

分 类 号 TP391

密 级

U D C

编 号 10486

武汉大学

硕 士 专 业 学 位 论 文

跨域知识图谱的知识  
表示学习研究

研 究 生 姓 名：杨世杰

学 号：2021282110124

指导教师姓名、职称：彭敏 教授

专 业 类 别（领 域）：计算机科学与技术

二〇二三年四月

# Cross-domain knowledge graph represents learning research

Candidate: Yang Shijie

StudentNumber: 2021282110124

Supervisor: Prof.Peng Min

Major: Computer Science and Technology

Speciality: Natural Language Processing



School of Computer Science  
Wuhan University

April, 2023

## 论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的研究成果。除文中已经标明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对本文的研究做出贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

学位论文作者（签名）：

年    月    日

## 摘 要

随着知识图谱在工业界的广泛应用，众多基于知识图谱的应用进入了人们的生活，如电商系统、医疗健康管理系统等。为了更充分利用图谱中的知识，知识表示学习将图谱中的实体和关系嵌入到低维连续空间，保留其语义信息的同时有助于各种下游任务的应用。通过知识表示学习，可以来预测用户的商品喜好或疾病诊断等。随着知识图谱数据规模的增长以及在个人电脑、手机等移动设备上的应用，知识图谱的处理和存储开始逐渐分布在多个不同的节点上。上述节点上的目标域知识图谱通常是源域知识图谱的一个子集，在用户使用过程中会不断引入其他的实体和关系。然而，这些引入的实体和关系可能包含源域知识图谱实体集合和关系集合中未定义的新的未见实体和关系，经典的知识表示方法无法获取这些未见实体和未见关系的向量表示，也就难以很好地完成知识表示任务。此外，考虑到合并的成本以及保护用户数据隐私的需求，目标域和源域知识图谱无法进行知识融合，从而形成了跨域知识表示问题。

为了解决这个问题，论文借助元学习的算法思想，在训练任务中通过标签模拟目标域知识图谱上的未见实体和关系，以帮助模型学习到处理目标域图谱新实体和关系的能力。同时引入本体信息为未见实体和关系的表示学习提供语义支持。主要解决了跨域知识表示学习中的两个问题：（1）当前处理跨域知识图谱的表示学习模型多采用图结构信息来学习未见实体和未见关系的表示，未能充分利用知识图谱其他语义信息来对表示学习进行补充。（2）面对跨域知识图谱时，如何在训练中学习对对新实体和新关系的表示能力，并且将模型的学习能力泛化到目标域知识图谱的表示学习任务上。

本文面向跨域知识图谱的知识表示学习，提出了一种基于本体信息和元学习的知识表示学习框架。主要包括：面向跨域知识表示学习中出现的新关系，采用关系图卷积网络，结合关系的拓扑信息和基于关系拓扑结构和描述文本的本体信息，进行联合学习，获得对未见关系的表示；面向跨域知识表示学习中新出现的实体，使用实体的邻接关系特征聚合获得初始化表示，再采用复杂图卷积网络，充分利用知识图谱中已知实体和关系的信息，学习和更新整个目标域的实体和关系的向量表示；在模型训练过程中，采用元学习方法划分任务并通过标签模拟跨域场景中的新实体和新关系，从而完成基于源域知识嵌入学习到目标域知识表示学习任务。

本文基于链接预测任务对跨域知识表示学习的有效性进行验证，并与多个基准模型进行比较。实验结果证明了本文提出的基于元学习本体增强的跨域知识表示学习模型的有效性。

**关键词：**知识表示学习；本体嵌入；元学习；图神经网络；归纳推理

## ABSTRACT

With the wide application of knowledge graphs in the industry, numerous knowledge-graph-based applications have entered people's lives, such as e-commerce systems, medical health management systems, etc. In order to fully utilize the knowledge in the graph, knowledge representation learning embeds entities and relationships in the graph into a low-dimensional continuous space while preserving their semantic information, which helps various downstream tasks applications. Through knowledge representation learning, it is possible to predict user preferences for products or disease diagnosis. With the growth of knowledge graph data and the application on personal computers, mobile phones, and other mobile devices, the processing and storage of knowledge graphs are gradually distributed across multiple nodes. The target domain knowledge graph on the above nodes is usually a subset of the source domain knowledge graph and will continuously introduce other entities and relationships during user use. However, these introduced entities and relationships may contain new unseen entities and relationships that are not defined in the source domain knowledge graph entity set and relationship set. Classic knowledge representation methods cannot obtain vector representations of these unseen entities and relationships, making it difficult to complete the knowledge representation task well. In addition, considering the cost of merging and the need to protect user data privacy, the target domain and source domain knowledge graphs cannot be merged, resulting in the problem of cross-domain knowledge representation.

To solve this problem, this paper uses the algorithmic idea of meta-learning to simulate unseen entities and relationships on the target domain knowledge graph through labels in the training task, helping the model learn the ability to handle new entities and relationships in the target domain graph. At the same time, ontology information is introduced to provide semantic support for the representation learning of unseen entities and relationships. This mainly solves two problems in cross-domain knowledge representation learning: (1) current representation learning models for cross-domain knowledge graphs mostly use graphic structural information to learn the representation of unseen entities and relationships and have not been able to fully utilize other semantic information in the knowledge graph to supplement the representation learning. (2) Faced with cross-domain knowledge graphs, how to learn the representation ability of new entities and new relationships in training and generalize the model's learning ability to the representation learning task of the target domain knowledge graph.

This paper proposes a knowledge representation learning framework based on ontology

information and meta-learning for cross-domain knowledge graphs. It mainly includes: for newly appeared relationships in cross-domain knowledge representation learning, using relationship graph convolutional networks to combine the topology information of relationships and ontology information based on relationship topology structure and description text to jointly learn and obtain the representation of unseen relationships; for newly appeared entities in cross-domain knowledge representation learning, using the adjacency relationship features of entities to aggregate the initial representation and then using a complex graph convolutional network to fully utilize the known entity and relationship information in the knowledge graph to learn and update the vector representation of the entire target domain entities and relationships. In the model training process, meta-learning is used to partition tasks and simulate new entities and new relationships in cross-domain scenarios through labels, thus completing the knowledge representation learning task learned from the source domain knowledge embedding to the target domain knowledge representation learning.

This paper verifies the effectiveness of cross-domain knowledge representation learning based on link prediction tasks and compares with multiple benchmark models. The experimental results prove the effectiveness of the proposed meta-learning ontology-enhanced cross-domain knowledge representation learning model.

**Key words:** knowledge presentation learning; ontology embedding; meta-learning; graph neural networks; inductive reasoning

# 目 录

摘 要	I
ABSTRACT	II
1 绪论	1
1.1 研究背景和意义	1
1.2 国内外研究现状及趋势	2
1.2.1 融合辅助信息的跨域表示学习方法	2
1.2.2 基于归纳推理的跨域表示学习方法	4
1.2.3 元学习在知识表示学习中的应用	6
1.3 本文主要研究内容	7
1.4 本文组织结构	8
2 跨域知识图谱的知识表示学习关键技术	9
2.1 跨域知识图谱的知识表示学习定义	9
2.2 本体嵌入关键技术	10
2.3 基于 GNN 的跨域知识表示学习关键技术	13
2.4 元学习训练方法	15
2.5 本章小结	16
3 基于关系拓扑结构及描述文本的本体信息嵌入	17
3.1 基于关系拓扑结构及描述文本的本体嵌入框架	17
3.2 关系本体图构建	18
3.2.1 关系本体三元组	18
3.2.2 关系位置元关系	19
3.3 基于描述文本的本体表示增强	21
3.3.1 本体结构信息嵌入	22
3.3.2 基于文本的本体嵌入	23
3.4 本章小结	23
4 基于元学习本体增强的跨域知识表示学习模型	25
4.1 模型整体架构	25
4.2 未见关系的特征嵌入	26
4.2.1 关系位置图构建	26

4.2.2	关系相关性系数聚合编码 .....	27
4.3	未见实体的特征嵌入 .....	28
4.3.1	基于关系聚合的实体表示 .....	28
4.3.2	基于 GNN 的实体关系联合嵌入 .....	29
4.4	基于元学习的训练任务设定 .....	29
4.5	基于链接预测任务的模型实现 .....	31
4.6	本章小结 .....	32
5	实验结果及分析 .....	33
5.1	数据集 .....	33
5.1.1	源数据集介绍 .....	33
5.1.2	任务数据集构建 .....	34
5.2	模型参数设置 .....	34
5.3	实验设计及评价指标 .....	35
5.4	实验结果及分析 .....	37
5.5	模型消融实验 .....	39
5.6	未见实体案例分析 .....	40
5.7	未见关系案例分析 .....	41
5.8	本章小结 .....	41
6	总结与展望 .....	43
6.1	总结 .....	43
6.2	未来工作 .....	44
	参考文献 .....	45
	致谢 .....	49
	攻硕期间取得的学术成果和参与的项目 .....	50



# 1 绪论

## 1.1 研究背景和意义

近几年,知识图谱(Knowledge Graph, KG)已经成为许多需要访问结构化知识的信息系统的基础<sup>[1]</sup>。知识图谱通过基于图的数据结构对人类知识进行存储,图中的节点和边代表了现实世界的实体和实体间的关系。典型的知识图谱有 Freebase<sup>[2]</sup>、NELL-995<sup>[3]</sup>、DBpedia<sup>[4]</sup>、YAGO<sup>[5]</sup> 等。由于知识图谱数据规模的增长以及在个人电脑、手机等移动设备上的应用,知识图谱的处理和存储开始逐渐分布在多个不同的节点上。例如,移动设备上的电商系统可以利用图谱对用户的商品喜好进行推理和预测。受到资源环境的限制,这些应用上的目标域知识图谱往往是覆盖范围最广、存储知识最多的源域知识图谱聚焦于某一部分的子集。并且随着图谱的应用,目标域知识图谱会不断引入源域图谱未定义的新的未见实体和关系。然而知识图谱中的知识非常庞杂,且往往是隐式或深层次的,难以直接利用或获取到有价值的信息。为了解决这个问题,知识图谱嵌入(Knowledge Graph Embedding, KGE)等的知识表示学习(Knowledge Representation Learning, KRL)方法的相关研究很快起步并获得了广泛的关注。知识图谱嵌入旨在将符号化实体和关系映射到低维稠密的向量空间,便于它们之间的计算<sup>[6]</sup>,以完成知识图谱上的众多下游任务,如知识图谱补全、三元组分类等。

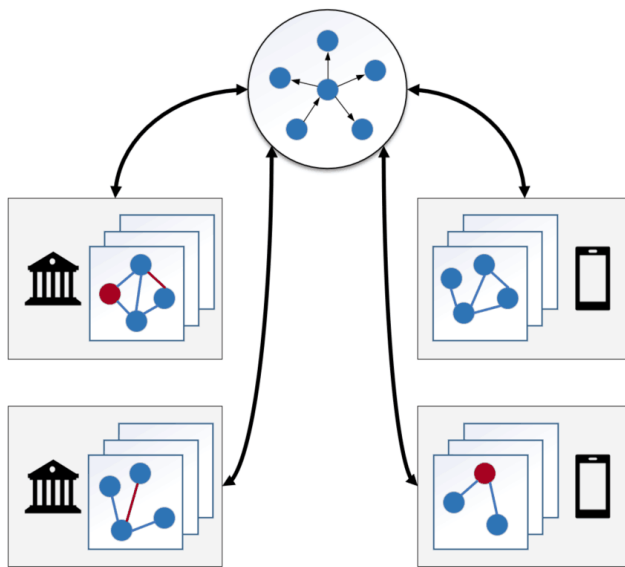


图 1.1 跨设备下的知识图谱分布

传统的知识表示学习任务在训练集上学习到知识图谱的嵌入表示,在测试集上通过获得实体或关系对应的向量直接进行计算。而在跨域知识图谱上,目标域知识图谱存在

源域知识图谱中未定义的实体和关系，无法直接从源域知识图谱的嵌入中获取到对应的向量表示。同时出于用户隐私、成本等方面的考量，无法将新的实体和关系在源域知识图谱中进行知识融合并重新训练。因此跨域知识表示学习任务要求表示学习方法，能够从源域知识图谱的训练中学习对目标域知识图谱中未见实体和未见关系的表示能力。现有的基于规则学习的跨域知识表示学习方法通过自主学习逻辑规则来实现图谱嵌入，但采用置信度评估规则质量，降低了规则的可信度，并且基于规则的模型也存在可扩展性以及大规模知识图谱训练效率的问题。基于图神经网络的跨域知识表示学习模型聚合邻域节点特征对目标域知识图谱上新的实体和关系进行表示，但这些模型没有考虑到实体与关系的联系，同时利用图结构信息进行嵌入的过程中，忽略了如本体信息这类高度抽象的语义信息对表示学习的作用。

为了解决上述在跨域知识图谱的表示学习上仍存在的问题，本文提出了一种基于本体信息和元学习的知识表示学习方法。通过对图谱结构信息和本体语义信息的特征学习，能够对目标域知识图谱上的新的实体和关系进行有效的嵌入表示，为跨域知识图谱的表示学习提供了一种新的解决方案。同时，本文通过采用元学习的方法在训练任务上模拟出未见的实体和未见的关系，使得模型能够通过多个训练任务的训练快速收敛，并在测试任务上取得良好的效果。在现实世界的应用中，可以采用本模型在大型的源知识图谱上进行训练，获得最优的模型参数，然后在各低资源的环境上，直接对目标知识图谱上进行小规模的知识图谱嵌入，可用于推荐系统、问答系统等下游应用等。能够有效解决跨域知识图谱的表示学习问题，同时减小整个训练流程的成本消耗，提高源知识图谱的样本利用率。

## 1.2 国内外研究现状及趋势

跨域知识表示学习一般涉及到融合辅助信息的知识表示学习方法和基于归纳推理的知识表示学习方法。近几年，元学习在知识表示学习中也获得众多学者的关注，用于在少样本、零样本等情境下对表示学习能力的加强。本节将阐述相关方法的研究现状及趋势。

### 1.2.1 融合辅助信息的跨域表示学习方法

经典的 KGE 方法主要包含基于翻译的表示学习方法和基于语义相似度的方法。基于翻译的方法通过测量低维向量间的距离来计算出三元组事实的可信度。这些方法通常会将关系的嵌入视为从头结点到尾结点的向量平移，代表模型有 TransE 及其有关的一些变体如 TransH<sup>[7]</sup>、TransR<sup>[8]</sup> 等。基于语义相似度的方法通过比较实体和关系在向量空间表示中隐藏的语义特征的相似程度，来评判三元组的合理性，代表模型有 RESCAL、DistMult 和 ComplEx 等。这两类方法的本质都是通过设定的打分函数对图谱实体和关系的语义表示向量进行更新，使得表示向量可以更多地贴近源知识图谱中的事实。因此，

在传统的、目标域知识图谱不存在源域未见实体和未见关系的知识图谱补全任务上，经典的 KGE 方法能够取得很好的效果。然而，在面向跨域知识图谱的补全任务上，经典的 KGE 方法无法从源知识图谱中学习到新的实体和关系有关的语义信息，不能很好完成对新三元组的预测任务。

为了对目标域知识图谱中存在的未见实体进行更有效的特征学习，一些学者尝试引入其他辅助信息，以增加表示向量中的隐含信息。知识图谱除了通过事实三元组对知识进行存储，同时还蕴含着丰富的其他信息可用于加强知识表示学习，如关系路径信息、图结构信息等。也有一些研究者通过使用图谱外的信息，如文本描述信息等增强传统的 KGE 模型的表示能力。

NTN<sup>[9]</sup> 模型最早在表示学习中引入实体的描述信息，该模型对实体进行嵌入时并没有采用随机初始化的嵌入，而是将实体的文本描述信息编码作为实体的表示向量。此后描述扩展的知识图谱表示模型 DKRL<sup>[10]</sup> 试图通过改进 TransE 模型，使其能够进一步处理实体描述。DKRL 通过将实体结构相关的向量和实体描述相关的向量联合表示，来作为实体嵌入。其中实体描述向量由描述文本经过连续词袋<sup>[11]</sup> 编码器或卷积神经网络编码器获得。另一种常见的引入辅助信息的方法借助实体的关系信息来进行表示学习，关系信息揭示了实体之间的一个或多个语义联系。PtransE<sup>[12]</sup> 首次将图谱的多跳关系信息作为知识图谱嵌入的辅助信息使用，提出了一个以关系路径为基础表示学习模型。该模型是在 TransE 的基础上进行改进的。新的模型打分函数由两个部分组成，其中一部分是针对头尾实体直接相连关系路径的打分；另一部分是针对头尾实体间其他多跳关系路径的打分。这一改进使得 TransE 的单步推理得以扩展为多步推理，提升了在链接预测任务上的表现。而冯俊等人则将知识图谱视为一个大的有向图，提出了基于图感知的表示模型 GAKE<sup>[13]</sup>，该模型利用知识图的结构信息生成实体和关系的表示。

除了引入知识图谱自身的隐含信息辅助表示学习外，部分学者也会引入图谱外的其他知识用于表示学习。例如上述的实体描述信息，也可以从新闻稿或者维基百科中获取。图谱外其他模态的信息也同样适用于辅助表示学习，如实体的图像信息。代表性的 IKRL<sup>[14]</sup> 模型实现了包含基于跨模态结构与基于图像的联合表示方法，该模型在遵循平移原则的基础上将图像编码到实体空间。IKRL 提出的跨模态表示可以确保基于结构的表示和基于图像的表示映射在同一个表示空间中。此外，还有其他辅助信息可以被用于加强表示学习，如 Ren<sup>[15]</sup> 团队提出的基于语义证据的表示学习方法。Ren 等人通过对语义证据的预测及实验验证，研究了 KGE 的外推问题，并且通过建模实现了对三种语义证据的加强，在存在有训练集未见的实体的知识图谱补全任务上，相比于传统的 KGE 方法取得了更好的表现效果。

### 1.2.2 基于归纳推理的跨域表示学习方法

为了能够解决传统 KGE 方法无法处理目标域知识图谱中存在的未见实体和未见关系的问题,学者们开始探索基于归纳推理的知识表示学习方法。基于归纳推理的知识表示方法通过对结构或者一般规则的学习尽量减小对实体和关系的依赖,从而能够将表示学习能力泛化到对目标域知识图谱的新三元组的预测任务上。基于归纳推理的知识表示学习方法主要分为基于规则抽取的表示学习方法和基于 GNN 的表示学习方法,本节将分别介绍两种方法的研究现状。

#### (1) 基于规则抽取的表示学习方法

基于规则的表示学习方法通过学习关系的共现模式来挖掘逻辑规则。规则从形式语言表达能力角度可分为“命题规则”和“一阶逻辑规则”<sup>[16]</sup>。命题规则不包含变量而是限定了具体的关系及实体,因此具有很强的限制性,无法应用到其他实体的推理任务上。一阶逻辑规则包含了关系及变量,可以视为对命题规则的抽象,与具体的实体无关。因此这些规则从本质上就获得了归纳推理的能力,摆脱了实体依赖的限制,可以对存在新实体和关系的三元组上的任务进行推理指导。

