arXiv:2301.01172v1 [cs.CL] 3 Jan 2023

结构知识与常识知识增强的少样本知识图谱补全综述

HAODI MA，美国佛罗里达大学

知识图谱（KG）是各种自然语言处理应用的关键组成部分。常识知识图谱（CKG）是一种特殊类型的知识图谱，其中实体和关系均由自由文本构成。然而，以往的知识图谱补全和常识知识图谱补全研究面临长尾关系和新加入关系的问题，这些关系没有足够的已知三元组用于训练。为此，研究者提出了少样本知识图谱补全（FKGC），结合图表示学习与少样本学习的优势，以应对标注数据有限这一问题。本文全面综述了此类任务的既有研究，并对一系列方法和应用进行梳理。具体而言，我们首先介绍少样本知识图谱补全面临的挑战、常用知识图谱和常识知识图谱；然后根据知识图谱的类型和方法，对现有研究进行系统分类和总结；最后，我们介绍少样本知识图谱补全模型在不同领域预测任务中的应用，并分享对其未来研究方向的见解。

CCS 概念： • **计算方法** → **知识表示与推理**；**关于信念与知识的推理**。

其他关键词和短语：知识图谱嵌入、链接预测、知识蒸馏、知识图谱

ACM 参考格式：
Haodi Ma。2022年。结构知识与常识知识增强的少样本知识图谱补全综述。1，1（2022年1月） - 月 2022年），27 页 ges. https://doi.org/XXXXXXX.XXXXXXX

1 引言

知识图谱（KGs）由三元组构成，其中每个三元组表示 r 头实体h与尾实体 t之间的关系。现实世界中的知识图谱示例包括 Freebase [5], Yago [58]和 NELL [10]。这些知识图谱包含数百万条事实，是问答、推荐系统和自然语言处理等应用的基础。

尽管如今的大规模知识图谱中存储了海量信息，但它们仍然高度不完整，这使得知识图谱补全（KGC）成为其下游应用面临的一项挑战。近期的研究趋势致力于学习实体和关系的低维表示，以进行缺失链接预测 [（Bordes 等，2013；Trouillon 等，2016；Dettmers 等，2017）]。这些方法的基本思想是基于知识图谱中的已知事实，对实体之间的各种关系模式进行建模和推理。例如，TransE 将关系建模为平移，旨在捕捉逆向和组合模式。RotatE 作为其中的代表方法，可以推理对称、非对称、逆向和组合模式。

然而，此类方法通常需要所有关系都有足够的训练三元组，才能学习嵌入。先前的研究 [78] 表明，知识图谱中的很大一部分关系属于长尾关系。换句话说，它们在知识图谱中只有少量实例

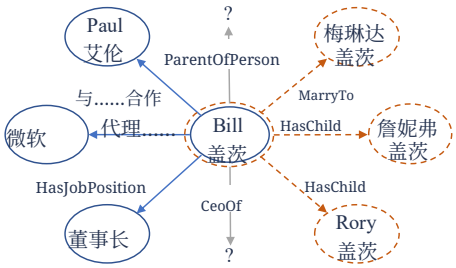
作者地址：Haodi Ma，ma.haodi@ufl.edu，美国佛罗里达州盖恩斯维尔，佛罗里达大学。

允许免费制作本作品全部或部分内容的数字或纸质副本，用于个人或课堂用途，前提是不得为营利或商业利益而制作或分发副本，且副本须附有本声明及首页上的完整引文。对于本作品中由 ACM 以外的其他方拥有版权的部分，必须尊重其版权。注明出处后允许摘录。若要以其他方式复制、重新出版、发布至服务器或向列表重新分发，须事先获得特别许可和/或支付费用。请向 permissions@acm.org 申请许可。
© 2022年 计算机协会。稿件已提交至 ACM

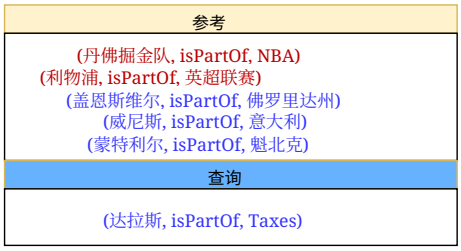
。例如，维基数据中约有 10% 的关系所对应的三元组不超过 10 个。此外，现实世界中的知识图谱通常是动态的，这意味着每当获取新知识时，都会添加新的关系和实体。为应对这些挑战，模型应能够仅凭少量示例预测新的三元组。

为应对上述挑战， [78] 提出了两个用于少样本知识图谱补全（FKGC）的基准数据集 NELL-One 和 Wiki-One，以及一个名为 GMatching 的基线模型。该模型引入局部邻居编码器，仅需针对每个查询关系提供少量样本，即可学习计算开销较大的实体表示。近期研究的一个分支 [54, 86] 采用了类似的方法，并通过利用局部图邻居提升嵌入质量，取得了显著的性能提升。他们进一步指出，与不同任务关系相关联的实体邻居应具有不同的影响。由于关系可能存在多义性，参考三元组对特定查询的贡献也应有所不同。例如，如果任务关系是`isPartOf`，如图1所示，该关系具有不同含义，例如（利物浦，isPartOf，英超联赛）中的组织相关含义，或（盖恩斯维尔，isPartOf，佛罗里达州）中的地点相关含义。显然，对于查询（达拉斯，isPartOf，Taxes）而言，地点相关的参考项应比其他参考项更具影响力。这些模型 [43, 54] 提出使用注意力网络来捕捉实体和参考项的动态属性。

图 1. 示例：(a) 具有多种角色的实体，致谢 [54]；(b) 展示对特定查询具有不同影响的参考项

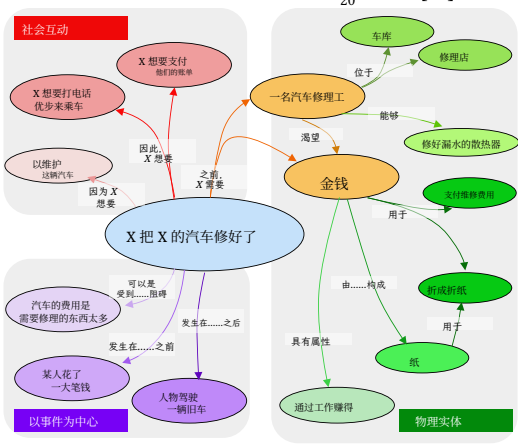


(a)



(b)

图 3. ATOMIC 的一个子集$_{20}^{20}$, 致谢 [28]



另一类少样本知识图谱补全模型 [12, 24, 37, 95] 基于模型无关元学习（MAML） [19]构建。这些模型利用元学习，仅通过少量实例来学习实体和关系表达性嵌入的学习过程。具体而言，它们利用训练集中的高频关系捕获元信息，其中包含不同任务关系之间的共有特征。借助元信息提供的良好参数初始化，这些模型能够快速适应测试任务，而每个任务关系仅提供少量实例。

另一方面，常识知识图谱（CKG）是一类特殊的知识图谱，例如 ATOMIC [50] 和 ConceptNet [57], ，其中实体和关系由自由文本构成，因此较少受到基于嵌入的模型关注。常识知识图谱具有动态性，因为不断有包含未见文本的实体被引入，这使其成为少样本知识图谱补全的天然基准数据集。此外，常识知识图谱中的实体和属性通常是自由文本。如图 3所示，与实体和关系名称具有结构化形式的一般知识图谱不同，常识知识图谱中的实体描述具有丰富的语义含义，隐含的语义关系可用于直接推断常识知识；但这一特点也使常识知识图谱比一般知识图谱更加稀疏，因为指代同一概念的实体可能是不同的节点。如 [67], 所示，ConceptNet 和 ATOMIC 的平均入度分别仅为 FB15K-237 的 1/15 和 1/8。由于常识知识图谱无法很好地套用通过关系连接两个实体的模式，基于嵌入的方法在捕获隐含常识知识方面能力有限。

与此同时，基于 Transformer 的上下文语言模型 [16, 48] 的最新训练进展激发了人们将语言模型（LM）用作知识库的兴趣。例如，近期研究聚焦于使用提示查询语言模型（例如，“披头士乐队成立于 __”）。COMET [8] 是一种基于变换器的知识图谱补全模型，在 ATOMIC [50] 上训练，旨在根据头实体和关系预测未见的尾实体。BertNet [26]更进一步，通过自动改写初始提示，直接从预训练语言模型中提取未见实体的三元组，以完成少样本知识图谱补全/知识图谱补全任务。

最后，本综述介绍少样本知识图谱补全模型在数据科学、视觉提取和医学领域中的典型应用。我们还基于当前模型表现出的不足，探讨少样本知识图谱补全在通用知识图谱和常识知识图谱上的未来研究方向。

2 预备知识

本节首先回顾不同的知识图谱。随后，我们正式定义知识图谱补全和少样本知识图谱补全。本节最后将简要介绍少样本学习和元学习，它们广泛应用于少样本知识图谱补全任务。

2.1 知识图谱

设 E 和 R 分别表示实体集和关系集，知识图谱 $G = \{(e_i, r_k, e_j)\} \subset E \times R \times E$ 是由事实三元组构成的集合，其中 E 表示实体集，R 表示关系集； e_i 和 r_k 分别是第 i 个实体和第 k 个关系。我们通常将 e_i 和 e_j 称为头实体和尾实体。知识图谱也可以表示为 $X \in \{0, 1\}^{|E| \times |R| \times |E|}$，称为 G 的邻接张量。当三元组 (e_i, r_k, e_j) 为真时，(i, j, k) 项 $X_{i,k,j} = 1$；否则 $X_{i,k,j} = 0$。常用知识图谱及其来源、规模和示例如表 1 所示

2.1.1 结构化知识图谱。
如前所述，以往的研究倾向于提取半结构化文本来构建知识图谱 [5, 39, 58]。这些知识图谱通常通过众包构建，或从众包数据源中提取。

FreeBase。 Freebase 是一个众包整理的知识图谱，首次发表于 2008 [5]，并已被用作许多任务（包括知识图谱补全）的标准基准知识图谱。Freebase 最新、最完整的版本共包含约 30 亿个三元组和约 5000 万个实体 [1]。Freebase 的一个广泛使用的子集 FB15K-237 排除了 Freebase 中的逆关系，包含 14541 个实体、237 种关系和 272,155 个训练三元组。Freebase 中包含的关系

[1]https://developers.google.com/freebase/

具有层次结构集合构成了定义明确的实体和关系空间，由此推动了嵌入模型的发展 .

