School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** In recent years, the rapid development of Internet technology and Web applications has triggered the explosion of various data on the Internet, which generates a large amount of valuable knowledge. How to organize, represent and analyze these knowledge has attracted much attention. Knowledge graph was thus developed to organize these knowledge in a semantical and visualized manner. Knowledge reasoning over knowledge graph then becomes one of the hot research topics and plays an important role in many applications such as vertical search and intelligent question-answer. The goal of knowledge reasoning over knowledge graph is to infer new facts or

\* 基金项目: 国家重点研发计划(2016YFB1000902, 2017YFB1002302); 国家自然科学基金(61772501, 61572473, 61572469, 91646120)

Foundation item: National Key Research and Development Program (2016YFB1000902, 2017YFB1002302); National Natural Science Foundation of China (61772501, 61572473, 61572469, 91646120)

本文由“本体工程与知识图谱”专题特约编辑欧阳丹彤教授推荐.

收稿时间: 2017-07-20; 修改时间: 2017-11-08; 采用时间: 2018-01-24; jos 在线出版时间: 2018-02-08

CNKI 网络优先出版: 2018-02-08 11:55:48, http://kns.cnki.net/kcms/detail/11.2560.TP.20180208.1155.007.html

identify erroneous facts according to existing ones. Unlike traditional knowledge reasoning, knowledge reasoning over knowledge graph is more diversified, due to the simplicity, intuitiveness, flexibility, and richness of knowledge representation in knowledge graph. Starting with the basic concept of knowledge reasoning, this paper presents a survey on the recently developed methods for knowledge reasoning over knowledge graph. Specifically, the research progress is reviewed in detail from two aspects: One-Step reasoning and multi-step reasoning, each including rule based reasoning, distributed embedding based reasoning, neural network based reasoning and hybrid reasoning. Finally, future research directions and outlook of knowledge reasoning over knowledge graph are discussed.

**Key words:** knowledge reasoning; knowledge graph; rule based reasoning; distributed embedding; neural network

在当今互联网、物联网、云计算等技术不断发展的环境下,各类应用层出不穷,因此产生了海量的数据资源<sup>[1-4]</sup>,其中包含大量有价值的知识.这吸引了许多研究人员对其进行深入挖掘和分析.如何组织表达这些知识,以便作进一步的计算和分析备受关注.知识图谱应运而生,将实体(包括概念、属性值)表示成图上的节点,节点之间的连边对应实体之间的关联关系,以一个网络化的结构表征所获得的知识,清晰、直观.

目前,已经涌现出一大批知识图谱,其中具有代表性的有 KnowItAll<sup>[5]</sup>、YAGO<sup>[6-8]</sup>、DBpedia<sup>[9]</sup>、Freebase<sup>[10]</sup>、NELL<sup>[11]</sup>、Probase<sup>[12]</sup>等.这些知识图谱从大量数据资源中抽取、组织和管理知识,希望为用户提供能够读懂用户需求的智能服务,例如理解搜索的语义,提供更精准的搜索答案.

这其中涉及面向知识图谱的知识推理,完成对数据的深度分析和推理.本质上,本文研究的面向知识图谱的知识推理是指根据知识图谱中已有的知识,采用某些方法,推理出新的知识或识别知识图谱中错误的知识.相应地,它包括两方面的内容:知识图谱补全(knowledge graph completion, knowledge base completion)<sup>[13-16]</sup>和知识图谱去噪(knowledge graph refinement, knowledge graph cleaning)<sup>[17,18]</sup>.知识图谱补全又包括连接预测(link prediction)<sup>[19-21]</sup>、实体预测(entity prediction)<sup>[22-24]</sup>、关系预测(relation prediction)<sup>[22-25]</sup>、属性预测(attribute prediction)<sup>[26]</sup>等任务.

近年来,由于社会各界对知识图谱的广泛关注,知识图谱研究取得了很大的进展,有不少关于知识图谱的综述性文献陆续发表.譬如,文献[27]对知识抽取的研究进展进行了综述,文献[28-31]分别对知识图谱构建、实体对齐、知识表示学习以及知识融合进行了综述,文献[32]对知识抽取、知识表示、知识融合和知识推理这4个方面的研究进展进行了总结和展望.

然而,尽管已有上述诸多知识图谱综述文献,但仍然缺乏对面向知识图谱的知识推理研究进行系统、深入地梳理与总结的工作.为此,本文对面向知识图谱知识推理的最新研究进展进行归纳总结,并展望未来发展方向和前景.具体地,本文首先引入知识推理简介;然后介绍传统知识推理方法用于面向知识图谱知识推理的具体实例;接着,将现有面向知识图谱的知识推理方法分为单步推理和多步推理两个类别,每类又从基于规则的推理、基于分布式表示的推理、基于神经网络的推理以及混合推理这4个角度逐一详细阐述研究进展;最后,本文总结并展望未来的研究方向.

## 1 知识推理简介

在介绍面向知识图谱的各类知识推理方法之前,本节首先介绍知识推理的基本概念、传统的知识推理以及面向知识图谱的知识推理.

### 1.1 知识推理的基本概念

关于知识推理的基本概念,学术界给出了各种类似的定义.王永庆<sup>[33]</sup>认为:推理是人们对各种事物进行分析、综合和决策,从已知的事实出发,通过运用已掌握的知识,找出其中蕴含的事实,或归纳出新的事实的过程.严格地说,就是按照某种策略由已知判断推出新的判断的思维过程.更具体地,Kompridis<sup>[34]</sup>定义推理为一系列能力的总称,包括有意识地理解事物的能力、建立和验证事实的能力、运用逻辑的能力以及基于新的或存在的知识改变或验证现有体系的能力.Tari<sup>[35]</sup>类似地定义知识推理为基于特定的规则和约束,从存在的知识获得新的知识.总的来说,知识推理就是利用已知的知识推出新知识的过程.

知识推理作为人类问题求解的主要方法,一直以来备受关注.一般来说,知识推理包括两种知识:一种是已

知的知识,另一种是由已知的知识推出的新知识.已知的知识可以是一个或多个句子,例如传统的三段论,由大前提、小前提和结论组成,其中,大前提和小前提是已知的知识,结论是推出的新知识.又如,在规则推理中,存在以下传递性规则:如果已知知识  $p_1$  的父亲是  $p_2$ ,  $p_2$  的母亲是  $p_3$ ,那么可以推出新知识  $p_1$  的祖母是  $p_3$ .已知的知识也可以是更复杂的案例,例如在案例推理中,利用案例库中的已有案例对新的案例进行推理.这里,案例库中的案例(包括问题描述部分和方案部分)是已知的知识,通常采用特征-值对的形式表达<sup>[36]</sup>,针对新案例推理得到的方案部分则是推出的新知识.表 1 展示了两个已知的案例,问题描述部分包含“Symptom”“Observations”等特征,最后一行为方案部分.已知的知识还可以用更为简洁的多元组形式表达,例如,在知识图谱中,通常以(头实体,关系,尾实体)的形式表达知识,推出的新知识也以三元组形式表达.

Table 1 Examples of case<sup>[36]</sup>表 1 具体案例实例<sup>[36]</sup>

Attributes labels	Case 1	Case 2
Symptom	Unusual car noise	Unusual car noise
Observations	Knocking engine	Knocking engine
Since last inspection (month)	4	14
Rhythmic pounding	Sometimes	No
Related to car speed	No	No
Oil pressure light flickering	No	Yes
Leaking oil	No	Rarely
Solution	Loose transmission torque converter	Oil burning

## 1.2 传统的知识推理

知识推理很早就受到了广泛关注,关于知识推理的研究最早可以追溯到古希腊哲学家,此后,其他一些学科的研究工作者,如人工智能专家与理论计算机科学家也对知识推理产生了兴趣<sup>[37]</sup>.

推理方法按新判断推出的途径划分,分为演绎推理、归纳推理和默认推理<sup>[33]</sup>.演绎推理是从一般到个别的推理.演绎推理发展历史悠久,涵盖自然演绎、归结原理、表演算等广泛使用的方法<sup>[38,39]</sup>.早在 1935 年, Gentzen<sup>[40,41]</sup>就提出自然演绎推理,将推理形式化为引入经典逻辑的推理规则的数学证明过程.归结原理由 Robinson<sup>[42]</sup>于 1965 年提出,是一种采用反证法的推理方法.它将证明原逻辑表达式(由给定子句推出目标子句)为永真转换为证明给定子句与目标子句取非的合取式存在矛盾的等价问题,然后进行归结.如果转换的逻辑表达式不可满足,推出矛盾,则原表达式为永真.表演算于 1991 年由 Schmidt-Schauß 和 Smolka<sup>[43]</sup>首次引入.该算法首先构建一个规则的完全森林,由带标签的有向图组成,每个节点用概念集标记,每条边用规则名标记,表示节点之间存在的规则关系,然后利用扩展规则,给节点标签添加概念或添加新的节点,进而基于这个完全森林进行演算推理.

归纳推理则可以追溯到 1964 年 Solomonoff<sup>[44,45]</sup>创立的普遍归纳推理理论,是一个基于观察的预测理论.概括地说,归纳推理是从足够多的事例中归纳出一般性结论的推理过程.

默认推理又称为缺省推理,1980 年, Reiter<sup>[46]</sup>正式提出缺省推理逻辑.缺省推理是在知识不完全的情况下,通过假设某些条件已经具备而进行的推理.

推理方法按所用知识的确定性划分,分为确定性推理和不确定性推理<sup>[33]</sup>.确定性推理所用的知识是精确的,并且推出的结论也是确定的.早在 1975 年, Shortliffe 和 Buchanan<sup>[47]</sup>就提出确定性理论——一种不确定性的推理模型.在不确定性推理中,知识和证据都具有某种程度的不确定性.不确定性推理又分为似然推理和近似推理(模糊推理),前者基于概率论,后者基于模糊逻辑<sup>[33]</sup>. Zadeh<sup>[48]</sup>在 1973 年首次提出模糊推理理论中最基本的推理规则,即模糊分离规则,随后,模糊推理不断发展完善.

将推理方法按推理过程中推出的结论是否单调增加来划分,分为单调推理和非单调推理<sup>[33]</sup>.单调推理中,随着推理的向前推进和新知识的加入,推出的结论单调递增,越来越接近最终目标.上述提到的演绎推理就属于单调推理<sup>[38]</sup>.非单调推理最早由 Minsky<sup>[49]</sup>正式提出.在推理过程中,随着新知识的加入,非单调推理需要否定已推出的结论,使得推理退回到前面的某一步,重新开始.

将推理方法按是否运用与问题有关的启发性知识来划分,分为启发式推理和非启发式推理<sup>[33]</sup>。启发式推理在推理过程中,运用解决问题的策略、技巧和经验,加快推理。而非启发式推理只按照一般的控制逻辑进行推理。

将推理方法从方法论的角度来划分,分为基于知识的推理、统计推理和直觉推理<sup>[33]</sup>。基于知识的推理根据已掌握的事实,通过运用知识进行推理。统计推理根据对事物的统计信息进行推理。直觉推理又称为常识性推理,是根据常识进行的推理。直觉推理依赖于感知经验和具体实例,当逻辑/规则与直觉不一致时,则不考虑这些逻辑/规则<sup>[50]</sup>。

推理方法还可以根据推理的繁简不同,分为简单推理和复合推理;根据结论是否具有必然性,分为必然性推理和或然性推理;根据推理控制方向,分为正向推理、逆向推理、混合推理和双向推理<sup>[33]</sup>。此外,还有时间推理、空间推理和案例推理等推理方法。时间推理是对与时间有关的知识进行的推理,早在1983年,Allen等人<sup>[51]</sup>就提出了基于时区的时间知识表示和时间推理方法。空间推理是指利用空间理论和人工智能技术对空间对象进行建模、描述和表示,并据此对空间对象之间的空间关系进行定性或定量分析和处理的过程<sup>[52]</sup>。案例推理起源于美国耶鲁大学的Schank教授,他的动态记忆理论<sup>[53]</sup>被认为是最早的案例推理的基础。案例推理通过使用或调整老问题的解决方案推理新的问题<sup>[54]</sup>。

上述传统的知识推理方法主要是基于逻辑、规则的推理,逐渐发展为最基本的通用推理方法。近年来,传统的知识推理继续发展更新,从内容上看主要是短语和句子的推理,包括基于词汇内容的推理<sup>[55,56]</sup>、基于数理逻辑的推理<sup>[57]</sup>、基于自然语言逻辑(自然语言与数理逻辑结合的一种逻辑)的推理<sup>[58-61]</sup>以及结合词汇内容和数理逻辑/自然语言逻辑的推理<sup>[62,63]</sup>。

除了一般的短语和句子级的推理之外,另一大类受到广泛关注的推理是本体推理。本体是共享概念的模型明确的形式化规范说明<sup>[64]</sup>。换句话说,本体类似于知识库中的模式,用来定义类和关系以及类层次和关系层次结构等。本体作为语义丰富的知识描述,其推理受到了广泛关注,从2012年开始,出现了一年一度的本体推理器评估(the OWL reasoner evaluation,简称ORE)竞赛。各种推理器层出不穷,例如 TrOWL<sup>[65]</sup>、Chainsaw<sup>[66]</sup>、jcel<sup>[67]</sup>、MORE<sup>[68]</sup>、ELepHant<sup>[69]</sup>、ELK<sup>[70]</sup>、HermiT8<sup>[71]</sup>、PAGODA<sup>[72]</sup>等。

### 1.3 面向知识图谱的知识推理

2012年,Google最早提出知识图谱的概念(<http://googleblog.blogspot.co.uk/2012/05/introducing-knowledge-graph-things-not.html>),随着知识图谱的出现,面向知识图谱的知识推理作为支撑上层应用的基础性服务引发了广泛关注。本文研究面向知识图谱的知识推理,后文中的“知识推理”如果未加特别说明,将特指“面向知识图谱的知识推理”。面向知识图谱的知识推理旨在基于已有的知识图谱事实,推理新的事实或识别错误知识,例如,在DBpedia中已知三元组 $(X, birthPlace, Y)$ ,可以在很大程度上推理出缺失的三元组 $(X, nationality, Y)$ <sup>[73]</sup>。面向知识图谱的知识推理,一方面随着知识图谱概念的出现与发展,作为补全知识图谱和知识图谱去噪的主要手段自然受到了广泛关注;另一方面,可以说是由传统的知识推理发展而来。

形式化地说,知识图谱通常用上述(头实体,关系,尾实体)的三元组形式表达事物的属性以及事物之间的语义关系,其中,事物和属性值作为三元组中的实体,属性和关系作为三元组中的关系。知识图谱补全实际上是给定三元组中任意两个元素,试图推理出缺失的另外一个元素。也即:给定头实体和关系(关系和尾实体),找出与之形成有效三元组的尾实体(头实体),也称为实体预测;同理,给定头实体和尾实体,找出与之形成有效三元组的关系,也称为关系预测。无论实体预测还是关系预测,最后都转化为选择与给定元素形成的三元组更可能有效的实体/关系作为推理预测结果。这种有效性可以通过规则的方式推理或通过基于特定假设的得分函数计算。而知识图谱去噪,实际上是在判断三元组的正确与否。因此,虽然知识图谱补全专注于扩充知识图谱,而知识图谱去噪专注于知识图谱内部已有三元组正确性的判断,但本质上都是评估三元组的有效性。后文在介绍各种推理方法时,本文注重的是方法的描述,而不刻意强调具体任务(知识图谱补全还是知识图谱去噪)。并且,相对于知识图谱去噪而言,知识图谱补全的工作更多得多,因此在介绍的推理工作中,知识图谱补全占了大部分。

知识图谱本质上是一种语义网络,可以对现实世界的事物及其相互关系进行形式化的描述<sup>[32]</sup>。语义网络最早由Quillian<sup>[74]</sup>提出,是一个带有标记的有向图,通过事物属性以及事物之间语义关系的直观表达,很容易找到



与节点有关的知识<sup>[33]</sup>.相比于传统知识中非结构化表达的形式,如一阶谓词<sup>[39]</sup>、产生式<sup>[75]</sup>等,知识图谱以结构化的方式表达知识,将事物的属性以及事物之间的语义关系显式地表示出来;相比于结构化表达的形式,如框架<sup>[49]</sup>、脚本<sup>[76]</sup>等,知识图谱中事物的属性以及事物之间的联系通常以三元组的形式刻画,更加简洁直观、灵活丰富.灵活体现在不需要采用框架<sup>[49]</sup>、脚本<sup>[76]</sup>等结构化表达方式中呆板笨重的槽等组成结构,只是简单的三元组形式.丰富体现在这种三元组的形式可以很容易找到与事物相关的所有知识.因此,面向知识图谱的知识推理不仅仅局限于以基于逻辑和规则为主的传统知识推理,可以有更多样化的推理方法.特别地,知识图谱相比于本体,关注的是大量的具体实例,即(头实体,关系,尾实体)三元组,实例可能偏向于某些具体的实体.而本体是模式导向的,虽然也创建实例<sup>[77]</sup>,但具体实例不是主要的关注点,以中立的方式主要描述概念、概念之间的关系和它们的属性.知识图谱可以理解为规模非常大的本体<sup>[78]</sup>,存储着大量的三元组.因此,相对于本体推理,面向知识图谱的知识推理由于知识图谱自身实例为主导的特征,不局限于本体主要的概念层面的抽象推理,可以有更具体化的推理方法.