为了对知识图谱事实三元组中隐含的规则进行挖掘,传统的规则挖掘方法如 AMIE<sup>[17]</sup>模型基于路径遍历的思想实现挖掘算法。这些模型将知识图的关系路径近似视为规则进行提取,通过采用统计度量或者固定的人工设定的模式进行规则的学习。但是由于知识图谱的复杂关系,遍历路径会使得规则提取的成本大大提高,无法应用于大型知识图谱。并且这些传统的规则提取方法存在无法扩展的问题。

随着知识表示学习方法的广泛应用,基于表示学习的规则挖掘方法通过对实体和关系的向量打分来挖掘规则,典型的模型有 RLvLR<sup>[18]</sup>、HARL<sup>[19]</sup>、IterE<sup>[20]</sup>等。RLvLR 模型通过联合表示学习和子图采样方法从图谱中学习规则,能够在大型的知识图谱上如 Freebase 和 YAGO 上进行有效的规则提取。HARL 在 RLvLR 的基础上考虑了实体的属性信息,添加了对实体属性的规则学习。IterE 模型为了能够获得对知识图谱中存在的大量的稀疏实体和关系更准确的表示,提出了表示学习和规则学习同时进行的迭代算法。该模型在表示学习的基础上添加了公理归纳和公理注入的模块,对稀疏的实体和关系通过公理注入来加强表示学习的效果,获得了规则和表示学习的共同提升效果。国内刘藤<sup>[21]</sup>等人在 IterE 模型的基础上进行了改进,专注于改进规则学习和规则融合模块。他们通过基于三元组的打分函数来对规则置信度计算方法进行改进,扩大了模型的适用性。此外,他们还迭代实现了利用表示更新规则以及通过规则增强表示,进一步提高了该方法的性能。

专注于跨域知识图谱的链接预测任务上,一些可微规则学习方法如 Neural-LP<sup>[22]</sup>模型和 DURM<sup>[23]</sup>模型,采用端到端的方法学习规则逻辑。此类方法通过设计可微的逻辑规则学习模型,并采用基于梯度的方法进行优化求解。然而这些方法不能解决知识图

谱中缺边的问题，同时在处理候选规则中的不合理部分仍存在不足。不同于直接从事三元组中学习逻辑规则的方法，一些方法利用子图隐式地表示逻辑规则。GraIL[9] 和 TACT<sup>[24]</sup> 通过对未见实体周围的封闭子图的学习逻辑规则，但是很显然随着实体邻居数目的指数增长，这些子图的规模可能会很庞大从而可能导致效率问题。

## (2) 基于 GNN 的表示学习方法

基于 GNN 的表示学习方法在传统的知识图谱嵌入方法上进行延伸，利用基于图神经网络的方法作为编码器学习图结构信息，然后采用 TransE、ComplEx 等传统知识图谱嵌入方法作为解码器。

Hamaguchi<sup>[25]</sup> 等人将图神经网络应用到知识图谱上。不同于先前学者利用 GNN 将整个图映射成一个图向量的方法，Hamaguchi 利用基于 GNN 的模型结构，将图中的实体和关系分别映射到单独的向量中，以此满足知识图谱补全等任务的需要。对于预测三元组中的新实体，该方法通过汇总所有已知实体的信息来生成新实体的嵌入向量。在聚合邻居节点特征时，该方法使用了简单的池化操作，没有对邻居节点进行区分或关注关系信息。尽管存在这些缺点，但该方法成功地展示了图神经网络在知识图谱上的有效应用，表明 GNN 通过编码节点信息和图的拓扑结构能够处理跨域知识图谱中的未见实体，并且催生了许多基于图神经网络的知识图谱嵌入方法。

图卷积网络<sup>[26]</sup> (Graph Convolutional Network, 简称 GCN) 是在卷积神经网络和图神经网络的基础上发展而来的。GCN 通过在节点特征和邻居信息上进行卷积操作来更新节点表示，以更充分地利用邻居节点之间的信息传递。然而，GCN 只适用于同构图，无法处理知识图谱中复杂的关系结构。而关系图卷积网络<sup>[27]</sup> (Relational Graph Convolutional Network, R-GCN) 则为不同类型的关系设置不同的权重矩阵，将 GCN 模型扩展到了关系图中。VR-GCN<sup>[28]</sup> 则考虑了知识图谱中关系方向以及不同实体类型可能具有不同特征空间这些问题。为了解决这些问题，VR-GCN 引入了 TransE 的思想来学习实体和关系的嵌入。而 TransGCN<sup>[29]</sup> 则在 VR-GCN 的基础上，通过利用 RotatE 思想对聚合函数进行扩展，提出了一种新的处理异构关系的方法。CompGCN<sup>[30]</sup> 通过集合 TransE、DistMult 和 HoIE 的思想设置了三种不同的聚合操作，对图谱中的实体和关系信息同时进行学习和表示。同时该模型通过在各层共享关系嵌入解决了过参数化的缺陷。

不同于上述模型对关系信息的关注和加强，受到注意力机制在基于序列任务中的优秀表现的启发，图注意力网络<sup>[31]</sup> (Graph Attention Networks, 简称 GAT) 将该机制引入到图卷积网络中并用于对图数据中的节点进行分类，以增加对邻接点的关注度的区分。该机制可以在节点-邻居对之间进行并行高效的计算，并且可以针对具有不同度数的图节点指定邻居的任意权重。相关学者的实验结果也表明，该机制的引入同样可以高效地作用于归纳学习问题以及处理其他任务。此外，r-GAT<sup>[32]</sup> (Relational Graph Attention Network) 采用多通道的方法学习实体的嵌入表示，其中每个通道对应不同的实体语义

信息。该方法还利用关系特征聚合邻居信息，从而能够有效处理复杂的关系图。

### 1.2.3 元学习在知识表示学习中的应用

元学习旨在帮助机器学习系统具备学习如何学习的能力。其重要应用场景包括零样本学习、小样本学习等数据稀少情况下，能够更快、更好地学习新的知识表示，并能够泛化至新的任务。

**MetaR**<sup>[33]</sup> 模型首次将元学习应用到少样本的链接预测任务上。该模型尝试通过元学习的方法捕捉源知识图谱事实三元组隐含的关系相关的元信息。这些关系相关的元信息通常是针对一类任务中常见且通用的信息，因此可以将所学到的元信息迁移到新的三元组中。并且，通过使用小批量的三元组进行元学习，可以加速整个学习过程。**MetaR** 采用 **TransE** 作为评分函数，并且结合元学习的方法解决少样本场景下的知识图谱补全任务，证明了元学习在跨域图谱上进行知识表示学习的有效性。**Meta-KGR**<sup>[34]</sup> 模型提出了一种将强化学习与元学习结合方法，以解决少样本的关系推理问题。该模型通过采用强化学习，训练了一个代理，用于搜索目标实体和推理路径。在测试任务中，根据查询的三元组中的关系划分任务，并利用元学习在高频任务上提取元信息进行知识迁移。此外，模型多跳路径的设置提高了可解释性。为了解决知识图谱补全过程中因为邻域稀疏导致构建邻域表示时邻域噪音过大的问题，**GANA**<sup>[35]</sup> 模型引入了图注意力机制和门控网络对噪音进行过滤。同时将采用 **TransH** 作为评分函数对 **MetaR** 进行改进，提高了模型对复杂关系的表示能力。

随着图神经网络在知识表示学习中的良好应用，结合 GNN 和元学习方法的研究也受到了学者们的关注。**Meta-iKG**<sup>[36]</sup> 包含基于子图的元学习器。它将链接预测任务翻译为子图建模问题，并采用局部子图传递子图特定信息。**Meta-iKG** 通过 GNN 网络对特定关系的子图进行编码，并使用元学习更快地学习到可迁移的子图信息。然而，该方法在对子图进行编码的过程中，仅提取了子图的结构语义，无法很好地处理反对称关系的三元组。最近提出的 **MorsE**<sup>[37]</sup> 同样结合了 GNN 和元学习方法，用于解决跨域知识图谱的知识表示学习中存在的新实体的问题。该模型将元知识学习分为实体初始化和图神经网络调制器两个模块。实体初始化器通过两个与实体无关的嵌入（即关系域嵌入和关系范围嵌入）初始化每个实体嵌入。而 GNN 调制器则利用实体的邻居结构信息增强实体嵌入。通过对与实体无关的元知识进行建模和学习，**MorsE** 可以为新实体生成高质量的嵌入。

整体而言，元学习方法更关注于获取关系信息。模型整体结构分为两个部分。第一部分负责融合信息，以获取任务关系的表示；第二部分利用元学习加速更新过程，以达到快速适应新关系的目的。为了满足具体任务的需要，这些方法往往会通过结合其他方法进行加强。

本文在基于元学习训练方法，结合 GNN 对关系进行结构信息和语义信息的加强，

以解决跨域知识图谱的知识表示存在未见实体和未见组件的问题。

### 1.3 本文主要研究内容

针对跨域知识图谱的知识表示学习，本文研究发现经典的知识图谱嵌入方法无法很好处理未见训练图谱中的新实体和新关系。现有的跨域知识表示学习方法没有充分利用到未见实体与关系的相关性，着重于采用图结构信息对实体和关系进行编码，且未考虑通用信息（如本体信息）对语义的补充。本文总结出以下两个主要难题：

- 1) 传统的知识图谱嵌入方法无法很好处理跨域知识图谱嵌入中未见实体和未见关系。基于归纳推理的模型，希望通过子图结构信息来学习新实体和关系的表示，但忽略了知识图谱其他语义信息对表示学习的补充。
- 2) 跨域知识图谱的知识表示学习要求模型能够在源域知识图谱上训练获得参数，并能够将学习能力迁移到存在新实体和关系的目标域知识图谱嵌入上。在跨域知识图谱场景下，如何使得模型在训练过程中提高对未见实体和关系的泛化学习能力。

许多学者已经尝试利用文本描述、属性描述、实体类型等多种辅助信息为表示学习添加更丰富的语义信息，但这些信息往往集中于局部的特征，例如实体类型强调类之间的层次关系，无法同时对关系和实体提供补充。相比之下，知识图谱本体可以提供更加丰富且完整的语义信息，包括实体类型、层次信息和关系信息。因此，本文设计了一种利用本体信息增强跨域知识表示学习的嵌入模型。为了更好地建模未见关系，模型使用本体信息加强关系的语义信息，同时利用关系间的相对位置构建关系位置图，以学习关系的结构信息。对于未见实体，本文通过聚合实体相连接的关系特征作为初始化表示。然后对图上所有的实体和关系采用 GNN 网络再次更新，充分利用到图谱中已知实体和关系的信息，得到了最终的表示向量。此外，为了模拟目标域知识图谱中存在未见关系和未见实体的场景，本文采用了元学习思想设计多个训练任务，以快速学习最佳模型参数。最后，本文进行了充分的实验和分析，验证了模型的有效性。

综上，本文的研究内容主要包括：

- 1) 针对本体嵌入中关系信息不充分的问题，通过构建关系本体图获得关系的位置元关系，补充了关系本体三元组。同时基于描述文本对本体嵌入进行加强。
- 2) 提出了一种嵌入学习模型框架，采用元学习的模型训练方法，分别训练多个单任务，并在各个训练任务中通过标签模拟未见实体和关系，融合本体信息和结构信息对新实体和新关系进行嵌入。在源域知识图谱上训练模型，并将训练的参数用于目标域的图谱嵌入上，提升了模型对新实体和关系的泛化学习能力。
- 3) 在多个数据集上进行了模型测试并与其他基准模型结果进行对比分析，验证了模型的有效性。同时通过一系列消融实验验证模型各部分的重要性，为之后的模型改进提供思路 and 支撑。

## 1.4 本文组织结构

本文的内容分为五章，以下主要概括各章的内容及安排：

第一章，绪论部分：主要介绍跨域知识图谱的知识表示学习相关研究，探讨其现实意义。然后综述了国内外研究者在融合辅助信息、基于归纳推理和元学习等方面的研究进展。最后概括了当前研究所面临的问题，并提出了解决思路。

第二章，跨域知识图谱的知识表示学习关键技术研究：分三个小结介绍了本文模型主要涉及到的三种技术的发展状况及原理分析。其中，本体嵌入作为模型主要的语义信息补充，介绍了现有的本体嵌入方法。本文模型作为基于 GNN 的表示学习方法的实现，介绍了跨域知识表示学习模型所涉及的关键技术。最后，本文简要介绍了元学习的主要技术及相关思想。

第三章，基于关系拓扑结构及描述文本的本体信息嵌入：介绍了如何捕捉本体的语义信息，并进行本体三元组的嵌入。为了补充本体中关系相关的三元组，本文采用了两种方法进行关系相关本体三元组的提取。在本体三元组结构嵌入的基础上，本章还详细说明了如何使用本体描述文本来增强本体嵌入的语义。

第四章，基于元学习本体增强的跨域知识表示模型：首先介绍了模型的元学习训练任务设置。随后，详细介绍了本文提出的跨域知识表示学习模型的各个组成部分，主要包括未见关系嵌入、未见实体嵌入以及对实体和关系的联合更新模块。在获取实体和关系的向量表示后，采用多种 KGE 得分函数来进行评分和更新。最后，介绍了链接预测任务和相应的训练流程。

第五章，实验结果及分析：介绍了本文中使用的跨域表示学习任务评测所需的两个数据集的构建方法。在目标数据集上，本文对本文中的模型和相关基准模型进行了充分的实验，并比较了实验结果，以验证模型的有效性。最后，本章进行了消融实验，对本文模型的各组成模块进行了验证，以验证各模块的重要性。

总结部分：综合全文的研究内容及实验结果，总结了模型的可取部分及仍存在的问题与可改进的方面，对未来的相关工作进行规划。



## 2 跨域知识图谱的知识表示学习关键技术

传统的知识图谱嵌入模型通过提取三元组中的事实特征来对实体和关系进行编码。这类模型设计了评分函数，如 TransE、RESCAL 等，来对嵌入后的三元组进行打分，从而改进了图谱嵌入的效果。然而，在跨域的混合知识图谱的实际应用下，图谱会引入训练数据中不存在的实体和关系。由于传统的嵌入模型不能很好地从训练集中学习到这些组件的特征表示，因此局限性较大。基于规则和归纳推理的方法通常仅集中于表示单个未见部分，而从知识图谱的子图或关系结构特征中学习到的向量表示则忽略了图谱中的隐含语义信息。

为了解决这一问题，本文运用元学习的思想，通过对训练任务的划分和目标场景的模拟，获得对未见实体和关系的特征学习能力。同时，本文将本体也嵌入到低维向量空间中用于辅助图谱表示学习，对未见部分的语义信息补充。在提取了未见关系和未见实体的特征后，为了充分利用已知实体和关系的特征信息，模型采用 GNN 网络对所有实体和关系统一进行特征更新。本章介绍本文模型涉及到的本体嵌入方法、基于 GNN 的表示学习方法以及元学习相关的关键技术和代表模型。

### 2.1 跨域知识图谱的知识表示学习定义

一个知识图谱由众多的事实三元组组成，通常可以定义为  $\mathcal{G} = (\mathcal{E}, \mathcal{R}, \mathcal{T})$ ，其中  $\mathcal{E}$  指代所有图谱实体的集合， $\mathcal{R}$  指代图谱所有关系的集合， $\mathcal{T}$  指代所有的实体三元组集合，三元组大多遵循 RDF 等标准进行存储。事实三元组集合中组成三元组的头尾实体和关系均来自于实体集  $\mathcal{E}$  和关系集  $\mathcal{R}$  中，即  $\mathcal{T} = \{(h, r, t) \subseteq \mathcal{E} \times \mathcal{R} \times \mathcal{E}\}$ 。对于传统知识图谱上的链接预测任务，通过给定一个三元组的头结点和关系  $(h, r, ?)$  或者尾结点和关系  $(?, r, t)$  来预测缺失的实体节点  $e \in \mathcal{E}$ ，使得该缺失的实体节点能构成一个事实三元组  $(h, r, t)$ ，以完成对现有知识图谱的知识补全。为了评估表示学习模型在链接预测任务上的效果，通常会设置两组三元组数据，一组训练三元组  $\mathcal{T}_{support}$  用于对表示学习模型的参数进行学习，另一组测试三元组  $\mathcal{T}_{query}$  包含了训练集中不存在的事实三元组用于对模型的学习效果进行测试。例如对一个尾结点的预测任务，给定在测试三元组中的一个事实  $(h, r, t) \in \mathcal{T}_{query}$ ，通过模型计算所有可能预测三元组  $\{(h, r, e) | e \in \mathcal{E}, (h, r, e) \notin \mathcal{T}_{support} \cup \mathcal{T}_{query}\}$ ，如果在所有预测三元组中  $(h, r, t)$  的得分越高则说明该表示模型的效果越好。

在跨域知识图谱的训练场景下，本文对传统的知识图谱链接预测任务的定义进行了场景适配。基于跨域知识图谱的现实场景，在源知识图谱上训练的模型需要直接应用在处于低资源环境的目标知识图谱上。由于成本、用户数据隐私等限制，无法将目标知识

图谱与源知识图谱进行合并重新训练，因此划分出跨域知识图谱，源域知识图谱包含大量已知事实三元组用于训练，目标域知识图谱作为源域知识图谱子集可能包含未定义的实体和关系用于测试。现给定一个用于训练的源域知识图谱  $\mathcal{G}^{train} = (\mathcal{E}^{train})$ ，训练的目标是在源域知识图谱上进行模型参数的学习，从而能够将该模型应用在包含未见实体和未见关系的目标域知识图谱上，即  $\mathcal{G}^{test} = (\mathcal{E}^{test}, \mathcal{R}^{test}, \mathcal{T}_{support}^{test}, \mathcal{T}_{query}^{test})$ 。跨域知识图谱中的实体集和关系集遵循  $(\mathcal{E}^{train} \neq \mathcal{E}^{test}, \mathcal{E}^{train} \cap \mathcal{E}^{test} \neq \emptyset)$  及  $(\mathcal{R}^{train} \neq \mathcal{R}^{test}, \mathcal{R}^{train} \cap \mathcal{R}^{test} \neq \emptyset)$ 。其中  $\mathcal{T}_{support}^{test}$  只用于标记测试集中实体与关系的结构，不用于对模型的训练。

## 2.2 本体嵌入关键技术

本体图是一种特殊的知识图谱，不同于一般的实例知识图谱。本体规定了一系列基本概念之间的语义关系。通常，本体以层次概念为骨架，通过属性来描述概念的语义关系，以表示通用或特定领域的知识。随着知识表示学习的不断发展，传统基于图结构信息进行表示学习的 KGE 方法在面对跨域知识图谱等场景下，存在一定局限性。因此，越来越多的学者尝试引入本体来补充传统 KGE 方法的不足，以提高表示学习的效果。例如，从实例视角出发，知识图谱中包含丰富的事实三元组，例如（奥巴马，政治家，美国）。而从本体的角度来看，知识图谱从大量的事实三元组中提取并构建出抽象概念的语义元关系，例如（政治家，领导，国家）。这些本体信息不仅描述实体或关系的限制条件，如属性域和取值范围等，也是对知识图谱语义信息的抽象和集中表示。