维基数据。 维基数据也是一个众包知识图谱，包含约7800万个数据项、约2.3万个类型和1600个关系 [45]。它创建之初被设计为一种管理维基百科中信息的替代方法。维基数据不仅提供事实信息，还通过存储事实的来源来提供事实的上下文。截至2014年，维基数据支持287种语言 [66]。2014年，谷歌将存储在Freebase中的数据转移到了维基数据 [46]。维基数据中的实体和关系通过属性-值对进行描述；

YAGO。 YAGO是一个由维基百科自动构建的大型知识库。该知识图谱将维基百科中10种不同语言的信息整合为一个整体，从而为知识提供多语言维度。它还为许多事实附加空间和时间信息，因此允许用户按空间和时间查询数据。由于由维基百科构建，YAGO继承了维基百科的层次结构，并且同样对实体和关系使用结构化文本。YAGO存在多个迭代版本，包括YAGO2和YAGO3。YAGO3包含8700万条事实、1080万个实体和7600万个关键词。

2.1.2 常识知识图谱。
常识知识图谱旨在为下游应用组织常识或特定领域知识。尽管现有的常识知识图谱 [50, 57] 也通常通过人工众包构建，但它们使用自由文本表示实体。

ATOMIC。 ATOMIC 数据集[2]由 [50], 发布，包含 87.7 万个元组，涵盖围绕特定事件提示（例如“X 去商店”）的各种常识社会知识。ATOMIC 包含以如果-那么关系组织的日常常识知识实体。该数据集总共包含超过 30 万个实体，实体由文本描述构成，平均包含 4.4 个词。具体而言，ATOMIC 从 9 个维度提炼常识知识，涵盖事件的原因（例如“X 需要开车去那里”）、事件对主体的影响（例如“获取食物”），以及事件对其他直接（或隐含）参与者的影响（例如“其他人将得到食物”）。

ConceptNet。 ConceptNet [57] 是一个多语言知识图谱，通过带标签的边将自然语言中的词语和短语连接起来。其知识收集自多个来源，包括专家创建的资源、众包项目和寓教于乐的游戏。它通过不同语言的词语和短语，表示理解语言所涉及的常识。这些“概念”有助于自然语言应用更好地理解人们使用的词语背后的含义。ConceptNet包含超过1300万个连接这些概念的链接。

Visual Genome. 除了使用自然语言来源外，Visual Genome [31] 还从图像中收集常识知识。它为每张图像收集对象、属性和关系的密集标注，以构建知识。具体而言，Visual Genome 总共包含超过 10 万张图像，每张图像平均包含 21 个对象、18 个属性，以及对象之间的 18 种关系。由于这些对象、属性和关系是从图像中提取的，因此该数据集使用 WordNet [41] 同义词集对其进行分类。本节首先回顾不同的知识图谱，随后正式定义知识图谱补全和少样本知识图谱补全。本节最后简要介绍少样本学习和元学习，它们广泛应用于少样本知识图谱补全任务中。

[2]https://homes.cs.washington.edu/msap/atomic/ 提交给 ACM 的稿件

表 1. 现有知识图谱及示例综述。

	来源	size	示例
Freebase	https://developers.google.com/freebase	4 个关系组，2M 节点，18M 边	/m/070xg，/sports/sports_team/colors，/m/01g5v（西雅图海鹰队）（略带蓝色） /m/06kxk2 ，/people/person/place_出生_地 ，/m/01_d4（卡尔·福尔曼）（美国/芝加哥）
维基数据	https://www.wikidata.org	1.2k 关系，75M 对象，9亿条边	奥地利:Q40 ，组成部分_：P361 ，欧盟：Q458 披头士乐队：Q1299 ，成立地点_：P740 ，利物浦：Q24826
YAGO	https://yago-knowledge.org/	8700万条事实，1080万个实体，ties，以及7600万条关键词	pl/Henryk_Pietras, 出生于，de/Debiensko fr/Chateau_de_Montcony，位于，勃艮第
Concept Net	https://conceptnet.io/	36 种关系，800 万节点，21M 条边	/c/en/go_to_bed，/r/具有前提条件，/c/en/get_ready_for_bed /c/en/章节_儿童_书籍，/r/位于，/c/en/书店 /c/pt/我们将使其麻木/v ，/r/词形 ，/c/pt/使晕眩
ATOMIC	https://allenai.org/data/atomic-2020	9种关系，30万节点，87.7万条边	金钱，用于，支付维修费用 金钱，由.....构成，纸 人物X接受工作，效果，快乐
视觉基因组	https://visualgenome.org/	42k 关系，3.8M 节点，230万条边，280万属性	男性.n.01 ，穿着.v.01 ，背包.n.01 椅子.n.01，具有.v.01，填充物.n.01 果汁瓶.n.01，位于.v.01，书桌.n.01

2.2 少样本知识图谱补全

2.2.1 知识图谱补全。

知识图谱补全（KGC）的目标是预测 G 中有效但未被观测到的三元组。形式化地说，给定一个带有关系 r_k 的头实体 e_i（尾实体 e_j），模型需要找到尾实体 e_j（头实体 e_i），从而在 G 中构成最合理的三元组（e_i, r_k, e_j）。知识图谱补全模型通常定义一个评分函数 f：$\mathrm{E} \times \mathrm{R} \times \mathrm{E} \rightarrow \mathbb{R}$，为每个三元组（$e_i, r_k, e_j$）$\in \mathrm{E} \times \mathrm{R} \times \mathrm{E}$ 分配一个分数 s（e_i, r_k, e_j），该分数表示三元组的合理性。

2.2.2 知识图谱嵌入。

知识图谱嵌入（KGE）旨在将实体和关系投影到一个定义良好的空间中，并以高维向量进行建模。知识嵌入（KGE）模型通常为每个实体 e_i 和关系 r_j 分配嵌入空间中的向量表示 $\mathbf{e}_i$、$\mathbf{r}_j$。随后，它们定义评分函数，以建模实体和关系之间的交互。

知识图谱嵌入模型通常可分为平移模型和双线性模型。平移模型的代表是 TransE [6], ，它将实体间的关系建模为其嵌入之差。这种方法在推理组合、反对称性和反演模式方面效果显著，但无法处理一对多、多对一和多对多关系。RotatE [59] 将关系建模为复数空间中的旋转，因此能够捕捉对称关系，但除此之外也与 TransE 一样受到限制。作为双线性模型的代表，ComplEx [64], 引入了一个由复数组成的对角矩阵来捕捉反对称性。其他模型，如 BoxE [1], 和 HAKE [91], ，能够利用复杂的知识图谱嵌入表示多种关系模式。

2.2.3 图神经网络模型。

近年来，图神经网络（GNN）在知识图谱补全任务上受到了广泛关注 [71, 84, 92]。凭借图神经网络的高表达能力，这些方法展现出了良好的性能。然而，在引入额外计算复杂度的同时，最先进的基于图神经网络的模型与知识图谱嵌入模型相比并未表现出明显优势 [92]。例如，NBFNet [97] 和 RED-GNN [90]在知识图谱补全基准上取得了具有竞争力的性能，但 Bellman-Ford算法需要在整个知识图谱中进行传播，这限制了它们在大规模图上的应用。

2.2.4 少样本学习。

少样本学习（FSL） [74] 旨在从现有任务中学习可迁移的通用先验知识，以应对带标签数据有限的新任务。它通常采用元学习框架 [19] ，将完整任务视为训练样本，使模型能够快速适应新任务。具体而言，给定一组任务 T 及其训练数据，在元训练阶段，模型的目标是学习在 T 中所有任务上都有效的全局参数 θ_r :

$$\theta' = \underset{\theta}{\operatorname{argmax}} \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}(\mathcal{D}_{\mathcal{T}_i}, \theta)$$

其中， p(T) 是任务的分布；D$_{\mathcal{T}_i}$是任务 T 的训练数据；L 是下游任务的损失函数。然后在元测试阶段，将 θ^* 作为初始化参数（先验知识），并在新任务 Tj上进行快速适应：

$$\theta^* = \mathcal{L}(\mathcal{D}_{\mathcal{T}_j}, \theta)$$

其中，Tj 仅拥有有限的带标签数据。以往的 FSL 方法通常可分为：（1）基于度量的方法 [56, 65] ，利用特定任务相似度度量，从支持集数据推广到查询数据；（2）基于优化的方法 [19, 20, 49] ，旨在寻找对任务变化敏感的模型参数，使基学习器能够通过少量梯度更新快速适应新的小样本任务。

2.2.5 少样本知识图谱补全。

根据知识图谱补全和少样本学习的定义，我们现在正式定义少样本知识图谱补全。考虑知识图谱 G $=\{(h,r,t)\} \subset$E $\times$R $\times$E，它是事实三元组的集合，其中 E 和 R 分别表示实体集和关系集。给定关系 $r \in$R 及其支持集 S$_r$ $=\{(h_k,t_k)|(h_k,r,t_k) \in$T $\}$，任务是补全尾实体 $t \in$ E 缺失的三元组 (h,r,t)。换言之，模型需要根据 (h,r) 从候选实体集 C 中预测 t。当 $|$S$_r|=K$ 和 K 非常小时，该任务称为 K 样本知识图谱补全。一个极端情形是当 $k=0$，这意味着没有支持三元组。此类任务也称为归纳式知识图谱补全、零样本知识图谱补全或图外知识图谱补全，模型需要为未见实体预测正确的关系。稿件已提交至 ACM

少样本知识图谱补全模型旨在将真实实体排在错误候选实体之前。在少样本知识图谱补全中，每个训练任务对应一个关系 $r \in \mathcal{R}$ 及其自身的支持/查询实体对，即 $\mathcal{T}_r = \{\mathcal{S}_r、\mathcal{Q}_r\}$。如前所述，$\mathcal{S}$ 包含 K 个样本的支持实体对。$\mathcal{Q}_r = \{(h_m, t_m)/\mathcal{C}h_m, r\}$ 由所有查询及其对应的候选项 $\mathcal{C}h_m$组成，r 这些候选项根据实体类型约束 [78] 进行选择。我们进一步将训练中的所有任务记为元训练集 $\mathcal{T}^{meta-training}$。

在元训练集上训练后，少样本学习模型将通过预测新关系$r' \in \mathrm{R}'$的事实进行测试。测试关系在元训练集中**未见过**，即 $\mathrm{R} \cup \mathrm{R}' = \varnothing$。测试阶段的每个关系也有自己的少样本支持集和查询集：$\mathrm{T}_{r'} = \{\mathrm{S}_{r'}, \mathrm{Q}_{r'}\}$，其定义与元训练阶段类似。我们将测试中的所有任务记为元测试集 $\mathrm{T}^{meta-testing}$。模型还可以访问背景知识图谱 G'，它是 G 的一个子集，包含除 $\mathrm{T}^{meta-training}$ 和 $\mathrm{T}^{meta-testing}$中的关系以外的所有关系。

3 少样本知识图谱补全模型

通常，结合结构知识的少样本知识图谱补全模型会将知识图谱补全模型与少样本学习结合起来，以适用于各种应用。除了知识图谱嵌入模型外，基于图神经网络的方法在少样本知识图谱补全中也展现出具有竞争力的性能，因为每个少样本任务的支持集中仅提供有限的带标签数据。另一方面，利用语义特征的模型会借助提示来结合结构信息与语义信息。

少样本知识图谱补全任务面临三个主要挑战 [30]:

- (1) **如何在少样本设置下学习三元组中最具代表性的信息?** 一般的机器学习算法需要大量数据来训练模型，而少样本场景中只有少量参考项。如何从有限的三元组中学习不同关系的代表性模式，是解决少样本知识图谱补全问题的关键。

- (2) **如何降低对背景知识图谱的过度依赖?** 大多数以往的少样本方法都依赖背景知识图谱，以获取实体邻域的信息或预训练实体嵌入。一些近期模型指出，全面的背景知识图谱并非总是可用，而且将其存储在内存中也会占用大量空间。

- (3) **如何利用负样本提升模型效能?** 最直观的匹配方法通常会比较查询与正样本参考之间的相似度，却忽略查询与负样本参考之间的相似度；而后者有助于提高三元组有效性度量的准确性。

本节根据近期少样本知识图谱补全模型采用少样本学习技术的方式，以及解决上述三个问题的方法，将融合结构知识的模型系统地分为基于度量的方法和基于优化的方法。随后，我们将从基于提示的结构模型讲到利用预训练语言模型的模型。表2列出了具有代表性的少样本知识图谱补全研究及其开源数据集/代码。