传统的知识推理方法(包括本体推理方法)可用于面向知识图谱的知识推理,本文第2节将给出具体阐述.特别地,演绎推理在传统的知识推理中,尤其是在本体推理中占据重要的位置,但在面向知识图谱的知识推理中,归纳推理成为主要方法(下文中介绍的方法也主要属于归纳推理).演绎推理由于只要前提有效,推出的结论必定可靠,在传统的知识推理中得到广泛应用.但在知识图谱中,由于实例数量多,涉及的内容也往往很广,需要大量的逻辑规则,在实例层面的演绎推理时间复杂度很高.而在抽象概念层面的演绎推理又面临大量的实例化问题,即:将抽象概念替换成具体实体,同样代价很高,并且很难获取覆盖面足够广的概念层面的推理规则.近年来,面向知识图谱的知识推理随着分布式表示、神经网络等技术的流行,已发展出独有的推理方法,根据推理类型划分,分为单步推理和多步推理.每类再根据方法划分,又包括基于规则的推理、基于分布式表示的推理、基于神经网络的推理以及混合推理.

## 2 基于传统方法的推理

传统的知识推理包括本体推理一直以来备受关注,产生了一系列的推理方法.面向知识图谱的知识推理可以应用这些方法完成知识图谱场景下的知识推理.本节将概述这些应用的实例,具体可分为两类:基于传统规则推理的方法和基于本体推理的方法,分别将传统的规则推理和本体推理方法用于面向知识图谱的知识推理.

### 2.1 基于传统规则推理的方法

基于传统规则推理的方法主要借鉴传统知识推理中的规则推理方法,在知识图谱上运用简单规则或统计特征进行推理.NELL 知识图谱内部的推理组件采用一阶关系学习算法进行推理<sup>[11]</sup>.推理组件学习概率规则,经过人工筛选过滤后,带入具体的实体将规则实例化,从已经学习到的其他关系实例推理新的关系实例.YAGO 知识图谱内部采用了一个推理机——Spass-YAGO 以丰富知识图谱内容<sup>[79]</sup>.Spass-YAGO 抽象化 YAGO 中的三元组到等价的规则类,采用链式叠加计算关系的传递性,叠加过程可以任意地迭代,通过这些规则完成 YAGO 的扩充.Wang 等人<sup>[80,81]</sup>提出了一阶概率语言模型 ProPPR(programming with personalized PageRank)进行知识图谱上的知识推理.ProPPR 构建有向证明图,节点对应“关系(头实体变量,尾实体变量)”形式的子句的连接或推理目标,其中,起始节点为查询子句,边对应规则,也即一个推理步,从一个子句归约到另一个子句.边的权重与特征向量相关联,当引入一个特征模板时,边的权重可以依赖于模板的部分实例化结果,如依赖于子句中某个变量的具体取值.同时,在图中添加从每个目标尾节点指向自己的自环以及每个节点到起始节点的自启动边.自环用于增大目标尾节点的权重,自启动边使得遍历偏向于增大推理步数少的推理的权重.最后,ProPPR 中的推理基于图上的个性化网页评分实现.Cohen<sup>[82]</sup>进一步提出了 TensorLog,用可微的过程进行推理.TensorLog 中,每个实体关联一个 one-hot 向量,为每个关系定义一个 $\{0,1\}$ 操作矩阵,如果第  $i$  个实体和第  $j$  个实体存在对应的关系, $(i,j)$ 位置上的值为 1,否则为 0.这些表达都固定,不更新.由此逻辑规则推理可以形式化为矩阵相乘.给定实体和关系,预测另一个实体时,对于每一条可能的路径,实体 one-hot 向量乘以路径上关系操作的乘积(给定尾实体时是乘积的转置),用待学习的置信度参数对所有路径的结果进行加权求和,即可得到所有实体对应的得分向量.由于采用

one-hot 表示,候选实体的得分可以通过该得分向量乘以其 one-hot 向量的转置得到.对于置信度参数,通过最大化知识图谱中三元组的得分学习.Paulheim 和 Bizer<sup>[18]</sup>提出两种算法 SDType 和 SDValidate,利用属性和类型的统计分布,补全类型(type)三元组以及识别错误的三元组.SDType 通过属性头尾实体的类型统计分布推理实体的类型,类似加权投票机制,对每个属性的投票赋予权重.SDValidate 首先计算关系-尾实体的频率,低频率的三元组进一步用属性和类型的统计分布计算得分,得分小于阈值的三元组被认为是潜在错误的.Jang 等人<sup>[83]</sup>则提出了基于模式的方法评估知识图谱三元组的质量,该方法直接在知识图谱中进行数据模式分析,根据假设——更频繁出现的模式更可靠,选择出现率高的模式,包括头实体模式、尾实体模式等,然后利用这些模式进行三元组质量分析.

## 2.2 基于本体推理的方法

基于本体推理的方法主要利用更为抽象化的本体层面的频繁模式、约束或路径进行推理.Bühmann 和 Lehmann<sup>[84]</sup>提出了基于模式的知识图谱补全,首先对多个本体库进行统计分析,发现频繁原子模式;然后在具体的知识图谱上查询这些原子模式和相关数据,得到候选原子集,即原子模式的实例,例如:原子模式“ $A \equiv B \cap \exists r.C$ ”对应的原子可以是“ $SoccerPlayer \equiv Person \cap \exists team.SoccerClub$ ”;最后,基于在知识图谱中的正确性统计,计算每个候选的得分,用大于阈值的候选作为规则补全知识图谱.Jiang 等人<sup>[17]</sup>关注于用启发式规则推理知识图谱中不确定和冲突的知识,提出了基于马尔可夫逻辑网(Markov logic network,简称 MLN)<sup>[85]</sup>的系统去噪抽取的 NELL 知识图谱.MLN<sup>[85]</sup>由带权的一阶逻辑规则组成,结合了一阶逻辑和概率图模型.用一组常量实例化后,MLN 最大化该实例化网络的概率分布.该概率分布基于推出实例多的强规则对应的权重大的原则,学习逻辑规则的权重.由此硬约束对应的权重无穷大.Jiang 等人<sup>[17]</sup>将来自原信息抽取系统的本体约束一阶逻辑规则作为 MLN 的硬约束,事实元组的置信度以及抽取模式的置信度作为软约束(带权重).软约束可以形式化为 MLN 的带权一阶逻辑规则,即带置信度的,事实元组或抽取模式推出其他事实元组正确,作为规则.置信度作为权重,是要学习到参数,通过逻辑回归在人工标注的事实元组上近似学习.其中,本体约束包括包含关系、互斥关系以及关系的头尾实体类型约束.判定不满足硬约束的事实元组是错误的,满足硬约束的事实元组进一步用软约束判定.然而,MLN 中,逻辑规则的句子(变量存在关系)的值必须取布尔量,难以引入子句置信度值,并且变量布尔指派的组合爆炸使得学习和推理难以精确优化.概率软逻辑(probabilistic soft logic,简称 PSL)<sup>[86,87]</sup>将子句取布尔量替换为取连续量,解决上述问题.连续量使得由与、或、非逻辑连接词连接子句形成的逻辑规则可以形式化为变量的简单操作.由此,学习问题转化为高效的凸优化问题.鉴于 PSL 的优化高效,连续量还能方便地表达不确定性,Pujara 等人<sup>[88]</sup>用 PSL 进行知识图谱上的知识推理.类似上述基于 MLN 的方法,该方法引入抽取模板的置信度以及本体约束,但形式化这些规则为变量公式进行学习优化.Pujara 等人<sup>[89]</sup>进一步权衡推理速度和质量,提出分块并行推理策略,便于扩展到大规模知识图谱.该方法首先将关系和实体类型作为节点,本体约束(包括头实体取值域和尾实体取值域等)作为边,建立本体图;然后,采用聚类技术,例如边最小割进行分块;最后,每块用 PSL 框架独立推理.Chen 等人<sup>[90]</sup>则提出了本体路径发现算法 OP(ontological pathfinding),用发现的规则扩充知识图谱.OP 首先构建头实体取值域-尾实体取值域模式图,节点代表实体域,边代表关系,连接头尾实体域,限制搜索空间为满足关系头尾实体域约束的规则;再通过继承上位词的关系得到封闭模式图(如图 1 所示,Actor 关联的虚线边继承自上位词 Person);然后,根据最大实体数限制和最大规则数限制将其划分为更小的独立但可能有重叠的部分,在每个部分构建候选规则;最后,根据经验:度数很大的节点,其邻居之间互相连接的可能性很小,去掉含有度超过阈值实体的规则.

总的来说,面向知识图谱的知识推理可以借鉴传统的知识推理方法,特别是本体推理方法.当规则、统计特征、本体频繁模式、本体约束/路径有效时准确率高,现有的典型知识图谱内部由于高准确率的要求,大多采用这些传统的推理方法.但无论是规则还是抽象层面的本体约束,都需要实例化,可计算性比较差,对于实例数量很大的知识图谱而言,代价很高.另一方面,有效并且覆盖率广的规则和本体约束难以获得,导致推理结果的召回率通常比较低.而统计特征过分依赖已有数据,不易迁移,难以处理样本稀疏的情况;并且,当数据存在噪声时,抽取的特征甚至误导推理.因此,面向知识图谱的知识推理逐渐发展出独有的具体推理方法.

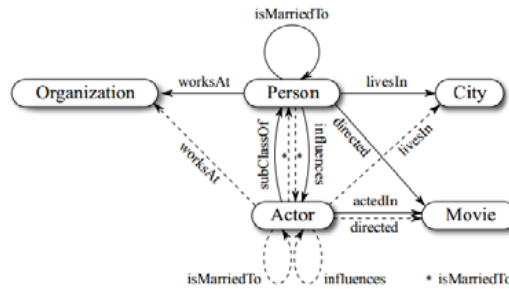


Fig.1 An example of schema closure graph, dashed arrows indicate inherited edges<sup>[90]</sup>

图 1 封闭模式图实例,虚线箭头为继承边<sup>[90]</sup>

### 3 单步推理

单步推理是指用直接关系即知识图谱中的事实元组进行学习和推理,根据所用方法的不同,具体可分为基于规则的推理、基于分布式表示的推理、基于神经网络的推理以及混合推理。

其中,基于规则的单步推理主要借鉴传统的规则推理方法,具体如第 2.1 节所示。鉴于这些方法的可计算性差、代价高等问题,基于分布式表示的推理、基于神经网络的推理以及混合推理得到了更广泛的关注和发展。

#### 3.1 基于分布式表示的推理

单步推理中,基于分布式表示的推理首先通过表示模型学习知识图谱中的事实元组,得到知识图谱的低维向量表示;然后,将推理预测转化为基于表示模型的简单向量操作。基于分布式表示的单步推理包括基于转移、基于张量/矩阵分解和基于空间分布等多类方法。

##### (1) 基于转移的表示推理

基于转移的表示推理根据转移假设设计得分函数,衡量多元组有效的可能性,得分越高,多元组越可能有效。也即正例元组的得分高,负例元组的得分低。由于关系数量相对较少,负例常常通过替换头实体或尾实体得到(部分工作也替换关系构建负例)。基于上述原则建模知识图谱中的事实元组及其对应的负例元组,最小化基于得分函数的损失,得到实体和关系的向量表示。推理预测时,选取与给定元素形成的多元组得分高的实体/关系作为预测结果。基本的转移假设将关系看成实体间的转移,后续发展出更复杂的转移假设,将关系看成经过某种映射后的实体之间的转移。基本转移假设的提出者 Bordes 等人<sup>[25]</sup>提出了第一个基于转移的表示模型 TransE,掀起了 Trans 系列的研究热潮。TransE 的主要思想是:如果三元组(头实体,关系,尾实体)成立,头实体向量  $h$  与关系向量  $r$  的和与尾实体向量  $t$  相近,否则远离。由上述基本转移假设得到得分函数  $-\|r + h - t\|_{L_1/L_2}$ ,即用  $L_1$  或  $L_2$  范数衡量距离。学习过程替换头实体或尾实体得到负例,类似支持向量机,最小化一个基于 Margin 的损失,使正例的得分比负例的得分至少高一个 Margin。在进行推理时,得分函数取值大的候选实体/关系即为推理结果。TransE 简单有效,但存在一些不足。

- 首先,TransE 严格要求有效的三元组满足头实体加关系在向量空间中与尾实体足够靠近,可以很好地处理一对一关系,但在处理多映射属性关系时,存在多个实体竞争一个点的现象。例如在一对多关系中,同一个头实体  $h$  与多个尾实体  $\{t_1, t_2, \dots, t_x\}$  存在关系  $r$ ,由于  $h + r \approx t_i, i = \{1, 2, \dots, x\}$  的约束,这些尾实体将竞争空间中的同一个点,即使它们语义差别很大,因而会造成向量空间的拥挤和误差;
- 其次,TransE 只专注于满足知识图谱中的三元组约束,然而知识图谱中存在大量层级关系,例如在知识图谱 WordNet<sup>[91]</sup>的子集 WN18 中,大约有 50%的层级关系,孩子关系就是一种层级关系,对应的实体和关系形成树结构。Li 等人<sup>[92]</sup>也指出,考虑层级结构有助于推理;
- 然后,TransE 没有考虑丰富的语义信息,缺乏对空间中向量分布位置的进一步调整;
- 再者,TransE 在单个知识图谱上进行学习推理,而单个知识图谱知识量有限;
- 此外,TransE 参数的选取与知识图谱数据独立,不能反映数据的特点;



- 同时,TransE 未考虑知识的时间约束。

针对 TransE 的上述缺陷与不足,学术界开展了一系列的改进工作。

针对处理多映射属性关系的不足,Wang 等人<sup>[19]</sup>提出了 TransH,在 TransE 的基础上为每个关系多学一个映射向量,用于将实体映射到关系指定的超平面;然后在该超平面,与 TransE 一样,关系表示向量看成映射后的实体之间的转移。映射向量使得对于不同关系,同一个实体在不同关系指定的超平面有不同的表示,一定程度上缓解了不能很好地处理多映射属性关系的问题。进一步地,Wen 等人<sup>[93]</sup>提出了  $m$ -TransH,直接建模多元关系,关系的损失函数不求助于分解为二元关系进行处理,而是将多元关系中的每一元看成一个角色,直接建模为角色函数的累加。Lin 等人<sup>[13]</sup>则提出了 TransR,在单独的实体空间和关系空间建立实体和关系的表示,每个关系对应一个空间,有一个映射矩阵,实体通过对应的映射矩阵映射到关系空间后,将关系向量看成实体向量间的转移。考虑到关系还可以进行更细致的划分,Lin 等人<sup>[13]</sup>提出了 CTransR,将关系划分为关系组,为每个关系组学习一个关系向量和映射矩阵。在此基础上,Ji 等人<sup>[94]</sup>认为,映射矩阵不仅与关系相关,还与实体相关,因此改进了 TransR,提出了更细粒度的 TransD,考虑实体和关系之间的交互。TransD 中,实体和关系均用两个向量表示,第 1 个指示含义,第 2 个动态构建映射矩阵。由此,实体的映射矩阵为对应关系和实体的第 2 个向量的乘积,同时考虑实体和关系的多样性。Fan 等人<sup>[95]</sup>提出了 TransM,直接根据关系的映射属性预先计算每个训练三元组的权重,用于加权损失函数。TransM 考虑了不同关系映射属性三元组的贡献差别,认为一对一映射属性的关系三元组贡献最大。然而,这些模型忽略了知识图谱的异质性(关系连接的实体对数目差异很大)和不均衡性(关系中头尾实体的数目差异很大)。为此,Ji 等人<sup>[96]</sup>提出了 TransSparse,同样关系看成实体之间的转移,但用自适应的稀疏转移矩阵替换一般的转移矩阵。为了克服关系之间的异质性,Ji 等人<sup>[96]</sup>提出了 TransSparse(share),转移矩阵的稀疏度由关系连接的实体对数目决定,头尾实体共享相同的转移矩阵。复杂关系的转移矩阵比简单关系更稠密。为了解决关系内部头尾实体的不均衡性问题,Ji 等人<sup>[96]</sup>提出了 TransSparse(separate),每个关系有两个单独的稀疏转移矩阵,头尾实体各一个,稀疏度由关系连接的头/尾实体数目决定。