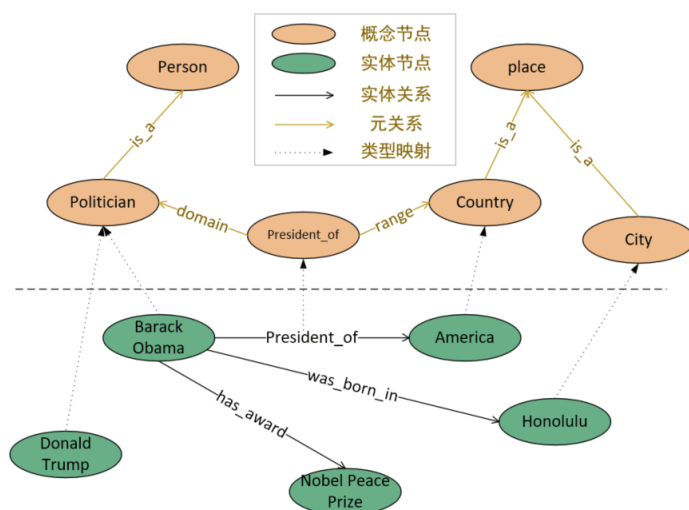


图 2.1 知识图谱的本体视角和实例视角

本体图包含概念和概念间的元关系，通常定义为，其中是概念的集合，是元关系的集合。类似于实体三元组，一个本体三元组  $(s, r, t)$  表示概念  $s, t$  通过元关系  $r$  进行联系。然而，与实例知识图谱复杂多样的关系不同，元关系可进一步分类为传递关系、对称关系、层次关系和其他简单关系<sup>[38]</sup>。借鉴于 KGE 方法的思想，On2Vec<sup>[38]</sup> 尝试将本体概念和元关系也映射为低维向量用于本体图的补全。但该模型认为由于本体关系大

多具有如传递、对称等特性，不能直接将 KGE 方法应用在一般的本体图上。例如对称关系  $r$  的两个三元组  $(c_1, r, c_1)$  和  $(c_1, r, c_1)$ ，当采用 TransE 等方法学习嵌入，无法同时兼顾三元组对应的向量满足  $c_{1,r} + r \approx c_{2,r}$  和  $c_{2,r} + r \approx c_{1,r}$ 。为了解决上述问题，On2Vec 通过设置两个特定于关系的投影，以区分同一概念在特定关系头尾的不同编码，如公式2.1所示：

$$S_d(T) = ||f_{1,r}(s) + r - f_{2,r}(t)||, \quad (2.1)$$

其中  $f_{1,r}$  和  $f_{2,r}$  分别代表了对特定关系  $r$  的三元组作为头本体和尾本体不同的投影操作。通过对头部本体和尾部本体分别进行不同的映射，可以很容易地解决上述传递元关系和对称元关系引起的矛盾问题。对于映头尾本体映射操作的选择，On2Vec 采用简单的线性变换处理，如公式2.2所示：

$$\begin{aligned} f_{1,r}(s) &= M_{1,r}s, \quad M_{1,r} \in \mathbb{R}^{k \times k}, \\ f_{2,r}(s) &= M_{2,r}s, \quad M_{2,r} \in \mathbb{R}^{k \times k}, \end{aligned} \quad (2.2)$$

特别对于层次关系，On2Vec 将层次关系进一步划分为  $R_r$  和  $R_c$ ， $R_r$  关系表示粗略概念被划分为更细致概念的细化关系，而  $R_c$  表示将更细致概念分组为更粗略概念的简略关系。该模型为了使得更细致概念的嵌入紧密地汇聚在一个紧密的邻域内，采用层次模型对层次关系进行单独的处理，对层次关系嵌入的评分函数设置如公式2.3所示：

$$\begin{aligned} S_{hm}(G) &= \sum_{r \in R_r} \sum_{s \in C} \sum_{t \in \sigma(s,r)} \omega(f_{1,r}(s) + r, f_{2,r}(t)) \\ &+ \sum_{r \in R_c} \sum_{t \in C} \sum_{s \in \sigma(t,r)} \omega(f_{2,r}(t) - r, f_{1,r}(s)), \end{aligned} \quad (2.3)$$

其中， $\omega$  是一个相对于两个参数向量的角度或距离单调递增的函数。On2Vec 采用余弦距离进行计算。其中  $\sigma$  是针对于相应层次关系的本体提取操作，对细化关系则寻找所有该关系下的所有尾本体，对简略关系则寻找该关系下的所有头本体。On2Vec 扩展了 TransE 方法，以捕获本体关系的关系属性和层次结构，实现对本体概念和元关系的表示。该方法证明了在本体图上应用知识图谱嵌入方法的有效性。

一些利用本体进行知识表示学习的方法，将本体信息视为实体和关系的附加类型信息。实体使用本体定义的类型进行表示，关系也使用语义类型进行表示。例如，将“马斯克”表示为本体定义的“人类”概念，“特斯拉”则表示为所属“公司”概念。SSE<sup>[39]</sup> 模型结合实体的语义类别，将属于同一类别的实体平滑地嵌入到语义空间中。为了更有效地捕捉这些高层本体类型信息带来的限制效果，该模型提出了两种在语义层面对嵌入变量的限制。首先该模型借用拉普拉斯特征映射的算法思想认为，同一个语义类别下的实体的嵌入应该更加相近。例如，对于同属于国家语义的“法国”实体和“意大利”的实体，他们对应的嵌入表示应该更为相似。为此模型为每一个实体设置了一个语义分类矩阵  $W_{ij}^{(1)}$ ，如果任意两个实体在类型上是相同的，那么该矩阵中对应的值即为 1，否则

为 0。基于该矩阵计算两个实体之间的语义平滑值，如公式2.4所示：

$$\mathcal{R}_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \|e_i - e_j\|_2^2 W_{ij}^{(1)}, \quad (2.4)$$

其中  $e_i$  和  $e_j$  指代两个实体  $i$  和实体  $j$  对应的嵌入表示，通过让两个同类实体的语义平滑值最小，来实现拉普拉斯特征映射的语义限制。其次，该模型还提出一种基于局部线性嵌入的语义约束思想。这种思想不同于拉普拉斯特征映射对数据对局部不变性的设定。基于局部线性嵌入的思想认为，一个实体可以通过其最接近的邻居节点经过一个线性的组合器来近似地表示。例如，“法国”的实体嵌入可以通过其语义相邻的节点，如“中国”、“意大利”等国家节点，通过线性组合的方式来近似地表示。同样，在该语义约束下 SSE 设置了一个类别邻接矩阵  $W_{ij}^{(2)}$ ，该矩阵指明实体的连接节点是否是该节点语义范畴内的最邻近节点。然后在该约束下计算实体的语义平滑值，如公式2.5所示：

$$\mathcal{R}_2 = \sum_{i=1}^n \left\| e_i - \sum_{e_j \in \mathcal{N}(e_i)} W_{ij}^{(2)} e_j \right\|_2^2, \quad (2.5)$$

其中  $e_i$  和  $e_j$  指代两个实体  $i$  和实体  $j$  对应的嵌入表示，通过使得实体  $i$  和所有最邻近节点的线性组合距离最短来实现该猜想的约束。最后，该模型通过在基于距离和基于语义相似度的嵌入模型上加入两种语义约束来提高嵌入表示的效果。实验证明，在语义类别的约束下，表示学习获得了很好的嵌入效果。虽然通过这两种类别信息的约束能够有效引入图谱除结构外的额外知识，但该模型认为单个实体只属于一个类别，忽略了多个类别之间可能的联系，没有完全充分利用好本体的类别语义知识。

TKRL<sup>[40]</sup> 模型基于传统的翻译模型，并结合本体的实体类型信息，提出了一个基于实体类型的嵌入方法。与 SSE 模型仅设定实体单一类别不同，TKPL 模型认为一个实体可能有多个不同语义层次的类别。并且这些不同层次的类别信息应该被转化到与类别绑定的不同的特征空间中。因此，三元组中的每个实体都应该包含多个实体类型相关的特征信息联合表示。为了获取实体的多类别语义信息，TKPL 模型采用所有实体类型转换矩阵的加权和来计算最后的转换矩阵，如公式2.6所示：

$$M = \frac{\sum_{i=1}^n \alpha_i M_{c_i}}{\sum_{i=1}^n \alpha_i}, \alpha_i = \begin{cases} 1, & c_i \in C_r \\ 0, & c_i \notin C_r \end{cases} \quad (2.6)$$

其中  $n$  是一个实体所有可能的实体类型数量， $c_i$  指代第  $i$  个实体类型， $M_{c_i}$  是该实体类型对应的转化矩阵， $\alpha_i$  是矩阵对应的权重值，最后的  $C_r$  指代三元组给定三元组  $(h, r, t)$  头实体所有与关系  $r$  相连接的可能的实体类型的集合。TKRL 引入实体类型信息，并对所有可能的类型信息进行聚合，来增强嵌入模型的学习效果。

为了提高知识表示学习的性能，现有的大多数本体模型在知识嵌入过程中都只包含单一的本体信息，无法实现对所有可用本体信息的无缝嵌入。在现实世界中，知识图谱

通常是不完整的, 单一的本体信息对补充图谱的作用有限。本文能够在知识嵌入过程中无缝地整合所有可用的本体信息, 充分补充图谱, 提高复杂场景下的决策能力。同时本文使用了一种简单的本体形式, 即 RDF Schema(RDFS) 中, 而那些更复杂的 OWL 本体可以按照一定的标准转换为 RDFS 本体。本体可以被用作知识图谱的模式, 定义实体类型、关系等。本文将本体表示为  $\mathcal{O} = \{\mathcal{C}, \mathcal{P}, \mathcal{T}_o\}$ , 其中  $\mathcal{C}$  是概念节点的集合 (即实体类型和实体关系),  $\mathcal{P}$  是属性的集合,  $\mathcal{T}_o$  是本体三元组的集合。

## 2.3 基于 GNN 的跨域知识表示学习关键技术

图神经网络模型是专门用于处理图数据的模型, 可以直接输入一个图网络, 并能够很容易地进行节点级别、边级别及图级别的预测任务。CNN 模型只能作用在具有相同结构的图像或者特定序列的语音和文字上, 而图数据没有固定的形式且邻居节点的也都是无序的, 因此 CNN 模型无法作用在复杂的知识图谱上。相比之下, GNN 通过聚合和更新操作, 能够学习到图谱结构和节点特征的有效信息。

为了有效地学习图的结构信息和节点特征, 并对特征进行嵌入, GNN 主要包含了两个部分: 聚合函数和更新函数。聚合函数可以将相邻节点的特征进行聚合, 并可以使用诸如 sum、mean 和 max 等常见操作。一次聚合操作可以提取邻接节点的信息, 这些节点距离当前节点为一跳。GNN 通常包含多层, 每一层都会用上一层的信息进行聚合和传递。因此,  $n$  层聚合后传递的信息包含了  $n$  层邻接节点的结构信息和节点本身的特征信息。每层的特征聚合函数, 如公式2.7所示:

$$h_v^k = \sigma(W_k \sum \frac{h_u^{k-1}}{|N(v)|} + B_k h_v^{k-1}) \quad \text{where } k = 1, \dots, k-1 \quad (2.7)$$

从公式可以看出, 每层包含两个部分信息。首先  $W_k \sum \frac{h_u^{k-1}}{|N(v)|}$  表示邻接点特征的聚合操作, 而后一部分是上一层聚合特征和权重参数的乘积。最后, 这两部分特征通过激活函数更新, 完成节点特征输出。在 GNN 模型中, 聚合操作没有区分邻接节点的重要性, 而是进行了简单的池化操作。

基于 GNN 的另外一种改进方向集中于对图结构关系表示的补充上。图神经网络关注节点特征的聚合和更新操作, 但在信息传递的过程中, 图的关系结构仅用于指明邻接点, 关系特征未参与节点的更新过程。为了增加关系信息对节点的影响, R-GCN 在每层节点特征计算中引入了邻接点间的对应关系, 更新函数如公式2.8所示:

$$h_i^{l+1} = \sigma \left( \sum_{r \in \mathcal{R}} \sum_{j \in \mathcal{N}_i^r} \frac{1}{c_{i,r}} W_r^l h_j^l + W_0^l h_i^l \right), \quad (2.8)$$

与 GNN 对邻接点特征聚合的操作不同, R-GCN 引入了关系特定的转换。这种转换取决于边的类型和方向。为了确保第 1 层的节点表示可以受到相应层次的表示的影响, R-GCN 在基础的图关系上为每个节点添加了一个自连接的特殊关系。如图2.2所示, 在

R-GCN 中每层对一个实体节点（红色块表示）进行特征生成的过程中，首先从邻接点获取特征（蓝色块表示）并根据该节点与邻接节点的关系类型进行特征转换得到该种关系对应的表示（绿色块表示），其中关系类型分别由入关系、出关系以及自循环关系组成。然后将所有关系转换后的邻接节点信息累加求和，并通过一个如 ReLU 的激活函数即可获得该节点本层的输出表示。相比于 R-GCN，近期提出的 CompGCN 在 R-GCN 模型的基础上进一步引入了注意力机制，针对每种边类型和方向分别进行了注意力计算以加强对重要信息的关注。而且在计算效率方面，它减少了每个节点的嵌入大小并减少了依赖于固定卷积核的计算量，因此更适合于大规模图数据。

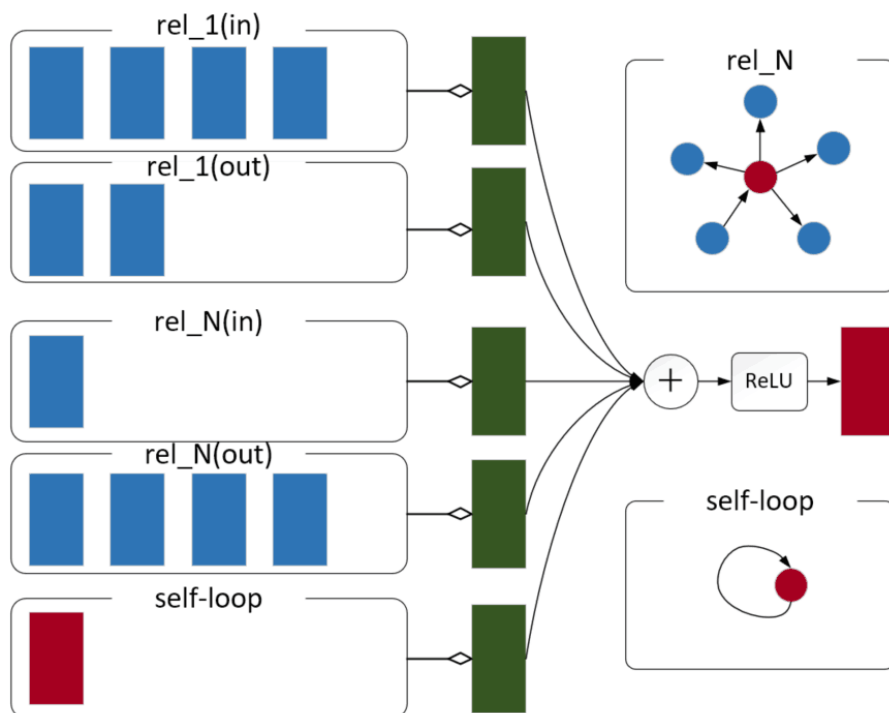


图 2.2 R-GCN 的特征传递

随着图卷积网络被广泛应用于图数据处理，它本身的能力也已经可以用于跨域知识表示学习。例如，INDIGO<sup>[41]</sup> 模型使用实体三元组与 GNN 的内层和外层的特征向量元素之间的一一对应关系对知识图谱进行编码，并避免了额外的打分函数，充分利用了 GNN 的特征聚合能力。其他方法如 Zhao<sup>[42]</sup> 等人通过基于注意力的图网络聚合未见关系的邻接结构特征，来作为未见关系的表示。但这些方法要么专注于传统的知识图谱表示学习领域，要么通过结构信息对未见的实体或关系进行嵌入，忽略了其他的语义信息。本文通过全局本体信息的嵌入表示，在加强知识引入的同时，借助关系的位置结构信息来训练一个兼顾两种未知部分的基于 GNN 的表示学习模型。

## 2.4 元学习训练方法

元学习最普适性的算法思想可以理解为“learning to learn”，它通过多个学习任务的训练来改进学习算法，而传统的机器学习算法则是在多个数据实例上进行模型的学习。

传统的机器学习方法会设置一个训练集  $\mathcal{D} = \{(x_1, y_1), \dots, (x_N, y_N)\}$ ，例如样本对 (输入的图片，图片的标签)。而元学习的目的是训练出一个模型函数  $y = f_\theta(x)$ ，通过训练来获得其中的参数  $\theta$ ，求解公式如公式2.9所示：

$$\theta^* = \arg \min_{\theta} \mathcal{L}(\mathcal{D}; \theta, \omega) \quad (2.9)$$

其中的  $\mathcal{L}$  是一个损失函数来计算真实标签与模型预测标签之间的误差， $\omega$  指代了模型如何学习的假设，例如如何为参数  $\theta$  选择合适的优化器或者为  $f$  选择函数类型等。传统的机器学习方法实现过程中，该部分由研究者手动设置。模型的泛化性能则通过评估模型在已知标签上的测试集测试来衡量。传统的机器学习假设是模型的优化是对每个训练集  $\mathcal{D}$  从头开始执行，模型如何学习的设定是预先指定的，这些设定将极大地影响模型的准确性和数据效率等性能指标。元学习试图通过学习学习算法本身来改进这些指标，而不是假设学习算法是预先指定或者固定的。此外元学习从任务的分布中学习，而不是从头开始。

元学习“learning to learn”的思想精髓可以看做一个包含内外两层的双层优化问题，双层优化<sup>[43]</sup>指代的是层次优化问题，其中一个优化包含另一个优化作为约束<sup>[44][45]</sup>。经典的内外双层模型的算法如 MAML，其算法流程如算法1所示：

```

Data:  $p(\mathcal{T})$ :distribution over tasks
Data:  $\alpha, \beta$  step size hyperparameters
randomly initialize  $\theta$ 
while not done do
    Sample batch of tasks  $\mathcal{T}_i \sim p(\mathcal{T})$  for  $\mathcal{T}_i$  do
        Evaluate  $\nabla_{\theta} \mathcal{L}_{\mathcal{T}_i}(f_{\theta})$  with respect to K examples
        Compute adapted parameters with gradient descent:  $\theta'_i = \theta - \alpha \nabla_{\theta} \mathcal{L}_{\mathcal{T}_i}(f_{\theta})$ 
    end
    Update  $\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}_i}(f_{\theta'})$ 
end
    
```

算法 1 Model-Agnostic Meta-Learning

在该视角下，元学习任务可以通过公式2.10所示来规范化：

$$\omega^* = \arg \min_{\omega} \sum_{i=1}^M \mathcal{L}^{\text{meta}}(\mathcal{D}_{\text{source}}^{\text{val}(i)}; \theta^{*(i)}, \omega), \quad (2.10)$$

$$s.t. \quad \theta^{*(i)}(\omega) = \arg \min_{\theta} \mathcal{L}^{\text{task}}(\mathcal{D}_{\text{source}}^{\text{train}(i)}; \theta, \omega), \quad (2.11)$$