3.1 基于度量的方法

现有基于度量的少样本知识图谱补全模型均采用匹配网络 [65] 或平移网络 [6] 的框架。对于基于匹配网络构建的模型，它们首先实现一个基于图神经网络的实体编码器，以生成实体嵌入。然后，在支持集中的实体对上应用聚合模块，计算每种关系的嵌入。最后，模型根据查询三元组与支持三元组的相似度，计算每个查询三元组被接受的概率。TransE [6] 和 ConvE [15] 等知识图谱嵌入模型也被广泛用于实体编码器，作为中间表示，以便进一步结合其他信息进行增强。

方法	学习任务	少样本学习技术	发表 venue	代码/数据链接
			结构化少样本知识图谱补全	
GMatching [78]	关系预测	基于匹配	EMNLP 2018	https://github.com/xwhan/One-shot-Relational-Learning
FSRL [86]	关系预测	基于匹配	AAAI'20	https://github.com/chuxuzhang/AAAI2020_FSRL
FAAN [54]	关系预测	基于匹配	EMNLP'20	https://github.com/JiaweiSheng/FAAN
GEN [3]	关系预测	基于匹配	NeurIPS'20	https://github.com/JinheonBaek/GEN
REFORM [70]	关系预测	基于匹配	CIKM'21	https://github.com/SongW-SW/REFORM
MetaR [12]	关系预测	基于匹配	EMNLP'19	https://github.com/AnselCmy/MetaR
GANA [43]	关系预测	基于匹配	SIGIR'21	https://github.com/ngl567/GANA-FewShotKGC
MetaP [30]	多跳关系预测	基于度量	SIGIR'21	https://github.com/jzystc/metap
Meta-KGR [37]	多跳关系预测	基于优化	EMNLP'19	https://github.com/THU-KEG/MetaKGR
ADK-KG [89]	多跳关系预测	基于优化	SDM'22	https://github.com/ADK-KG/ADK-KG
ZS-GAN [47]	多跳关系预测	基于优化	NeurIPS'21	https://github.com/Panda0406/Zero-shot-knowledge-graph-relational-learning
			常识 FKGC	
ConMask [55]	关系预测	基于文本融合	AAAI'18	https://github.com/bxshi/ConMask
MIA [44]	关系预测	基于文本融合	WWW'21	--
InductiveE [54]	实体预测	基于语言模型	IJCNN'21	https://github.com/BinWang28/InductivE
BERTRL [85]	关系预测	基于语言模型	AAAI'22	https://github.com/zhw12/BERTRL
OntoPrompt [83]	实体预测	基于提示	WWW'22	https://github.com/zjunlp/KnowPrompt
ZS-SKA [12]	关系预测	基于提示	arXiv	--
COMET [8]	关系预测	基于语言模型	AAAI'21	https://github.com/atcbosselut/comet-commonsense
BERTNet [26]	三元组预测	基于语言模型	–	https://github.com/tanyuqian/knowledge-harvest-from-lms

表 2. 提供开源代码/数据的代表性少样本知识图谱补全方法。

遵循这一框架，GMatching [78] 是首个解决单样本知识图谱补全问题的工作。该工作首先提出了一种邻居编码器，利用局部图结构生成更优的实体嵌入。其动机在于，尽管以往知识图谱嵌入模型的实体嵌入能够编码关系信息，但先前的研究 [77]表明，显式建模路径等结构模式仍有助于关系预测。GMatching 中的邻居编码器仅对每个给定实体的一跳邻居进行编码，即对一组（关系，实体）元组进行编码，以确保其适用于大规模知识图谱。具体而言，GMatching 首先为一跳邻居集合中的每个元组使用预训练知识图谱嵌入，然后应用前馈层对每个元组中关系与实体之间的交互进行编码。随后，邻居编码器分别应用于支持实体对和查询实体对，以生成各自的表示。接着，模型利用基于 LSTM 的循环处理模块 [65] 在参考实体对与每个查询实体对之间执行多步匹配。最后，模型利用匹配分数对每个查询的候选集合中的所有实体进行排序。除了提出首个少样本知识图谱补全任务基线模型外，该工作还提出了两个广泛使用的基准数据集：NELL-One 和 Wiki-One [78]。二者均按照少样本知识图谱补全任务的设置构建。更多统计数据和详情见表 3 和第 4.2节。

秉持相同思路，FSRL [86] 将 GMatching 扩展到少样本设置。它进一步提出了一种关系感知异构邻居编码器，基于异构图结构和注意力机制增强实体嵌入，使模型能够对不同邻居对任务关系的不同影响进行编码。这里的主要观点是，不同邻居对任务关系的影响应有所不同，而 GMatching 等模型忽略了这一点。例如，以 ParentOfPerson 作为任务关系时，与 (CeoOf, Microsoft) 相比，邻居 (MarryTo, Melinda Gates) 应具有更高权重。为解决这一问题，FSRL 引入了一个注意力模块，在对所有邻居进行编码时分配不同的注意力权重，从而生成实体嵌入。

通过应用注意力邻居编码器，FSRL 获取支持集中的每个实体对的表示。随后，它采用基于 RNN 的聚合器，对每个任务中的支持实体对之间的交互进行建模，以生成能够表征整个支持集的信息丰富的表示。受使用循环神经网络聚合节点嵌入的启发 [25], FSRL 对所有实体对应用循环自编码器聚合器。为了构建参考集的嵌入，它聚合编码器的所有隐藏状态，并通过添加残差连接 [27]和注意力权重对其进行扩展。

借助参考集的聚合表示，FSRL 使用匹配网络来发现参考集中相似的实体对。它不再将参考集中的每个实体对与查询实体对逐一比较，而是使用一种具有长短期记忆网络单元的类似循环匹配处理器，直接计算参考集与查询实体对之间的相似度，以进行最终的答案排序。在训练阶段，模型每次采样一个任务关系，并针对该任务优化模型。模型会采样少样本实体对作为支持集，并采样一批查询实体对。通过污染查询实体对中的尾实体来构造负训练集。在梯度下降步骤中，模型利用元学习进行参数优化，使 FSRL 能够很好地迁移到测试少样本关系上。

尽管 FSRL [86] 提出根据邻居与中心实体的相关性区别对待邻居，但它在所有任务关系中仍为所有邻居分配固定权重。这种方法会导致实体在不同任务中的嵌入保持静态，从而损害系统的有效性。FAAN [54], 更进一步，认为实体邻居在不同任务关系下应产生不同影响。例如，SteveJobs 与 HasJobPosition 和 HasChild 这两个任务关系相关。直观而言，如果任务关系是 CeoOf，模型就应更加关注实体 SteveJobs 的工作职位角色，而不是其家庭角色。

此外，任务关系在不同上下文中可能具有不同含义。例如，如果任务关系是`isPartOf`，如图 1所示，该关系可能有不同含义，例如 （利物浦，isPartOf，英超联赛）中的组织相关含义，或 （盖恩斯维尔，is PartOf，佛罗里达州）中的位置相关含义。显然，对于查询（达拉斯，isPartOf，Taxes），位置相关的参考项应比其他参考项更具影响力。因此，参考（支持）三元组也应对不同查询产生不同的贡献。

为应对上述挑战，FAAN 提出了一种自适应注意力邻居编码器，利用一跳实体邻居对实体嵌入进行建模。他们还采用 TransE [6] 将任务关系嵌入建模为头实体嵌入与尾实体嵌入之间的平移，即 $\mathbf{r} \approx \mathbf{h} - \mathbf{t}$。随后，为了进一步对参考实体的不同角色进行建模，FAAN 根据实体邻居关系与任务关系的相关性训练注意力度量，从而为参考集中的每个实体进一步获得角色感知邻居嵌入。该编码器能够根据不同任务关系动态调整注意力分数。自适应机制有助于根据邻居的不同贡献捕捉实体的多样化角色。每个实体的最终表示同时编码了预训练嵌入及其角色感知邻居嵌入。

借助编码器提供的增强实体表示，FAAN 进一步对支持集和查询集中的三元组应用一系列 Transformer 块，以捕获任务关系的各种含义。它借鉴了 [69] 中学习动态知识图谱嵌入的思路。对于每个元素，它将元素嵌入和位置嵌入输入若干 Transformer 块，以获得有意义的实体对嵌入。

随后，FAAN 不再使用静态表示来预测不同查询，而是通过聚合支持集中所有参考项及其对任务关系的注意力得分，获得支持集的一般自适应表示。FAAN 也采用与 FSRL 相同的元训练方式，即在元训练集中的不同任务关系上训练模型，以生成一组能够在所有任务中表现良好、并能快速适应

测试集中的小样本任务。通过上述方法，FAAN 能够捕捉实体和参考表示的细粒度含义，从而提升其质量。与 FSRL 的匹配得分相近，FAAN 在少样本知识图谱补全任务上的表现优于以往模型。

另一方面，HARV 专注于捕捉邻居关系与实体之间的差异，以及关系之间此前被忽略的交互。它通过分离头实体与关系之间（关系级）以及关系与尾实体之间（实体级）的信息，引入了一种用于中心实体表示的分层邻居聚合器。关系级注意力权重基于头实体和关系嵌入计算得到。通过结合这种注意力，聚合头实体的邻居关系，从而生成关系级嵌入h。随后，将关系级嵌入与每个尾实体进行拼接，用于生成实体级注意力权重。两级权重最终生成三元组级权重，并据此计算增强实体表示。关系编码器将关系之间的交互纳入考虑。该编码器是 FSRL 中 LSTM聚合器 结合双向长短期记忆网络（Bi-LSTM）的扩展，用于更新所有支持实体对的表示。支持实体对的嵌入与双向长短期记忆网络编码器输出的嵌入进行拼接，作为每个实体对的最终表示；而支持集则通过对所有支持实体对进行基于注意力的聚合来表示。

此外，GEN [3] 研究了图外少样本知识图谱补全场景中的关系预测，涉及未见实体之间以及已见实体与未见实体之间的关系。GEN 通过元学习，将知识从已见实体推广至未见实体，并将知识从拥有大量链接的实体迁移至链接较少的实体。GEN 还针对传导式推理开发了随机嵌入层，用于建模未见实体之间链接预测的不确定性。GEN 与任何图神经网络兼容。具体而言，在元训练阶段，采用两个 GEN 分别进行归纳式和传导式链接预测。第一个 GEN 是归纳式 GEN。它学习对未见实体进行编码，即使这些实体尚未被观测到，并预测已见实体与未见实体之间的链接。第二个则是传导式 GEN，它进一步学习预测未见实体彼此之间的链接。为了实现传导式推理，GEN 中的元学习框架能够在元训练期间模拟未见实体，而传统学习方案无法观测到这些实体。此外，由于未见实体的链接预测本质上并不可靠，而在每个实体仅有少量三元组的少样本场景中，这种情况会更加严重，因此 GEN 学习未见实体表示的分布，通过随机嵌入来应对这种不确定性。此外，我们采用迁移学习策略来建模长尾分布。这些方法使 GEN 能够表示与已见实体良好对齐的未见实体。如前所述，由于缺少支持三元组，未见实体的表示存在不确定性，因此朴素的 GEN 可能会受到图外少样本链接预测固有的不可靠性的影响。为解决这一问题，随机层通过学习某个实体嵌入上的分布来嵌入未见实体。GEN 还利用蒙特卡洛Dropout，对传导式 GEN 输出嵌入中的不确定性来源进行建模。