针对未考虑层级关系的不足,Hu 等人<sup>[97]</sup>提出了实体的层级表示。具体地,基于知识图谱构建实体层级结构,实体作为叶节点,上层概念作为内部节点,构成有向无环图。其中,每个内部节点关联一个需要学习的距离度量矩阵。基于目标:为每个实体  $e$  找到一个有助于预测上下文实体(文本语料中出现在  $e$  的上下文中的实体)的表示,优化一个基于实体之间距离的目标函数。实体层级结构的应用正体现在距离的计算上。具体地,在实体层级结构图上找到两个实体之间的所有路径,进而得到所有相关的内部节点集,距离正比于所有内部节点距离度量矩阵的加权和。进一步地,Xie 等人<sup>[23]</sup>提出了 TKRL(type-embodied knowledge representation learning)学习知识图谱实体和关系的表示,将层级类型信息用于映射矩阵、训练时负例的选择和评估时候选的过滤。TKRL 同样根据知识图谱构建层级结构,内部节点也关联一个矩阵,称为映射矩阵。实体的映射矩阵通过组合实体对应的层级结构上一级一级的内部节点的映射矩阵得到,组合可以是连乘形式或者加权和的形式。头尾实体再根据各自对应的组合映射矩阵映射到对应的关系类型空间,与关系向量进行转移。训练过程中,在替换正例的实体构造负例时,提高与正例实体属于同一个内部节点(类型)的实体的选取概率。测试时,移除不属于关系特定的实体类型的候选实体。

针对未考虑丰富语义信息的不足,Guo 等人<sup>[98]</sup>提出了 SSE(semantic smooth embedding)学习知识图谱实体和关系的表示,利用实体语义类信息强制表示空间几何结构语义平滑。直观地,属于相同语义类的实体在表示空间中邻近,实体语义类信息作为损失函数的正则项。其中,实体的语义类为知识图谱中以该实体为头实体,“Generalization”关系对应的尾实体。该方法可用于任何表示模型。正则项的引入是比较浅层的考虑语义平滑方式,更深层次地,Nguyen 等人<sup>[24]</sup>提出了 TransE-NMM(TransE-neighborhood mixture modeling),在 TransE 的基础上定义基于邻居的实体表示,引入邻居实体信息进行实体和关系的表示学习。具体地,TransE-NMM 引入逆关系三元组,即,对知识图谱中的每个  $(h, r, t)$  三元组添加  $(t, r^{-1}, h)$  三元组。由此定义实体的邻居为以该实体为尾实体的所有头实体-关系对。基于邻居的向量为每个头实体-关系对的头实体向量加上关系向量的带权和。该向量与带权实体向量相加之和即为实体的最终表示,用于 TransE 模型。



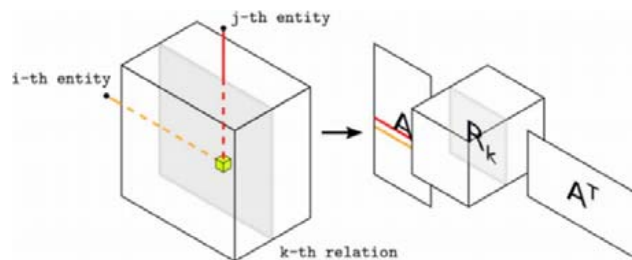
针对只在单个知识图谱上进行学习推理的不足,Wang 等人<sup>[99]</sup>在相同的连续向量空间共同表示知识图谱的实体和文本语料的词,提出包含 3 部分(知识模型、文本模型和对齐模型)概率和的模型,同时保持知识图谱中实体之间的关系和文本语料中词的共现.知识模型与文本模型都通过概率模型实现,即:在知识模型中,最大化知识图谱三元组存在的概率,转化为给定其中的任意两个元素,另一个元素的概率和.类似地,在文本模型中,最大化词的共现概率.其中,对齐通过实体名或维基百科锚文本实现.由于实体名通常很短并且有歧义,而锚文本依赖于数据源并且数量有限,该对齐机制效果受到限制.Zhong 等人<sup>[100]</sup>进一步扩展了该模型,基于实体的描述信息提出一个新的对齐模型,要求实体的表示向量不仅满足三元组约束,而且与从实体描述计算得到的表示向量相等.然而在实际知识图谱中,许多实体缺失实体描述.因此,Wang 等人<sup>[101]</sup>提出了 TEKE(text-enhanced knowledge embedding),引入文本语料中丰富的上下文信息扩展知识图谱的语义结构,学习知识图谱实体和关系的表示. TEKE 通过标注文本语料中的知识图谱实体,构建实体和词的指定共现窗口的共现网络,将文本语料和知识图谱联系起来.基于共现网络,可以得到实体和关系的文本上下文.其中,实体的文本上下文为实体的邻居(出现在共现窗口内的词),关系的文本上下文为头实体和尾实体的共同邻居.实体和关系的表示形式化为对应文本上下文的线性函数.由此,每个关系对不同头尾实体允许有不同的表示.不同于以上引入外部文本的工作,Cai 等人<sup>[96]</sup>提出了跨知识图谱的表示方法 cross-KG,同时学习两个不同知识图谱的表示.通过映射语义相关的两个知识图谱中的实体和关系到统一的语义空间,cross-KG 借助更大和更稠密知识图谱的知识,促进稀疏知识图谱的表示学习,提升连接预测效果.

针对参数与时间因素方面的不足,Jia 等人<sup>[102]</sup>提出了 TransA,解决目前表示方法存在的两个问题:最优损失函数的参数在一个给定的候选集中选择;并且,不同实体和关系的两个知识图谱共享相同的候选集,忽略了不同知识图谱的局部性.TransA 将 Margin 参数分为实体和关系指定的局部 Margin 的加权和.Margin 和实体、关系的表示相关,不同知识图谱自适应学习,而不是从指定的候选集选取.鉴于大部分三元组只在给定的时间区间有效,Jiang 等人<sup>[103]</sup>提出新颖的时间感知知识图谱补全模型 TAE(time-aware embedding),用三元组和时间信息预测知识图谱中的连接,即:给定三元组中的两个元素与时间区间,预测另一个元素.为了引入三元组的发生时间,Jiang 等人<sup>[103]</sup>编码时间序信息为表示空间几何结构上的正则.具体定义时间序关系对 $(r_1, r_2)$ 为共享相同头实体的时间序关系,例如 $\langle wasBornIn, diedIn \rangle$ ,认为其中更早发生的关系  $r_1$  可以通过一个时间演化矩阵演化到更晚发生的关系  $r_2$ .由此,时间序关系对的得分函数为  $r_1$  对应向量乘以时间演化矩阵得到的向量与  $r_2$  对应向量之间的相似度.对于三元组,采用基于转移的表示模型,如 TransE 的得分函数.学习过程同时考虑三元组和时间序关系对.为了引入有效时间,Jiang 等人<sup>[103]</sup>提出了基于整型线性规划(integer linear programming,简称 ILP)的推理模型.目标函数基于三元组的得分函数,用时间一致性约束作为约束条件.进一步集成以上两个模型,即 ILP 的目标函数替换为基于 TAE,获得更好的推理效果.Tay 等人<sup>[104]</sup>提出了自适应的鲁棒转移模型 puTransE(parallel universe TransE)进行知识图谱实体和关系的表示学习,puTransE 产生多个表示空间,每个表示空间对应一个采样的关系和先后通过语义感知与结构感知选择机制得到的三元组集合.每个表示空间分别用 TransE 学习,其中,Margin 和学习率等参数随机选取.puTransE 采用分治策略结合每个平行表示空间学习的局部信息进行全局表示学习.通过增加平行表示空间或使某些空间失效,puTransE 可以快速处理动态变化知识图谱的增或删除操作.

## (2) 基于张量/矩阵分解的表示推理

基于张量/矩阵分解的表示推理将(头实体,关系,尾实体)三元组看成张量/矩阵中的元素构建张量/矩阵,通过张量/矩阵分解方法进行表示学习.分解得到的向量表示相乘重构成张量/矩阵,元素值即为对应三元组有效与否的得分,可以认为得分大于特定阈值的三元组有效,或候选预测按照得分排序,选择得分高的候选作为推理结果.该类方法的典型代表是 Nickel 等人<sup>[105]</sup>提出的 RESCAL,基于三阶张量进行表示学习,模型如图 2 所示.如果三元组成立,三阶张量上对应的元素值为 1,否则为 0.例如,图 2 中第  $i$  个实体和第  $j$  个实体存在第  $k$  种关系,对应的位置 $(i, j, k)$ 元素值为 1.该三阶张量对每种关系进行切片分解,其中,对第  $k$  种关系的切片,如图 2 所示,可以分解为实体表示矩阵  $A$ 、第  $k$  种关系对应的非对称矩阵  $R_k$  与实体表示矩阵的转置  $A^T$  的连乘.通过最小化重构误差学习实体和关系的表示.张量分解模型虽然推理准确率高,但内存占用量大,计算速度慢.为此,Chang 等

人<sup>[106]</sup>提出了 TRESICAL,在 RESCAL 的基础上引入实体类型信息这一关系域知识,在损失函数的计算中排除不满足关系特定的实体类型约束的三元组,加速计算.Nickel 等人<sup>[107]</sup>则提出了新的张量分解模型 ARE(additive relational effects)学习知识图谱三元组的隐性和观察到的模式,用一个附加项增广 RESCAL 模型(隐性模式),对应观察到的模式.这里,观察到的模式是指用可观察的关系学习方法,例如规则方法等,得到的预测结果构成的三阶张量.附加项为该三阶张量乘以一个权重向量,权重向量衡量关系学习方法对各个关系的预测能力.该附加项通过减少不连接部分,降低 RESCAL 分解需要的阶.不同地,He 等人<sup>[108]</sup>直接处理知识图谱的稀疏性问题,并集成不同知识图谱的补全,同时进行,利用不同知识图谱的互补性进行各自的推理.方法将知识图谱补全问题看成大的矩阵分解任务,类比推荐系统的用户-电影评分矩阵,构建关系-实体对矩阵,矩阵元素值对应某实体对存在某关系的概率.当实体对存在关系时,对应的元素值初始化为 1.由于关系对应的平均三元组数相比于总的实体对数很小,矩阵很稀疏.为了减少稀疏性,利用关系和实体间的类型一致性约束初始化矩阵中的错误元素,即,设置类型不一致三元组对应的元素为 0.同时,在矩阵分解中引入集成的不同知识图谱中关系的相似性,构成关系相似性正则,即:如果来自不同知识图谱的两关系相似,则两关系在特征空间的隐性特征表示接近,充分利用不同知识图谱的互补性.为了更好地处理对称和反对称关系,Trouillon 等人<sup>[21]</sup>提出了基于复数值向量表示的矩阵分解方法进行连接预测,关系矩阵的分解形式化为对称矩阵和反对称矩阵的和.这些张量/矩阵分解模型不能快速处理动态变化的知识图谱.为此,Tay<sup>[109]</sup>提出了基于张量分解模型 RESCAL<sup>[105]</sup>的可扩展集成框架 RSTE(random semantic tensor ensemble).RSTE 采用分治策略,从知识图谱中采样多个多样的更小规模子图张量,通过集成子图张量的 RESCAL 分解进行连接预测.RSTE 极大地降低了内存占用和运行时间,同时,通过增加子图张量的分解或使某些子图张量分解得到的结果无效,可以快速处理动态变化知识图谱的增或删除操作.

Fig.2 RESCAL model<sup>[105]</sup>图2 RESCAL 模型<sup>[105]</sup>

### (3) 基于空间分布的表示推理

基于空间分布的表示推理建立模型拟合知识图谱中实体和关系的空间分布特征,使得在向量表示空间中,实体和关系的空间分布尽可能地与原知识图谱一致.该类方法通过设计对应的能够反映空间分布特征的得分函数,与简单地采用基于转移假设得分函数的基于转移方法区别开来,但采用与基于转移方法类似的学习和推理过程. Xiao 等人<sup>[110]</sup>提出了高斯混合模型 TransG,第一次从产生式的角度看待表示学习.例如,TransE 的产生式过程为尾头实体向量之差符合以关系为均值、单位阵为协方差的高斯分布.TransG 自动发现关系的隐性语义,利用关系不同隐性语义向量的混合转移头尾实体对,建模三元组,对应的产生式过程为尾头实体向量之差符合以每个关系隐性语义向量为均值、单位阵为协方差的高斯分布的加权和.由此,TransG 中三元组的得分函数为关系对应的每个隐性语义向量用 TransE 的得分函数作用后的加权和.基于得分函数的指数目标函数使得 TransG 可以学习出关系在对应三元组中最主要的隐性语义向量,指数作用拉大了其与关系的其他语义向量对三元组的贡献差距.He 等人<sup>[111]</sup>提出了另一个高斯模型 KG2E,转向基于密度的表示,直接建模实体和关系的确定性,在多维高斯分布的空间中学习知识图谱的表示.每个实体/关系用一个高斯分布表示,均值指示位置,协方差恰当地传达了确定性.三元组的得分函数基于实体转移分布和关系分布的 KL 散度计算.KL 散度的非对称性使得 KG2E 可以有效处理多种类型关系.Xiao 等人<sup>[20]</sup>提出了基于流形的表示模型 ManifoldE(manifold-based

embedding),扩展了三元组的位置从向量空间中的一个点到一个流形结构,扩展了基于转移的头实体向量加关系向量等于尾实体向量到基于轨道的流形函数,取得了很好的连接预测效果,是目前除了神经网络方法以外最有效的方法。

总体上看,基于分布式表示的单步推理研究工作比较全面,然而建模时通常只考虑满足知识图谱事实元组的约束,未考虑更深入的组合语义信息,推理能力受限。其中,基于转移的表示推理以 TransE<sup>[25]</sup>为代表,简单、有效,计算效率高,这也引发了广泛关注,许多研究者相继深入其中,涌现出一系列的方法,形成了 Trans 系列。因此,基于转移的表示推理研究工作相对比较成熟,相应地,其继续发展的空间也就比较小,趋于饱和状态。基于张量/矩阵分解的表示推理以 RESCAL<sup>[105]</sup>为代表,虽然可解释性强,性能相对较好,但时间复杂度较高,难以推广应用,这成为一个很大的瓶颈。学术界对这一类方法的关注和研究也相对较少。为此,基于张量/矩阵分解的表示推理还有一定的发展空间,特别是复杂度相对较低的基于矩阵分解的方法,但高复杂度的本质问题对发展有很大的局限性。基于空间分布的表示推理很好地拟合了知识图谱中三元组的空间分布,推理能力强,一定程度上还能解决 TransE 的多个实体竞争空间中一个点的拥挤问题。并且,目前该类方法的研究还比较少,处于研究初期阶段,鉴于其相对较低的复杂度和相对较高的有效性,还有较大的发展空间。但该类方法的创新性要求比较高,Trans 系列基于问题改进的发展路线不太适用于该类方法,需要引入更多创新性的方法,如数学领域的方法,建模空间分布。

### 3.2 基于神经网络的推理

单步推理中,基于神经网络的推理利用神经网络直接建模知识图谱事实元组,得到事实元组元素的向量表示,用于进一步的推理。该类方法依然是一种基于得分函数的方法,区别于其他方法,整个网络构成一个得分函数,神经网络的输出即为得分值。Socher 等人<sup>[112]</sup>提出了神经张量网络 NTN(neural tensor network),用双线性张量层代替传统的神经网络层,在不同的维度下,将头实体和尾实体联系起来,刻画实体间复杂的语义联系。其中,实体的向量表示通过词向量的平均得到,充分利用词向量构建实体表示。具体地,每个三元组用关系特定的神经网络学习,头尾实体作为输入,与关系张量构成双线性张量积,进行三阶交互,同时建模头尾实体和关系的二阶交互。最后,模型返回三元组的置信度,即:如果头尾实体之间存在该特定关系,返回高的得分;否则,返回低的得分。特别地,关系特定的三阶张量的每个切片对应一种不同的语义类型,一种关系多个切片可以更好地建模该关系下不同实体间的不同语义关系。Chen 等人<sup>[113]</sup>引入类似的神经张量网络模型预测知识图谱中新的关系,通过用从文本无监督学到的词向量初始化实体表示提升模型,甚至可以预测知识图谱中未出现实体的关系。近期,Shi 和 Wenginger<sup>[114]</sup>提出了共享变量神经网络模型 ProjE。ProjE 通过简单的组合操作组合三元组的已知部分(头实体与关系或尾实体与关系)建立目标向量空间,并映射未知部分(尾实体或头实体)的候选集到相同的空间,用基于候选的排序损失学习模型。相比于普遍采用的转移矩阵,组合操作减少了大量参数。进一步通过候选采样,处理大规模知识图谱。

总而言之,基于神经网络的单步推理试图利用神经网络强大的学习能力建模知识图谱事实元组,获得很好的推理能力和泛化能力。然而,神经网络固有的可解释性问题也依然存在于知识图谱的应用中,如何恰当地解释神经网络的推理能力是一大难点。目前,基于神经网络的单步推理研究工作还比较少,但神经网络的高表达能力及其应用于其他领域,包括图像处理、文本处理,特别是和知识图谱结构比较类似的社交网络等图结构数据领域的突出表现和高性能,使得该方向的研究前景广阔。如何扩展其他领域中更多基于神经网络的方法到知识图谱领域,成为未来要深入研究的问题。一般图结构数据,例如社交网络的表示和推理学的是节点,而知识图谱的表示和推理关注的是节点(实体)和边(关系)。因此,从一般图结构数据基于神经网络的方法迁移到知识图谱将是一个相对比较简单突破口。与此同时,关于神经网络可解释性问题的研究也有待进一步开展。

### 3.3 混合推理

单步推理中,混合推理通过混合多种单步推理方法,充分利用不同方法的优势,例如基于规则推理的高准确率、基于分布式表示推理的强计算能力、基于神经网络推理的强学习能力和泛化能力。混合单步推理包括混合



规则与分布式表示以及混合神经网络与分布式表示的推理.