其中  $\mathcal{L}^{\text{meta}}$  和  $\mathcal{L}^{\text{task}}$  分别指代外层的优化目标和内层的优化目标，例如在分类任务下的交叉熵。但是这两层的优化级别并不对称，内层优化在基于外层参数  $\omega$  的优化过程中不能对  $\omega$  进行修改。公式中  $\omega$  可以指代如非凸优化<sup>[46]</sup>的内层模型的初始化参数或其他可学习的超参数。因此，元学习的整个训练流程分为两层优化：内层模型首先接收外层模型的参数  $\omega$ ，然后根据自己的任务在该任务的训练集上进行训练，最终在该任务的测试集上计算出损失函数。接下来，外层模型使用内层模型计算出的损失函数对参数  $\omega$  进行更新，使得内层函数的损失达到最优。元学习的思想即通过外层模型的训练，学习到内层模型一个更好的设定，以便让内层模型更好地完成各种任务。

如前所述，内层模型需要特定训练集和测试集以针对面向的问题进行训练。以任务为训练单位的设定是元学习方法不同于传统机器学习方法的一个特点。从训练任务的角度来看，元学习的目标是学习出一种通用的学习算法，这些算法能够在新任务上获得更好的表现。内层模型可以视为使用外层模型参数的传统机器学习算法，其数据集为  $\mathcal{D} = (\mathcal{D}^{\text{train}}, \mathcal{D}^{\text{val}})$ ，针对单个任务的损失函数即为  $\mathcal{L}(\mathcal{D}; \omega) = \mathcal{L}(\mathcal{D}^{\text{val}}; \omega^*(\mathcal{D}^{\text{train}}, \omega), \omega)$ 。在实际应用中，通常只有一个训练集和测试集。因此，一般会从源训练集中抽样出一组任务用于进行训练。这些任务的训练集和测试集被称为 support 集和 query 集，以避免与最终模型训练后进行评估的测试集混淆。

本文旨在跨域知识图谱上进行知识表示的相关研究，并尝试解决在含有未见部分的目标域知识图谱上的链接预测任务。经典的知识图谱表示学习的链接预测任务采用的数据集会设置一个训练集和测试集，测试集中不包含新的实体和关系。因此，本文采用现有数据集中符合问题条件的测试集进行采样，并借鉴元学习“learning to learn”的思想，将这些数据抽取为任务进行训练和测试。从这些训练任务上学习到能够处理未知组件的知识表示方法。

## 2.5 本章小结

本章首先介绍了跨域知识图谱及问题的定义，强调了跨域知识表示学习中新的实体和关系对传统知识表示学习的影响。为了能够将本体信息用于到知识表示中，介绍了使用本体嵌入的基本方法和代表性的一些应用模型。本体的向量表示，能够为实例图谱的实体和关系提供较为完整的语义信息补充。同时，本文希望通过结合 GNN 模型对邻接实体和关系的特征进行学习，在第三部分介绍了 GNN 在知识图谱的知识表示学习上的应用及基于 GNN 的改进的一些模型方法。最后介绍了元学习相关的思想、原理及方法，为后续模型的任务划分及训练流程提供理论支撑和参考。



### 3 基于关系拓扑结构及描述文本的本体信息嵌入

本体是领域内公认的概念的集合。它通过对实例的抽象描述语义层次上的知识。本体知识不依赖于具体的关系和实体，因此能够对目标域知识图谱中的未见关系和未见实体提供语义信息。为了能够从本体中学习到关系和实体抽象的语义信息，并且将本体信息进行高效的表示学习用于后续的知识图谱嵌入。本章将介绍如何补充本体三元组，尤其是对本体中关系相关的三元组进行补充。为了对本体中关系相关的三元组进行补充，本文通过对源知识图谱三元组中的关系头尾节点信息和关系间的拓扑信息提取，得到了关系的定义域、值域三元组以及关系位置元关系三元组，以此构建了关系加强的本体图。然后在使用传统 KGE 方法得到的本体初步嵌入的基础上，本文又通过本体概念的描述文本信息对初步嵌入进行增强，学习到了融合本体三元组结构信息和文本信息的本体嵌入。

#### 3.1 基于关系拓扑结构及描述文本的本体嵌入框架

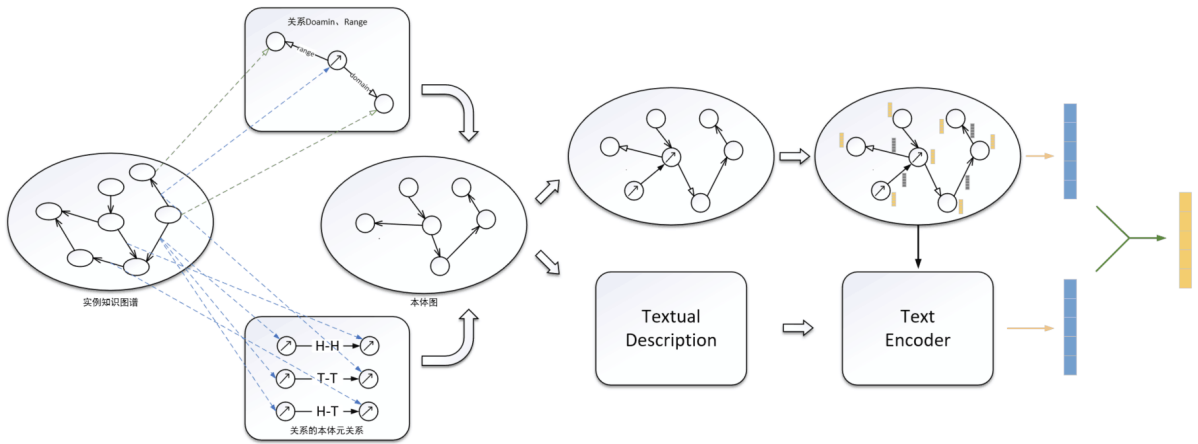


图 3.1 本体嵌入整体框架

基于描述文本加强的本体嵌入框架主要包含两个部分，分别是关系本体构建和描述文本增强的本体表示。首先，作为本体嵌入的基础，本文了解到大多数引入本体信息的模型仅通过实体的类型和层次信息作为本体信息的体现，忽略了关系在本体中的体现。因此，本文首先对实例知识图谱中的关系进行处理，将关系的首尾实体抽象为实体类型，构建了关系的 domain、range 相关的本体三元组。考虑到知识图谱中关系的语义相关性，本文提出四种关系的位置元关系进一步对本体中的关系信息进行加强。而后在本体三元组的基础上学习对本体信息的嵌入。在通过传统 KGE 方法对本体三元组结构信

息嵌入的基础上,本文又引入了本体节点的描述文本对本体嵌入的语义进行加强,获得了最终的本体向量表示。

## 3.2 关系本体图构建

结合本体的知识表示模型往往仅使用了类型、层次等信息,忽略了关系在本体中的重要体现。本节将介绍本文如何在现有的本体三元组基础上,对关系相关的本体三元组进行抽取。同时,根据关系的相对位置定义了四种关系的位置元关系,对本体中的关系信息进行补充和加强。

### 3.2.1 关系本体三元组

在 RDF 的设置下,知识图谱中的知识总是以三元组的形式出现,通过 RDF 的主语、谓语和宾语来描述事实。RDF 通过类和属性描述个体之间的关系。这些类和属性由模式定义。RDF 模式 (RDF Schema, RDFS) 提供了对类和属性的简单描述,一般用于对本体的简单定义。RDFS 提供了最基本的对类和属性的描述元语:

- `rdfs:type`: 用于指定个体的类;
- `rdfs:subClassOf`: 用于指定类的父类;
- `rdfs:subPropertyOf`: 用于指定属性的父属性;
- `rdfs:domain`: 用于指定属性的定义域;
- `rdfs:range`: 用于指定属性的值域;

通过定义元语的方式来描述元数据之间的关系。同时,通过元语可以将知识分为两类:一类是数据层面的知识,如 (Obama, type, Person) 说明 Obama 是 Person 的一个实例;另一类是模式层面的知识,如 (speaker, domain, Person) 说明 speaker 属性的定义域是 Person 类。从简单意义上讲,数据层面的知识更多作用于实体,而模式层面的知识更多作用于关系。但当下将本体信息引入知识图谱嵌入的方法大多数仅采用数据层面的知识,忽略了模式层面知识对实例关系的知识补充。例如 TransT<sup>[47]</sup> 模型根据头尾实体的类型计算相似度作为知识图谱的先验知识改进三元组评分、JOIE<sup>[48]</sup> 模型通过实体的本体类型信息将本体图和实例图进行跨视图的表示学习。这些模型主要通过本体的数据层面的知识,如实体类型对表示学习进行加强,更多偏向于实体的知识补充,忽略了对关系的补充。而跨域知识图谱中存在未见关系,本体中对关系模式的补充是非常必要的。

同样借助于实体的类型,本文通过对实例图谱中事实三元组的处理提取关系的本体三元组。如图3.2所示,对一个事实三元组 (Obama, President\_of, America), 其中头尾实体的类型 (Obama, type, Person) 和 (America, type, Country), 可知对于关系 President\_of 在本体中的定义域即头本体应该是类型 Person, 值域即尾本体, 因此构建出相应的关系本体三元组 (President\_of, domain, Person) 和 (President\_of, range, country)。

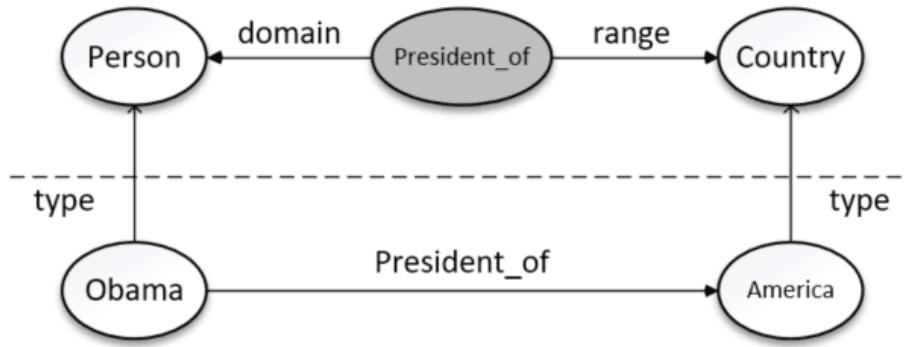


图 3.2 关系本体三元组

但从源知识图谱直接抽取会产生大量的 **domain** 和 **range** 相关的关系本体三元组，其中包含了所有关系可能的值域和定义域。为了能够将抽取的关系的值域和定义域的本体三元组能够尽可能表示最普适的元信息，对于抽取所有的关系本体三元组，本文通过统计关系本体三元组在实例知识图谱中出现的频率，设置阈值对三元组进行筛选，以出现频率代表关系本体三元组的普适程度。除此之外，除了关系的 **domain** 和 **range** 的模式三元组，本文任务关系与关系之前存在相似性的关联，因此通过对关系描述文本的相似度匹配，计算关系与关系间的相似程度，本文还引入了关系与关系的联系和 **type** 本身的本体三元组中的 **isa**、**synonym** 关系共同设置为 **generalizations** 关系加入到本体三元组数据中。其中对关系描述文本的相似度匹配，本文采用 **word2vec** 进行文本嵌入和相似度计算。

### 3.2.2 关系位置元关系

上述关系本体三元组的抽取从值域和定义域的层面对本体中的关系信息进行了补充。除此之外，本文希望通过知识图谱中关系的语义相关性进一步对关系相关的语义信息进行捕捉。知识图谱中的实体具有隐含的语义相关性，例如北京和武汉在作为类型城市的实例具有相似的语义。而关系的语义相关性也非常常见，例如关系 **“/people/person/nationality”** 和 **“/people/ethnicity/languages\_spoken”** 具有很强的语义相关性，因为一个人说的语言很大程度上与他的国籍有关，在知识图谱上则可能直接表示为上述两个关系与同一个实体相关联。相反的，上述关系与语义差别很大的其他关系，如 **“/film/film/country”**，则没有很强的语义相关性。并且这种语义相关性同关系的方向密切相关有关，如图3.3的 **“sister\_of”** 和通过 **e1** 节点连接的 **“has\_gender”**，及通过 **e2** 节点连接的 **“sister\_gender”** 因为拓扑关系的不同，具备了完全不同的隐含信息。

为了对两个关系之间的相关性进行建模，本文将关系与关系之间的拓扑关系建模为四种关系的位置元关系，分别为 **tail-head**、**head-tail**、**tail-tail**、**head-head**。位置元关系的头结点和尾结点都代表了两个相邻关系的指向，比如 **(relation1, tail-head, relation2)** 代

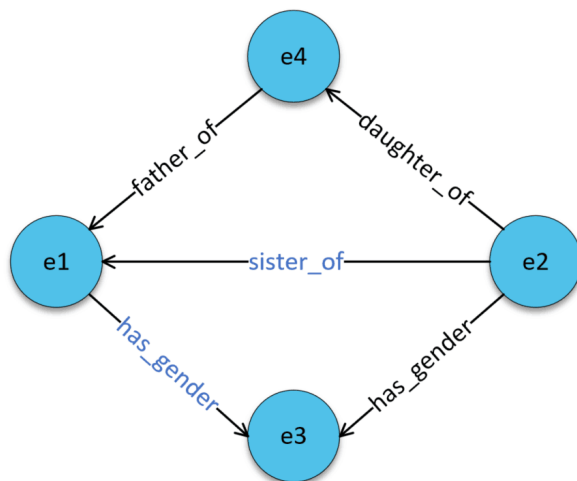


图 3.3 关系位置实例

表同一个实体连接的两个相邻关系 1 和关系 2，且关系 1 指向该实体而关系 2 则从该实体指向其他实体。对于在训练三元组中的两个关系，如果它们符合其中一种的相对未见关系，则提取一个位置元关系补充到本体三元组中。

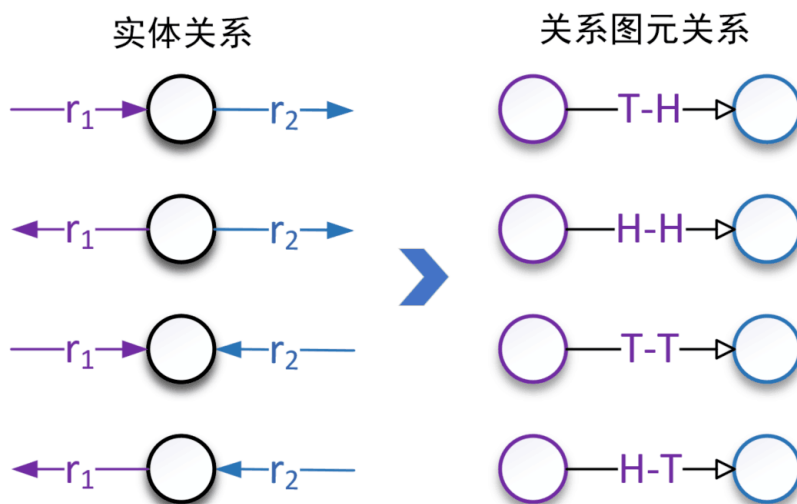


图 3.4 实体关系与关系图元关系的映射

将位置元关系补充到本体三元组中后，本文希望通过传统的知识图谱嵌入方法对本体概念学习向量表示。但对位置元关系分析可知，“head-head”元关系自身、“tail-tail”元关系自身以及“head-tail”和“tail-head”元关系对都具有对称性。如图3.3中通过 e1 连接的“sister\_of”和“has\_gender”关系存在有 (sister\_of, head-tail, has\_gender) 与 (has\_gender, tail-head, sister\_of)，通过本文第 2 章对 On2Vec 模型的分析可知无法采用传统的 KGE 方法对有对称的元关系直接进行嵌入。因此，本文在补充位置元关系时仅保留了“head-tail”、“head-head”和“tail-tail”元关系，并且对“head-head”、“tail-tail”

进行去重，避免了具有对称性位置元关系的存在。

通过对上述两种关系本体信息的补充，结合实体类型相关的本体三元组，本文在本体三元组中保留了 domain、range、generalizations 以及三个位置元关系总计 6 种元关系。本体三元组统计数据如下表3.1所示，最终构建出的本体图局部如图3.5所示。

表 3.1 本体三元组统计信息

	关系三元组数	元关系三元组数	其他三元组	总计
NELL_Ext	1816	2135	332	4326
DB_Ext	1727	1104	464	3295

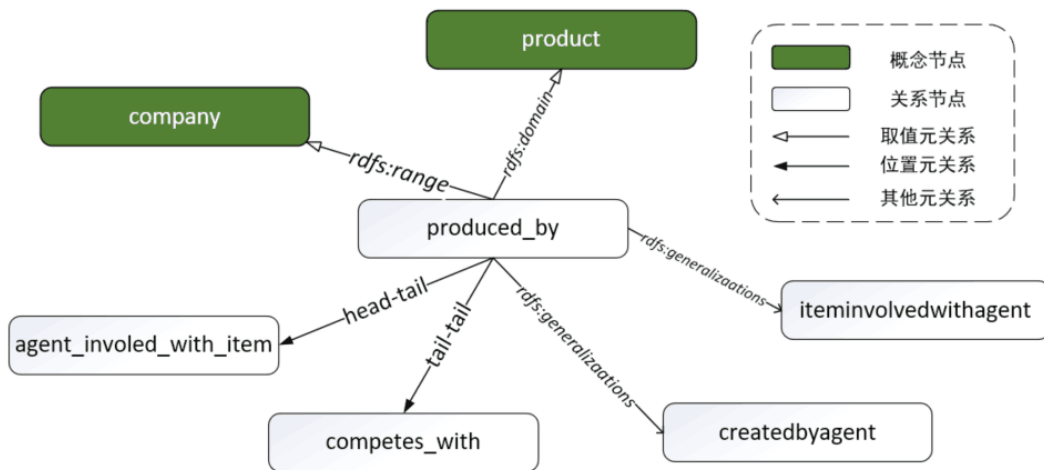


图 3.5 关系本体图

### 3.3 基于描述文本的本体表示增强

为了将符号化的本体三元组用于后续对未见实体和未见关系的语义补充，需要将本体三元组转化为低维向量表示。本节先从基础的本体三元组数据中学习到结构化的嵌入表示，然后从三元组的概念节点的描述信息中使用词嵌入学习到概念节点的描述文本嵌入。为了将描述文本的嵌入补充到本体信息的结构化中去，本文使用一个共享参数的线性层将结构化嵌入表示和描述文本嵌入表示映射到同一个表示空间中。在映射后的线性层中，参照 TransE 的评分思想，本文采用三个距离打分函数将映射后的两种嵌入表示联合更新，学习到兼顾结构信息和描述文本信息的嵌入表示。最后将两种表示拼接作为本体信息进行后续操作。基于描述文本的本体表示增强结构图3.6所示：