最近，REFORM [70] 提出了一种误差感知模块，用于控制影响少样本知识图谱补全的错误所带来的负面影响。它与原始的少样本知识图谱补全略有不同：它根据少样本关系类别，预测查询实体对缺失的关系类别。由于大多数现实世界中的知识图谱都是自动构建的，许多错误未经人工验证便被纳入知识图谱。这些错误会显著降低以往方法在少样本知识图谱补全上的性能，尤其是在可依赖的支持三元组数量很少时。REFORM 的邻居编码器通过注意力机制，专注于选择最可靠的邻居，以增强实体表示。注意力权重矩阵使用预训练嵌入（在 REFORM 中为 TransE）进行训练，以确保正确的邻居获得较高权重。随后，利用 softmax函数对该矩阵进行归一化，从而为每个实体获取鲁棒的嵌入。参考实体对 稿件已提交至 ACM

通过拼接其头实体和尾实体的嵌入来表示。随后，为了生成支持集中的关系的鲁棒嵌入，REFORM 包含一个基于 Transformer编码器的跨关系聚合模块，用于捕捉关系相关性和支持实例。该 Transformer编码器基于多头注意力机制，使每个输入嵌入都参与对所有其他输入嵌入的编码。然后，在误差缓解模块中，REFORM 利用图卷积网络（GCN）为每个查询任务生成各类关系的置信权重。置信权重可视为注意力权重，用于限制错误的影响。具体而言，REFORM 构建一个面向查询的图，以衡量不同支持实例对特定查询关系的影响。GCN 通过最小化查询关系被划分到错误类别的损失进行训练。

采用平移网络的一个代表性方法是 MetaR。其核心思想并非编码邻居信息，而是将单个任务中共有的共享信息从参考实例迁移至查询三元组。MetaR 将这类信息称为关系元信息。关系元信息学习器根据支持集中的头实体和尾实体嵌入，生成实体对的表示。给定支持集中的头实体和尾实体对后，学习器首先通过一个以 LeakyReLU 作为激活函数的全连接神经网络，提取特定于实体对的关系元信息。 [38] 任务的最终关系元信息是当前支持集中所有特定于实体对的关系元信息的平均值。

MetaR 还利用元学习来加速学习过程，这一方法称为梯度元信息。如第2.2.5节所述，模型应能够快速更新一个新的少样本任务。继承 TransE 的思想， [6], MetaR 使用类似的评分函数 $||\mathbf{h}_i + \mathrm{R}\mathcal{T}_r - \mathbf{t}_i||$，结合关系元信息计算每个实体对的分数。然后，通过使用所有正负三元组的分数，最小化支持集上的损失，参数梯度便可指示参数应如何更新。遵循这一梯度更新规则，MetaR 能够快速更新关系元信息，并使用更新后的关系元信息和相同的评分函数对查询集进行评分。模型通过最小化一个批次中所有任务的查询损失之和进行训练。与依赖背景知识图谱的 GMatching [78],相比，我们的 MetaR 不依赖背景知识图谱，因此更加稳健，因为在真实场景中，少样本链接预测可能无法获取背景知识图谱。

GANA [43], 更进一步，通过注意力机制和LSTM聚合器改进嵌入与关系的元计算，从而扩展了MetaR。其动机在于，当邻居稀疏，甚至没有合适的邻居可用于表示少样本关系时，噪声邻居信息可能会损害模型。GANA提出了一种全局-局部框架。在全局阶段，构建了一个门控注意力邻居聚合器，以准确整合少样本关系邻域的语义，即使知识图谱中的邻域极其稀疏，也有助于过滤噪声邻居。将与少样本关系相关联的头实体和尾实体及其邻域结合起来，以消除由邻域稀疏导致的噪声邻居信息。门控机制可以确定邻域表示在表示少样本关系时的重要性。具体而言，在全局阶段开发了一种基于图注意力网络（GAT）的邻居编码器，以捕捉邻居的不同影响，从而提升实体嵌入的质量。该编码器基于可训练的线性变换矩阵，为每个邻居生成注意力权重。GANA采用带线性变换的门控值，消除由邻域稀疏造成的噪声邻居，并自动确定实体邻居对少样本任务关系的影响。随后，通过结合实体嵌入及其邻居表示来表示实体。支持集的最终三元组邻居表示是头实体表示与尾实体表示的拼接。对支持集进行编码后，GANA采用注意力双向LSTM编码器，整合支持集中查询关系的多个邻域表示。查询关系表示是双向长短期记忆网络最终隐藏状态的加权

和，该表示通过结合支持集中的所有邻居嵌入得到。对于局部阶段，设计了一种基于元学习的TransH（MTransH）方法，用于建模复杂关系，并以少样本学习的方式训练模型。采用TransH的原因是它能够建模复杂关系。采用与MAML方法类似的损失函数，以学习元训练集中所有少样本（查询）关系的良好初始化参数。

另一种类似的少样本知识图谱补全模型 HiRe [76], 可视为 GANA 的扩展。它提出联合捕获三个层面的关系信息：实体级、三元组级和上下文级。该模型采用对比学习，将三元组中头实体和尾实体的邻居并集编码在一起，以获取更广泛的上下文。HiRe 提出了一种目标三元组的上下文编码器，基于自注意力机制学习其真实和虚假上下文的嵌入，从而为上下文中的重要邻居分配更高权重。此外，模型采用对比损失，使三元组靠近其实际上下文，并远离虚假上下文。随后，在三元组级关系学习阶段，HiRe 不再使用长短期记忆网络，而是开发了一种基于Transformer的元关系学习器，用于捕获参考三元组之间的交互，并生成目标关系的元关系表示。最后，HiRe 采用基于 TransD 的 [29] 元评分函数，以捕获实体和关系的多样性。它也采用了与 GANA 类似的基于 MAML 的训练策略。借助这三个层面的关系信息，与当前最先进的模型相比，HiRe 在 NELL-One 和 Wiki-One 上表现更佳。消融研究进一步证明，这三个层面的关系信息对 HiRe 的性能都至关重要，未来的模型可以进一步加以利用。

Meta-iKG 是该研究方向上另一项近期工作，提出利用局部子图传递子图特有的信息，并通过元学习中的元梯度快速学习可迁移的模式。近期，图神经网络被引入归纳式关系推理，以捕获目标三元组周围的多跳信息。例如，GraIL [63] 提出了一种基于子图的关系推理框架，用于处理未见实体。CoMPILE [40]通过引入节点-边通信式消息传递机制对这一思想进行了扩展，以建模有向子图。Meta-iKG 可视为 CoMPILE 方法在少样本知识图谱补全上的扩展。Meta-iKG 不再局限于传导式设置，也不再无法处理未见实体，而是面向少样本归纳式知识图谱补全任务，其中包括测试集中的新实体。该模型根据关系实例数量是否达到阈值 K，将关系划分为小样本关系和大样本关系；随后，使用大样本关系进行元训练，以寻找良好初始化的参数，并遵循 MAML 框架，在包含小样本关系的三元组上对模型进行适配。Meta-iKG 继承了 MetaR 的结构，在关系特定学习阶段，首先提取目标实体与尾实体之间的直接包围子图。随后，应用归纳式节点标记函数来识别子图中实体的不同角色。节点嵌入通过其与目标实体之间的距离进行初始化，以表示子图中每个节点的相对位置。接着，Meta-iKG 沿用 CoMPILE 的思路，使用通信式消息传递神经网络为每个子图评分，将其对目标三元组的可信度编码为任务损失。常规的元学习步骤能够确保模型在小样本关系上的性能。然而，由于任务关系的查询集只会更新最终参数，这些步骤可能会使更新后的参数产生偏差。为了确保 Meta-iKG 在大样本关系上也能取得良好性能，该方法引入了大样本关系更新流程，使用支持集以较低的学习率进一步更新最终参数。这一操作使 Meta-iKG 能够在整个归纳式数据集上实现良好的泛化。

为了解决知识图谱依赖问题，并在训练阶段进一步利用负样本，本文提出了一种元模式学习框架 MetaP [30], 。数据中的模式是用于分类数据的代表性规律。知识图谱中的三元组也遵循关系特定模式，可用于衡量三元组的有效性。关系模式指头实体、关系和尾实体的特征共现规律。MetaP 设计了一种基于卷积滤波器的模式学习器，可直接提取三元组的模式。它能够从有限的参考项中学习关系特定模式的潜在表示，稿件已提交至 ACM

因此不依赖背景知识图谱。此外，通过利用负样本参考，MetaP 能够更准确地衡量查询三元组的有效性。本文提出了一种带有效性平衡机制（VBM）的模式匹配器，用于预测查询三元组的模式为正或负的概率。

3.2 基于优化的方法

基于优化的少样本知识图谱补全模型依赖 MAML 进行模型优化。换言之，这类模型通过使用 MAML 优化图神经网络，解决少样本关系预测问题。

由于 GMatching 和 FSRL 等方法只关注事实预测，且仅利用一跳邻居，因此会遗漏知识图谱中包含的更多结构信息，导致结果缺乏可解释性。为此，研究者提出了多跳关系推理，利用多跳推理路径推断事实。例如：(披头士乐队, 位于, 利物浦) ∧(利物浦, 属于, 英国) → (披头士乐队, 总部位于, 英国)。近期模型 [14, 34]提出了多跳推理方法，利用知识图谱中关系的符号组合信息，实现可解释的推理结果。这些研究还将多跳推理视为一个序列决策过程，并利用强化学习解决此类任务。随后，有研究提出了一种基于元学习的多跳推理算法 Meta-KGR，其理念与上述方法相似，旨在实现可解释且有效的少样本关系推理。具体而言，Meta-KGR 引入了一个强化学习框架来建模多跳推理过程，并使用循环神经网络对搜索路径进行编码。随后，该方法采用 MAML，从高频关系中学习有效的元参数，从而能够快速适应少样本关系。

具体而言，将从 [34], 多跳推理中推导出的同策略强化学习（RL）建模为一个马尔可夫决策过程（MDP）：给定查询关系r_q，模型从源实体 e_s出发，依次遍历若干关系和实体，直到到达目标实体 e_o。MDP 模块包括：（1）状态，即当前步骤所处的实体。所有状态都共享源实体和任务关系作为全局上下文。（2）动作：状态 s_t 下的动作空间包括当前实体的所有出边关系—实体元组。（3）奖励：如果模型到达正确的目标实体，就会获得值为 1 的终止奖励。否则，将根据预训练知识图谱嵌入计算目标实体与当前实体之间的相似度，并据此给予奖励。随后，策略网络用于确定每个状态下采取的动作。接着，Meta-KGR 会结合在背景知识图谱上的搜索历史来应用策略网络。当前步骤之前的搜索历史由长短期记忆网络（LSTM）编码。动作空间通过堆叠当前步骤动作空间中的所有动作嵌入来表示。策略网络经过训练，以最大化所有三元组上的期望奖励。在元学习中，Meta-KGR 使用类似 MAML 的元策略网络，使其能够利用在这一阶段学到的良好初始化参数，针对每个查询关系快速适配关系感知策略网络。

FIRE [87] 通过异构邻居聚合器和搜索空间剪枝策略扩展了 Meta-KGR。具体而言，FIRE 利用同策略强化学习对多跳推理路径进行建模，并使用多跳异构结构信息对实体嵌入进行编码。随后，它利用知识图谱嵌入剪枝推理搜索空间，以提高推理效率。在优化过程中也应用了元学习，使学习到的参数能够针对少样本任务关系快速适配。

具体而言，由于 Meta-KGR 中原始的 RL 模块并未将异构结构信息编码到实体嵌入中，FIRE 保留了 FSRL 中的结构编码模块，以增强实体嵌入。需要注意的是，知识图谱中的一些实体拥有大量邻居，导致特定步骤中的动作空间冗余。与基于中心性分数剪除边的 [14, 34] 不同，FIRE 将状态之间的结构相关性视为指导动作搜索的重要