#### (1) 混合规则与分布式表示的推理

在分布式表示辅助规则发现方面,一些传统的推理规则发现方法通过计算关系之间的分布式相似度实现,其中,关系表示为对应实例的特征向量.然而,这些方法没有考虑特定的上下文,并且独立建模关系,忽略了关系之间的依赖.为了解决这两个问题,Han 等人<sup>[115]</sup>提出了上下文敏感的推理规则发现方法.首先构建关系图(如图3所示),抽象关系元组(如  $X \text{ buy } Y$ )和实例化特征(如  $X=\text{Facebook}$  和  $Y=\text{WhatsApp}$ )作为节点,边代表抽象关系元组和实例化特征的共现(如  $X \text{ buy } Y$  和  $X=\text{Facebook}$  之间的连边)或抽象关系元组之间的语义依赖关系(同义词、上位词等,如  $X \text{ buy } Y$  和  $X \text{ purchase } Y$  之间的连边).然后,基于关系图学习特定上下文的关系表示(如图3中关系  $\text{acquire}$  对特定上下文( $\text{Google}, \text{YouTube}$ )和( $\text{children}, \text{skill}$ )的表示将不同),形式化为抽象关系元组与对应特征上下文敏感的相关性得分的拼接向量,相关性得分通过关系元组到特征的重启型随机行走计算.最后,计算关系向量之间的相似度,大于阈值的关系对及对应上下文形成推理规则.这里,上下文信息的利用体现在随机行走重启概率的计算上,对于上下文无关(相关性通过词相似度计算)的特征节点设置高的重启概率,反之设置低的重启概率.

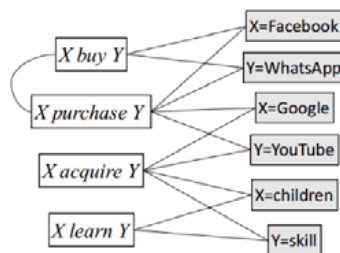


Fig.3 A small example of relation graph<sup>[115]</sup>

图3 关系图实例<sup>[115]</sup>

在规则辅助基于分布式表示的推理方面,Wang 等人<sup>[116]</sup>在表示模型中无缝嵌入逻辑规则和物理规则,形式化推理为整型线性规划问题(ILP).物理规则包括关系对应的实体类型限制、关系对应的实体数目限制等.逻辑规则:若关系  $r_1$  可以推出关系  $r_2$ ,则关系  $r_1$  连接的两个实体也被关系  $r_2$  连接.通过 ILP 集成表示模型和规则.目标函数基于三元组的得分函数,可以用上述任何表示模型的得分函数,而每个规则形式化为一个需要满足的条件.Guo 等人<sup>[117]</sup>编码领域知识为简单的规则,进行增强规则的 TransE 知识图谱补全.方法首先训练 TransE 模型,然后结合规则(如首都这一关系的头实体必须是国家实体)调整 TransE 的推理结果.这些方法中的规则需要实例化,即,替换规则中的变量为知识图谱中的实体.为了避开高代价的实例化,Demeester 等人<sup>[118]</sup>提出直接在关系表示中强加一阶限制作为隐性正则项,基于序表示的概念,在向量表示中直接捕捉偏序关系.具体地,强制关系向量的每一维非负,如果关系  $r_1$  能够推出  $r_2$ ,则令  $r_1$  向量的每一维小于  $r_2$  向量的对应维.该偏序限制将得到全局一致的关系表示,即:如果进一步  $r_2$  能够推出  $r_3$ ,则  $r_1$  能够推出  $r_3$ ,刚好对应  $r_1$  向量小于  $r_2$  向量,也小于  $r_3$ .

#### (2) 混合神经网络与分布式表示的推理

混合神经网络与分布式表示的推理有两种混合模式:一种是用神经网络方法建模外部信息,例如相关的外部文本、实体描述等,表示模型建模知识图谱中的三元组;另一种是用神经网络方法建模知识图谱,其输出进一步用于表示模型.Toutanova 等人<sup>[119]</sup>捕捉文本关系的组成结构,共同建模文本关系三元组和知识图谱三元组.其中,文本关系为对实体对描述进行主、谓、宾等的标注得到的依赖路径,与对应的实体对构成文本关系三元组.采用一个卷积神经网络学习文本关系,得到其向量表示.这些文本关系三元组和知识图谱三元组通过相同的表示学习模型一起训练,允许信息源之间更深层的交互.模型将实体对描述作为三元组的附加信息,只在训练过程中进行学习,调整向量表示,不能处理预测时实体只有实体描述而不存在知识图谱三元组的情况.Xie 等人<sup>[120]</sup>则提出了 DKRL(description-embodied knowledge representation learning),结合三元组和实体描述学习知识表示,直接从实体描述建立向量表示,可以处理实体只有实体描述的情况.三元组的得分函数由两部分组成:第1部分和



TransE 中三元组的得分函数一样;第 2 部分是在 TransE 得分函数基础上替换头实体或尾实体或同时替换头尾实体向量为基于对应实体描述的实体向量,得到的 3 个得分函数之和.基于实体描述的实体向量用卷积神经网络学习.由于得分函数由两部分 4 项组成,当实体未出现在知识图谱的训练集中时,可以忽略需要该直接实体表示的项,并根据实体描述得到基于实体描述的实体表示,带入相应的项,计算得分,进行推理预测.Schlichtkrull 等人<sup>[121]</sup>引入关系图卷积神经网络(relational graph convolutional network,简称 R-GCN)建模知识图谱.模型可以看成自动编码器,包括一个编码器和一个解码器.编码器是一个 R-GCN,实体与知识图谱中的邻居实体进行卷积学习,产生实体的隐性特征向量表示,输入解码器;解码器是一个分解模型——DistMult<sup>[122]</sup>,得分函数为头实体向量的转置,关系特定的对角矩阵和尾实体向量的乘积,由此引入关系向量的建模.最后,模型通过交叉熵损失学习.

大体上,混合单步推理通过混合不同单步推理方法实现优势互补.然而,目前的混合单步推理还停留在两种方法的浅层混合,即:以一种方法为主,另一种为辅的推理,尚缺乏更深层次的混合模式以充分利用各方法的优势.同时,混合推理方法不同于其他具体的某类方法,它更像是一种策略,主要基于对各类方法的透彻分析找到优势互补的方法,进行混合,可创新的点在选择哪些方法进行混合以及混合模式上.其中,混合规则与分布式表示的单步推理,混合策略主要是以一方为主,以另一方为辅,同等考虑各个混合方法的更深层次的混合模式有待进一步研究.而混合神经网络与分布式表示的单步推理主要有两种思路:一种是用现有的表示模型学习知识图谱的三元组,用神经网络学习相关的文本语料;另一种是直接神经网络建模知识图谱,同时集成表示学习模型.第 2 种思路的研究目前尚少,也比较新,鉴于神经网络在各方面的突出性能,这将成为未来研究的一大热点.

## 4 多步推理

多步推理是在单步推理建模直接关系的基础上进一步建模间接关系,即多步关系.多步关系是一种传递性约束,例如以下两步关系的例子: $a$  和  $b$  存在关系  $r_1$ ,  $b$  和  $c$  存在关系  $r_2$ ,该两步路径对应的直接关系是  $a$  和  $c$  存在关系  $r_3$ .多步关系的引入,建模了更多信息,往往比单步推理效果更好.多步推理按不同的推理方法划分,同样分为基于规则的推理、基于分布式表示的推理、基于神经网络的推理以及混合推理.

### 4.1 基于规则的推理

多步推理中,基于规则的推理不同于基于规则的单步推理,后者用到的是类似关系  $r_1$  推出关系  $r_2$  的简单经验规则或一些基于统计的频繁模式.而多步推理用到的规则更为复杂,如传递性规则.鉴于人工获取有效且覆盖率广的传递性规则代价比较高,这些规则一般通过挖掘的实体间路径来近似.根据是否引入局部结构,基于规则的多步推理可分为基于全局结构和引入局部结构的规则推理两种.

#### (1) 基于全局结构的规则推理

基于全局结构的规则推理在整个知识图谱上进行路径挖掘,将一些路径近似地看成规则.实体间的路径进一步作为判断实体间是否存在指定关系的特征训练学习模型.这类工作的一大部分内容是基于随机行走的规则挖掘.Lao 等人<sup>[123]</sup>提出了 PRA 算法(path ranking algorithm),将路径作为特征,预测实体间是否存在指定关系. PRA 首先确定要学习的目标关系;然后找出目标关系的正例三元组,替换头/尾实体得到负例三元组;再构造特征集合,将这些三元组中两个实体之间的一条路径作为一个特征;接着,根据随机行走的思想计算路径的特征值,构成每个三元组的特征向量,每维对应一个特征的特征值;最后,用这些正负例三元组对应的特征向量训练 logistic 回归分类器.Lao 等人<sup>[124]</sup>进一步提出了知识图谱上基于受限和加权随机行走的推理,与上述基本 PRA 算法的不同之处在于路径的产生过程.基本 PRA 路径通过枚举产生,该方法提出数据导向的路径发现算法,试图只产生潜在的有助于推理的路径,解决枚举对大规模知识图谱的不适用性.具体地,路径途径的实体节点至少出现在一定比例的训练数据中,同时限制路径的长度,认为短路径更有助于推理,进一步减少路径的数量.同样,两个实体之间所有路径的加权概率和作为得分衡量两个实体存在某种关系的可能性.当实体之间没有路径关联时, PRA 失效.为此,赵泽亚等人<sup>[125]</sup>结合内容信息和结构信息,提出了内容-结构推理方法 CSRI(content-structural relation inference),预测与给定实体具有给定关系的目标实体. CSRI 首先通过广度优先找到从给定实体出发三

步以内的实体,和与给定实体具有相同属性的实体一起构成候选实体集;然后,对于每个候选实体,计算概率得分,选出得分前  $N$  的实体作为推理结果.概率得分由内容影响因子和结构影响因子两部分组成.内容影响因子取以下两部分的最大值:给定实体与候选实体的属性相似度以及二者组成的实体对与已知存在给定关系的实体对之间的属性相似度.结构影响因子通过 PRA 算法计算.两部分的加权系数通过训练集上的最大似然估计得到.当知识图谱稀疏时,实体之间的路径很少,甚至不存在路径,不能直接用 PRA 算法预测关系是否存在,另一种处理方式是先用外部知识增广知识图谱.为此,Kotnis 等人<sup>[126]</sup>提出用从 5 亿网络文本语料挖掘的桥梁实体和边(关系)增广知识图谱,然后用 PRA 在增广的知识图谱上进行知识推理.具体地,通过限制深度的深度优先搜索挖掘连接两个实体的名词短语作为桥梁实体,进而得到相应的关系连边.图 4 给出了一个利用桥梁实体和附带的边帮助推理的实例,原知识图谱中只包含两个实体:Yankees 和 Base Ball,没有边,通过桥梁实体 Brian McCain 和附带的两条边,帮助 PRA 算法推理缺失的三元组(Yankees,teamPlaysSport,Base Ball).这些基于 PRA 的工作通常采用单任务学习方式,每个关系用自己的训练数据独立学习,忽略了特定关系之间的丰富关联.并且,对于不频繁的关系,无法得到足够的训练数据.为了解决这些问题,Wang 等人<sup>[127]</sup>提出了多任务学习框架 CPRA(coupled PRA).CPRA 首先引入基于对应的共同路径数的关系相似度度量,相应地,设计凝聚聚类策略,自动发现相互高关联的关系.然后,采用多任务学习策略有效耦合聚到同一组关系的预测.具体地,通过半共享参数的分类器耦合关系,即,通过耦合系数结合关系特定的参数和共享参数.由此引入关系关联,允许它们之间隐性信息的共享.基于 PRA 的方法可以自动挖掘路径,但却不能很好地确保路径的有效性,甚至引入了噪声,误导了推理.为此,Wei 等人<sup>[128]</sup>提出了目标导向的推理公式挖掘算法,通过指定每步的推理目标引导随机行走.算法更倾向于访问有助于推理目标的结构,因此可以提高随机行走的效率并且避免噪声.具体地,算法最大化可以推理头实体到尾实体路径的概率,同时,不能推理到尾实体或引入噪声的路径概率要小.然而,这些方法都基于静态知识图谱进行推理,没有考虑丰富的时空信息,为此,Jia 等人<sup>[129]</sup>提出了开放知识网络 OpenKN,是一种面向网络大数据的、开放的、自适应的且可演化的知识网络.OpenKN 将知识图谱形式化为八元组,除了大多数知识图谱存储的三元组信息外,还包含实体类型信息以及实体和关系的时空信息.在此基础上,针对时序动态知识推理,赵泽亚等人<sup>[130]</sup>基于 PRA 提出了时间差关系路径方法 TDLP(time-difference-labeled path),将关系的时间信息融入结构路径,预测给定实体对之间是否会产生给定关系,以及何时产生.时间差关系路径是指时间差增广的抽象关系路径,即,替换一般关系路径上的实体为抽象概念,路径上的每一步对应一个时间,取值为关系路径对应的直接关系三元组的产生时间与路径上该步三元组的产生时间之差.训练时,针对每一个待预测的关系及对应的实体对,TDLP 首先从网络中找出所有时间差关系路径,将其看成该待预测关系的不同特征,与 PRA 一样,通过 logistic 回归学习特征权重.测试时,对于每一个候选时间,匹配目标关系特征集中的时间差关系路径,若匹配成功,则根据随机行走思想计算特征值;否则,对应的特征值为 0,通过逻辑回归模型计算最终得分.如果最大得分大于阈值,则推理出该实体对会在最大得分对应的时间产生待预测的关系,否则不产生该关系.

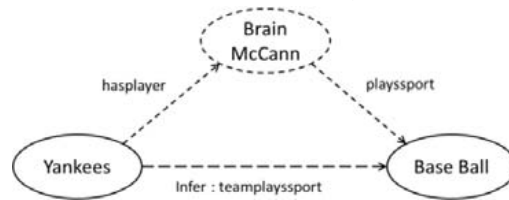


Fig.4 An example showing how bridging entity and the edges incident on it can help the reasoning<sup>[126]</sup>

图 4 桥梁实体和附带的边帮助推理的实例<sup>[126]</sup>

基于全局结构的规则推理除了随机行走系列,还可以用知识图谱上的其他规则挖掘模型获取规则,包括整型线性规划、关联规则挖掘等.Berant 等人<sup>[131]</sup>提出了利用传递约束学习全局最优蕴含规则集.算法定义蕴含图,节点为关系元组(如  $p(t_1, t_2)$ ,  $p$  为关系类型,  $t_1$  和  $t_2$  为头尾实体),边表示蕴含规则,形式化表达蕴含图为整型线性规划(ILP).其中,变量指示节点之间是否存在边,传递约束作为条件,最大化变量的加权和.同时引入图分解

和增量 ILP 扩展技术扩展到更大规模的图.图分解得到子图,独立学习,最后合并子图得到的结果.增量 ILP 不提前指定传递约束,而是在传递约束被违背时添加传递得到的边.如此迭代,直至没有传递约束被违背.Galárraga 等人<sup>[132]</sup>提出了**知识图谱关联规则挖掘算法 AMIE(association rule mining under incomplete evidence)**,挖掘**传递性规则**.AMIE 维护一个规则队列,初始化为空规则,迭代地每次从规则队列中取出一个规则,如果规则是封闭规则且未被删除,则输出该规则.同时,该规则用定义的规则挖掘操作算子作用,将未被删除的规则插入规则队列.如此循环,直至规则队列为空.这里的删除条件是:如果  $B_1 \wedge \dots \wedge B_n \wedge B_{n+1} \Rightarrow H$  的置信度小于  $B_1 \wedge \dots \wedge B_n \Rightarrow H$ ,则删除前者,即,长规则并且置信度不够高将被丢弃.封闭规则指的是规则中的每个变量至少出现两次的规则,例如封闭规则  $hasChild(p,c) \wedge isCitizenOf(p,s) \Rightarrow isCitizenOf(c,s)$ ,变量  $p, c$  和  $s$  都至少出现两次.规则挖掘操作算子包括增加规则中的变量、实例化变量等.由于知识图谱的开放世界假设(open world assumption,简称 OWA),不在知识图谱中的元组不能看成是负例,算法主要的挑战是提供负例.为此,算法提出了部分完整假设(partial completeness assumption,简称 PCA),如果知识图谱存在实体的某个属性,那么知识图谱存在该实体的所有属性.由此,置信度用 PCA 计算,算法可以进一步返回置信度超过阈值的规则.在此基础上,Galárraga 等人<sup>[133]</sup>提出了**AMIE+**,优化 AMIE,采用剪枝和近似策略挖掘更大的知识图谱,利用类型信息和联合推理提高预测精度.利用类型信息是指考虑关系对应的头尾实体的取值类型,过滤掉不符合类型约束的候选.联合推理是指对于给定元素,如果多个规则都能推理到某个候选,则结合这几个规则的 PCA 置信度计算候选的得分,而不是取最大 PCA 置信度.当现有方法用规则推理知识图谱中新的事实元组时,没有考虑规则的例外情况,即:用该例外实体实例化规则,规则不成立,引入错误.为此,Gad-Elrab 等人<sup>[134]</sup>提出**通过在规则中添加例外情况,相当于负例,以减少错误的引入,提高基于规则的预测精度**.算法将知识图谱和用规则挖掘算法得到的规则集作为输入,基于观察到的三元组产生一个用例外情况丰富后的规则集,用于预测.针对非时序动态知识推理,赵泽亚等人<sup>[135]</sup>则提出一种在大规模演化知识网络 OpenKN<sup>[129]</sup>上,**基于背包问题的知识关联推理方法 KP-LIM(knapsack constrained link inference method)**.KP-LIM 首先引入连接延展(link extendable,简称 LE)模式,即,任意两实体都可以通过一条边(关系)进行关联的子网络,考虑包含 3 个实体的这样子网络.在进行关联推理时,将 LE 模式进行分解,即:取出任意一条边作为推理结果,其他两条边作为推理源(传递性规则).利用背包问题的思想,从分解的 LE 模式中选出置信度较高且涵盖关系类型更广泛的模式子集,并用子集中的所有模式进行关联推理.模式的置信度与重量关联,置信度高的模式,重量低.时间信息的利用则体现在模式的价值这一参数上,模式的价值为对应的所有实例关系时间之和.