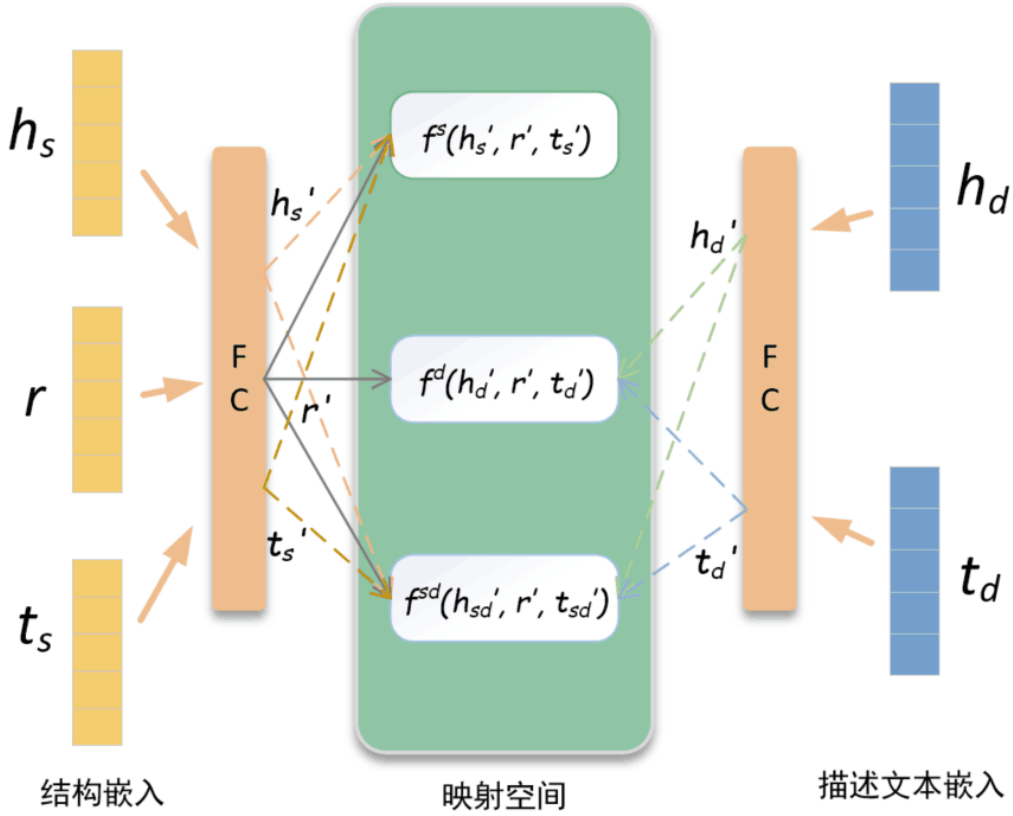


图 3.6 本体嵌入架构

### 3.3.1 本体结构信息嵌入

对本体定义的结构化三元组进行嵌入的方法有很多种，因为在对本体三元组进行构建的时候本文已经尽可能减小了对称关系对嵌入效果的影响，因此可以通过多种传统的知识图谱嵌入方法学习本体表示。与传统的知识图谱嵌入表示类似，对于一个本体三元组  $(c_i, p, c_j)$ ，本体语义信息编码的目的就是设计出一个打分函数  $f(c_i, p, c_j)$  作为编码模型的激活函数，本文采用 RotatE 对本体三元组进行编码。例如按照 RotatE 模型的设定，本体三元组的元关系属性是头尾两个实体节点的转移量，打分函数如公式3.1所示：

$$f_{RotatE}(c_i, p, c_j) = -\|\mathbf{c}_i \circ \mathbf{p} + \mathbf{c}_j\|, \quad (3.1)$$

其中  $\mathbf{c}_i$ ,  $\mathbf{p}$ ,  $\mathbf{c}_j$  是一个本体三元组相应的概念编码和元关系编码， $\circ$  指代了向量的旋转操作。同时为了提高所有本体三元组的嵌入效果，本文采用自对抗负抽样损失函数来计算损失更新模型，如公式3.2所示：

$$\mathcal{L}_{\mathcal{O}} = \frac{1}{|\mathcal{T}_{\mathcal{O}}|} \sum_{(c_i, p, c_j) \in \mathcal{T}_{\mathcal{O}}} [\gamma_o + f(c_i', p, c_j') + f(c_i, p, c_j)], \quad (3.2)$$

其中  $\gamma_o$  是控制正负样本得分的参数，同时  $c_i$ ,  $c_j$  是不存在于本体三元组中的负样本，为了生成这些负样本，本文在所有的本体概念中分别遮盖住已存在的本体三元组的头节点

和尾结点，然后从所有本体节点中随机筛选出其他节点组成负样本。

### 3.3.2 基于文本的本体嵌入

除了结构化的本体三元组外，本体信息还有许多对本体进行详细描述文本，如本体概念“companyceo”的描述文本“specifies that a particular CEO is the CEO of a particular company”。这些描述文本可以为本体信息提取提供额外的语义信息，因此本文通过使用文本描述来加强对本体三元组的语义嵌入。然而，描述文本的建模与一般的三元组建模因模型的差异而无法直接融合。因此对于给定的本体三元组  $(c_i, p, c_j)$ ，本文首先获得了三元组的结构嵌入  $h_s/r/t_s \in \mathbb{R}^{d_1}$  和每个本体节点描述文本的向量表示  $h_d/t_d \in \mathbb{R}^{d_2}$ ，表示文本描述信息。为了融合这两个不同层面的嵌入，本文引入了一个全连接层，将两个不同的嵌入同时映射到同一表示维度上。映射后的结构嵌入和文本嵌入分别表示为  $h'_s$  和  $h'_d$ ，在统一表示空间中，使用 TransE 对三元组结构的嵌入进行打分，打分函数如公式3.3所示：

$$f^s = -\|h'_s + r' - t'_s\|, \quad (3.3)$$

对描述文本的嵌入进行打分，打分函数如公式3.4所示：

$$f^d = -\|h'_d + r' - t'_d\|, \quad (3.4)$$

同时为了为了使这两种类型的表示相互兼容和互补，本文遵循 DKRL 模型的设定来定义交叉和相加得分函数，打分函数如公式3.5所示：

$$f^{sd} = -\|h'_s + r' - t'_d\| - \|h'_d + r' - t'_s\|, \quad (3.5)$$

所有的四个得分函数综合可以保证两个层面的嵌入表示可以在相同空间里进行学习和更新，最后本体嵌入的得分函数如公式3.6所示：

$$f'(c_i, p, c_j) = f^s + f^d + f^{sd}, \quad (3.6)$$

因此在本体嵌入的损失函数也响应的转化为：

$$\mathcal{L}_{\mathcal{O}}^{ont} = \frac{1}{|\mathcal{T}_{\mathcal{O}}|} \sum_{(c_i, p, c_j) \in \mathcal{T}_{\mathcal{O}}} [\gamma_o + f'(c'_i, p, c'_j) + f'(c_i, p, c_j)], \quad (3.7)$$

经过训练后每个本体节点都有两个层面的嵌入：三元组结构嵌入和描述文本嵌入。本文将这两种映射后的嵌入进行拼接作为本体节点的最终的向量表示。上述的描述文本的向量表示，本文采用了词袋模型进行生成。

## 3.4 本章小结

本章通过对本体三元组的关系部分构建，强调了关系在本体信息中的体现。尤其是在处理跨域知识图谱中存在的新的关系，本体中的信息能够对关系表示进行有效的语义

补充。通过本体三元组的结构信息和本体的描述文本信息，能够学习到好的本体嵌入，为下一章未见关系的建模提供了良好的先验知识。



## 4 基于元学习本体增强的跨域知识表示学习模型

跨域知识图谱的知识表示学习的关键在于如何对目标域知识图谱中的未见关系和未见实体进行嵌入。借鉴于元学习“learning to learn”的算法思想，本文通过设置多个训练任务，且在任务中模拟存在新的实体和新的关系的跨域场景。基于训练任务进行模型训练，使得模型能够学习到对未见实体和未见关系的嵌入能力。在每一个训练任务上，模型通过对关系的拓扑信息和本体信息进行学习获得对未见关系的嵌入表示，通过对未见实体邻接的关系信息聚合获得未见实体的嵌入表示。然后为了充分利用到已知关系和已知实体的特征信息，本文采用 CompGCN 在实例知识图谱上对所有实体和关系的初始化嵌入再进行一次更新，获得所有实体和关系最终的向量表示。最后通过多个 KGE 方法计算损失并进行模型优化。该模型的主要创新点如下：1) 关系和实体的表示能够同时学习到事实三元组结构上的拓扑信息和本体层面上的语义信息。2) 能够通过元学习的训练流程模拟出目标问题的学习任务对模型参数进行更新。3) 能够同时对测试集中未见的关系和未见的实体嵌入编码。

### 4.1 模型整体架构

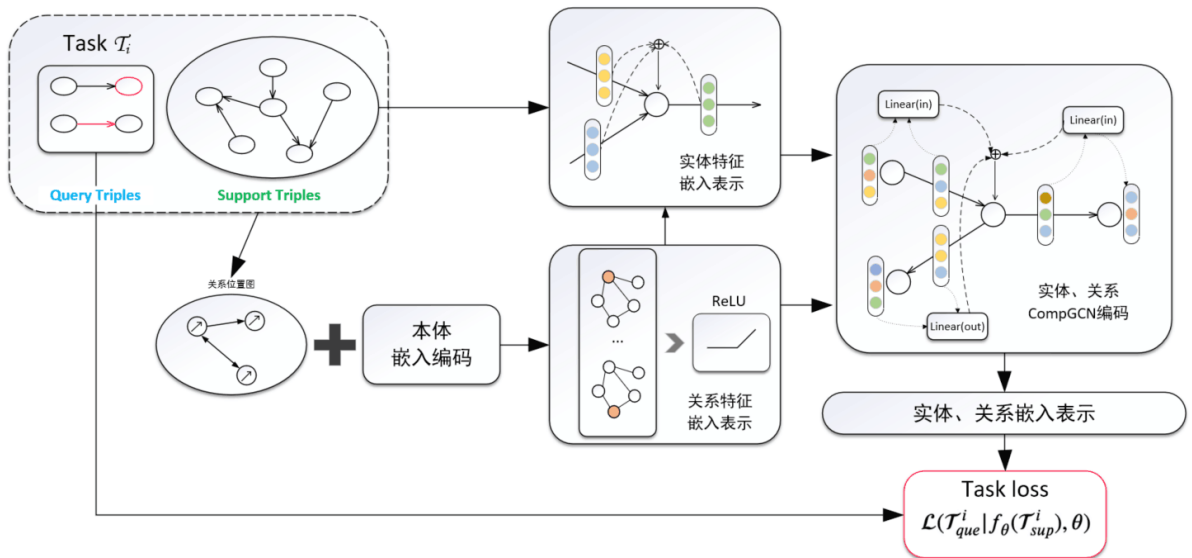


图 4.1 模型架构图

本文提出的模型整体结构如图4.1所示，包含三个主要的部分：未见关系的特征嵌入、未见实体的特征嵌入以及基于元学习训练设定。为了提取关系特征，本文首先根据实例图谱中关系的相对位置结构构建了关系位置图，其中节点为实例图谱中的关系。将本体信息和关系位置图结合后，通过两层关系图卷积网络对关系的语义信息和结构信息

进行更新和学习，获得关系特征的嵌入表示。为了对未见实体进行表示学习，本文对未见实体的邻接关系进行聚合，并使用关系方向特定的调整矩阵，提取所有关系的特征，以获得实体特征的初始化嵌入表示。为了利用实例图谱中的已知关系和已知实体，本文在实例图谱上采用两层复杂图卷积网络聚合所有实体和关系及其邻域特征，并对它们的嵌入表示进行更新。模型修改了复杂图卷积网络的输出层，使得关系和实体的维度不求一致，从而可以采用多种 KGE 模型作为打分函数计算任务损失，来对模型进行调优。而为了能够让模型在训练过程中获得对目标域知识图谱未见实体和关系表示学习的能力，本文设置了多个训练任务对跨域场景进行模拟，在训练任务上获得最优参数，并将训练的参数用于目标域的图谱嵌入上。

对于训练集中可见的关系和实体，本文采用了传统的知识图谱表示学习的学习流程，即分别设置关系特征矩阵和实体特征矩阵，在初始化阶段对这两个矩阵按照各自目标维度随机向量初始化。在模型训练过程中通过 TransE 等打分函数对嵌入结果打分后进行更新，可以得到可见部分的特征嵌入。

## 4.2 未见关系的特征嵌入

### 4.2.1 关系位置图构建

为了能够与目标域知识图谱中的未见关系进行有效的特征学习，由于在源知识图谱中不存在相关的事实三元组可以提供知识，必须要从其他层面获取到对未见关系编码有效的语义信息。通过第三章，本文在本体中添加了关系的本体三元组，学习到了关系在本体层面的语义信息。同样，基于第三章的关系位置元关系的设定，可以将源域知识图谱所有的关系在关系位置图中表示，关系与关系之间通过已定义的四种关系位置元关系 (tail-head, tail-tail, head-head, head-tail) 进行连接。在关系位置图中，关系与关系间的位置联系可以作为关系的一种拓扑特征信息。这些结构性的信息会减少对具体实体和关系的依赖，在对关系特征进行学习的时候，通过聚合相连其他关系的信息来对关系进行特征表示，能够作用在未见的关系上。

与第三章抽取关系的位置元关系不同，在位置关系图中，本节没有对图中具有对称性的位置元关系进行筛选和去除。为了能够防止对图中对称的元关系重复的进行信息传递，在对关系节点进行特征聚合的时候，本节仅考虑每个关系节点的入向关系，以此屏蔽了位置元关系的对称性的影响。

因此根据输入的实体三元组，可以将原始的图结构转变为关系相关图，同时在构建图的过程中没有使用到任何额外的实体属性或关系属性，可以作用到任何未见或已知的关系上。在关系相关图中的节点代表关系，边表示在原始三元组任意两个关系对应的元关系。

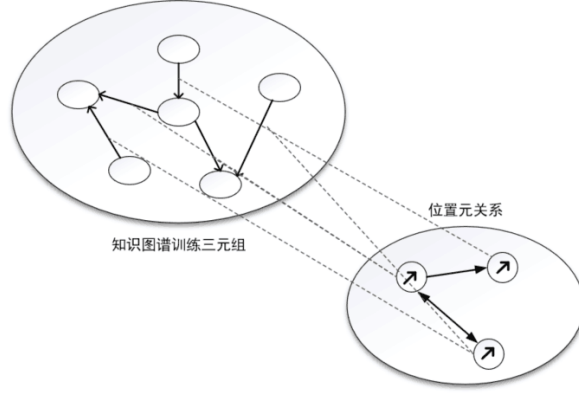


图 4.2 从实例图构建关系图

#### 4.2.2 关系相关性系数聚合编码

在关系位置图中，结点为原三元组中的关系，可以通过聚合关系结点的其他相邻关系结点的特征作为该关系的特征。为了能够将本体的嵌入作为对关系嵌入的语义补充，在构建好的关系位置图上，本文通过关系节点到本体概念的映射获取到各关系节点的初始化特征表示。然后通过多层图卷积网络聚合邻接节点的特征来对关系节点的特征学习，对于节点的更新公式如公式4.1所示：

$$h_v = f \left( \sum_{(u,r) \in \mathcal{N}(v)} W h_u \right) \quad (4.1)$$

其中  $h_v$  是相邻关系节点的特征。每层图卷积会有两步的操作，首先左乘归一化后的邻接矩阵对邻居节点的特征进行聚合，然后右乘一个可学习的线性转换矩阵  $W_r$  将输入的特征映射到目标特征空间中，最后使用一个非线性的激活函数来获取本层的特征输出。

为了在聚合关系本体信息和拓扑信息时区分不同元关系的的重要性，本文在对不同元关系连接的关系特征进行聚合时设置一个可学习的权重参数，根据任务的表现来学习不同元关系对应的重要程度，计算公式转化如公式4.2所示：

$$h_v = f \left( \sum_{(u,r) \in \mathcal{N}(v)} W_{dir(r)} h_u \right) \quad (4.2)$$

其中  $W_{dir(r)}$  是两个关系节点相连的元关系类型相对应的参数，根据本文设定的四种不同的元关系，该系数由四个不同的参数控制，如公式4.3所示：

$$W_{dir(r)} = \left\{ \begin{array}{ll} W_{t-h}, & r \in R_{tail-head} \\ W_{h-t}, & r \in R_{head-tail} \\ W_{t-t}, & r \in R_{tail-tail} \\ W_{h-h}, & r \in R_{head-head} \end{array} \right\} \quad (4.3)$$

### 4.3 未见实体的特征嵌入

#### 4.3.1 基于关系聚合的实体表示

在仅存在源域知识图谱的一般表示学习中，知识表示学习通过对知识图谱三元组的结构信息学习，能够使得表示向量尽可能贴近事实。但在目标域知识图谱中存在未见实体的情况下，由于缺乏事实三元组的支撑，传统的 KGE 方法无法学习到未见实体的特征信息。

为了解决上述的问题，本文借鉴人类对未见实体的推理过程来对未见实体进行编码。如图4.3所示，传统的 KGE 方法能够通过对事实三元组的学习，对已知的 Tom 节点进行有效的学习。对于存在未见实体的目标域知识图谱，常人虽然无法获得节点 X、Y、Z、A 的具体内容，但是可以通过对左右两个图谱的推理可知：节点 X 应该是一个类似于 Tom 的一个学生类型的节点。因为 X 具有和 Tom 类似的邻域结构信息，如 student\_of、advisor\_of 及 lives\_in 等。这些结构性的信息可以帮助人去理解一个新的实体。同时这些实体的邻域结构信息与实体本身具体信息是无关且通用的。

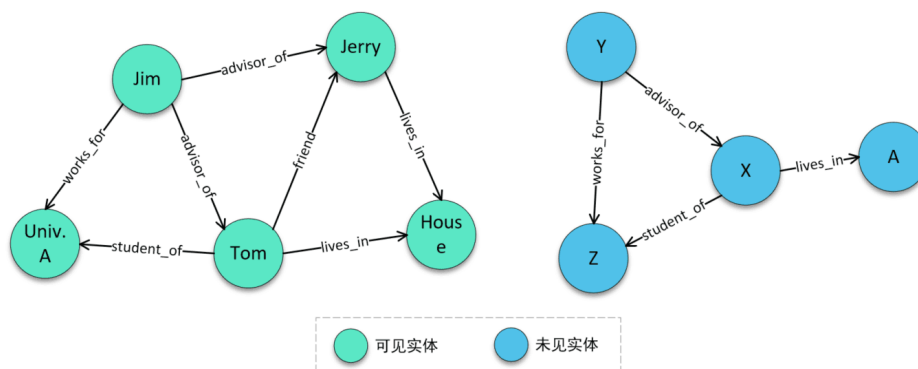


图 4.3 未见实体的关系特征

本节通过模拟人类对于新实体的推理过程，对实体邻域的结构信息进行建模来获得实体的特征编码。实体邻域结构信息最直接的体现便是与实体直接联系各个关系，因此本节通过使用实体相关的关系信息获得未见实体的嵌入。同时考虑到实体关系的方向性，本文通过公式4.4来进行实体关系的聚合：

$$\mathbf{h}_e = \frac{1}{|\mathcal{I}(e)| + |\mathcal{O}(e)|} \left( \sum_{r \in \mathcal{I}(e)} \mathbf{w}_{in}^{ent} \mathbf{h}_r + \sum_{r \in \mathcal{O}(e)} \mathbf{w}_{out}^{ent} \mathbf{h}_r \right) \quad (4.4)$$