特征，并采用知识感知搜索空间剪枝策略。模型根据步骤 t 的实体与步骤 $t+1$ 的可能候选项之间的结构相关性，使用 TransE 等预训练的知识图谱嵌入，仅保留每一步中相关性最高的前m 个实体。随后，模型利用元学习进行快速适应，学习经过良好初始化的参数，使其能够快速泛化到少样本关系。

最近，ADK-KG [89]通过开发一种文本增强异构图神经网络来编码节点嵌入，进一步改进了 FIRE。该网络使用内容信息对实体和关系嵌入进行预训练，并通过任务权重增强 MAML。这是首项利用预训练语言模型在少样本知识图谱补全任务中捕获内容信息的工作。其强化学习模块与 FIRE 和 Meta-KGR 中的模块类似，能够生成每个实体的编码。问题在于，强化学习仅对推理过程进行编码，却忽略了知识图谱中的内容信息和结构信息。因此，ADK-KG 开发了一种文本增强异构图神经网络，以丰富实体嵌入。首先，对于知识图谱中的每个实体和关系，ADK-KG 提取其文本信息作为内容特征。随后，它合并实体和关系的所有文本特征，并将其输入预训练 BERT 语言模型 [16]，以获得相应的内容特征向量。对于枚举型内容（例如维基数据中的实体和关系 ID），ADK-KG 使用独热编码将其转换为二进制特征向量。之后，利用神经网络对每种关系类型下所选邻居的内容嵌入进行编码和聚合。所选邻居包括一阶邻居及其关系，还包括通过随机游走获得的、与一阶邻居共享相同关系的高阶邻居。最后，由于不同关系类型的邻居对最终实体表示的贡献不同，该模型采用注意力机制，利用基于关系类型的邻域嵌入生成每个实体的最终嵌入。随后，这些嵌入被用于基于强化学习的推理，以替代预训练实体嵌入。在元学习步骤中，由于知识图谱中的关系通常具有不同的含义，AKD-KG 使用自注意力机制为它们分配不同的权重。

另一项近期拓展 Meta-KGR 和 FIRE 的工作是 THML [94]。THML 认为，基于强化学习的模型的泛化能力通常受限于其在困难关系（训练损失较高的关系）上的推理表现较差。THML 通过在每个训练批次中识别困难关系，并在关系层面和关系簇层面上，使用由此有效生成的新难度感知训练批次进一步训练模型，从而解决少样本知识图谱补全中的这一问题。与 Meta-KGR 和 FIRE 一样，THML 也在难度感知元强化学习模块中将推理过程建模为马尔可夫决策过程。主要区别在于，为了解决错误推理路径导致的稀疏奖励问题以及效率问题，THML 将奖励分为三部分。终止奖励与 FIRE 相同，但 THML 使用 ConvE [15] 生成预训练嵌入。路径奖励对推理链的长度进行编码，鼓励模型通过尽可能短的关系链找到目标实体，因为较短的路径通常比更长的路径能提供更可靠的推理 [14]。如果路径长度超过 3，则可能会被舍弃。知识图谱推理的另一个问题是，模型在训练阶段倾向于推断语义含义相似的路径，这可能会使模型陷入局部最优路径。因此，THML 提出使用多样性奖励鼓励模型寻找不同的路径。此外，在训练阶段，THML 不再随机采样三元组查询，而是采用两级难度感知采样策略。关系层面的难度感知采样会在线对一个批次中所有关系的推理准确率进行排序，以便为下一个批次选择困难关系。在关系簇层面的难度感知采样中，THML 获取所有关系的预训练 TransE 嵌入，然后执行 k均值算法来形成关系簇。接着，它将当前批次中包含推理准确率最低的困难关系的簇选为困难簇，并将其加入下一个批次。

除了上述讨论的方法外，少样本知识图谱补全领域还有其他类型的基于优化的解决方案。例如，ZSGAN [47] 通过生成对抗网络建立文本与知识图谱之间的联系，研究零样本知识图谱补全。其动机在于，新类别的语义特征可以通过其稿件提交给 ACM

文本描述反映出来。此外，文本描述包含丰富且明确的信息，这对大规模识别任务至关重要。ZAGAN 的核心是设计一种条件生成模型，从原始文本描述中学习合格的关系嵌入。

ZSGAN 利用特征编码器对真实数据进行表示，从知识图谱嵌入中生成合理的真实数据分布。特征编码器预先使用训练集进行训练，并在对抗训练过程中保持固定。邻居编码器仅考虑每个实体的一跳邻居，并使用图卷积网络 [51] 生成每个实体基于结构的表示。然后，使用前馈层作为实体编码器，从每个头实体与尾实体对中提取信息。最后，将头实体和尾实体的邻居嵌入与实体对嵌入拼接，得到关系事实表示。

给定文本表示后，生成器旨在生成合理的关系嵌入，以捕获知识图谱特征空间中相应的关系语义信息。基于此，查询关系的预测被转化为监督分类任务。具体而言，文本嵌入是移除停用词和标点后，对词嵌入按词频－逆文档频率加权并求和得到的向量。文本嵌入被输入生成对抗模型（生成对抗网络，GAN），以生成关系嵌入。相反，判别器旨在将伪数据与真实数据分布区分开来，同时识别关系类型。首先，输入特征经带有 LeakyReLU [38] 的全连接（FC）层进行转换。随后分为两个网络分支。第一个分支是一个 FC 层，充当二元分类器，用于区分真实数据与伪数据。另一个分支负责分类。在训练过程中，为了稳定训练并消除模式崩溃，ZSGAN 还采用梯度惩罚；当梯度范数偏离目标值 1 时，该惩罚会对模型施加约束。类似地，RAN [88] 提出了一个通用的基于特征生成的框架，用于处理具有未见关系的少样本知识图谱补全，以及从文本中进行少样本关系抽取。

不同于以往采用实体对匹配的工作，P-INT [79] 利用从头实体到尾实体的路径来表示实体对，并计算路径之间的交互，以解决少样本知识图谱补全问题。其动机仍然是纳入更多结构信息和表达性信息，并在少样本推理中排除噪声邻居。为了提取每个支持实体对的支持子图，P-INT 采用双向广度优先搜索算法 [77] 来剪枝搜索空间。左右路径的相交邻居用于生成从头实体到尾实体、长度各异的路径。随后，这些路径中的关系被合并为一组支持关系。模型使用预训练的 TransE 嵌入，计算知识图谱中每对关系之间的相似度。随后，为了推理查询子图，P-INT 通过固定步数扩展数量有限的邻居，并对每个邻居计算其与支持关系的最大相似度，返回排名前L 的邻居。经过 T 跳后，模型扩展查询子图，最多包含 $T \times L$ 个实体。在推理的同时，P-INT 可以追踪从查询的头实体到子图中每个扩展实体的所有路径。在匹配组件中，P-INT 计算任意两条路径之间的相似度。随后，使用 [77] 中的径向基函数聚合函数从相似度矩阵中提取相似度特征，作为路径之间的交互。受 FAAN [54], 启发，P-INT 根据路径与查询关系的相关性，为不同路径计算关系感知注意力，以建模它们对匹配结果的影响。

3.3 基于本体的方法

除了上述提到的少样本学习方法外，随着预训练/转换微调模型已成为自然语言处理的默认范式，从大语言模型中提取知识也日益受到关注。因此，如何在结构化关系知识与自然语言知识之间进行有效迁移，已成为一项具有挑战性的任务。近期的研究尝试融入本体等结构知识，以

增强语言理解。该方向的代表性方法之一是 OntoZSL [21]，它提出了一种新颖的零样本学习框架，不仅利用本体模式增强类别语义，还采用与 ZSGAN 类似的基于本体的生成对抗网络（GAN），为未见类别合成训练样本。OntoZSL 首先设计了一个本体编码器，通过考虑概念间的结构关系及其在文本描述中的相关性，从本体模式中学习关系表示。它使用预训练的 TransE 为本体模式中的所有概念生成默认嵌入。随后，模型通过将结构表示和文本表示投影到同一嵌入空间，并使用与 TransE 相同的评分函数同时学习二者，构建一个文本感知语义嵌入模型。利用本体编码器学得的两种表示，OntoZSL 接着采用生成对抗网络 [47] 来学习并训练真实的关系嵌入，这些嵌入来自包含任务关系所有一跳邻居三元组的包。包中所有实体对的嵌入共同构成任务关系的真实嵌入，其中包含该任务关系的语义和结构特征。最后，OntoZSL 使用训练好的生成器，根据文本描述为每个未见关系生成合理的关系嵌入。在预测时，模型计算关系嵌入与候选项的头实体、尾实体对之间的相似度。

3.4 利用常识知识进行少样本知识图谱补全

以往模型主要关注知识图谱中的结构信息，如一跳邻居和路径，却相对忽视语义信息。ADK-KG 等模型 [89] 考虑了内容信息，但仍依赖邻居来生成实体和关系的嵌入。如前所述，近年来对 BERT [16]和 GPT [9],等预训练语言模型的广泛研究，促使一些方法通过微调这些模型来利用文本信息进行少样本/零样本知识图谱补全。本节将介绍这一类别中的若干代表性方法，探讨如何有效利用知识图谱中的常识/文本信息。

ConMask [55] 是最早提出的、旨在解决零样本知识图谱补全问题的模型之一，它通过实体名称和文本描述对未见实体进行编码。一般而言，该模型将三元组中实体和关系的文本嵌入输入一个模型，该模型由基于注意力的关系依赖文本掩码模块和基于卷积神经网络的目标融合模块组成。为捕捉与任务关系相关的文本信息，ConMask 使用关系依赖内容掩码模块来减少给定描述中的噪声。该组件首先对输入描述进行预处理，根据任务关系筛选出相关的短片段。ConMask 利用注意力机制屏蔽与任务无关的内容，并为词语分配依赖于关系的相似度分数。这里的一个常见问题是，得分最高的词语并非目标实体，而是与任务关系语义含义相近的词语。由于实际目标词总是出现在这些指示词附近，ConMask 会根据每个词语的上下文调整其相似度分数。随后，模型通过目标融合提取基于关系的实体嵌入。该模型为此设计了三个不含反卷积操作的全卷积神经网络（FCN）层。由于直接通过目标融合模块生成实体嵌入可能代价高昂，ConMask 采用语义平均函数聚合词嵌入，以表示实体名称，并为每个实体生成其他文本特征的表示。

尽管 ConMask 成功学习了实体名称及其部分文本描述的嵌入，以将未见实体连接到知识图谱，但它并未充分利用文本描述中丰富的特征信息。此外，ConMask 提出的关系依赖的内容掩码方法可能很快就无法找到目标词。为应对这些问题，本文提出了多重交互注意力（MIA）模型 [44]，通过获取头实体描述、头实体名称、关系名称与候选尾实体描述之间的交互，生成更具表达力的表示。此外，MIA 同样利用头实体描述中的额外文本特征来增强头实体表示，并将注意力机制应用于候选尾实体，以增强其表示。稿件已提交至 ACM

实体描述。具体而言，MIA 首先将头实体描述、问题和候选尾实体描述中的每个词转换为若干连续表示，包括 GloVe、词性、命名实体识别和 BERT 嵌入，并将它们拼接起来，形成每个词的输入表示。接着，为了通过与所有相关文本信息的交互来增强实体表示，鉴于同一描述中的词语重要性各不相同，且相关描述通常会相互提及，MIA 在每种交互中都采用相同的词级序列对齐注意力机制。模型的第三个组成部分是一个循环神经网络层，它使用双向长短期记忆网络对文本上下文进行编码。模型还在多个候选尾实体描述之间应用相同的注意力机制，以增强其表示。MIA 还探索了不同的评分函数，以提升模型的收敛速度；与其他最先进的模型相比，该模型取得了显著改进。不过，这种方法的问题在于它严重依赖实体描述，只有在必要信息可用时才能发挥作用。