## (2) 引入局部结构的规则推理

在整个知识图谱上进行推理,一方面代价往往比较高,而知识的推理一般主要与局部区域的知识更相关,因此可以在局部结构上进行推理;另一方面,全局结构信息粒度比较粗,结合细粒度的局部结构信息将有助于推理.Gardner 等人<sup>[15]</sup>提出了比 PRA<sup>[123]</sup>更简单、有效的从知识图谱产生特征矩阵的算法:**子图特征抽取(subgraph feature extraction,简称 SFE)**.给定图中的实体对集合,SFE 首先进行广度优先搜索获取每个实体周围的图;然后在局部子图进行多特征抽取,得到每个实体对的特征向量进行学习推理.Liu 等人<sup>[136]</sup>则提出了**层次随机行走推理算法 HiRi(hierarchical random-walk inference)**,进行大规模知识图谱的关系学习.HiRi 是一个两层的随机行走模型:上层对应关系路径发现和基于 PRA 全局视角的学习过程;下层用于以局部视角从关系特定的子图抽取信息.具体考虑给定关系对应子图的 3 跳转移概率矩阵,即,实体之间通过子图上的随机行走 3 跳内可达的概率作为得分.对于任意给定的候选三元组,线性组合两层机制得到的得分结果,选取得分高的作为推理结果.

总的来说,基于规则的多步推理不同于基于规则的单步推理采用简单经验知识或统计特征进行推理,所用规则为以传递性约束为主的更复杂的规则或特征,一般通过知识图谱上的规则挖掘算法得到,避开了人工构建规则的高代价.但是,挖掘的规则即使通过一定的过滤机制进行筛选,仍然不可避免地引入了一些信息量很低、作用很小、甚至是噪声的规则,误导了推理.并且,目前使用的规则主要是传递性约束,有待进一步挖掘更多样、更有效的复杂规则.这本质上是一个很大的难点,人工挖掘复杂规则代价很高,很难达到足够广的覆盖率,而根据数据集自适应挖掘的算法更难以实现复杂规则的挖掘,特别是难以保证挖掘规则的可靠性和可解释性.而局



部结构的引入,一方面在局部结构上推理,加快速度,一定程度上缓解了规则方法的高复杂度问题;另一方面,挖掘更丰富的局部结构特征,补充规则。因此,单纯的基于规则的多步推理有以下可继续探索的研究点:首先,自适应挖掘复杂规则或更多有效特征,该研究点难度比较大,需要复杂算法的支撑;其次,引入局部结构,该研究点一定程度上在弥补规则方法缺陷的同时,引入了新的优势,即,关联更密切的局部结构信息的充分学习,并且目前这方面的工作还不多,值得进一步深入研究。

## 4.2 基于分布式表示的推理

多步推理中,基于分布式表示的推理与基于分布式表示的单步推理类似,都是通过向量化知识图谱进行推理。不同的是,多步推理在学习向量表示的过程中引入了多步关系约束,使得学到的向量表示更有助于实体和关系的推理预测。Guu 等人<sup>[137]</sup>在 TransE 等模型的基础上提出了新的组合训练目标,将知识图谱向量空间模型看成边(关系)的遍历操作,直接建模路径的中间步,正则化实体向量的空间分布。组合训练用于不同的表示模型有不同的形式,例如在 TransE 中,组合训练体现为路径上关系向量的加和。模型首先在路径长度为 1 的数据(三元组)上训练,然后在所有路径数据上训练,保证在组合边形成路径前把握基本的边。对于实体间存在多条路径的情况,上述模型同等地对待这些路径,没有区分路径的贡献差别,显然存在实际意义很小的路径。Lin 等人<sup>[22]</sup>提出了 PTransE,区别对待实体间的不同路径。PTransE 在 TransE 的基础上多建模了关系路径约束,通过关系的组合操作建模路径。如图 5 所示,依次经过关系 BornInCity、CityInState 和 StateInCountry 的路径,通过关系的组合操作建模,组合操作可以是路径上关系的加和、连乘等形式,然后将组合后的路径看成头实体 Steve Jobs 和尾实体 United State 之间的转移。对于实体间的多条路径,Lin 等人提出了路径约束的资源分配算法<sup>[22]</sup>权衡关系路径的权重,用于对路径建模结果进行加权。

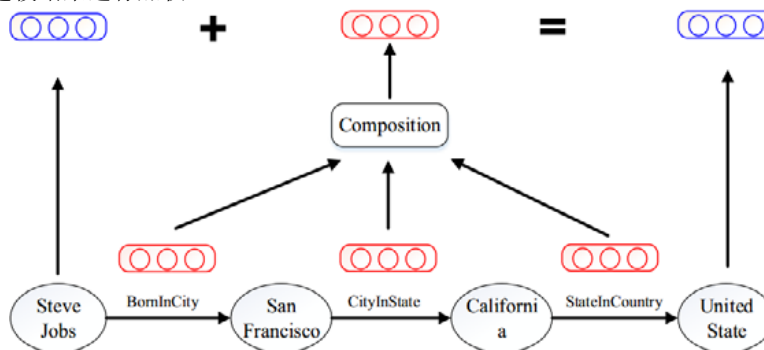


Fig.5 An example of path representation of the PTransE model<sup>[22]</sup>

图 5 PTransE 路径建模示例<sup>[22]</sup>

Lin 等人<sup>[138]</sup>则提出了关系路径表示的组合学习模型 RPE(relation paths embedding),不同于 PTransE 的另一种建模路径的方式,推理效果略优于 PTransE。RPE 中,每个实体通过关系映射矩阵和路径映射矩阵同时映射到对应的关系空间和路径空间,关系/路径看作映射后的实体之间的转移。路径映射矩阵通过路径上关系对应的映射矩阵加和或连乘得到。由此,三元组的得分函数包括头尾实体映射到关系空间的转移得分与头尾实体映射到对应的各个路径空间的转移得分之和。类型限制可以从传统的关系指定的类型限制方式扩展到新提出的路径指定的类型限制方式。RPE 中,类型限制用于两个方面:一是训练时负例的选取,选取和要替换的实体属于同一类型的实体替换,提高模型的学习能力;二是测试时,候选实体从同一类型的实体集选取,过滤不满足类型限制的实体。目前的知识表示方法大多单独学习每个三元组,PTransE 等也仅仅引入了多跳关系信息。为此,Feng 等人<sup>[139]</sup>提出了图感知的知识表示方法 GAKE(graph aware knowledge embedding),利用知识图谱更多的结构信息学习实体和关系的向量表示。GAKE 引入了 3 类图上下文信息:邻居上下文、路径上下文和边上上下文,从不同角度反映知识属性,同时设计 attention 机制,即实体和关系的权重学习,学习有代表能力的实体或关系。其中,邻居上下文反映三元组关系,实体的邻居上下文为以该实体为头实体的所有三元组中的关系和尾实体对,关系的邻



居上下文为该关系对应的所有三元组中的头实体和尾实体对.路径上下文为多步路径上的实体和关系.实体的边上下文为与该实体相连的所有关系,关系的边上下文为该关系连接的所有实体.每类上下文的目标函数为给定上下文,实体/关系的概率函数和,GAKE 最大化 3 类目标函数的加权和.

总体上看,基于分布式表示的多步推理主要是在基于分布式表示的单步推理基础上加入多步关系的建模,采用补充建模或共同建模的方式.补充建模的方式以单步关系(直接关系)为主,多步关系辅助学习,用于调整空间中向量的位置.而共同建模的方式同等地对待直接关系和间接关系,获得更好的向量表示,但可能引入级联误差,即,路径上中间步骤关系建模引起的误差将一直传递累计到最终结果.对于共同建模的方式,其中某些高质量路径和对应的直接关系实际上构成了传递性规则.在这种情况下,基于分布式表示的多步推理,一定程度上可以看成在建模知识图谱中三元组的同时,以方便计算的向量操作形式额外建模了传递性规则,提高了推理性能.这就需要确定多步路径与对应直接关系是否具有强关联性.然而,PTTransE<sup>[22]</sup>也只泛泛地计算多步路径在整个知识图谱中的可靠性.实际上,某条路径可能对于预测关系  $r_1$  来说是不可靠的路径,但对于预测关系  $r_2$  是决定性的路径.因此,将路径的可靠性与对应的直接关系关联考虑,将得到更有可能有效的传递性规则,在此基础上的建模将进一步提升效果.这也将成为一个小的研究点.

#### 4.3 基于神经网络的推理

多步推理中,基于神经网络的推理旨在用神经网络建模学习多步推理过程,包括建模多步路径以及模拟计算机或人脑的推理.

##### (1) 神经网络建模多步路径的推理

该类方法用神经网络建模路径,充分学习多步路径的向量表示,得分函数关联于路径的表示与直接关系表示的相似度,希望正例对应的相似度大,即乘积大,负例小.Neelakantan 等人<sup>[14]</sup>提出为每个关系类型或所有关系训练一个 RNN(recurrent neural network),得到任意长度路径的组合表示.每一步,输入路径上的关系向量和当前路径向量一起产生下一个组合向量,最后一步的输出作为路径的向量表示.训练过程中,最大化正例(知识图谱中的实体对和对应关系组成的实例)的概率且最小化负例(未观察到的实体对、关系实例)的概率.正例的概率也即得分函数,为关系的向量与连接对应实体对的所有路径向量乘积的最大值取 *sigmoid*,负例的概率即为 1 减去可能为正例的概率.选取与关系乘积最大的路径,也即认为与关系最相似的那条路径可以预测实体对是否存在该关系.该方法对于每个三元组只用了一条路径的向量计算概率得分,未考虑其他路径,并且忽略了路径上实体的建模,只是关系的推理.为此,Das 等人<sup>[140]</sup>训练了一个单一的高能力 RNN,在路径建模中,除了建模每步关系,还建模对应的实体,共同推理所有的关系、实体和实体类型.其中,实体向量表示通过实体类型向量表示的简单加和得到,实体类型向量表示在训练中学习.特别地,文献[140]用神经关注模型(neural attention model),一种选择机制,引入关联的多条路径而不是单条路径进行推理.选择机制包括选取与关系相似度排在前三的路径对应的相似度求平均取 *sigmoid*,作为实体对存在关系的概率也即得分函数,或选取关系与所有路径相似度的平均取 *sigmoid* 等.

##### (2) 神经网络模拟计算机或人脑的推理

神经网络模拟计算机或人脑的推理,利用神经网络强大的学习能力,模拟计算机或人脑的知识存储和处理方式.一般用一个存储结构模拟人脑的存储记忆,用一个控制器模拟人脑的控制处理中心.通过对知识图谱中已知三元组的学习记忆,希望神经网络能够具有人脑的推理能力,推理出新的三元组.Graves 等人<sup>[141]</sup>提出了 DNC(differentiable neural computer),包含一个 LSTM 神经网络控制器和可以读写的外部存储矩阵.训练时,DNC 以知识图谱三元组向量作为输入,通过神经网络进行外部存储矩阵的读写,模拟人脑利用已有的经验知识学习推理新知识,并更新已有的知识.测试时,需要推理预测的三元组对应字段留空(例如预测头实体,则头实体字段留空),输入训练好的 DNC,控制器不断与外部存储矩阵交互,多步推理,最后输出补全的三元组.Shen 等人<sup>[142]</sup>提出了隐性推理网(IRN),通过一个控制器和共享内存,在神经空间隐性地多步推理. IRN 拼接三元组的已知部分向量输入 RNN 神经网络控制器,控制器判断当前状态向量是否已经编码了足够的信息:如果没有编码足够信息,则根据当前状态向量和通过 attention 机制从共享内存得到的关注向量(attention vector)共同产生下一个状态,实

现多步推理;否则,停止推理,产生输出向量,与目标向量进行比较,梯度更新参数和共享内存,进行模型学习.测试时,用产生的输出向量找到与之相似度大的实体向量作为预测结果.

总而言之,基于神经网络的多步推理直接建模多步路径或推理过程,相比于基于神经网络的单步推理,研究工作更丰富,可解释性更强,效果更好.然而,可解释性需要进一步增强.其中,直接建模多步路径的方式,由于多步路径可以看成是关系或关系和实体的一个序列,主要通过 RNN 建模.鉴于该方向研究空间比较小,并且已有相关的工作,其继续发展的空间比较小.直接建模推理过程的方式可以模拟人的学习和推理过程,而人具备强大的学习推理能力,因此这将会是一个很热门的研究方向.并且,人的学习和推理过程相对复杂,使其也成为具有很大挑战性的问题.基于现有的模型,进一步研究如何更好地建模推理过程,将成为未来需要致力解决的难题.

#### 4.4 混合推理

多步推理中,混合推理通过混合不同多步方法进行推理,实现优势互补.分布式表示方法由于其计算的便捷性,通常被用于与其他方法混合.混合多步推理具体包括混合 PRA 与分布式表示、混合规则与分布式表示以及混合规则与神经网络的推理.

##### (1) 混合 PRA 与分布式表示的推理

混合 PRA 与分布式表示的推理主要用分布式表示方法从相关文本语料中学习连接实体的关系隐性表示,增广知识图谱,降低知识图谱的不连通性和稀疏性,帮助路径发现过程,从而提升推理效果.Gardner 等人<sup>[143]</sup>扩展了 PRA 算法中的图谱结构,提出了  $PRA_{latent}$ ,将从大规模文本语料中挖掘的关系隐性特征向量作为边标签,有效提升了知识推理任务的性能,并且相比于直接用关系字符串,降低了稀疏性.具体地,通过大规模文本语料构建关系-头尾实体对矩阵,矩阵元素值对应频次,采用主成分分析(PCA)得到关系低维矩阵,映射到具体的关系向量表示.在此基础上,Gardner 等人<sup>[144]</sup>提出在随机行走推理中引入向量空间相似度.从文本语料中建立的子知识图谱,与现有知识图谱通过别名关系连接起来.从语料中抽取的文本关系三元组,其关系字符串同样用通过 PCA 得到的关系向量替代.算法引入了向量空间相似度,在根据路径计算特征值时,允许行走与下一步要走的边(关系)语义相似的边.在这个过程中,如果遇到没有向量表示的原知识图谱中的关系,则按原来的 PRA 算法选择下一条边.Gardner 等人<sup>[145]</sup>进一步解释在符号逻辑系统引入向量空间表示,用向量空间语义增广逻辑推理可以归约为在规则中关系的位置引入不同形式的关系析取.基于文本语料关系聚类的 PRA,允许走与下一步要走的边在同一簇中的其他边,可以形式化为给每个规则添加关系的析取.同样地,基于文本语料关系向量空间相似度的 PRA,允许行走与下一步要走的边语义相似的边,可以形式化为依赖于相似度函数的关系的析取.

##### (2) 混合规则与分布式表示的推理

在基于规则分布式表示学习的推理方面,Wang 等人<sup>[146]</sup>提出通过矩阵分解直接学习一阶逻辑规则的表示,框架如图 6 所示.从三元组训练集中产生推理规则集后,首先构建 ProPPR 有向证明图<sup>[80,81]</sup>,图 6 左边部分展示了从某查询子句到目标实体节点 1 和节点 2 的证明图.然后将证明图转化为一个矩阵,映射训练实例到二值矩阵的行,映射规则到列,矩阵元素对应实例推理(倾向于短的推理路径)是否用到特定的规则:是,则元素值为 1;反之,为 0.例如,图 6 中训练实例 1(以 Query Node 为查询子句,Solution 1 为推理目标的某关系三元组)用到了规则 1、规则 3 和规则 4,因此训练实例 1 对应的第 1 行,规则 1、规则 3 和规则 4 对应位置的元素值为 1.用一种可扩展的矩阵分解方法,通过低阶近似学习实例和逻辑规则的隐性连续表示.测试时,用候选规则集构建测试三元组实例的有向证明图进行推理.