其中  $\mathcal{I}(e)$  是实体  $e$  所有连接关系中指向实体  $e$  的集合， $\mathcal{O}(e)$  是实体  $e$  所有连接关系中由实体  $e$  出发指向其他实体的关系集合， $\mathbf{w}_{in}^{ent}$  和  $\mathbf{w}_{out}^{ent}$  分别是用于在聚合实体邻接关系入关系和出关系是对关系嵌入的权重矩阵。

然而，这种对实体通过关系聚合的方式只能对实体嵌入进行初始化表示，因为实体的邻域结构只传递类型级别的信息，而不是实例级别的信息。例如，对于节点  $X$ ，我们只能推断出该节点是一个类似 Tom 的学生类型节点，但无法得知节点  $X$  具体是谁。为了解决这个问题，本文在下一节引入了基于 GNN 的实体关系联合嵌入模块，根据实体的多跳邻域结构调制每个实体和关系的初始化嵌入。

#### 4.3.2 基于 GNN 的实体关系联合嵌入

在上述两个章节中，本文通过关系的本体信息和拓扑信息学习到了未见关系的初始化嵌入，通过实体的邻域结构信息的聚合学习到了未见实体的初始化嵌入。但是这些初始化的嵌入仅传递了部分的信息，如实体偏向于类型的初始化嵌入，没有充分利用到已知实体和已知关系的信息。因此，本节通过在实例知识图谱上，借鉴于 CompGCN 的模型结构，设计了一个能够同时对实体和关系进行邻域信息聚合的模块。该模块通过对每一个实体和关系聚合多跳邻域结构信息，能够充分利用到所有已知关系和已知实体的信息，学习到未见实体和未见关系更充分的语义表示。

该 GNN 网络也以 CompGCN 模型为基础，但是 CompGCN 原模型在对关系和实体进行聚合操作的时候采用了 TransE、HoLE 和 DistMul 模型的评分函数，限制了关系和实体的维度必须一致。为了能够使得该 GNN 网络能够输出维度不同的实体换关系的编码，本文借鉴于 MaKEr<sup>[37]</sup> 模型的设置，将 CompGCN 模型原有的实体关系聚合器修改为了一个线性转化层，可以让模型更好适应以多种传统 KGE 方法如 RotatE 来作为解码器进行下游任务，GNN 层的实体嵌入更新公式如公式4.5所示：

$$\mathbf{h}_e^{l+1} = f \left( \frac{\sum_{(r,e) \in \mathcal{N}(e)} \mathbf{W}_{dir(r)}^l [\mathbf{h}_r^l; \mathbf{h}_e^l]}{|\mathcal{N}(e)|} + \mathbf{W}_{self}^l \mathbf{h}_e^l \right) \quad (4.5)$$

其中  $\mathcal{N}(e)$  是实体  $e$  的所有相连的关系集合， $\mathbf{W}_{dir(r)}^l$  是对集合中关系不同方向特定的权重参数，对入方向和出方向时分别记为  $\mathbf{W}_{in}^l$  和  $\mathbf{W}_{out}^l$ ， $[\cdot]$  而指代两个向量的连接操作。 $\mathbf{W}_{self}^l$  是针对实体  $e$  本身特征自循环更新的模型学习参数， $f$  指代模型 GNN 模型的激活函数。同时，关系也在该层 GCN 中也进行了更新操作，如公式4.6所示：

$$\mathbf{h}_r^{l+1} = \mathbf{W}_{rel}^l \mathbf{h}_r^l \quad (4.6)$$

经过两层 GNN 网络完成了对关系嵌入和实体嵌入的更新，该模块即可输出该训练任务下所有可见的实体和关系以及未见的实体和关系的向量表示，用于本次任务的 loss 的计算及模型参数更新。

#### 4.4 基于元学习的训练任务设定

为了能够对知识图谱中的未见关系和未见实体进行表示学习，借鉴于元学习 “learning to learn” 的思想，本文设置了训练元任务和测试元任务。训练元任务是用于训练元

学习算法的任务，即在训练过程中让算法从元任务中获取经验并调整参数，以便于在测试任务中能够快速地适应。测试元任务则是在训练完成后用于测试元学习算法性能的任务。不同于元学习一般划分多个不同元任务的设定，本文采用单任务设定，即训练元任务和测试元任务均视为单一的任务。单任务设定可以使得元学习算法更加适应目标任务，并提高模型的泛化能力。同时，模型可以更加专注于学习适应固定任务的策略，还可以减少模型的计算和存储成本，并使得模型更加容易调整 and 解释。

本文从训练集中抽取了一系列包含未见实体和关系的训练任务来模拟跨域环境，并且在该训练环境中对模型进行训练。每个训练任务  $T^i = (\mathcal{E}^i, \mathcal{R}^i, \mathcal{T}_{sup}^i, \mathcal{T}_{que}^i)$  包含训练实体集、关系集、support 训练三元组集以及 query 测试三元组集。为了模拟未见的组件，将部分实体和关系标记为未见，每个训练任务被重新定义为如4.7所示：

$$T^i = (\mathcal{E}^i = (\hat{\mathcal{E}}^i, \tilde{\mathcal{E}}^i), \mathcal{R}^i = (\hat{\mathcal{R}}^i, \tilde{\mathcal{R}}^i), \mathcal{T}_{sup}^i, \mathcal{T}_{que}^i) \quad (4.7)$$

其中  $\hat{\mathcal{E}}^i \in \mathcal{E}^{train}$  指代实体集中的可见实体而  $\tilde{\mathcal{E}}^i \notin \mathcal{E}^{train}$  指代未见的实体； $\hat{\mathcal{R}}^i \in \mathcal{R}^{train}$  指代关系集中的可见关系  $\tilde{\mathcal{R}}^i \notin \mathcal{R}^{train}$  指代未见的关系。对训练集中三元组抽取训练任务和测试任务的流程如算法2所示：

```

Data: 读取所有三元组，构建大图  $\mathcal{G}_{train}$ 
for  $i < num\_subgraph$  do
    选取  $num\_root$  个根节点
    while 子图节点  $< 50$  do
        | 从根节点随机游走  $num\_step$  步，构建子图  $\mathcal{G}_i$ 
    end
    从子图三元组中抽取 10% 作为 query 三元组
    对 query 三元组中的关系和实体随机标记为未见
    构建 support 三元组中的实体和关系联系矩阵，构建关系位置图
    将 support 三元组、query 三元组、关系位置图存储
end
    
```

算法 2 任务子图构建流程

所有训练的总目标是基于每个任务的 support 集上进行对未见组件的训练使得在 query 集上的评估得分最高，计算函数如公式4.8所示：

$$\max_{\theta} \mathbb{E}_{T^i \sim p(T)} \left[ \sum_{(h,r,t) \in \mathcal{T}_{que}^i} \frac{1}{|\mathcal{T}_{que}^i|} \mathcal{M}(h, r, t | \mathcal{T}_{sup}^i) \right] \quad (4.8)$$

其中  $\mathcal{M}$  通过在 support 集上学习到实体和关系的嵌入表示，然后在 query 集上对三元组

计算得分。本文采用自对抗负抽样损失函数来对计算损失更新模型，如公式4.9所示：

$$\begin{aligned} \mathcal{L}(T^i) = & \frac{1}{|\mathcal{T}_{que}^i|} \sum_{(h,r,t) \in \mathcal{T}_{que}^i} -\log \sigma(\gamma + s(h, r, t)) \\ & - \sum_{i=1}^n p(h'_i, r, t'_i) \log \sigma(-\gamma - s(h, r, t)) \end{aligned} \quad (4.9)$$

$n$  指代负采样的数量， $p(h'_i, r, t'_i)$  指代负采样的权重系数， $s(h, r, t)$  是用于计算嵌入向量在 query 三元组上的得分函数。

## 4.5 基于链接预测任务的模型实现

知识图谱的知识表示学习通过学习实体和关系的向量表示来捕捉隐含的语义关系和规律，跨域知识图谱的知识表示学习在传统知识表示学习的基础上，能够学习到目标域知识图谱上新的实体和关系的特征表示。为了验证知识表示学习的有效性，本文结合下游链接预测任务对提出的知识表示学习方法进行评判。

链接预测任务通常涉及知识图谱中实体对之间的关系推理。该任务基于现有事实来推断出未知的三元组，以帮助补全知识图谱中的缺失信息、推断未知实体之间的关系及预测新实体属性等。链接预测任务可分为头实体预测和尾实体预测两种类型。对于头实体预测，已知关系和一个实体，需要预测与该实体关联的另一个实体。对于尾实体预测，则是已知实体和它的某个关系，需要预测与该实体存在该关系的其他实体。为了评估模型在链接预测方面的能力，常用的指标是基于排名的，如 MRR 和 Hits@k 等。其中 MRR 被定义为测试数据集中每个三元组的预测排序的平均排名的倒数，而 Hits@k 则是指在测试数据集中每个三元组的预测概率排名中前 k 名的正确率，即命中率。这些指标可以帮助确定预测结果的准确性和可靠性，同时有助于对不同模型进行比较和评估。

本文在训练任务的 support 上学习到实体和关系的嵌入，并在 query 集上进行损失计算。在进行每次任务损失的过程中，通过正负样本的联合计算出损失，并以此进行梯度下降求解。负样本损失计算公式如公式4.10所示：

$$\mathcal{L}_{neg} = \sum_{i=1}^n (\text{softmax}(W[s_{tail}; s_{head}]) \log \sigma(-[s_{tail}; s_{head}])) \quad (4.10)$$

其中  $s_{tail}$  和  $s_{head}$  是对一个测试三元组对应的尾结点负采样得分和头结点负采样得分， $n$  代表负采样的数量， $[\cdot]$  操作将正负样本的得分矩阵进行简单的连接操作， $W$  是对得分进行微调的权重参数，本文默认设置为 1。之后和正样本的得分求得算术平均即为该测试三元组的单个损失。

对于模型的训练流程，本文按照元学习单任务训练方法进行，每个任务的目标都是使该任务下的连接预测效果最优，单个任务的训练流程包含有 5 步：

- 1) 初始化模型参数，主要包含了训练数据集中可见实体和关系的嵌入初始化、元关系的权重参数以及两个不同阶段的 GCN 层和 CompGCN 层的传递参数。

- 2) 获得任务的支持集和查询集，主要从原训练集中抽取单个任务的子图，并划分为支持集和查询集，同时对关系和节点的标签进行设置以模拟出未见组件。
- 3) 单次任务训练获得实体关系的嵌入，通过 KGE 模型计算查询集损失，通过损失函数进行梯度下降求解。
- 4) 更新模型参数，重复下一个任务的训练直至效果不再更优。
- 5) 对测试集进行模型测试，输出任务得分。

## 4.6 本章小结

本章从本文模型的总体架构开始，详细介绍了模型的各个组成部分，包括主要的未见关系嵌入、未见实体嵌入以及基于元学习的训练任务设定和基于链接预测任务的模型实现的说明。对未见关系嵌入，说明了本文模型如何对关系的拓扑结构进行提取以及如何通过图卷积层对关系的本体信息和拓扑信息进行联合学习，从而获得未见关系的有效表示；对未见实体嵌入，通过未见实体的邻接关系的特征聚合，并且在实例知识图谱上通过 GNN 来聚合其他关系和实体信息，完成最后的实体和关系的特征更新。之后的章节中将对本文模型在测试数据集上进行充分的实验，检验本文模型的有效性。



## 5 实验结果及分析

在前文中，本文介绍了一种面向跨域知识图谱的知识表示学习方法，通过结合元学习和本体嵌入技术，获得了对跨域知识图谱中新的实体和关系的处理能力。该方法主要包括基于关系拓扑信息和描述文本的本体嵌入框架以及基于元学习的图谱嵌入模型，并从理论层面解释了如何利用本体信息对未见实体和关系进行建模。同时，采用元学习的训练方法来模拟跨域知识图谱场景，以进一步提高模型对新实体和新关系的泛化学习能力。为了验证模型的有效性，本章主要验证和评估了本文模型（MOKER）的实验效果。实验主要针对面向跨域情景的测试数据集进行，与现有的处理类似问题的归纳知识图谱表示模型进行对比。结果表明，引入了多层次的特征信息后，本文模型在实验数据上明显优于其他现有模型。最后的多个消融实验也证明了模型各个组成部分的重要性。

### 5.1 数据集

由于传统的知识图谱数据集通常基于“closed-world”设定，测试三元组中的实体和关系在训练数据集中已知，不存在跨域场景中的未见实体和未见关系。因此，为测试模型在包含未知实体和关系的跨域场景下的有效性，本文利用包含本体层次三元组的 DBpedia 和 NELL-995 知识图谱构建了新的数据集 DB\_Ext 和 NELL\_Ext。这两个数据集是从 DBpedia 和 NELL-995 数据集抽取出的子集。

#### 5.1.1 源数据集介绍

**DBpedia:** DBpedia 是一个基于维基百科的语义知识库。该数据集由开源社区维护并使用维基百科文章和其他在线网络资源进行扩展。DBpedia 从 Wikipedia 中提取结构化的数据，转化为事实三元组进行存储，包括图像、标签、描述文本等结构化属性，最新的数据集快照含有 8.5 亿条事实三元组数据。此外，DBpedia 注重本体论的构建，本体类型数共 768 个，主要包括人物、地点、工作、组织等概念。这些实体、属性和关系通常以 RDF 图的形式存储，标准查询语言为 SPARQL。DBpedia 开源社区提供网页公布数据集的最新之版本和统计数据，同时支持各数据集版本的检索和下载，为研究者使用提供了便利。

**NELL-995:** NELL-995 数据集是根据 NELL 知识库抽取用于知识图谱补全任务的基础数据集。NELL (Never Ending Language Learning) 是由卡内基梅隆大学领导的自动化学习系统，旨在从互联网的非结构化文本中自动提取知识。基于 NELL 知识库，NELL-995 数据集包含 995 个实体和 129 个关系，覆盖地理、医学、体育、音乐、文化等广泛领域。NELL-995 数据集的实体和关系被细粒度分类，每个实体被分配到一组多层次的关系类

型中。同样地，与其他实体之间的关系也分类。因此，NELL-995 数据集是学习细粒度知识表示和关系推理模型的理想数据集。

### 5.1.2 任务数据集构建

为了抽取出包含未见实体和关系的数据子集，本文首先构建一个包含所有三元组的实例知识图谱，在其中随机选取 100 个根实体节点并将每个节点的 10 个相邻接点组成测试子图，将子图中 1/10 的实体和关系标记为测试集，并从训练集中剔除这些实体和关系相关的三元组作为测试三元组。最终每个数据集中都包含了两个基本的子知识图谱，训练知识图谱  $\mathcal{G}^{train}$  和测试图谱  $\mathcal{G}^{test}$ ，后者包含了训练图谱中不存在的实体和关系。此外，为了测试模型对未见关系及未见实体的单独学习性能，在测试集中的 query 集设置中，本文将测试三元组分为三类：1、所有测试三元组中只包含了未见的实体（unseen\_ent）；2、所有测试三元组中只包含了未见的关系（unseen\_rel）；3、所有测试三元组中同时包含了未见的实体和关系（unseen\_both）。两个数据集中的统计数据如表5.1所示。其中，DB\_Ext 数据集包含 243 个仅含未见实体的测试三元组、10 个仅含未见关系的三元组和 243 个同时包含两个未见组件的三元组；NELL\_Ext 数据集包含 565 个仅含未见实体的测试三元组、12 个仅含未见关系的三元组和 115 个同时包含两个未见组件的三元组。各数据集包含三元组的统计数量如表5.1所示。

表 5.1 数据集统计数据 (括号中为未见组件数量)

	训练图谱			测试图谱			
	实体数	关系数	三元组数	实体数	关系数	support 三元组数	query 三元组数
NELL_Ext	1583	153	5269	851(753)	140(30)	2160	692
DB_Ext	795	115	1508	913(884)	128(46)	1930	496

## 5.2 模型参数设置

对于用于对比的 baseline 模型的参数设置，本文遵循了相关论文给出的最优超参数设置，对于本文提出的模型的相关参数设置如表5.2所示：

其中主要包含一下三个方面的参数设置：

- 1) 在本地三元组进行嵌入时模型学习率 lr、本体三元组概念节点结构嵌入维度 ent\_str\_dim、本体概念节点文本嵌入维度 ent\_text\_dim、线性隐藏层维度 mapping\_size 以及训练的 epoch 数量和 batch 的大小。
- 2) 对知识图谱表示元学习训练相关的任务学习率 lr、单任务支持集的 batch 数量 train\_bs、单任务查询集的 batch 数量 eval\_bs、元训练总任务数 num\_step、元训

表 5.2 模型超参数设置

本体嵌入参数	设置值	元学习训练参数	设置值	嵌入参数	设置值
lr	0.00005	lr	0.001	dim	300
ent_str_dim	150	train_bs	64	num_gcn	2
ent_text_dim	300	eval_bs	16	num_comgcn	2
mapping_size	300	num_step	100000	gcn_dim	300
training_epochs	1000	early_stop_patience	20	hid_drop	0.3
batch_size	100	num_sample_size	10	-	-

训练提前结束无效训练任务计数 `early_stop_patience` 以及单个 `batch` 采样的根节点数。

- 3) 对实体和关系嵌入基本维度的设置 `dim`、关系位置图对关系进行 GCN 的层数 `num_gcn`、GCN 中间传递维度 `gcn_dim` 以及 GCN 层的丢弃率 `hid_drop`、对关系和实体进行联合学习的 ComGCN 的层数。

在对三元组进行链接预测的评定过程中，本文通过对 CompGCN 第二层输出的改进可支持多种 KGE 模型作为评分函数，实际采用的 KGE 模型包含 TransE、DistMult、ComplEx 及 RotatE。实体和关系的维度根据采用的 KGE 模型在基础嵌入维度上进行调整，具体如下表 5.3 所示：

表 5.3 评分函数

模型名	实体维度	关系维度	评分函数
TranE	dim	dim	$F = -\ h + r - t\ $
DistMult	dim	dim	$F = h^T \text{diag}(r) t$
ComplEx	$2 * \text{dim}$	$2 * \text{dim}$	$F = \text{Re}(h^T \text{diag}(r) t)$
RotatE	$2 * \text{dim}$	dim	$F = -\ h \circ r - t\ $

其中评分函数中的  $h$ 、 $r$ 、 $t$  分别指代头实体、关系和尾实体的嵌入表示， $\text{Re}$  表示复向量的实部分量， $\circ$  操作表示旋转操作。

### 5.3 实验设计及评价指标

本实验包含本体嵌入表示学习和图谱表示学习两个阶段，第一个阶段主要学习到融合描述文本信息的本体嵌入表示，第二阶段使用本体信息进行跨域知识图谱的表示学习。

第一阶段首先采用传统的表示学习方法对本体三元组数据进行初步嵌入表示。其

次，从预训练词嵌入 glove 中获取本体节点的描述信息中各描述单词的初始化词向量。使用 TF-IDF 统计方法识别单词的重要程度，并对单词的词向量进行加权聚合计算，以获得本体节点描述文本的初始化向量嵌入。最后，将本体三元组的结构化表示嵌入和本体描述文本的表示嵌入映射到同一个空间，以进行评分和更新，从而获得拼接后的最终本体嵌入。