InductiveE [67] 提出了一种常识知识图谱链接预测方法，通过利用文本实体描述来处理未见实体。该方法直接根据实体描述构建表示，而非利用文本实体表示作为训练初始化，从而实现归纳学习。具体而言，它首先使用 fastText 词嵌入模型，将实体的文本嵌入拼接起来表示实体 [4] ，并结合预训练 BERT [*CLS*] 最后一层中 token 的表示 [16]。为进一步结合实体邻居信息，增强实体的语义表示，InductiveE 为未见实体添加相似性链接，以初始化其邻居信息。随后，它将稠密化图中的实体表示输入由门控关系图卷积网络编码器和简化版 ConvE [15] 解码器组成的模型，以预测每个三元组的得分。门控编码器用于确保能够自适应地控制从邻接连接融合到中心节点的信息量。对于解码器，Conv-TransE [53] 已被证明在知识图谱补全中能够有效且高效地为三元组评分。InducctiveE 的解码器进一步对其进行了改进，在卷积之前添加洗牌操作，以促进嵌入之间的交互，并提升卷积模型的表达能力。

InductiveE 仅探索了针对未见实体的归纳学习，而针对未见/新关系的归纳学习对于现实世界中的常识少样本知识图谱补全任务也很有价值。为应对这一挑战，开发了一种知识图谱补全模型 KG-BERT [82], 。该模型将三元组中的头实体、关系和尾实体转换为文本序列，然后将三元组预测作为下游文本分类任务，并使用给定的训练三元组对 BERT 进行微调。对于具有名称信息的未见实体和关系，可以将与其相关的候选三元组直接转换为文本序列并进行预测。

沿用这一思路，BERTRL [85] 也提出将三元组预测视为 BERT 的下游文本分类任务，并利用实体和关系的文本信息。然而，它使用单个三元组以及连接两个实体的可能路径来微调 BERT，并显式地进行推理。给定一个查询三元组（$h, ?\ t$），为了利用头实体和尾实体的邻域知识并出于效率考虑选择合适的邻居，BERTRL 完全依赖 BERT 对此类信息进行编码。为了将知识图谱中的结构知识形式化，使其适用于 BERT 模型，BERTRL 收集查询三元组中头实体与尾实体之间所有长度为 k 的推理路径。随后，它将每条路径分别作为 BERT 的输入。每条路径都会单独对查询三元组进行推断，并给出一个置信度分数。由此，问题被转化为二分类问题，其中 [*cls*] 上线性层的分数表示在给定推理路径的情况下，查询三元组是否正确。推理阶段采用对分数集合进行最大聚合的方式，以综合所有推理路径的分数。得分最高的路径可用于解释预测的推理过程。

尽管这些工作都尝试同时考虑文本信息和结构信息，但仍将它们视为两种不同类型的知识。另一方面，GPT 中提出的提示调优3 [9]作为一种新兴方法，已应用于关系抽取、实体识别等任务。近期研究尝试将外部知识融入提示设计。OntoPrompt [83] 是这一方法的代表之一。它利用提示为预训练语言模型（LM）中的常识知识与知识图谱中的结构知识搭建桥梁，以完成少样本知识图谱补全任务。OntoPrompt 首先通过本体转换来丰富结构知识并将其转换为文本格式，并利用预定义模板将知识转换为文本提示。具体而言，对于知识图谱补全，模型将本体表示中的头实体类型和尾实体类型用作约束。它使用从外部维基数据查询中获取的相应条目作为本体来源，并提取文本描述。它沿用 KG-BERT 的做法，将知识图谱补全视为三元组分类任务，并将实体和关系拼接成输入序列。它还使用可学习虚拟标记来增强提示表示。接下来，OntoPrompt 提出跨度敏感知识注入，以选择有信息量的知识。考虑到无关且嘈杂的知识可能改变原句的含义，OntoPrompt 利用基于跨度的可见矩阵，限制相应知识对知识注入的影响。这样，输入句子中的并非所有标记都会受到外部知识的影响。第三，OntoPrompt 开发了一种联合训练算法，以联合优化表示。需要注意的是，注入的外部知识应与周围上下文相关；可学习标记以随机初始化，并在固定语言模型下与注入的本体标记一同优化。受先前研究 [23] 的启发，该研究指出低数据场景中的提示调优不稳定，且可能导致性能不佳，因此模型进一步优化所有参数，以联合训练本体文本和输入文本的表示。借助上述组件，OntoPrompt 能够利用预训练语言模型丰富与任务相关的知识，避免负面知识融合，并将常识知识整合到结构知识中，从而解决知识缺失、知识噪声和知识异质性等难题，并在少样本知识图谱补全任务中取得了良好表现。

ZS-SKA [22]也利用提示来处理少样本知识图谱补全任务。与利用本体知识不同，他们直接通过语义知识增强来实现零样本关系分类。

ZS-SKA首先遵循词级句子翻译规则，根据已见关系的实例生成包含未见关系的增强实例。为了对每个实例进行编码，ZS-SKA会对句子中的所有词进行分词，并将其输入BERT，为每个词元生成上下文表示。获得分词后的输入句子后，使用卷积神经网络层。然后，他们基于ConceptNet [57] 设计提示，以整合从已见关系中学习到的语义知识信息，并利用这些知识推断未见关系的特征。他们考虑多种语义知识，包括关系描述和命名实体，以有效地学习未见关系。输入序列采用自然语言片段模板进行封装。提示模板中不使用实际标签集，而是根据每个标签的知识图谱自动构建加权虚拟标签词。提示通过嵌入输入词的超类以及未见关系的虚拟标签嵌入来表示。通过原型网络 [56], ，利用增强实例和提示生成已见关系与未见关系的表示，并计算距离以预测未见关系。

上述研究已经表明，预训练语言模型作为外部来源，能够帮助少样本知识图谱补全任务中的实体学习和关系学习。更进一步，除了借助预训练语言模型的能力来补全零样本/少样本查询元组之外，另一个热门方向是通过为未见实体构造三元组，直接补全知识图谱。该任务的代表性模型包括 COMET [8] 和 BERTNet [26]。

COMET [8]是首个针对自动常识知识图谱补全的全面研究。它是一种基于常识知识的生成式变换器模型，能够学习以自然语言生成详尽且多样的常识描述。如第 2.1.2 节所述，常识知识图谱推理面临的一个主要挑战是，以开放文本形式表示的常识知识通常无法适配固定模式。一般而言，COMET 通过在现有元组上训练变换器来解决这一问题，这些元组作为知识种子集，用于借助预训练语言模型学习常识知识的自适应表示。随后，该语言模型可用于生成包含未见实体的新元组。模型通过生成能够在语义上补全现有种子短语和关系类型的短语来识别关系。具体而言，COMET 中的变换器语言模型遵循 GPT [48], 的结构，由多个多头注意力和全连接层变换器块组成，用于编码输入文本。模型的输入是由知识元组中的词构成的拼接序列。为编码变换器忽略的词元间顺序信息，COMET 为序列中的每个位置添加位置嵌入。每个词的位置嵌入与词嵌入相加，得到最终表示，并作为第一个变换器层的输入。COMET 的训练目标是根据元组的头实体和关系生成尾实体，这遵循知识图谱补全任务的设置，但模型预期能够生成训练集中不存在的新元组。

COMET 的局限在于，它只能为具有已见关系的新实体生成三元组。理想的零样本知识图谱补全模型应当能够构建包含未见实体和关系的元组。随后提出了 BERTNet [26] ，用于从预训练语言模型中挖掘包含隐式知识的知识图谱。BERTNet 只需输入一个少样本种子集，其中包含初始提示和每种关系对应的种子实体，即可提取未见关系的知识。通常，该模型会自动生成不同的提示，并在给定的语言模型中搜索新颖知识。

BERTNet 解决了知识图谱补全/知识图谱构建中的两项挑战。首先，研究表明，即使提示略有不同，语言模型的输出也会不一致，因此很难从语言模型中可靠地提取知识。一个直观的解决方案是自动学习最优提示。这类方法需要大量训练数据，而少样本或零样本设置下无法获得此类数据。为此，BERTNet 对初始提示进行无监督改写，生成一组多样化的提示及其置信度。随后，提取能够持续满足这些提示的实体对，以生成新颖三元组。另一个挑战在于，由于语言模型中的候选空间十分庞大，搜索合适的实体对并非易事。BERTNet 设计了一种高效的搜索与重新评分策略，在知识准确率与覆盖率之间取得平衡。模型设计了一种提示/实体对兼容性函数，在每个知识搜索步骤中动态地为候选提示和实体分配权重。具体来说，在搜索实体时，BERTNet 首先使用单独兼容分数；该分数更便于设置阈值和筛选，再通过对所有提示进行加权平均，生成大量候选实体对。随后，根据总兼容分数对这些候选项重新排序，以选出最终的实体对。

由于 BERTNet 除了预训练语言模型外，只需要一组种子提示和关系对应的少样本实体，因此即使面对结构复杂或包含多个实体的关系，该模型也能灵活地提取新颖知识。此外，生成的三元组可以视为对相应黑箱语言模型的解释。BERTNet 的另一项创新在于，它不只是关注 hits@10 和 BLUE 分数等指标，还会将生成的元组直接整合到背景知识图谱中，并将新的知识图谱用于下游任务。这些任务上的表现表明，BERTNet 能够生成新颖的高质量元组。

4 应用与资源

4.1 应用

少样本知识图谱补全模型已应用于知识图谱预测任务以外的问题。一个主要应用是利用少样本链接预测技术解决其他图相关任务。例如，元图 [7]研究了不同网络上的少样本链接预测，包括蛋白质-蛋白质相互作用（PPI）网络 [98], 3、三维点云数据 [42] 和学术社交网络 [62]。MetaTNE [32]也利用少样本知识图谱补全任务中常用的元变换器来预测蛋白质-蛋白质相互作用。同样，分子属性预测一直是一项要求较高的任务，因为人工预测成本高且效率低。Meta-MGNN [24]和 Pre-PAR [73] 已被提出，旨在利用少样本链接预测方法解决此类任务。Meta-MGNN 将每个分子视为一张图，并使用图级图神经网络对每个分子进行编码。此外，它还引入注意力机制，使 MAML 能够感知分子属性之间的差异，从而进一步改进模型编码。此外，Pre-PAR 通过捕捉不同分子属性之间的关系结构，对 Meta-MGNN 进行了改进，从而在相似分子之间有效、高效地传播有限标签。同样，GEN [3]也被用于药物-药物相互作用预测。

另一个与链接预测相关且具有挑战性的现实任务是推荐问题。例如，正如 Yelp 挑战赛 [2],中所提出的，用户评论已成为 Yelp 等网络服务的重要组成部分。用户可以通过由自由文本和数字星级评分构成的评论，发布对商家、产品和服务的看法，因此用户、服务/产品与评论之间的交互直观地形成了结构化和文本知识。近期收集用户相关信息的技术（如支持 GPS 的设备）有助于构建基于位置的社交网络，为推荐系统提供有价值的位置信息。此类系统面临的一项挑战是，新用户加入时，平台上除基本信息和位置外几乎没有关于该用户的现有知识。因此，SEATLE [33] 等少样本链接预测模型以及基于异质信息网络的模型 [36, 93] ，旨在通过元学习解决图上的冷启动推荐问题。其他模型 [17, 18, 35, 61, 68, 80, 81]则采用属性匹配（将用户与以往用户进行匹配）、利用跨网络元学习从其他平台迁移知识、通过动态元图推理基于更新后的信息对用户进行建模等方法来处理推荐问题。