在图分布式表示作为先验的推理方面,Wei 等人<sup>[146]</sup>提出在选择的实例上进行网络采样推理(infering via grounding network sampling,简称 INS).首先,采用一个表示模型,如 TransE,学习实体和关系的表示,计算候选和查询的相似度得分,选择得分前  $N$  的候选,形成更小的候选集;与此同时,过滤掉一部分噪声实例.然后,通过马尔可夫逻辑网<sup>[85]</sup>上数据驱动的推理算法进行推理.特别地,替换随机行走的均匀转移概率为基于向量相似度的转移概率函数,提出 INS-ES(INS merging embedding similarity priori),引导随机行走,进一步提高推理准确率.

在共同建模规则和知识图谱的推理方面,Guo 等人<sup>[147]</sup>提出了共同表示知识图谱和逻辑规则的模型 KALE(embeddings by jointly modeling knowledge and logic).其主要思想是:在统一的框架中,形式化建模三元组和规

则.三元组看成原子公式,通过基于转移的方法,如 TransE,建模;规则形式化为复杂公式,通过  $t$  阶模糊逻辑建模,规则的真值由成分原子的真值合成.KALE 最小化原子和复杂公式的全局损失,获得实体和关系的向量表示.

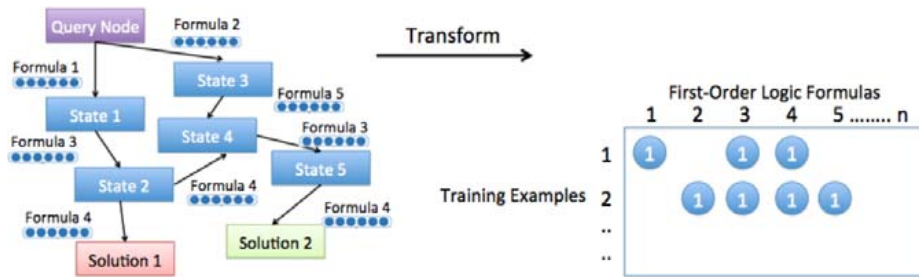


Fig.6 Matrix factorization framework for learning first-order logic embeddings<sup>[146]</sup>

图 6 学习一阶逻辑规则表示的矩阵分解框架<sup>[146]</sup>

### (3) 混合规则与神经网络的推理

混合规则与神经网络的推理主要将规则转化为向量操作,应用于强学习能力的神经网络方法中,实现一个可微的模型.前文中的 TensorLog<sup>[82]</sup>仅局限于学习一阶逻辑规则的参数(每个规则关联的置信度),Yang 等人<sup>[148]</sup>进一步提出一个完全可微的系统,第一次结合一阶逻辑规则的参数和结构(知识图谱特定的规则集)学习到一个端对端可微模型 Neural LP(neural logic programming).与 TensorLog 一样,Neural LP 中,每个实体关联一个固定的 one-hot 向量,每个关系关联一个固定的  $\{0,1\}$  操作矩阵,逻辑规则推理形式化为矩阵相乘.实体的得分函数与所有对应路径上关系操作矩阵乘积的置信度加权和相关.Neural LP 设计了一个带 attention 机制和存储的神经控制器系统,学习过程顺序组合 TensorLog 中使用的可微操作,需要学习的参数有规则集及其对应的置信度.由于每个置信度关联一个特定的规则,枚举规则是一个离散的任务,直接学习很难形式化为一个可微的过程.为了解决这个问题,交换得分函数的乘和加操作,也即先加再乘.先从关系的角度对操作矩阵进行置信度加权求和,然后再乘.然而这样得到的路径都一样长,关联于乘操作的次数.为此,引入辅助存储向量、存储 attention 向量和操作 attention 向量,采用循环形式进行学习.其中,辅助存储向量初始化为给定实体,后续为每一步得到的中间推理结果;存储 attention 向量存储到当前步为止每步存储向量的权重;操作 attention 向量存储每个 TensorLog 操作(关系操作矩阵)的权重.最后,模型计算所有存储向量的一个加权平均,权重为存储 attention 向量.这样,使用 attention 选择合适的规则长度.整个过程用 LSTM 实现,每步推理首先使用 attention 机制选择(操作 attention 向量)TensorLog 操作子集,然后从存储选择(存储 attention 向量)内容用这些操作作用,作用后的结果为该步的存储向量,加入存储.

大体上,混合多步推理相比于混合单步推理内容更为丰富,其中,混合规则和图分布式表示的推理得到了比较全面的研究,产生了更有效的混合模式:共同建模,取得了更好的效果.但目前的混合推理依然局限于两种方法的混合.各类混合方法中,混合规则和神经网络的推理具有很大的发展空间,规则方法的高准确率和可解释性以及神经网络方法的强学习和高泛化能力,使得二者的结合可以得到高准确率的可微模型,避开了传统规则方法的计算难题,一定程度上也增加了神经网络方法的可解释性.未来,混合规则与神经网络的推理将值得更深入的研究,可以从混合模式以及规则和神经网络的具体表现形式等角度进一步加以研究.

## 5 知识推理的典型应用

前述的知识图谱补全和去噪是知识推理的两大基础应用.现有的知识图谱由于数据来源的不全面以及知识获取的遗漏,不可能构建完备的知识图谱.通常,代表性的知识图谱中有 69%~99%的实体缺少至少一个属性信息三元组<sup>[149]</sup>.例如,在 Freebase 中,93.8%的人没有出生地信息,78.5%的人没有国籍信息<sup>[150]</sup>.解决办法之一就是通过知识推理方法,利用知识图谱中已有的知识去推理出新的事实(即隐含的知识),从而尽可能地对知识图



谱进行补全.另一方面,就知识图谱自身而言,由于数据来源的噪声以及抽取过程的不准确,内部也存在噪声知识和知识矛盾现象<sup>[151]</sup>.例如在 NELL 中,采集到的知识正确率随着时间推移不断下降,第 1 个月,正确率达到 0.9;第 2 个月,正确率降为 0.71<sup>[17]</sup>.主要原因是知识导出的抽取模版不可靠,导致抽取错误的知识,错误的知识用于产生更多不可靠的模版,如此循环.NELL 也周期地使用人工监督移除不正确的事实元组,然而人工标注代价很高.这就需要通过知识推理方法,自动且高效地完成这一过程.

除了知识图谱补全与去噪,知识推理在垂直搜索、智能问答、机器翻译等领域也发挥了重要作用,在疾病诊断、金融反欺诈、数据异常分析等诸多不同的领域已展示出良好的应用前景.

在垂直搜索领域,在国外,Google 公司提出了 Knowledge Graph 与 Knowledge Vault,Facebook 推出了 Graph Search,微软推出了 Bing Satori;在国内,搜狗提出了知立方,百度推出了中文知识图谱搜索.知识推理能够更好地理解用户的搜索意图,提供接近“专、精、深”的垂直搜索,回答复杂的推理问题.例如 Google 的搜索,在 Google 搜索引擎输入查询,搜索引擎在利用知识图谱直接给出推理得到的精确回答的同时,在搜索结果的右侧显示该词条的深层信息.百度的搜索在知识图谱的支持下,也能更好地理解用户的搜索意图,类似地返回推理的精确答案,附带信息来源.

在智能问答领域,IBM 的 Watson、Google 的 Google Now、苹果公司的 Siri、亚马逊的 Alexa、微软的小娜和小冰以及百度的度秘等是近期代表性的智能问答系统.这些系统基于知识图谱的知识推理,提供精确、简洁的答案.例如,IBM 研发的超级计算机 Watson,2011 年, Watson 在美国知识竞赛节目“危险边缘 Jeopardy!”中上演人机问答大战,战胜了人类冠军选手 Ken 和 Brad.节目中的问题覆盖面广,参赛者需具备各个领域的知识,能够解析和推理隐晦含义、反讽与谜语等.这其中,面向知识图谱的知识推理发挥了重要作用.

## 6 知识推理总结与展望

近年来,随着知识图谱的快速发展,面向知识图谱的知识推理作为知识图谱补全和去噪的主要途径,引发了广泛关注,得到了很好的发展.然而,面向知识图谱的知识推理与实际应用还有很长的距离.本节将对面向知识图谱的知识推理进行总结,并对未来发展方向进行展望.

### 6.1 知识推理总结

知识推理根据推理类型可分为单步推理和多步推理两大类.每类又包括基于规则的推理、基于分布式表示的推理、基于神经网络的推理以及混合推理.各类方法的汇总见表 2.

总的来说,单步推理基于知识图谱中的事实元组建模,而多步推理在单步推理的基础上建模了多步路径的约束,表达能力往往比单步推理更强,推理预测效果更好.基于分布式表示的单步推理中细粒度建模空间分布的方法,由于充分考虑了知识图谱的空间分布特征,表达推理能力强.针对该类方法的研究还比较少,有待进一步细致挖掘知识图谱的空间分布特征,探索更多的建模方法,同时扩展到多步推理.在单步推理和多步推理的各个子类方法中,基于神经网络的推理和混合推理是推理效果相对较好的方法,有待推进进一步的研究工作.基于神经网络的推理利用神经网络的强大学习能力进行建模推理,其中已有一些模拟计算机或人知识存储和推理的研究工作.随着神经网络的持续发展,更深入的研究依然有待开展.进一步增强神经网络针对推理这一任务的解释性也是一大难点.混合推理试图利用各种推理方法的优势获得更好的推理性能.通常,集成学习的效果要好于单个模型<sup>[152]</sup>.然而,目前的混合推理局限于两种方法的混合,如何混合多种互补的方法,进一步提高推理能力,有待进一步研究.同时,目前的混合模式大多比较浅层,未来需要探索更多更深层次的混合模式.其中,混合基于规则的推理方法局限于简单规则和传递性约束,有待研究如何引入更多样的有效规则.



Table 2 Classification summary table of knowledge reasoning  
表 2 知识推理方法分类汇总表

类别	子类	方法描述	方法优点	方法缺点
单步推理	基于规则的推理	利用简单规则、统计特征等进行推理	可解释性强; 规则正确时,准确率高	规则不易获得,不全面,导致推理结果召回率低;统计特征不易迁移,当数据存在噪声时,可能误导推理;需要一个一个地推理候选实例,可计算性比较差
	基于分布式表示的推理	利用知识图谱表示模型,得到知识图谱的低维向量表示,通过向量操作进行推理预测	计算方便快捷	通常只考虑满足知识图谱事实元组的约束,未考虑更深入的组合语义信息,推理能力受限
	基于神经网络的推理	利用神经网络直接建模知识图谱事实元组	推理能力较强	相比基于分布式表示的推理,复杂度更高,可解释性更弱
	混合推理	混合多种单步推理方法	不同推理方法优势互补	目前还停留在浅层的混合,即:以一种方法为主,另一种为辅的推理,缺乏更深层次的混合模式
多步推理	基于规则的推理	通过挖掘全局或局部复杂规则,主要是传递性规则进行推理	自动挖掘复杂规则,而不是通过专家构建	挖掘的规则不可避免地存在噪声,误导推理
	基于分布式表示的推理	向量化事实元组中实体和关系的同时,建模多步关系约束	获得知识图谱更好的向量表示,推理预测能力提高	可能存在级联误差
	基于神经网络的推理	利用神经网络学习多步推理,直接建模多步路径或推理过程	表达能力丰富,推理能力强	需要进一步增强可解释性
	混合推理	混合多种多步推理方法	充分利用不同多步推理方法的优势,出现了更有效的混合模式:共同建模,推理能力强	局限于两种方法的混合

6.2 面向多元关系的知识推理

相比于一元与二元关系,多元关系结构多样、语义复杂并且难以处理.因此,现有的知识处理方面的工作主要集中在二元关系上.面向多元关系的研究工作很少,并且已有的工作通常将其简化为二元关系进行处理,损失了大量的语义信息.然而多元关系在知识图谱中不是少数派,例如在 Freebase 中,超过三分之一的实体卷入了多元关系<sup>[93]</sup>.

未来有必要研究如何进行多元关系的形式化表达(representation)与表示学习(embedding)以及如何兼顾表达能力与推理能力.表达能力越强,表达结构越复杂,灵活性越差,推理能力相应地越弱.例如  $m$ -TransH<sup>[93]</sup>将多元关系形式化表达为角色集合(如{演员,角色,电影})到实体集的映射,每个关系类型的每个具体映射称为该关系类型的实例,用事实 ID 标识.也即,组成事实的每个要素看成一个角色和角色对应的实体.这种表达方法不会损失结构信息,基本与原事实等价.然而,它需要预先确定关系类型及其对应的角色序列.多元关系数据严格按照相应的关系类型角色序列顺序地存储对应的实体,不够灵活.同时,对角色的建模和推理能力受限(角色仅被映射到一个实数).因此,如何权衡表达能力与推理能力,一方面使得表达尽可能少地损失多元关系的结构信息,另一方面使得推理能够灵活进行,成为需要解决的一大问题,也成为未来研究的一大挑战.

6.3 融合多源信息与多种方法的知识推理

融合多源信息的知识推理,通过结合文本语料或其他知识图谱,利用更多的额外信息降低知识图谱的不连通性和稀疏性,进行有效推理.融合多种方法的知识推理,通过在更深层次混合不同方法,优势互补,提升推理性能.例如,共同建模规则和知识图谱的推理<sup>[147]</sup>是目前性能突出的方法.此外,鉴于神经网络在各个领域包括知识图谱领域的突出性能,融合神经网络与其他互补的方法将成为未来研究的主要热点.例如 Neural LP<sup>[148]</sup>,结合神经网络的强学习和泛化能力与规则方法的高准确率和可解释性,获得了极好的推理结果,在两个常用的数据集 WN18(知识图谱 WordNet<sup>[91]</sup>的子集)和 FB15k(知识图谱 Freebase<sup>[10]</sup>的子集)中,对于实体预测任务,有效的实体排在前 10 的比例分别达到了 99.8%和 91.6%.而如何同时融合多源信息与多种方法,进一步提升推理性能,也将成为未来的一大研究方向.其中,融合模式,即以什么方式融合,是一大难点.

#### 6.4 基于小样本学习的知识推理

现有的知识推理模型往往需要大量高质量的样本进行训练学习,这需要耗费很大的代价去获取样本.在实际应用中,甚至难以获得大量的训练样本,极大地限制了现有知识推理模型的应用范围.另一方面,人往往凭借相关先验知识只需少量样本就能快速学习推理.在此过程中,大脑感知外部环境,对感兴趣或待学习的信息保持关注,并通过与已有先验知识的结合快速建立起新的知识,而后,经过神经元的加工整理,形成难以被遗忘的长时记忆.由此,人不断地从生活经验中建立并整合知识,从而学会处理日益复杂的任务.在持续不断的学习过程中,对以往知识检索利用,使得人们只需要少量的训练就能快速地学会新的任务.目前已有的一些用神经网络模拟人脑的学习和推理的研究<sup>[141,142]</sup>,但复杂度较高,依然需要大量训练样本的支撑.为此,基于小样本学习的知识推理将成为未来研究的重要方向,即:如何以已建立的知识图谱(如 YAGO<sup>[6-8]</sup>、Freebase<sup>[110]</sup>和 NELL<sup>[111]</sup>)等为先验,通过少量高质量样本进行新的学习和推理.

#### 6.5 动态知识推理

目前的知识推理方法主要针对静态知识图谱进行.然而,知识图谱随着时间的推移往往动态变化,包括增、删和改操作,例如,“谁是现任中华人民共和国主席?”在 2008 年 3 月~2013 年 3 月期间,该问题的答案是胡锦涛,2013 年 3 月至今,该问题的答案是习近平.因此,如何进行动态知识推理成为未来的一个研究热点.目前有简单引入时间要素的工作<sup>[103]</sup>、新模型的尝试<sup>[104,109]</sup>以及基于新提出的演化网络的推理研究<sup>[130,135]</sup>.例如,Tay 等人<sup>[104]</sup>提出了 puTransE,通过分治策略实现知识图谱的分块学习和集成推理,可以有效处理知识图谱的动态增和删.对于新增的三元组,通过建立新的平行空间学习;对于删除的三元组,通过在预测时使对应的表示空间无效实现.然而,puTransE 不能直接处理知识图谱的动态修改(当然,可以通过简单地将修改操作看成先删再增,进行间接处理),并且删操作往往不是真的删除事实元组,而是超过了事实元组的有效时间,事实元组失效,在有效时间内,事实元组依然成立.另一方面,基于演化网络 OpenKN<sup>[129]</sup>的动态推理<sup>[130,135,153]</sup>还不够成熟.为此,有待开展更深层次的研究,有效地进行动态知识推理,灵活处理知识图谱的动态增、删和改.