第二阶段对未见组件进行表示学习，为了在元学习中模拟出存在未见组件的场景，在每个元学习任务的设置中都人为抹除了一些实体和关系的标签，使得这些实体和关系必须通过的模型学习得到嵌入表示，而不是从传统嵌入方法的嵌入矩阵中取得。

对于模型在链接预测任务上的评价指标，本文选取了 MRR 和 Hit@10 作为评判的标准。其中 MRR 通过计算预测三元组排名的倒数来进行计算，即对测试三元组中的所有事实三元组，如果该三元组在预测排名中得分高那么排名就靠前，对应的倒数也会比较大，因此 MRR 作为评价指标越大则证明总体链接预测的排名较好。而 Hits@n 描述的是在所有预测三元组排名中前 n 的三元组所占的平均比例，其计算公式如公式 5.1 所示：

$$HITS@n = \frac{1}{|S|} \sum_{i=1}^{|S|} \mathbb{I}(\text{rank}_i \leq n) \quad (5.1)$$

假如 n 设置为 10，那么统计事实三元组在预测三元组中前 n 名的个数，最后再除以总个数就得到了 Hits@10 的结果，其中  $\mathbb{I}(\cdot)$  为 indicator 函数（若条件真则函数值为 1，否则为 0）。

参与比较的模型如下：

Neural-LP<sup>[22]</sup>（2017）：该模型基于知识库构建了一个可学习逻辑规则的可微模型。逻辑规则是独立于实体和关系的，因此理论上该模型可在任何未见的实体上应用，并在归纳图谱补全任务中相比传统方法（即对实体进行结构信息表示学习）有明显提升。

DURM<sup>[23]</sup>（2019）：提出了一种可微的、可同时学习规则逻辑及其置信度得分的方法。该方法可以使用梯度优化来处理归纳逻辑编程任务，并可用于处理含有未知实体的链接预测任务。

GraIL<sup>[49]</sup>（2020）：该模型不直接学习实体节点嵌入，也没有使用任何节点的属性。相反，它在测试三元组候选关系的周围构建子图，并利用子图的结构和结构化的节点特征来预测三元组。这使得该模型能够很好地应用于未知的实体三元组预测任务。

CoMPILE<sup>[50]</sup>（2021）：该模型对 GraIL 模型子图归纳模型进行了改进。改进包括加强对子图关系方向性的限制，并在未知节点特征聚合的信息传递过程中增加了先前模型中忽略的关系特征。

MaKEr<sup>[33]</sup>（2022）：该模型通过学习关系结构的特征来聚合关系邻接关系的特征，进而对关系进行表示，并聚合关系特征对实体进行编码。模型利用拓扑结构的信息，在一定程度上实现对未见实体和未见关系的表示。

## 5.4 实验结果及分析

各模型在 NELL\_Ext 上的链接预测任务实验结果如下表5.4所示，各模型在 DB\_Ext 上的链接预测任务实验结果如表5.5所示。按照测试数据集的三元组对未知实体和关系的包含情况，分为了只包含未见实体的结果 (u\_ent)、只包含未见关系的结果 (u\_rel) 以及同时包含未见实体和未见关系的结果 (u\_both)。表格中加粗的为最优的实验效果，带有下列划线则是该类基本模型中表现最优的得分，模型括号后面的指代的是在评分阶段采用的 KGE 评分函数。

表 5.4 NELL\_Ext 数据集结果

NELL_Ext						
	u_ent		u_rel		u_both	
	MRR	Hits@10	MRR	Hits@10	MRR	Hits@10
Neural-LP	30.48	47.96	-	-	-	-
DRUM	31.82	48.32	-	-	-	-
GraIL	71.62	92.92	-	-	-	-
CoMPILE	<u>75.94</u>	<u>93.62</u>	-	-	-	-
MaKEr(TransE)	70.82	92.00	24.56	54.17	21.53	51.74
MaKEr(DistMult)	70.63	91.33	27.02	60.00	<b>41.39</b>	57.65
MaKEr(ComplEx)	72.24	91.91	18.27	34.17	29.39	59.65
MaKEr(RotatE)	<u>77.09</u>	<u>94.64</u>	<u>31.53</u>	<u>55.00</u>	31.45	<u>62.35</u>
MOKER(TransE)	78.28	94.86	20.72	53.34	27.11	55.85
MOKER(DistMult)	75.98	92.46	19.30	22.50	31.37	55.65
MOKER(ComplEx)	73.61	90.60	24.44	38.33	29.70	54.96
MOKER(RotatE)	<b>79.92</b>	<b>94.73</b>	<b>45.07</b>	<b>75.63</b>	<u>40.33</u>	<b>67.06</b>

上面两个表格展示了各模型在 NELL\_Ext 和 DB\_Ext 上的链接预测结果。其中 Hits@10 在计算的过程中本文选取了 50 个候选进行评估，并根据不同类型的查询三元组（即 u\_ent、u\_rel 和 u\_both）显示了不同模型的得分。而且对于 GraIL、Neural-LP、DRUM 和 CoMPILE 在对应的论文中只在包含未知实体的数据集上做了实验，因此在对存在未见关系的测试集结果上进行留空处理，但是仍在未见实体的测试集中比较相对应的效果。上述结果均为模型运行 4 次后取平均值的结果。

结果表明，本文提出的 MOKER 模型相比其他基准模型有所改进，并在不同的 KGE 评分模型上有不同程度的提升。最优结果基本上由 MOKER 与 RotatE 模型组合得出。相比其他 KGE 模型，RotatE 模型采用了更加复杂的关系和实体的映射关系，因此能表现

表 5.5 DB\_Ext 数据集结果

	DB_Ext					
	u_ent		u_rel		u_both	
	MRR	Hits@10	MRR	Hits@10	MRR	Hits@10
Neural-LP	57.15	73.46	-	-	-	-
DRUM	59.88	73.25	-	-	-	-
GraIL	59.44	<u>80.86</u>	-	-	-	-
CoMPILE	<u>60.66</u>	79.93	-	-	-	-
MaKEr(TransE)	54.4	83.7	31.13	54.00	38.66	66.5
MaKEr(DistMult)	46.24	81.07	16.43	11.00	32.16	56.71
MaKEr(ComplEx)	53.79	82.47	19.95	29.00	36.88	59.26
MaKEr(RotatE)	<u>59.55</u>	<u>86.09</u>	<u>32.93</u>	<u>55.00</u>	<b>41.27</b>	<u>66.54</u>
MOKER(TransE)	64.63	89.60	<b>44.77</b>	70.25	34.92	<b>69.42</b>
MOKER(DistMult)	56.73	80.25	13.68	11.00	33.94	61.74
MOKER(ComplEx)	52.22	77.34	14.40	15.00	31.72	59.49
MOKER(RotatE)	<b>66.43</b>	<b>89.67</b>	41.80	<b>74.00</b>	<u>35.11</u>	63.81

更多没有重叠的特征信息。这也证明了本文模型的有效性。

在处理未见的实体的测试集上，本文比较了基于规则学习处理未见实体的 Neural-LP 模型和 DRUM 模型以及基于子图推理的 GraIL 模型和 CoMPILE 模型。从结果上可以明显看出基于规则的模型依赖于从数据集学习出的规则，模型效果受数据集影响较大，且需要有大量的数据集或者样本平衡要求十分严格。总体效果上没有基于子图的模型效果好，在 NELL\_Ext 数据集上结果差距较大。CoMPILE 模型在 GraIL 模型的基础上强调了关系的重要性，总体效果上要比 GraIL 表现更好。但上述四个模型首先无法处理未见关系，而且基于子图的归纳推理模型强调了测试三元组中头尾实体间的局部子图信息，没有完全利用到实体周围的结构特征信息以及关系信息。MaKEr 模型虽然考虑到了关系对实体的重要性，但是在关系表示的特征学习上效果比本文模型要差，因此通过邻接关系聚合特征表示未知实体的效果在 MRR 的评价指标上比本文模型平均低了 4.42%，在 Hits@10 的表现上本文模型在对应 RotatE 评分函数上也领先了 1.85%，说明了本文模型在未见实体表示上的有效性。

在处理未见关系和同时处理两种未知组件的测试集上，本文比较了通过结构信息对关系进行编码并通过聚合关系信息对未见实体和未见关系进行编码的 MaKEr 模型。对于未见的关系，本文通过引入额外的本体知识作为关系语义信息的补充，同时通过关系的位置结构利用关系图卷积对关系的表示进行更新，以学习到未见关系周围的结构拓扑

信息。从两个数据集上的链接预测任务结果来看,相比于 MaKEr 模型仅学习结构信息对关系表示的方法,本文引入的本体信息能够有效对关系的表示进行语义补充。在处理未见关系时 MOKER 模型中表现出明显的优势,其 MRR 得分比 MaKEr 的得分高 14.54%。而在 Hits@10 的得分方面,70.63 的得分比 MaKEr 的得分高出 20% 左右。表明本文模型能够更好地捕捉关系的语义和结构信息,在考虑局部和全局关系时都具有不错的性能。在同时含有两种未知组件的测试集上,相比于 MaKEr 模型,本文模型在 NELL\_Ext 数据集的 Hits@10 评分上有平均 1.82% 的提升,在 DB\_Ext 数据集的 MRR 和 Hits@10 分别有平均 1.18% 和 0.53% 的提升。

除此之外,本文实验发现,TransE 和 RotatE 的模型效果比 DistMult 和 ComplEx 模型更好,尤其在关系处理方面。在本体嵌入实验中,采用 RotatE 作为评分函数,对实体知识图谱进行表示。DistMult 和 ComplEx 模型因其复杂性不适合与 RotatE 的本体嵌入配合,而效果反而下降。相比之下,TransE 模型简洁易操作,在进行特征提取时更加适合。而 RotatE 作为近几年提出的 KGE 模型,一方面本身与本体嵌入方法相同,另一方面个更全面的对实体和关系进行低维嵌入表示,因此取得了本文模型中的最优效果。

## 5.5 模型消融实验

本节将介绍针对模型的几个重要模块的多项消融实验,以展示本文模型的不同部分的重要性,主要设置了 5 项不同的消融设置实验:(1) 去除元学习的设置;(2) 去除本体的设置;(3) 去除实体聚合表示的设置;(4) 去除关系 GCN 聚合的设置;(5) 同时去除本体和元学习的设置。获得的实验结果如表 5.6 所示:

表 5.6 在 NELL\_Ext 上的消融实验结果

	NELL_Ext					
	u_ent		u_rel		u_both	
	MRR	Hits@10	MRR	Hits@10	MRR	Hits@10
MOKER(TransE)	<b>78.28</b>	<b>94.86</b>	20.72	53.34	<b>27.11</b>	<b>55.85</b>
no_meta_TransE	29.31	43.82	12.65	30.83	10.18	22.26
no_ont_TransE	76.89	94.07	14.96	27.50	14.06	28.78
no_ent_TransE	76.85	94.07	18.00	55.83	19.43	50.7
no_gcn_TransE	74.68	92.27	<b>23.88</b>	<b>61.67</b>	19.85	46.26
no_meta_ont_TransE	26.57	38.69	11.15	25.00	12.62	27.74

其中个消融实验的具体设置如下:

- 1) 去除元学习的设置: 在模型训练阶段,本文采用了元学习的设置在训练集上提取任务子图并通过随机标签来模拟出未见的组件从而训练出模型处理未见组件的能

力。在该消融实验下，取消掉了对子图标签的模拟效果即按照传统 KGE 模型的训练方法，在训练集上对所有三元组进行嵌入和损失计算。

- 2) 去除本体的设置：本文模型的一个创新点即在对关系嵌入表示的时候通过关系位置图加入本体信息进行学习，在该消融实验设置下，将本体信息替换为随机表示对未见关系进行特征学习。
- 3) 去除实体聚合表示的设置：本文模型在处理未见实体时，采用该实体周围的关系信息聚合表示，该消融实验设置下对未见实体进行随机化嵌入设置，考察未见实体的表示模块。
- 4) 去除关系 GCN 聚合的设置：在构建完关系的位置图并引入本体嵌入后，本非模型通过 GCN 来加强关系对邻接结构的特征学习，该消融实验下去除两层 GCN 层，直接使用本体信息。
- 5) 同时去除本体和元学习设置：同时结合消融实验 (1) 和 (2) 的设置。

消融实验结果与 MOKER(TransE) 模型在 NELL\_Ext 上的实验结果进行比较。从消融实验结果可以看出，基本所有的消融设置都会导致性能下降，表明这些设计的重要性。但是在去除关系 GCN 聚合的设置上，在仅含有未见关系的测试集上效果反而提升，分析可知该层 GCN 的作用是在关系本体嵌入的基础上学习关系邻接关系的结构信息。这些结构信息的引入一定程度上会影响仅对关系的表示效果导致在仅包含未见关系的测试集上的效果下降。但是这些结构信息的引入在中未见实体的表示中，因为实体需要聚合邻接关系进行表示，因此引入的结构信息会对实体的表示效果进行提升，可以发现去掉 GCN 在包含未见实体和同时包含未见实体和未见关系的测试集上效果都有明显的下降，因此也侧面证明了 GCN 模型对实验效果提升的必要性。此外，本文观察到元学习设置对模型性能至关重要，表明在推广到测试知识图谱的任务上对模型进行元训练的有效性。

## 5.6 未见实体案例分析

在图5.1中展示了本体提出的 MOKER 模型和去除元学习和本体嵌入的传统 TransE-KGE 模型产生的在 NELL-Ext 数据集的测试集上实体嵌入的可视化。在这张图中，不同颜色展示了不同类型的实体，圆点代表未见的实体，叉号则代表了该类型下的已知实体。MOKER 在实体获得初始化的向量表示后，通过两层 GCN 来学习邻域节点和相连关系的特征，能够使得同一类型的实体在表示空间上尽可能聚集在较紧密的邻域中。从可视化中也可以看出 MOKER 产生的嵌入分布与对应类型更加一致，在嵌入映射的距离上更近紧凑，而 TransE-KGE 产生的嵌入则混合了不同实体类型。MOKER 将嵌入映射到了不同的聚类中，而 TransE-KGE 中不同实体类型的嵌入则混合在一起。此外，本文的模型在对未见的实体进行表示学习的过程中，采用聚合实体关系的特征来初始化实体向量，相邻关系的结构特征可以表现出实体的类型信息，因此模型可以将未见过的实



体的嵌入与同一类型的已知实体聚类。不同类型实体的聚类表明，MOKER 能够用包含合理语义和信息性知识的嵌入来表示未见过的实体。

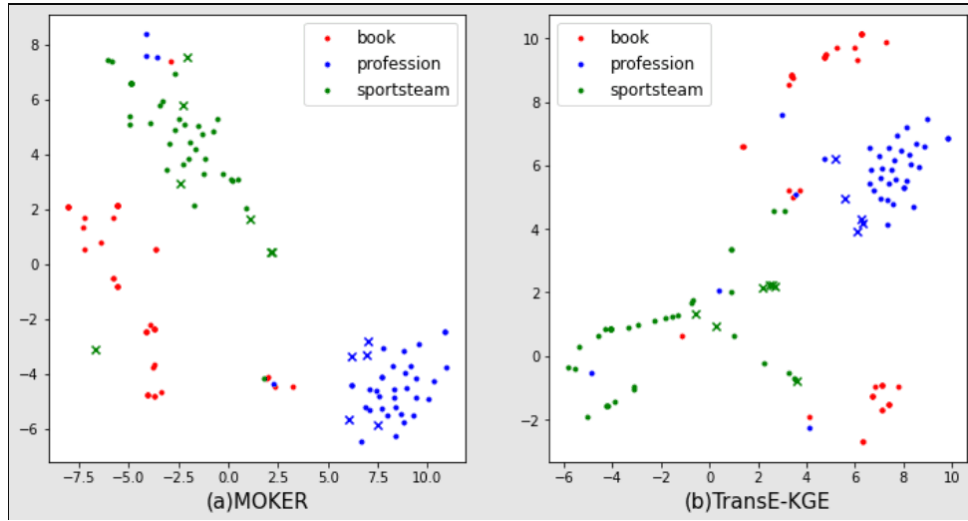


图 5.1 实体嵌入可视化分析

## 5.7 未见关系案例分析

对于未见关系，本文选取了部分关系同样映射到了二维空间上进行可视化分析，如图5.2所示，其中圆点代表通过训练集中可见的关系，叉号代表测试集中的未见关系。本文模型通过采用关系位置图根据关系的相对位置对关系局部的邻接关系进行建模学习到了关系的拓扑结构信息，同时引入本体信息作为语义补充，因此同一实体的具有相近语义的相邻关系在距离上应该表现为更加相近，从图上可以看出，对于未见关系 `has_office_in_city` 贴近于具有类似语义的关系 `has_office_in_coutry`，未见关系 `person_born_in_city` 更贴近于已知关系 `person_born_in_location`。而对于已知关系 `parent_of_person` 和 `father_of_person` 和未见关系 `mother_of_person`，本文可知一个实体如果存在 `parent_of_person` 的关系那么该实体节点的邻接关系中只能存在其中一个 `father` 或者 `mother` 的关系，因此在图中本文可以观察到 `mother_of_person` 在距离上更接近于 `parent_of_person` 关系，而远离 `father_of_person` 关系。由此可见，本文模型通过关系的位置图上联合本体语义信息有效的学习到了对应关系的语义关系，且其中相似的关系在向量空间中靠近，显示了本文提出的 MOKER 在嵌入未知关系方面的有效性。

## 5.8 本章小结

本章将本文提出的模型在测试数据集上 NELL\_Ext 和 DB\_Ext 上进行了链接预测任务的相关实验，测试数据集从 NELL-995 和 DBpedia 上抽取子集并添加了测试集中的未见组件。与多个基准模型相比，本文提出的模型在任务得分上均有不同程度的提升，并通过对实验结果的分析验证了模型对于表示学习效果增强的有效性。同时通过对各个模