如 BERTNet [26], 中所述，少样本知识图谱补全模型有助于提升原始知识图谱的质量和覆盖率，因此其输出可用于下游任务。例如，可供性推理与提取 [96]可以利用少样本知识图谱补全模型，借助预训练语言模型为未见实体生成可供性，或将其与背景知识图谱中的已见实体进行相似度匹配。近来，在 EPIC [13] 和 STAR [75]等大规模视频提取/识别数据集以及动作规划任务中，由于推理场景和动态环境知识变化迅速，需要进行程序性推理。少样本知识图谱补全模型在此类任务中大有可为，因为它们能够轻松捕捉新的少样本信息，并迅速调整和优化自身，以适应每个推理步骤。

4.2 资源

随着零样本知识图谱补全方法的蓬勃发展，人们提出了各种基准数据集用于评估。这些基准数据集通常基于现有的常用典型知识图谱补全数据集构建，包括 FB15k [5],FB15k-237 [46], NELL-995 [77], 以及从维基数据 [66]等热门知识图谱中提取的一些其他子知识图谱。这类基准数据集通常从原始知识图谱中获取实体信息。例如，DBpedia50k、FB20k 和 Wiki- data5M 分别从 DBpedia 和维基百科收集实体的文本描述。下面介绍几种零样本/少样本知识图谱补全的代表性资源：

表3. 少样本知识图谱补全基准数据集统计 s

（a）关系预测基准数据集

	实体数	关系数	三元组数	任务关系 #（训练/验证/测试）	来源
NELL-One [78]	68,545	358	181,109	51/5/11	NELL
Wiki-One [78]	4,838,244	822	5,859,240	133/16/34	维基数据
NELL-ZS [47]	65,567	181	181,109	139/10/32	NELL
Wiki-ZS [47]	605,812	537	724,967	469/20/48	维基数据

(b) 实体预测基准

	实体数 #	关系数 #	三元组数（训练/验证/测试）#	未见实体 # (train/valid/test) #	来源
WN18RR-sub [3]	4,478	11	93,003 (-)	3000	WN188RR
FB15K-237-sub [3]	10,938	237	72,065/6,246/9,867	2,500/1,000/1,500	FB15K-237
NELL-995-sub [3]	5,694	200	22,345/3,676/5,852	1,500/600/900	NELL-995
FB15K-237-OWE [52]	14,405	235	242,489/10,963/36,250	2081 (-)	FB15K-237
Wikidata5M [72]	4,594,485	1222	20,496,514v/6,699/6,894	4,579,609/7,374/7,475	维基数据

• **FB15k-237-OWE** [52] 基于 FB15k-237 构建，用于零样本知识图谱补全。他们首先抽取一组尾实体，并从 FB15K-237 中随机选取与之相关的头实体。随后，将所有头实体属于相关实体集的三元组移至测试集，构成用于尾实体预测的测试集。同时，移除尾实体属于相关实体集的三元组。头实体预测的测试集以类似方式生成。进一步移除关系未包含在训练集中的三元组后，这两个集合构成最终测试集。然后将该测试集进一步划分为验证集和最终测试集。该数据集包含 2,081 个未见实体、12,324 个已见实体和 235 种关系。训练集/验证集/测试集中的三元组数量分别为 242,489/10,963/36,250。

• **Wikidata5M** [72]最初用于评估文本感知知识图谱嵌入模型，如今已被广泛用于零样本知识图谱补全。该数据集基于维基数据和英文维基百科转储构建。每个实体都使用其维基百科页面的第一部分作为描述。该数据集排除了没有维基百科页面或描述少于 5 个单词的实体。随后，他们从维基数据转储中提取所有三元组。数据集保留了具有合格头实体和尾实体的三元组，最终包含 4,594,485 个实体、822 个关系和 20,624,575 个三元组。为支持零样本设置，他们随机提取两个子知识图谱作为验证集和测试集，并将剩余部分用作训练集。他们确保训练集、验证集和测试集中的实体与三元组彼此不相交。每个集合的详细信息见表 3。

• **WN18RR、FB15k-237 和 NELL-995 的子集**由 [3] 构建，用于已见实体与未见实体之间，以及未见实体彼此之间的图外补全。这些子集均从各自的原始基准数据集中系统地抽取。研究随机抽取一组关联三元组少于 100 个的实体作为未见实体。随后，这些三元组被进一步划分，组成三个元数据集（元训练集、元验证集和

元测试集）。其余原始基准数据集被视为背景知识图谱/图内部分，其中的实体则分别视为已见实体。之后对元数据集进行清理，以确保每个三元组至少包含一个未见实体，且所有三元组均位于图内部分之外。表3中列出了有关已见/未见实体的更多统计信息。

• **NELL/Wiki-ZS**是两个零样本知识图谱补全基准数据集。每个基准数据集都包含三个关系不相交集合：训练集包含已见关系，验证集/测试集包含未见关系。相应的三元组也按此划分，而测试集/验证集中的所有实体均在训练集中出现。NELL-ZS 的训练集/验证集/测试集分别包含 139/10/32 个关系，而 Wiki-ZS 的训练集/验证集/测试集分别包含 469/20/48 个关系。GEN [47] 使用关系文本描述作为文本信息，而 OntoZSL [21] 则构建本体模式；这些模式不仅包含文本信息，还包含其他关系知识，包括两个基准数据集中的关系层级和关系域。

• **NELL/Wiki-One** 最初由 GMatching [78] 开发，用于评估仅有一个支持实例、且关系未见过的小样本知识图谱补全。为构建测试集，研究者从原始基准数据集中提取关联三元组数量少于 500 条但多于 50 条的关系，并将其用作任务关系。经过这一预处理，NELL-One 提取出 67 个关系，并进一步将其划分为 51/5/11 个关系，用于进一步提取相关三元组并构建训练集/验证集/测试集。类似地，Wiki-One 提取出 183 个关系，并将其划分为 133/16/34 个关系，以构建训练集/验证集/测试集中的三元组。在实体方面，NELL-One 提取了 68,545 个实体，Wiki-One 则提取了 4,838,244 个实体。此外，还提取了另外 291 个和 639 个关系作为背景关系，以便为这些实体构建更多三元组。训练集/验证集/测试集中的关系互不重叠，因此这两个基准数据集既可用于零样本知识图谱补全任务，也可用于小样本知识图谱补全任务。

5 分析

从 GMatching [78], 开始，由于问题设定源自知识图谱补全问题，少样本知识图谱补全模型直观地通过元学习框架增强原本用于捕获结构信息的知识图谱嵌入或图神经网络模型。早期的度量类模型，如 MetaR、FSRL [86] 和 FAAN [54], ，侧重于通过为不同邻居分配静态或任务感知注意力，有效利用邻居信息。优化模型，如 Meta-KGR [37], ADK-GK [89],ZS-GAN [47], ，致力于借助多跳路径信息改进实体和关系嵌入。这些路径也可用于解释推理逻辑，从而提升模型的可解释性。元学习使这些模型能够在测试阶段快速适应小样本任务。

然而，这些模型忽略了背景知识图谱中的语义信息，例如实体/关系名称和文本描述。为结合这两类信息，研究者提出了 OntoPrompt [83] 和 ZS-SKA [22] 等基于提示的模型。这些模型利用提示将知识图谱中的三元组转换为类句子知识。随后，使用 BERT 等预训练语言模型生成实体和关系嵌入。实体和关系的文本描述等语义信息有助于丰富嵌入。

所有这些模型都依赖大规模知识图谱，因为它们为实体和关系嵌入提供结构信息。另一方面，GPT模型 [9, 48] 表明，预训练语言模型本身也包含结构知识；有些模型更进一步，完全依赖预训练语言模型。COMET [8] 和 BERTNet [26]是这一类模型的代表。COMET 遵循 GPT 模型的框架，而 BERTNet 使用提示来 交由 ACM 的稿件

	GMatching [78]	FSRL [86]	MetaR [12]	OntoPrompt [83]	InductiveE [54]	COMET![8]	BERTNet [26]
NELL-One [78]	✓	✓	✓	✗	✗	✗	✗
FB15K-237-sub [3]	✓	✓	✓	✓	✗	✗	✗
WN18RR-sub [3]	✓	✓	✓	✓	✗	✗	✗
NELL-ZS [47]	✗	✗	✗	✗	✗	✗	✗
ConceptNet [57]	✗	✗	✗	✗	✓	✓	✓
ATOMIC [50]	✗	✗	✗	✗	✓	✓	✗

表4. 代表性少样本知识图谱补全模型的评估

生成新颖的三元组。与采用结构知识的模型相比，这些模型通常既适用于结构知识图谱，也适用于常识知识图谱。这种方法使少样本知识图谱补全模型摆脱了对背景知识图谱的依赖，并将其应用扩展到更广泛的下游任务。

6 未来方向

如前所述，早期的少样本知识图谱补全模型主要是将元学习和迁移学习方法扩展到少样本知识图谱补全领域。元学习中的另一类知识是学习过程。未来，将关于学习的知识（例如，先前推理过程 [60]）表示为知识图谱，并将其与元学习或迁移学习算法相结合，可能催生出适用于不同少样本学习任务的更通用的神经符号范式。基于传播的方法，例如 GEN [3] 和 REFORM [70]，通过利用少样本样例来解决少样本知识图谱补全问题；也就是说，三元组对未见实体与已见实体的分布和相关性之间的关联进行建模。利用这些少样本链接，以及未见实体相关性的辅助信息（如文本描述、属性和模式），将是一个很有前景的解决方案。

另一类少样本知识图谱补全模型（如 OntoPrompt [83] ）证明，利用本体/规则结构化知识，是推断未见实体/关系的符号知识（如三元组）的一种有前景的方法。另一方面，与广泛研究的基于映射的方法 [11] 相比，基于生成的模型（如 OntoZSL [21] ）在预测时不会偏向已见或未见类别。因此，以知识图谱嵌入为条件的基于生成的零样本学习方法，可能成为少样本知识图谱补全任务的未来方向。

直到最近，语义信息才开始受到少样本知识图谱补全模型的关注。GANA [43]首次提出整合少样本关系邻域的语义信息，而 ZSGAN [47] 则利用任务关系的文本表示生成合理的关系嵌入。COMET [8]、KG-BERT [82],和 BERTNet [26] 等研究进一步展示了学习未见/少样本实体和关系表示的有效性。这些模型的表现表明，多模态知识也有益于少样本知识图谱补全任务。如何将来自更多模态的数据高效整合到当前的少样本知识图谱补全模型中，是一个值得探索的方向，不仅有助于解决知识图谱补全任务，也有助于其他领域的零样本/少样本推理。

根据表4，目前尚无足够的评估来比较基于大规模知识图谱的少样本知识图谱补全模型与利用预训练语言模型的模型的性能。BERTNet 和 COMET 等模型对 NELL 和 Freebase 等结构化知识图谱有效，目前仍停留在理论层面。此类评估也将有助于结合多模态信息的少样本知识图谱补全研究。