#### 6.6 其他研究方向

除了以上 4 个主要研究方向,还有大量关于知识推理的研究工作亟待开展,例如面向大规模知识图谱的快速推理.信息时代,随着数据的增长,知识图谱的规模越来越大将成为未来的发展趋势.因此,如何优化推理模型,提高推理速度,扩展到大规模知识图谱,保证推理的时效性,将成为未来需要致力解决的问题.

### 7 结束语

伴随着信息时代的发展,知识图谱作为知识的一种结构化和语义化的表达方式,引发了广泛的关注.面向知识图谱的知识推理作为知识图谱补全和去噪的主要方式,已经在垂直搜索、智能问答、机器翻译等领域发挥了重要作用,也在疾病诊断、金融反欺诈、数据异常分析等诸多不同的领域展示出良好的应用前景.本文详细阐述了面向知识图谱的知识推理,对各类推理方法的研究现状与面临的问题进行了总结.在此基础上,本文对未来研究方向进行了分析和展望.

#### References:

- [1] Wang YZ, Jin XL, Cheng XQ. Network big data: Present and future. Chinese Journal of Computers, 2013,36(6):1125–1138 (in Chinese with English abstract).
- [2] Cheng XQ, Jin XL, Wang YZ, Guo JF, Zhang TY, Li GJ. Survey on big data system and analytic technology. Ruan Jian Xue Bao/ Journal of Software, 2014,25(9):1889–1908 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4674.htm> [doi: 10.13328/j.cnki.jos.004674]
- [3] Wang YZ, Jia YT, Liu DW, Jin XL, Cheng XQ. Open Web knowledge aided information search and data mining. Journal of Computer Research and Development, 2015,52(2):456–474 (in Chinese with English abstract).
- [4] Jin XL, Wah BW, Cheng XQ, Wang YZ. Significance and challenges of big data research. Big Data Research, 2015,2(2):59–64.

- [5] Etzioni O, Cafarella M, Downey D, Kok S, Popescu AM, Shaked T, Soderland S, Weld DS, Yates A. Web-Scale information extraction in knowitall (preliminary results). In: Proc. of the 13th Int'l Conf. on World Wide Web. New York: ACM Press, 2004. 100–110.
- [6] Suchanek FM, Kasneci G, Weikum G. YAGO: A core of semantic knowledge. In: Proc. of the 16th Int'l Conf. on World Wide Web. New York: ACM Press, 2007. 697–706.
- [7] Hoffart J, Suchanek FM, Berberich K, Weikum G. YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. *Artificial Intelligence*, 2013, 194: 28–61.
- [8] Mahdisoltani F, Biega J, Suchanek F. Yago3: A knowledge base from multilingual Wikipedias. In: Proc. of the 7th Biennial Conf. on Innovative Data Systems Research (CIDR). 2014.
- [9] Auer S, Bizer C, Kobilarov G, Lehmann J, Cyganiak R, Ives Z. DBpedia: A nucleus for a Web of open data. In: Proc. of the 6th Int'l Semantic Web Conf. Berlin, Heidelberg: Springer-Verlag, 2007. 722–735.
- [10] Bollacker K, Evans C, Paritosh P, Sturge T, Taylor J. Freebase: A collaboratively created graph database for structuring human knowledge. In: Proc. of the 2008 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM Press, 2008. 1247–1250.
- [11] Carlson A, Betteridge J, Kisiel B, Settles B, Hruschka Jr ER, Mitchell TM. Toward an architecture for never-ending language learning. In: Proc. of the 24th AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI, 2010. 1306–1313.
- [12] Wu W, Li H, Wang H, Zhu KQ. Probase: A probabilistic taxonomy for text understanding. In: Proc. of the 2012 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM Press, 2012. 481–492.
- [13] Lin Y, Liu Z, Sun M, Liu Y, Zhu X. Learning entity and relation embeddings for knowledge graph completion. In: Proc. of the 29th AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI, 2015. 2181–2187.
- [14] Neelakantan A, Roth B, McCallum A. Compositional vector space models for knowledge base completion. In: Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics. 2015. 156–166.
- [15] Gardner M, Mitchell TM. Efficient and expressive knowledge base completion using subgraph feature extraction. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2015. 1488–1498.
- [16] Wei Z, Zhao J, Liu K, Qi Z, Sun Z, Tian G. Large-Scale knowledge base completion: Inferring via grounding network sampling over selected instances. In: Proc. of the 24th ACM Int'l Conf. on Information and Knowledge Management. New York: ACM Press, 2015. 1331–1340.
- [17] Jiang S, Lowd D, Dou D. Learning to refine an automatically extracted knowledge base using markov logic. In: Proc. of the 12th IEEE Int'l Conf. on Data Mining. New Jersey: IEEE, 2012. 912–917.
- [18] Paulheim H, Bizer C. Improving the quality of linked data using statistical distributions. *Int'l Journal on Semantic Web and Information Systems (IJSWIS)*, 2014, 10(2): 63–86.
- [19] Wang Z, Zhang J, Feng J, Chen Z. Knowledge graph embedding by translating on hyperplanes. In: Proc. of the 28th AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI, 2014. 1112–1119.
- [20] Xiao H, Huang M, Zhu X. From one point to a manifold: Knowledge graph embedding for precise link prediction. In: Proc. of the 25th Int'l Joint Conf. on Artificial Intelligence. AAAI, 2016. 1315–1321.
- [21] Trouillon T, Welbl J, Riedel S, Gaussier E, Bouchard G. Complex embeddings for simple link prediction. In: Proc. of the 33rd Int'l Conf. on Machine Learning. New York: ACM Press, 2016. 2071–2080.
- [22] Lin Y, Liu Z, Luan H, Sun M, Rao S, Liu S. Modeling relation paths for representation learning of knowledge bases. *arXiv Preprint arXiv: 150600379*, 2015.
- [23] Xie R, Liu Z, Sun M. Representation learning of knowledge graphs with hierarchical types. In: Proc. of the 25th Int'l Joint Conf. on Artificial Intelligence. Menlo Park: AAAI, 2016. 2965–2971.
- [24] Nguyen DQ, Sirts K, Qu L, Johnson M. Neighborhood mixture model for knowledge base completion. *arXiv Preprint arXiv: 160606461*, 2016.
- [25] Bordes A, Usunier N, Garcia-Duran A, Weston J, Yakhnenko O. Translating embeddings for modeling multi-relational data. In: Proc. of the Advances in Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2013. 2787–2795.
- [26] Lin Y, Liu Z, Sun M. Knowledge representation learning with entities, attributes and relations. In: Proc. of the 25th Int'l Joint Conf. on Artificial Intelligence. Menlo Park: AAAI, 2016. 2866–2872.

- [27] Zhao J, Liu K, Zhou GY, Cai L. Open information extraction. *Journal of Chinese Information Processing*, 2011,25(6):98–110 (in Chinese with English abstract).
- [28] Liu Q, Li Y, Duan H, Liu Y, Qin ZG. Knowledge graph construction technique. *Journal of Computer Research and Development*, 2016,53(3):582–600 (in Chinese with English abstract).
- [29] Zhuang Y, Li GL, Feng JH. A survey on entity alignment of knowledge base. *Journal of Computer Research and Development*, 2016, 53(1):165–192 (in Chinese with English abstract).
- [30] Liu ZY, Sun MS, Lin YK, Xie RB. Knowledge representation learning: A review. *Journal of Computer Research and Development*, 2016,53(2):247–261 (in Chinese with English abstract).
- [31] Lin HL, Wang YZ, Jia YT, Zhang P, Wang WP. Network big data oriented knowledge fusion methods: A survey. *Chinese Journal of Computers*, 2017,40(1):1–27 (in Chinese with English abstract).
- [32] Xu ZL, Sheng YP, He LR, Wang YF. Review on knowledge graph techniques. *Journal of University of Electronic Science and Technology of China*, 2016,45(4):589–606 (in Chinese with English abstract).
- [33] Wang YQ. *Principles and Methods of Artificial Intelligence*. Xi'an: Xi'an Jiaotong University Press, 1998 (in Chinese).
- [34] Kompridis N. So we need something else for reason to mean. *Int'l Journal of Philosophical Studies*, 2000,8(3):271–295.
- [35] Tari L. *Knowledge Inference*. New York: Springer-Verlag, 2013. 1074–1078.
- [36] Richter MM, Weber RO. *Case-Based Reasoning*. Springer-Verlag, 2016.
- [37] Lu RQ, Ying MS. A model of knowledge inference. *Science in China (Series E): Technical Science*, 1998,28(4):363–369 (in Chinese with English abstract).
- [38] *Handbook of Automated Reasoning*. Elsevier, 2001.
- [39] *Handbook of Knowledge Representation*. Elsevier, 2008.
- [40] Gentzen G. Untersuchungen über das logische Schließen. I. *Mathematische Zeitschrift*, 1935,39(1):176–210.
- [41] Gentzen G. Untersuchungen über das logische Schließen. II. *Mathematische Zeitschrift*, 1935,39(1):405–431.
- [42] Robinson JA. Automatic deduction with hyper-resolution. *Int'l Journal of Computing and Mathematics*, 1965,1:227–234.
- [43] Schmidt-Schauß M, Smolka G. Attributive concept descriptions with complements. *Artificial Intelligence*, 1991,48(1):1–26.
- [44] Solomonof RJ. A formal theory of inductive inference. Part I. *Information and Control*, 1964,7(1):1–22.
- [45] Solomonof RJ. A formal theory of inductive inference. Part II. *Information and Control*, 1964,7(2):224–254.
- [46] Reiter R. A logic for default reasoning. *Artificial intelligence*, 1980,13(1-2):81–132.
- [47] Shortliffe EH, Buchanan BG. A model of inexact reasoning in medicine. *Mathematical Biosciences*, 1975,23(3):351–379.
- [48] Zadeh LA. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. on Systems, Man, and Cybernetics*, 1973,SMC-3(1):28–44.
- [49] Minsky M. A framework for representing knowledge. In: *The Psychology of Computer Vision*. 1975. 211–277.
- [50] Norenzayan A, Nisbett RE, Smith EE, Kim BJ. Cultural preferences for formal versus intuitive reasoning. *Cognitive Science*, 2002, 26(5):653–684.
- [51] Allen JF. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 1983,26(11):832–843.
- [52] Liu YB, Liu DY. A survey on spatial inference and geographic information systems. *Ruan Jian Xue Bao/Journal of Software*, 2000, 11(12):1598–1606 (in Chinese with English abstract). [http://www.jos.org.cn/jos/ch/reader/create\\_pdf.aspx?file\\_no=20001206&journal\\_id=jos](http://www.jos.org.cn/jos/ch/reader/create_pdf.aspx?file_no=20001206&journal_id=jos)
- [53] Schank R. *Dynamic Memory: A Theory of Learning in People and Computers*. New York: Cambridge University Press, 1982.
- [54] Riesbeck CK, Schank RC. *Inside Case-Based Reasoning*. Psychology Press, 2013.
- [55] Jijkoun V, Rijke M. Recognizing textual entailment using lexical similarity. In: *Proc. of the 1st PASCAL Challenges Workshop on Recognising Textual Entailment*. Stroudsburg: ACL, 2005. 73–76.
- [56] Hickl A, Williams J, Bensley J, Roberts K, Rink B, Shi Y. Recognizing textual entailment with LCC's GROUNDHOG system. In: *Proc. of the 2nd PASCAL Challenges Workshop on Recognising Textual Entailment*. Stroudsburg: ACL, 2006.
- [57] Muggleton S, De Raedt L. Inductive logic programming: Theory and methods. *The Journal of Logic Programming*, 1994,19(20): 629–679.



- [58] MacCartney B, Manning CD. Natural logic for textual inference. In: Proc. of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing. Stroudsburg: ACL, 2007. 193–200.
- [59] MacCartney B, Manning CD. Modeling semantic containment and exclusion in natural language inference. In: Proc. of the 22nd Int'l Conf. on Computational Linguistics. Stroudsburg: ACL, 2008. 521–528.
- [60] Angeli G, Manning CD. NaturalLI: Natural logic inference for common sense reasoning. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2014. 534–545.
- [61] Zhao JX, Pan ZH, Wang SS. Fuzzy knowledge inference and search processing based on FLcom. *Computer Engineering and Applications*, 2015,51(19):37–42 (in Chinese with English abstract).
- [62] Bos J, Markert K. When logical inference helps determining textual entailment (and when it doesn't). In: Proc. of the 2nd PASCAL Challenges Workshop on Recognising Textual Entailment. Stroudsburg: ACL, 2006. 26–31.
- [63] Angeli G, Nayak N, Manning CD. Combining natural logic and shallow reasoning for question answering. In: Proc. of the 54th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2016. 442–452.
- [64] Studer R, Benjamins VR, Fensel D. Knowledge engineering: Principles and methods. *Data & Knowledge Engineering*, 1998,25(1): 161–197.
- [65] Thomas E, Pan J, Ren Y. TrOWL: Tractable OWL 2 reasoning infrastructure. In: Proc. of the 7th Int'l Conf. on Semantic Web: Research and Applications. Berlin, Heidelberg: Springer-Verlag, 2010. 431–435.
- [66] Tsarkov D, Palmisano I. Chainsaw: A metareasoner for large ontologies. In: Proc. of the OWL Reasoner Evaluation Workshop. Princeton: Citeseer, 2012. 19–27.
- [67] Mendez J. jcel: A modular rule-based reasoner. In: Proc. of the OWL Reasoner Evaluation Workshop. Princeton: Citeseer, 2012. 130–135.
- [68] Armas Romero A, Cuenca Grau B, Horrocks I. MORE: Modular combination of OWL reasoners for ontology classification. In: Proc. of the The Semantic Web—ISWC 2012. 2012,7649:1–16.
- [69] Sertkaya B. The ELepHant reasoner system description. In: Proc. of the OWL Reasoner Evaluation Workshop. Princeton: Citeseer, 2013. 87–93.
- [70] Kazakov Y, Krötzsch M, Simančík F. The incredible ELK. *Journal of Automated Reasoning*, 2014,53(1):1–61.
- [71] Glimm B, Horrocks I, Motik B, Stoilos G, Wang Z. HermiT: An OWL 2 reasoner. *Journal of Automated Reasoning*, 2014,53(3): 245–269.
- [72] Zhou Y, Nenov Y, Grau BC, Horrocks I. Pay-as-You-Go OWL query answering using a triple store. In: Proc. of the 28th AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI, 2014. 1142–1148.
- [73] He W, Feng Y, Zhao D. Improving knowledge base completion by incorporating implicit information. In: Proc. of the Joint Int'l Semantic Technology Conf. Berlin, Heidelberg: Springer-Verlag, 2015. 141–153.
- [74] Quillian MR. Semantic Memory. Bolt Beranek and Newman Inc., 1966.
- [75] Post EL. Formal reductions of the general combinatorial decision problem. *American Journal of Mathematics*, 1943,65(2):197–215.
- [76] Schank RC, Abelson RP. Scripts, plans, and knowledge. In: Proc. of the 4th Int'l Joint Conf. on Artificial Intelligence. 1975. 151–157.
- [77] Noy NF, McGuinness DL. *Ontology Development 101: A Guide To Creating Your First Ontology*. 2001.
- [78] Ehrlinger L, Wöß W. Towards a definition of knowledge graphs. In: Proc. of the Semantics. 2016.
- [79] Suda M, Weidenbach C, Wischniewski P. On the saturation of YAGO. In: Proc. of the 5th Int'l Joint Conf. on Automated Reasoning. Berlin, Heidelberg: Springer-Verlag, 2010. 441–456.
- [80] Wang WY, Mazaitis K, Cohen WW. Programming with personalized pagerank: A locally groundable first-order probabilistic logic. In: Proc. of the 22nd ACM Int'l Conf. on Information and Knowledge Management. New York: ACM Press, 2013. 2129–2138.
- [81] Wang WY, Mazaitis K, Lao N, Cohen WW. Efficient inference and learning in a large knowledge base: Reasoning with extracted information using a locally groundable first-order probabilistic logic. *Machine Learning*, 2015,100(1):101–126.
- [82] Cohen WW. TensorLog: A differentiable deductive database. arXiv Preprint arXiv: 160506523, 2016.
- [83] Jang S, Megawati M, Choi J, Yi M. Semi-Automatic quality assessment of linked data without requiring ontology. In: Proc. of the Int'l Semantic Web Conf. (ISWC) NLPDBpedia. Berlin, Heidelberg: Springer-Verlag, 2015. 45–55.