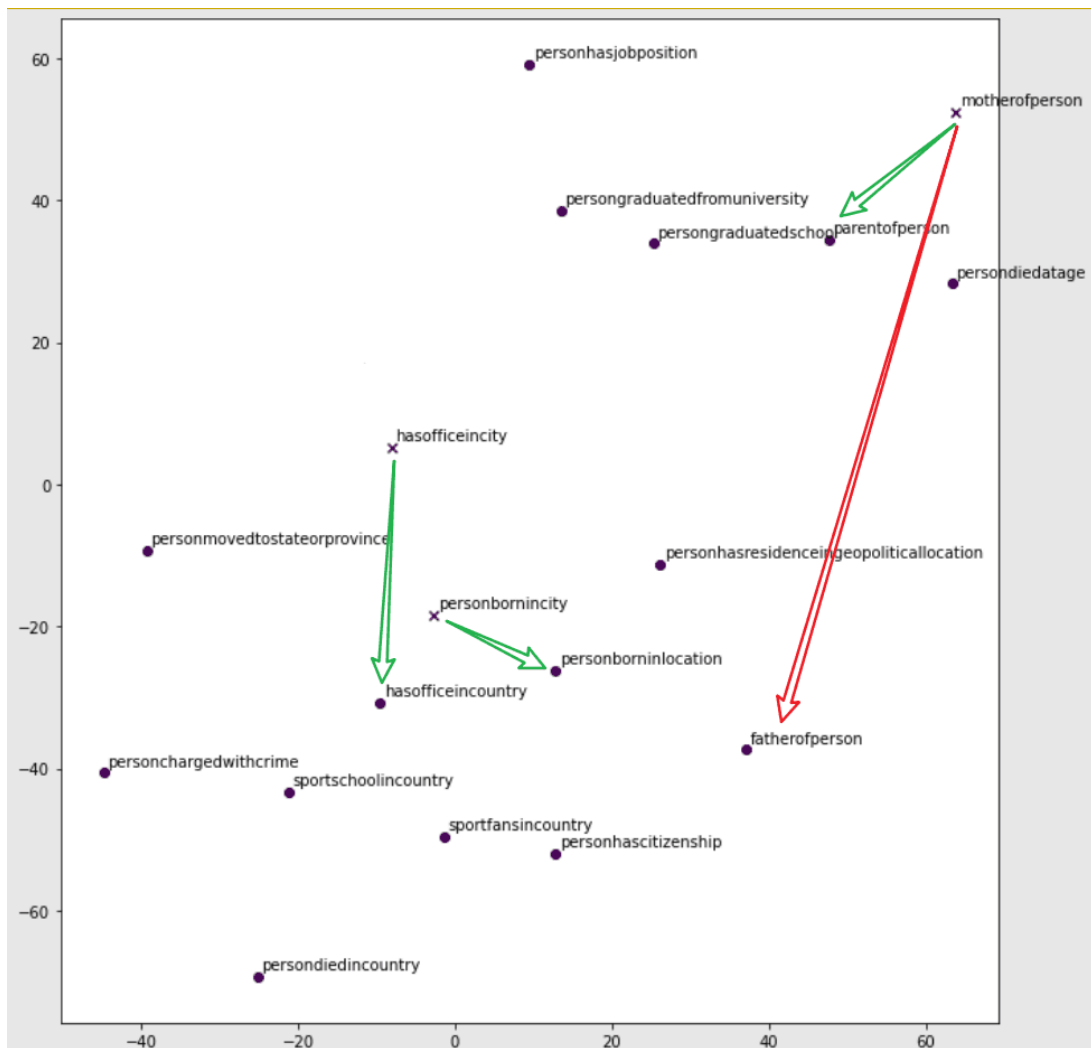


图 5.2 实体嵌入可视化分析

型组件的消融实验发现明显的效果下降，证明了模型模块的重要性；最后对未见实体和未见关系的案例分析可知得到的嵌入表示符合模型原理设定，再次证明了该模型对未见组件表示上的突出能力。

## 6 总结与展望

### 6.1 总结

当下知识图谱技术的应用日益渗透到每个人的日常生活中，从搜索引擎到推荐系统，不断探索知识图谱知识的应用与补充是未来发展不可或缺的一个研究方向。但是出于数据隐私、成本等多方面的考量，大到公司多个图谱存储服务器，小到人们每日使用的移动终端，我们无法将所有分散知识图谱新添加的实体和关系完全覆盖。如何在源域知识图谱上训练，并且能够将表示学习能力泛化到应用于包含未见实体和关系的目标域知识图谱的研究是无可避免的需求。

面向跨域知识图谱的知识表示学习问题，本文采用了元学习的方法，在训练任务中模拟了跨域场景下的未见关系和未见实体，从而获得了模型处理未见组件的能力。元学习方法具有训练效率高的优点，可以在算力珍贵的时代大幅降低成本消耗。对于未见关系的表示能力学习，本文依据关系间的相对位置关系构建了一个以关系为结点的图，并通过预定义的元关系将其连接起来，以体现出关系的拓扑信息。此外，我们将本体信息嵌入作为关系语义信息的体现，再经过图卷积网络对关系节点进行拓扑信息和语义信息的联合学习，获得了对未见关系的表达。针对未见实体的表示能力学习，本文通过分析总结出相似类型实体在邻接结构上表现出高度相似性这一结论，并采用了实体的邻接关系特征聚合作为未见实体的初始化表示，避免了对未见实体周围节点的苛刻性要求。但聚合关系的表示仍然不足以支撑对实体特征的体现，为了能够充分利用到已知实体和关系的特征信息，我们使用 CompGCN 对实体和关系进行邻接信息的学习和更新，从而获得了各组件的嵌入。通过实验、基准模型比较和案例分析，本文证明了所提出模型的有效性。归纳起来，本文的工作主要包括：

- 1) 为了充分捕捉到本体中实体和关系对应的语义信息，对本体三元组进行嵌入，学习到本体的向量表示。同时为了补充本体信息中的关系信息，构建了关系的位置元关系，补充了本体中的关系本体三元组。在本体三元组结构信息嵌入的基础上，基于描述文本进一步对本体嵌入加强。
- 2) 在对跨域知识图谱的知识表示学习中引入了本体信息作为知识图谱的语义信息补充，通过对拓扑信息和语义信息的结合能够有效对未见实体和关系进行表示学习，同时采用复杂图卷积层对表示学习的效果进行加强，有效学习到了实体和关系的特征表示。模型总体采用元学习的任务设定对模型进行训练，在元学习的单任务中对跨域场景进行模拟，使得模型能够学习到处理未见实体和关系的能力，同时提高训练的效率，降低成本。
- 3) 在两个知识图谱链接预测基准数据集的基础上构建满足跨域场景要求下的测试数

据集，并与多个基准模型进行实验效果比较，通过对结果分析可知本文模型相比于其他基准模型均有不同程度的提升，证明了本文模型的有效性。

## 6.2 未来工作

虽然本文在测试数据集上的效果相比于其他基准模型已经获得了明显的提升，但是在本文的整个研究过程中仍旧发现了当下模型在应用方面值得继续探究的几个研究方向：

- 1) 本体嵌入的评分函数对模型效果的影响：在对本体的嵌入过程中，本文采用了目前传统 KGE 方法中效果最好的 RotatE 模型作为评分函数。但是从实验结果中我们可以发现对于 DistMult 和 ComplEx 评分函数的得分会比更简单的 TransE 更低，那么不同的本体嵌入方法可能对模型的效果会产生影响，是否本体嵌入评分函数和模型评分函数相对应会更获得更好的模型效果值得探究。
- 2) 数据集本体信息的提取：本文模型研究本体信息对知识图谱表示学习的可用性，本体三元组是必不可少的一环。但当下的一些基准知识图谱有些并没有相应的本体三元组，如被广泛使用的 FB15K-237 数据集，因为其源数据集已经停止维护，本体类型信息也都比较杂乱无法使用。虽然本文在处理测试数据集时采用了对实例三元组类型信息补充了本体三元组中的关系信息，但面对一个数据集如何更充分获取到其本体信息仍旧是待解决的难题。
- 3) 规则信息的引入：从 DB\_Ext 的基准模型测试信息可以看出基于规则提取的模型在未见的实体上也能够不错的表现效果，而且规则理论上可以同时作用于实体和关系，是对跨域知识图谱表示学习进行约束的良好设定。如何将规则信息融入到本文模型更好的学习到未见组件的表示也是未来工作的一个值得期待的方向。

## 参考文献

- [1] Zou X. A survey on application of knowledge graph[A]. Journal of Physics: Conference Series : Vol 1487[C], 2020 : 012016.
- [2] Bollacker K, Evans C, Paritosh P, et al. Freebase: a collaboratively created graph database for structuring human knowledge[A]. Proceedings of the 2008 ACM SIGMOD international conference on Management of data[C], 2008 : 1247 – 1250.
- [3] Xiong W, Hoang T, Wang W Y. Deeppath: A reinforcement learning method for knowledge graph reasoning[J]. arXiv preprint arXiv:1707.06690, 2017.
- [4] 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.
- [5] Suchanek F M, Kasneci G, Weikum G. Yago: a core of semantic knowledge[A]. Proceedings of the 16th international conference on World Wide Web[C], 2007 : 697 – 706.
- [6] 张正航, 钱育蓉, 行艳妮, et al. 基于 TransE 的表示学习方法研究综述 [J]. 计算机应用研究, 2021, 38(03) : 656 – 663.
- [7] Wang Z, Zhang J, Feng J, et al. Knowledge graph embedding by translating on hyperplanes[A]. Proceedings of the AAAI conference on artificial intelligence : Vol 28[C], 2014.
- [8] Lin Y, Liu Z, Sun M, et al. Learning entity and relation embeddings for knowledge graph completion[A]. Proceedings of the AAAI conference on artificial intelligence : Vol 29[C], 2015.
- [9] Socher R, Chen D, Manning C D, et al. Reasoning with neural tensor networks for knowledge base completion[J]. Advances in neural information processing systems, 2013, 26.
- [10] Xie R, Liu Z, Jia J, et al. Representation learning of knowledge graphs with entity descriptions[A]. Proceedings of the AAAI Conference on Artificial Intelligence : Vol 30[C], 2016.
- [11] Valverde-Rebaza J C, de Andrade Lopes A. Link prediction in complex networks based on cluster information[A]. Advances in Artificial Intelligence-SBIA 2012: 21th Brazilian Symposium on Artificial Intelligence, Curitiba, Brazil, October 20-25, 2012. Proceedings[C], 2012 : 92 – 101.
- [12] Lin Y, Liu Z, Luan H, et al. Modeling relation paths for representation learning of knowledge bases[J]. arXiv preprint arXiv:1506.00379, 2015.
- [13] Feng J, Huang M, Yang Y, et al. GAKE: Graph aware knowledge embedding[A]. Proceed-

- ings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers[C], 2016 : 641 – 651.
- [14] Xie R, Liu Z, Luan H, et al. Image-embodied knowledge representation learning[J]. arXiv preprint arXiv:1609.07028, 2016.
- [15] Li R, Cao Y, Zhu Q, et al. How does knowledge graph embedding extrapolate to unseen data: a semantic evidence view[A]. Proceedings of the AAAI Conference on Artificial Intelligence : Vol 36[C], 2022 : 5781 – 5791.
- [16] 刘洪波, 陈越, 卢记仓, et al. 面向知识图谱的规则挖掘研究综述 [J]. 计算机工程与应用, : 1 – 11.
- [17] Galárraga L A, Teflioudi C, Hose K, et al. AMIE: association rule mining under incomplete evidence in ontological knowledge bases[A]. Proceedings of the 22nd international conference on World Wide Web[C], 2013 : 413 – 422.
- [18] Omran P G, Wang K, Wang Z. Scalable Rule Learning via Learning Representation.[A]. IJCAI[C], 2018 : 2149 – 2155.
- [19] Omran P G, Wang Z, Wang K. Learning Rules With Attributes and Relations in Knowledge Graphs.[A]. AAAI Spring Symposium: MAKE[C], 2022.
- [20] Zhang W, Paudel B, Wang L, et al. Iteratively learning embeddings and rules for knowledge graph reasoning[A]. The world wide web conference[C], 2019 : 2366 – 2377.
- [21] 刘藤, 陈恒, 李冠宇. 联合 FOL 规则的知识图谱表示学习方法 [J]. 计算机工程与应用, 2021, 57(04): 100 – 107.
- [22] Yang F, Yang Z, Cohen W W. Differentiable learning of logical rules for knowledge base reasoning[J]. Advances in neural information processing systems, 2017, 30.
- [23] Sadeghian A, Armandpour M, Ding P, et al. Drum: End-to-end differentiable rule mining on knowledge graphs[J]. Advances in Neural Information Processing Systems, 2019, 32.
- [24] Chen J, He H, Wu F, et al. Topology-aware correlations between relations for inductive link prediction in knowledge graphs[A]. Proceedings of the AAAI Conference on Artificial Intelligence : Vol 35[C], 2021 : 6271 – 6278.
- [25] Hamaguchi T, Oiwa H, Shimbo M, et al. Knowledge transfer for out-of-knowledge-base entities: A graph neural network approach[J]. arXiv preprint arXiv:1706.05674, 2017.
- [26] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[J]. arXiv preprint arXiv:1609.02907, 2016.
- [27] Schlichtkrull M, Kipf T N, Bloem P, et al. Modeling relational data with graph convolutional networks[A]. The Semantic Web: 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, Proceedings 15[C], 2018 : 593 – 607.
- [28] Ye R, Li X, Fang Y, et al. A Vectorized Relational Graph Convolutional Network for

- Multi-Relational Network Alignment.[A]. IJCAI[C], 2019 : 4135 – 4141.
- [29] Cai L, Yan B, Mai G, et al. TransGCN: Coupling transformation assumptions with graph convolutional networks for link prediction[A]. Proceedings of the 10th international conference on knowledge capture[C], 2019 : 131 – 138.
- [30] Vashishth S, Sanyal S, Nitin V, et al. Composition-based multi-relational graph convolutional networks[J]. arXiv preprint arXiv:1911.03082, 2019.
- [31] Nathani D, Chauhan J, Sharma C, et al. Learning attention-based embeddings for relation prediction in knowledge graphs[J]. arXiv preprint arXiv:1906.01195, 2019.
- [32] Chen M, Zhang Y, Kou X, et al. r-GAT: Relational Graph Attention Network for Multi-Relational Graphs[J]. arXiv preprint arXiv:2109.05922, 2021.
- [33] Chen M, Zhang W, Zhang W, et al. Meta relational learning for few-shot link prediction in knowledge graphs[J]. arXiv preprint arXiv:1909.01515, 2019.
- [34] Lv X, Gu Y, Han X, et al. Adapting meta knowledge graph information for multi-hop reasoning over few-shot relations[J]. arXiv preprint arXiv:1908.11513, 2019.
- [35] Niu G, Li Y, Tang C, et al. Relational learning with gated and attentive neighbor aggregator for few-shot knowledge graph completion[A]. Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval[C], 2021 : 213 – 222.
- [36] Zheng S, Mai S, Sun Y, et al. Subgraph-aware few-shot inductive link prediction via meta-learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2022.
- [37] Chen M, Zhang W, Zhu Y, et al. Meta-knowledge transfer for inductive knowledge graph embedding[A]. Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval[C], 2022 : 927 – 937.
- [38] Chen M, Tian Y, Chen X, et al. On2vec: Embedding-based relation prediction for ontology population[A]. Proceedings of the 2018 SIAM International Conference on Data Mining[C], 2018 : 315 – 323.
- [39] Guo S, Wang Q, Wang B, et al. SSE: Semantically smooth embedding for knowledge graphs[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 29(4): 884 – 897.
- [40] Xie R, Liu Z, Sun M, et al. Representation learning of knowledge graphs with hierarchical types.[A]. IJCAI : Vol 2016[C], 2016 : 2965 – 2971.
- [41] Liu S, Grau B, Horrocks I, et al. Indigo: Gnn-based inductive knowledge graph completion using pair-wise encoding[J]. Advances in Neural Information Processing Systems, 2021, 34 : 2034 – 2045.
- [42] Zhao M, Jia W, Huang Y. Attention-based aggregation graph networks for knowledge

- graph information transfer[A]. Advances in Knowledge Discovery and Data Mining: 24th Pacific-Asia Conference, PAKDD 2020, Singapore, May 11–14, 2020, Proceedings, Part II 24[C], 2020 : 542–554.
- [43] Stackelberg H v, others. Theory of the market economy[J], 1952.
- [44] Franceschi L, Frasconi P, Salzo S, et al. Bilevel programming for hyperparameter optimization and meta-learning[A]. International Conference on Machine Learning[C], 2018 : 1568–1577.
- [45] Sinha A, Malo P, Deb K. A review on bilevel optimization: From classical to evolutionary approaches and applications[J]. IEEE Transactions on Evolutionary Computation, 2017, 22(2) : 276–295.
- [46] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks[A]. International conference on machine learning[C], 2017 : 1126–1135.
- [47] Ma S, Ding J, Jia W, et al. Transt: Type-based multiple embedding representations for knowledge graph completion[A]. Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2017, Skopje, Macedonia, September 18–22, 2017, Proceedings, Part I 10[C], 2017 : 717–733.
- [48] Hao J, Chen M, Yu W, et al. Universal representation learning of knowledge bases by jointly embedding instances and ontological concepts[A]. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining[C], 2019 : 1709–1719.
- [49] Teru K, Denis E, Hamilton W. Inductive relation prediction by subgraph reasoning[A]. International Conference on Machine Learning[C], 2020 : 9448–9457.
- [50] Mai S, Zheng S, Yang Y, et al. Communicative message passing for inductive relation reasoning[A]. Proceedings of the AAAI Conference on Artificial Intelligence : Vol 35[C], 2021 : 4294–4302.



## 致谢

岁月流转，转眼已经陪伴武汉大学六个春秋。随着毕业论文的完结，我的研究生生涯也即将告一段落。

在度过两年平平无奇的大学生涯后，我很庆幸能够鼓起勇气敲开了彭敏教授办公室的门，向她表达了希望能够通过实验室的实习来提升自己的意愿。更让我高兴的是彭老师原意给我尝试的机会，最终我也如愿留在了实验室从事科研和项目的工作。我从心底里非常感谢我的指导老师彭敏教授，是您一直精益求精的教学精神和耐心细致的指导，让我能够在探索学术领域的路上有所收获。也是在实验室项目的不断磨砺中，让我从一个什么都不太懂的项目小白成功收获了不错的工作 offer，您的教诲和启发，对我而言永久难忘，必将深深烙印在我的心中。

同时，我要感谢我所在的实验室团队。在他们的帮助下，我体会到科研和工作的融洽和乐趣。尤其是刘奔博士在我毕业论文撰写的过程中，不辞辛苦地为我解忧答惑，对我遇到的每一个问题都能够给我指明一个正确的解决方向。也是在刘奔博士的悉心帮助下，我能够对论文的结构进行优化和调整，完成了整个论文的修改工作。除此之外其他的实验室成员也都是我求学路上重要的伙伴，正是在他们的支持下，我能够勇往直前并应对每一个挑战。

最后，我还要感谢我的家人，是他们一直支撑着我继续求学，不断给予我鼓励和支持。我还要感谢我的女友月月，是她在我心情最失落、最困难的时候陪伴着我，让我能够有继续走下去的勇气。

山水一程，愿我们各自未来的旅程都继续迸发出新的精彩！

## 攻硕期间取得的学术成果和参与的项目

- [1] 岩土工程监测管理及预测预警系统
- [2] 武汉城管应急管理系统
- [3] 草莓智慧园区管理系统

## 武汉大学学位论文使用授权协议书

本学位论文作者愿意遵守武汉大学关于保存、使用学位论文的管理办法及规定，即：学校有权保存学位论文的印刷本和电子版，并提供文献检索与阅览服务；学校可以采用影印、缩印、数字化或其它复制手段保存论文；在以教学与科研服务为目的前提下，学校可以在校园网内公布部分或全部内容。

- 1、 在本论文提交当年，同意在校园网内以及中国高等教育文献保障系统（CALIS）、高校学位论文系统提供查询及前十六页浏览服务。
- 2、 在本论文提交 ☐ 当年/ ☐ 一年/ ☐ 两年/ ☐ 三年以后，同意在校园网内允许读者在线浏览并下载全文，学校可以为存在馆际合作关系的兄弟高校用户提供文献传递服务和交换服务。（保密论文解密后遵守此规定）

论文作者（签名）： \_\_\_\_\_

学 号： \_\_\_\_\_

学 院： \_\_\_\_\_

日期：            年    月    日