- [84] Böhmann L, Lehmann J. Pattern based knowledge base enrichment. In: Proc. of the 12th Int'l Semantic Web Conf. Berlin, Heidelberg: Springer-Verlag, 2013. 33–48.
- [85] Richardson M, Domingos P. Markov logic networks. *Machine Learning*, 2006,62(1):107–136.
- [86] Bröcheler M, Mihalkova L, Getoor L. Probabilistic similarity logic. In: Proc. of the 26th Conf. on Uncertainty in Artificial Intelligence. AUAI, 2010. 73–82.
- [87] Kimmig A, Bach S, Broecheler M, Huang B, Getoor L. A short introduction to probabilistic soft logic. In: Proc. of the NIPS Workshop on Probabilistic Programming: Foundations and Applications. 2012. 1–4.
- [88] Pujara J, Miao H, Getoor L, Cohen W. Knowledge graph identification. In: Proc. of the 12th Int'l Semantic Web Conf. Berlin, Heidelberg: Springer-Verlag, 2013. 542–557.
- [89] Pujara J, Miao H, Getoor L, Cohen WW. Ontology-Aware partitioning for knowledge graph identification. In: Proc. of the 3rd Workshop on Automated Knowledge Base Construction. New York: ACM Press, 2013. 19–24.
- [90] Chen Y, Goldberg S, Wang DZ, Johri SS. Ontological pathfinding: Mining first-order knowledge from large knowledge bases. In: Proc. of the 2016 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM Press, 2016. 835–846.
- [91] Miller GA. WordNet: A lexical database for English. *Communications of the ACM*, 1995,38(11):39–41.
- [92] Li M, Jia Y, Wang Y, Li J, Cheng X. Hierarchy-Based link prediction in knowledge graphs. In: Proc. of the 25th Int'l Conf. on World Wide Web. New York: ACM Press, 2016. 77–78.
- [93] Wen J, Li J, Mao Y, Chen S, Zhang R. On the representation and embedding of knowledge bases beyond binary relations. *arXiv Preprint arXiv: 160408642*, 2016.
- [94] Ji G, He S, Xu L, Liu K, Zhao J. Knowledge graph embedding via dynamic mapping matrix. In: Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics, Vol.1. Stroudsburg: ACL, 2015. 687–696.
- [95] Fan M, Zhou Q, Chang E, Zheng TF. Transition-Based knowledge graph embedding with relational mapping properties. In: Proc. of the 28th Pacific Asia Conf. on Language, Information and Computation. Stroudsburg: ACL, 2014. 328–337.
- [96] Ji G, Liu K, He S, Zhao J. Knowledge graph completion with adaptive sparse transfer matrix. In: Proc. of the 30th AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI, 2016. 985–991.
- [97] Hu Z, Huang P, Deng Y, Gao Y, Xing EP. Entity hierarchy embedding. In: Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics, Vol.1. Stroudsburg: ACL, 2015. 1292–1300.
- [98] Guo S, Wang Q, Wang B, Wang L, Guo L. Semantically smooth knowledge graph embedding. In: Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics, Vol.1. Stroudsburg: ACL, 2015. 84–94.
- [99] Wang Z, Zhang J, Feng J, Chen Z. Knowledge graph and text jointly embedding. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2014. 1591–1601.
- [100] Zhong H, Zhang J, Wang Z, Wan H, Chen Z. Aligning knowledge and text embeddings by entity descriptions. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2015. 267–272.
- [101] Wang Z, Li J. Text-Enhanced representation learning for knowledge graph. In: Proc. of the 25th Int'l Joint Conf. on Artificial Intelligence. Menlo Park: AAAI, 2016. 1293–1299.
- [102] Jia Y, Wang Y, Lin H, Jin X, Cheng X. Locally adaptive translation for knowledge graph embedding. In: Proc. of the 30th AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI, 2016. 992–998.
- [103] Jiang T, Liu T, Ge T, Sha L, Chang B, Li S, Sui Z. Towards time-aware knowledge graph completion. In: Proc. of the 26th Int'l Conf. on Computational Linguistics. Stroudsburg: ACL, 2016. 1715–1724.
- [104] Tay Y, Luu AT, Hui SC. Non-Parametric estimation of multiple embeddings for link prediction on dynamic knowledge graphs. In: Proc. of the 31st AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI, 2017. 1243–1249.
- [105] Nickel M, Tresp V, Kriegel HP. A three-way model for collective learning on multi-relational data. In: Proc. of the 28th Int'l Conf. on Machine Learning. New York: ACM Press, 2011. 809–816.
- [106] Chang KW, Yih WT, Yang B, Meek C. Typed tensor decomposition of knowledge bases for relation extraction. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg, 2014. 1568–1579.
- [107] Nickel M, Jiang X, Tresp V. Reducing the rank in relational factorization models by including observable patterns. In: Proc. of the Advances in Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2014. 1179–1187.

- [108] He W, Feng Y, Zou L, Zhao D. Knowledge base completion using matrix factorization. In: Proc. of the 17th Asia-Pacific Conf. on Web Technologies and Applications. Springer-Verlag, 2015. 256–267.
- [109] Tay Y, Luu AT, Hui SC, Brauer F. Random semantic tensor ensemble for scalable knowledge graph link prediction. In: Proc. of the 10th ACM Int'l Conf. on Web Search and Data Mining. New York: ACM Press, 2017. 751–760.
- [110] Xiao H, Huang M, Hao Y, Zhu X. TransG: A generative mixture model for knowledge graph embedding. arXiv Preprint arXiv: 150905488, 2015.
- [111] He S, Liu K, Ji G, Zhao J. Learning to represent knowledge graphs with gaussian embedding. In: Proc. of the 24th ACM Int'l Conf. on Information and Knowledge Management. New York: ACM Press, 2015. 623–632.
- [112] Socher R, Chen D, Manning CD, Ng A. Reasoning with neural tensor networks for knowledge base completion. In: Proc. of the Advances in Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2013. 926–934.
- [113] Chen D, Socher R, Manning CD, Ng AY. Learning new facts from knowledge bases with neural tensor networks and semantic word vectors. arXiv Preprint arXiv: 13013618, 2013.
- [114] Shi B, Weninger T. ProjE: Embedding projection for knowledge graph completion. In: Proc. of the 31st AAAI Conf. on Artificial Intelligence. Menlo Park, 2017. 1236–1242.
- [115] Han X, Le S. Context-Sensitive inference rule discovery: A graph-based method. In: Proc. of the 26th Int'l Conf. on Computational Linguistics. Stroudsburg: ACL, 2016. 2902–2911.
- [116] Wang Q, Wang B, Guo L. Knowledge base completion using embeddings and rules. In: Proc. of the 24th Int'l Joint Conf. on Artificial Intelligence. San Francisco: Morgan Kaufmann Publishers, 2015. 1859–1866.
- [117] Guo S, Ding B, Wang Q, Wang L, Wang B. Knowledge Base Completion via Rule-Enhanced Relational Learning. Springer-Verlag, 2016. 219–227.
- [118] Demeester T, Rocktäschel T, Riedel S. Regularizing relation representations by first-order implications. In: Proc. of the 5th Workshop on Automated Knowledge Base Construction. Stroudsburg: ACL, 2016. 75–80.
- [119] Toutanova K, Chen D, Pantel P, Poon H, Choudhury P, Gamon M. Representing text for joint embedding of text and knowledge bases. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2015. 1499–1509.
- [120] Xie R, Liu Z, Jia J, Luan H, Sun M. Representation learning of knowledge graphs with entity descriptions. In: Proc. of the 30th AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI, 2016. 2659–2665.
- [121] Schlichtkrull M, Kipf TN, Bloem P, Berg RVD, Titov I, Welling M. Modeling relational data with graph convolutional networks. arXiv Preprint arXiv: 170306103, 2017.
- [122] Yang B, Yih WT, He X, Gao J, Deng L. Embedding entities and relations for learning and inference in knowledge bases. arXiv Preprint arXiv:14126575, 2014.
- [123] Lao N, Cohen WW. Relational retrieval using a combination of path-constrained random walks. *Machine Learning*, 2010,81(1): 53–67.
- [124] Lao N, Mitchell T, Cohen WW. Random walk inference and learning in a large scale knowledge base. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2011. 529–539.
- [125] Zhao ZY, Jia YT, Wang YZ, Cheng XQ. Content-Structural relation inference in knowledge base. In: Proc. of the 28th AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI, 2014. 3154–3155.
- [126] Kotnis B, Bansal P, Talukdar PP. Knowledge base inference using bridging entities. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2015. 2038–2043.
- [127] Wang Q, Liu J, Luo Y, Wang B, Lin C. Knowledge base completion via coupled path ranking. In: Proc. of the 54th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2016. 1308–1318.
- [128] Wei Z, Zhao J, Liu K. Mining inference formulas by goal-directed random walks. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2016. 1379–1388.
- [129] Jia YT, Wang YZ, Cheng XQ, Jin XL, Guo JF. OpenKN: An open knowledge computational engine for network big data. In: Proc. of the 2014 IEEE/ACM Int'l Conf. on Advances in Social Networks Analysis and Mining. IEEE, 2014. 657–664.
- [130] Zhao ZY, Jia YT, Wang YZ, Jin XL, Cheng XQ. Temporal link prediction based on dynamic heterogeneous information network. *Journal of Computer Research and Development*, 2015,52(8):1735–1741 (in Chinese with English abstract).

- [131] Berant J, Dagan I, Goldberger J. Global learning of typed entailment rules. In: Proc. of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: ACL, 2011. 610–619.
- [132] Galárraga LA, Teflioudi C, Hose K, Suchanek F. AMIE: Association rule mining under incomplete evidence in ontological knowledge bases. In: Proc. of the 22nd Int'l Conf. on World Wide Web. New York: ACM Press, 2013. 413–422.
- [133] Galárraga L, Teflioudi C, Hose K, Suchanek FM. Fast rule mining in ontological knowledge bases with AMIE+. VLDB Endowment, 2015,24(6):707–730.
- [134] Gad-Elrab MH, Stepanova D, Urbani J, Weikum G. Exception-Enriched rule learning from knowledge graphs. In: Proc. of the 15th Int'l Semantic Web Conf. Berlin, Heidelberg: Springer-Verlag, 2016. 234–251.
- [135] Zhao ZY, Jia YT, Wang YZ, Jin XL, Cheng XQ. Link inference in large scale evolutional knowledge network. Journal of Computer Research and Development, 2016,53(2):492–502 (in Chinese with English abstract).
- [136] Liu Q, Jiang L, Han M, Liu Y, Qin Z. Hierarchical random walk inference in knowledge graphs. In: Proc. of the 39th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. New York: ACM, Press 2016. 445–454.
- [137] Guu K, Miller J, Liang P. Traversing knowledge graphs in vector space. arXiv Preprint arXiv: 150601094, 2015.
- [138] Lin X, Liang Y, Guan R. Compositional learning of relation paths embedding for knowledge base completion. arXiv Preprint arXiv: 161107232, 2016.
- [139] Feng J, Huang M, Yang Y, Zhu X. GAKE: Graph aware knowledge embedding. In: Proc. of the 26th Int'l Conf. on Computational Linguistics. Stroudsburg: ACL, 2016. 641–651.
- [140] Das R, Neelakantan A, Belanger D, McCallum A. Chains of reasoning over entities, relations, and text using recurrent neural networks. arXiv Preprint arXiv: 160701426, 2016.
- [141] Graves A, Wayne G, Reynolds M, Harley T, Danihelka I, Grabska-Barwińska A, Colmenarejo SG, Grefenstette E, Ramalho T, Agapiou J, Badia AP, Hermann KM, Zwols Y, Ostrovski G, Cain A, King H, Summerfield C, Blunsom P, Kavukcuoglu K, Hassabis D. Hybrid computing using a neural network with dynamic external memory. Nature, 2016,538(7626):471–476.
- [142] Shen Y, Huang PS, Chang MW, Gao J. Modeling large-scale structured relationships with shared memory for knowledge base completion. In: Proc. of the 2nd Workshop on Representation Learning for NLP. ACL, 2017. 57–68.
- [143] Gardner M, Talukdar PP, Kisiel B, Mitchell T. Improving learning and inference in a large knowledge-base using latent syntactic cues. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2013. 833–838.
- [144] Gardner M, Talukdar PP, Krishnamurthy J, Mitchell T. Incorporating vector space similarity in random walk inference over knowledge bases. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2014. 397–406.
- [145] Gardner M, Talukdar P, Mitchell T. Combining vector space embeddings with symbolic logical inference over open-domain text. In: Proc. of the AAAI 2015 Spring Symp. Series. Menlo Park: AAAI, 2015. 61–65.
- [146] Wang WY, Cohen WW. Learning first-order logic embeddings via matrix factorization. In: Proc. of the 25th Int'l Joint Conf. on Artificial Intelligence. Menlo Park: AAAI, 2016. 2132–2138.
- [147] Guo S, Wang Q, Wang L, Wang B, Guo L. Jointly embedding knowledge graphs and logical rules. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2016. 1488–1498.
- [148] Yang F, Yang Z, Cohen WW. Differentiable learning of logical rules for knowledge base reasoning. arXiv Preprint arXiv: 170208367, 2017.
- [149] Suchanek FM, Gross-Amblard D, Abiteboul S. Watermarking for ontologies. In: Proc. of the 10th Int'l Semantic Web Conf. Berlin, Heidelberg: Springer-Verlag, 2011. 697–713.
- [150] Min B, Grishman R, Wan L, Wang C, Gondek D. Distant supervision for relation extraction with an incomplete knowledge base. In: Proc. of the 2013 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: ACL, 2013. 777–782.
- [151] Dong XL, Gabrilovich E, Heitz G, Horn W, Murphy K, Sun S, Zhang W. From data fusion to knowledge fusion. VLDB Endowment, 2014,7(10):881–892.
- [152] Dietterich TG. Ensemble methods in machine learning. In: Proc. of the 1st Int'l Workshop on Multiple Classifier Systems. Berlin, Heidelberg: Springer-Verlag, 2000. 1–15.



- [153] Jia Y, Wang Y, Jin X, Zhao Z, Cheng X. Link inference in dynamic heterogeneous information network: A knapsack-based approach. IEEE Trans. on Computational Social Systems, 2017,4(3):80–92.

#### 附中文参考文献:

- [1] 王元卓,靳小龙,程学旗.网络大数据:现状与展望.计算机学报,2013,36(6):1125–1138.
- [2] 程学旗,靳小龙,王元卓,郭嘉丰,张铁赢,李国杰.大数据系统和分析技术综述.软件学报,2014,25(9):1889–1908. <http://www.jos.org.cn/1000-9825/4674.htm> [doi: 10.13328/j.cnki.jos.004674]
- [3] 王元卓,贾岩涛,刘大伟,靳小龙,程学旗.基于开放网络知识的信息检索与数据挖掘.计算机研究与发展,2015,52(2):456–474.
- [27] 赵军,刘康,周光有,蔡黎.开放式文本信息抽取.中文信息学报,2011,25(6):98–110.
- [28] 刘峤,李杨,段宏,刘瑶,秦志光.知识图谱构建技术综述.计算机研究与发展,2016,53(3):582–600.
- [29] 庄严,李国良,冯建华.知识库实体对齐技术综述.计算机研究与发展,2016,53(1):165–192.
- [30] 刘知远,孙茂松,林衍凯,谢若冰.知识表示学习研究进展.计算机研究与发展,2016,53(2):247–261.
- [31] 林海伦,王元卓,贾岩涛,张鹏,王伟平.面向网络大数据的知识融合方法综述.计算机学报,2017,40(1):1–27.
- [32] 徐增林,盛泳潘,贺丽荣,王雅芳.知识图谱技术综述.电子科技大学学报,2016,45(4):589–606.
- [33] 王永庆.人工智能原理与方法.西安:西安交通大学出版社,1998.
- [37] 陆汝铃,应明生.知识推理的一个模型.中国科学 E 辑:技术科学,1998,28(4):363–369.
- [52] 刘亚彬,刘大有.空间推理与地理信息系统综述.软件学报,2000,11(12):1598–1606. [http://www.jos.org.cn/jos/ch/reader/create\\_pdf.aspx?file\\_no=20001206&journal\\_id=jos](http://www.jos.org.cn/jos/ch/reader/create_pdf.aspx?file_no=20001206&journal_id=jos)
- [61] 赵洁心,潘正华,王姗姗.基于 FLcom 的模糊知识推理与搜索处理.计算机工程与应用,2015,51(19):37–42.
- [130] 赵泽亚,贾岩涛,王元卓,靳小龙,程学旗.基于动态异构信息网络的时序关系预测.计算机研究与发展,2015,52(8):1735–1741.
- [135] 赵泽亚,贾岩涛,王元卓,靳小龙,程学旗.大规模演化知识网络中的关联推理.计算机研究与发展,2016,53(2):492–502.



官赛萍(1991—),女,福建三明人,学士,主要研究领域为知识图谱.



王元卓(1978—),男,博士,研究员,博士生导师,CCF 杰出会员,主要研究领域为网络大数据分析,开放知识计算,社交网络演化计算,网络与信息安全.



靳小龙(1976—),男,博士,研究员,CCF 高级会员,主要研究领域为知识图谱,社会计算,大数据.



程学旗(1971—),男,博士,研究员,博士生导师,CCF 会士,主要研究领域为网络科学与社会计算,互联网搜索与挖掘,网络信息安全,分布式系统,大型仿真平台.



贾岩涛(1983—),男,博士,副研究员,CCF 专业会员,主要研究领域为开放知识计算,数据挖掘.