7 结论

知识图谱是多个领域中各类任务的重要来源，因此，少样本知识图谱补全模型为高效地将新知识整合到现有知识图谱中提供了机会，从而提升知识图谱的质量和覆盖率。在本综述中，我们首先介绍少样本知识图谱补全的主要挑战和基础知识，随后全面回顾相关研究。我们将这些方法分为两类：基于结构知识的少样本知识图谱补全，以及基于语义知识的少样本知识图谱补全。这种分类展现了少样本知识图谱补全方法的发展趋势：从元学习和基于注意力的模型，转向利用语义信息和文本信息，甚至直接从预训练语言模型中提取结构三元组。随后，我们讨论少样本知识图谱补全模型如何迁移或应用于不同领域。最后，我们总结不同少样本知识图谱补全模型仍面临的挑战及未来可能的发展方向。我们希望本综述能为关注少样本知识图谱补全的研究者提供有价值的参考，并推动该领域未来的研究。

参考项

[1] Ralph Abboud、Ismail Ceylan、Thomas Lukasiewicz 和 Tommaso Salvatori。2020年。Boxe：一种用于知识库补全的盒嵌入模型。神经信息处理系统进展33（2020年），9649–9661。[2] Nabiha Asghar。2016年。Yelp 数据集挑战赛：评论评分预测。*arXiv*预印本 *arXiv:1605.05362*（2016年）。[3] Jinheon Baek、Dong Bok Lee 和 Sung Ju Hwang。2020年。学习外推知识：传导式小样本图外链接预测。神经信息处理系统进展33（2020年），546–560。[4] Piotr Bojanowski、Édouard Grave、Armand Joulin 和 Tomáš Mikolov。2017年。利用子词信息丰富词向量。计算语言学协会学报5（2017年），135–146。[5] Kurt Bollacker、Colin Evans、Praveen Paritosh、Tim Sturge 和 Jamie Taylor。2008年。Freebase：一个协作创建的图数据库，用于组织人类知识。载于*2008*年*ACM SIGMOD*数据管理国际会议论文集。1247–1250。[6] Antoine Bordes、Nicolas Usunier、Alberto Garcia-Duran、Jason Weston 和 Oksana Yakhnenko。2013年。用于建模多关系数据的平移嵌入。神经信息处理系统进展26（2013年）。[7] Avishek Joey Bose、Ankit Jain、Piero Molino 和 William L Hamilton。2019年。元图：通过元学习进行少样本链接预测。*arXiv*预印本 *arXiv:1912.09867*（2019年）。[8] Antoine Bosselut、Hannah Rashkin、Maarten Sap、Chaitanya Malaviya、Asli Celikyilmaz 和 Yejin Choi。2019年。COMET：用于自动构建知识图谱的常识变换器。*arXiv*预印本 *arXiv:1906.05317*（2019年）。[9] Tom Brown、Benjamin Mann、Nick Ryder、Melanie Subbiah、Jared D Kaplan、Prafulla Dhariwal、Arvind Neelakantan、Pranav Shyam、Girish Sastry、Amanda Askell 等。2020年。语言模型是少样本学习者。神经信息处理系统进展33（2020年），1877–1901。[10] Andrew Carlson、Justin Betteridge、Bryan Kisiel、Burr Settles、Estevam R Hruschka 和 Tom M Mitchell。2010年。迈向永不停歇的语言学习架构。载于第二十四届*AAAI*人工智能会议。[11] Jiaoyan Chen、Yuxia Geng、Zhuo Chen、Jeff Z Pan、Yuan He、Wen Zhang、Ian Horrocks 和 Huajun Chen。2021年。知识图谱低资源学习：全面综述。*arXiv*预印本 *arXiv:2112.10006*（2021年）。[12] Mingyang Chen、Wen Zhang、Wei Zhang、Qiang Chen 和 Huajun Chen。2019年。知识图谱中少样本链接预测的元关系学习。载于*2019*年自然语言处理实证方法会议与第九届自然语言处理国际联合会议（*EMNLP-IJCNLP*）论文集。4217–4226。[13] Dima Damen、Hazel Doughty、Giovanni Maria Farinella、Sanja Fidler、Antonino Furnari、Evangelos Kazakos、Davide Moltisanti、J onathan Munro、Toby Perrett、Will Price 等。2018年。扩展自我中心视觉：EPIC-KITCHENS 数据集。载于欧洲计算机视觉会议（*ECCV*）论文集。720–736。[14] Rajarshi Das、Shehzaad Dhuliawala、Manzil Zaheer、Luke Vilnis、Ishan Durugkar、Akshay Krishnamurthy、Alex Smola 和 Andrew McCallum。2018年。踏上旅程，找到答案：使用强化学习对知识库中的路径进行推理。载于国际学习表征会议。[15] Tim Dettmers、Pasquale Minervini、Pontus Stenetorp 和 Sebastian Riedel。2018年。卷积二维知识图谱嵌入。载于*AAAI*人工智能会议论文集，第32卷。[16] Jacob Devlin、Ming-Wei Chang、Kenton Lee 和 Kristina Toutanova。2018年。BERT：用于语言理解的深度双向变换器预训练。*arXiv*预印本 *arXiv:1810.04805*（2018年）。[17] Kaize Ding、Jianling Wang、Jundong Li、Kai Shu、Chenghao Liu 和 Huan Liu。2020年。用于属性网络少样本学习的图原型网络。载于*CIKM*。[18] Kaize Ding、Qinghai Zhou、Hanghang Tong 和 Huan Liu。2021年。通过跨网络元学习进行少样本网络异常检测。载于*WWW*。

[19] Chelsea Finn、Pieter Abbeel 和 Sergey Levine。2017年。用于深度网络快速适应的模型无关元学习。载于国际机器学习会议。PMLR，1126–1135。[20] Chelsea Finn、Kelvin Xu 和 Sergey Levine。2018年。概率模型无关元学习。神经信息处理系统进展31（2018年）。[21] Yuxia Geng、Jiaoyan Chen、Zhuo Chen、Jeff Z Pan、Zhiquan Ye、Zonggang Yuan、Yantao Jia 和 Huajun Chen。2021年。OntoZSL：本体增强零样本学习。载于2021年网络会议论文集。3325–3336。[22] Jiaying Gong 和 Hoda Eldardiry。2021年。利用语义知识增强的基于提示的零样本关系分类。arXiv预印本 arXiv:2112.04539（2021年）。[23] Yuxian Gu、Xu Han、Zhiyuan Liu 和 Minlie Huang。2022年。PPT：用于少样本学习的预训练提示调优。载于计算语言学协会第60届年会论文集（第一卷：长篇论文）。8410–8423。[24] Zhichun Guo、Chuxu Zhang、Wenhao Yu、John Herr、Olaf Wiest、Meng Jiang 和 Nitesh V Chawla。2021年。用于分子属性预测的少样本图学习。载于WWW。[25] Will Hamilton、Zhitao Ying 和 Jure Leskovec。2017年。大图上的归纳式表示学习。神经信息处理系统进展30（2017年）。[26] Shibo Hao、Bowen Tan、Kaiwen Tang、Hengzhe Zhang、Eric P Xing 和 Zhiting Hu。2022年。BERTNet：从预训练语言模型中获取知识图谱。arXiv预印本 arXiv:2206.14268（2022年）。[27] Kaiming He、Xiangyu Zhang、Shaoqing Ren 和 Jian Sun。2016年。用于图像识别的深度残差学习。载于IEEE计算机视觉与模式识别会议论文集。770–778。[28] Jena D Hwang、Chandra Bhagavatula、Ronan Le Bras、Jeff Da、Keisuke Sakaguchi、Antoine Bosselut 和 Yejin Choi。2021年。（comet-）ATOMIC 2020：符号与神经常识知识图谱。载于AAAI人工智能会议论文集，第35卷。6384–6392。[29] Guoliang Ji、Shizhu He、Liheng Xu、Kang Liu 和 Jun Zhao。2015年。通过动态映射矩阵进行知识图谱嵌入。载于计算语言学协会第53届年会与自然语言处理国际联合会议第七届会议论文集（第一卷：长篇论文）。687–696。[30] Zhiyi Jiang、Jianliang Gao 和 Xinqi Lv。2021年。MetaP：单样本知识图谱补全的元模式学习。载于第44届ACM信息检索研究与发展国际会议论文集。2232–2236。[31] Ranjay Krishna、Yuke Zhu、Oliver Groth、Justin Johnson、Kenji Hata、Joshua Kravitz、Stephanie Chen、Yannis Kalantidis、Li-Jia Li、David A Shamma 等。2017年。Visual Genome：利用众包密集图像标注连接语言与视觉。国际计算机视觉期刊123卷第1期（2017年），32–73。[32] Lin Lan、Pinghui Wang、Xuefeng Du、Kaikai Song、Jing Tao 和 Xiaohong Guan。2020年。通过元变换网络嵌入对具有少样本新标签的图进行节点分类。载于NeurIPS。[33] Ruirui Li、Xian Wu、Xian Wu 和 Wei Wang。2020年。基于位置的社交网络中新用户推荐的少样本学习。载于WWW。[34] Xi Victoria Lin、Richard Socher 和 Caiming Xiong。2018年。通过奖励塑形进行多跳知识图谱推理。载于2018年自然语言处理实证方法会议论文集。3243–3253。[35] Zemin Liu、Yuan Fang、Chenghao Liu 和 Steven CH Hoi。2021年。用于图上少样本节点分类的相对与绝对位置嵌入。载于AAAI。[36] Yuanfu Lu、Yuan Fang 和 Chuan Shi。2020年。异质信息网络上的元学习：冷启动推荐。载于第26届ACM SIGKDD知识发现与数据挖掘国际会议论文集。1563–1573。[37] Xin Lv、Yuxian Gu、Xu Han、Lei Hou、Juanzi Li 和 Zhiyuan Liu。2019年。利用元知识图谱信息对少样本关系进行多跳推理。载于EMNLP/IJCNLP（1）。[38] Andrew L Maas、Awni Y Hannun、Andrew Y Ng 等。2013年。整流非线性函数改进神经网络声学模型。载于ICML会议录，第30卷。美国佐治亚州亚特兰大，3。[39] Farzaneh Mahdisoltani、Joanna Biega 和 Fabian Suchanek。2014年。YAGO3：来自多语言维基百科的知识库。载于第七届两年一度的创新数据系统研究会议。CIDR会议。[40] Sijie Mai、Shuangjia Zheng、Yuedong Yang 和 Haifeng Hu。2021年。用于归纳式关系推理的通信式消息传递。载于AAAI。4294–4302。[41] George A Miller。1995年。WordNet：英语词汇数据库。ACM通讯38卷第11期（1995年），39–41。[42] Marion Neumann、Plinio Moreno、Laura Antanas、Roman Garnett 和 Kristian Kersting。2013年。用于任务相关机器人抓取中物体类别预测的图核。载于第十一届图挖掘与学习研讨会在线论文集。0–6。[43] Guanglin Niu、Yang Li、Chengguang Tang、Ruiying Geng、Jian Dai、Qiao Liu、Hao Wang、Jian Sun、Fei Huang 和 Luo Si。2021年。利用门控注意力邻居聚合器进行关系学习，以实现少样本知识图谱补全。载于第44届ACM信息检索研究与发展国际会议论文集。213–222。[44] Lei Niu、Chenpeng Fu、Qiang Yang、Zhixu Li、Zhigang Chen、Qingsheng Liu 和 Kai Zheng。2021年。利用多重交互注意力进行开放世界知识图谱补全。万维网24卷第1期（2021年），419–439。[45] Heiko Paulheim。2017年。知识图谱精炼：方法与评估方法综述。语义网8卷第3期（2017年），489–508。[46] Thomas Pellissier Tanon、Denny Vrandečić、Sebastian Schaffert、Thomas Steiner 和 Lydia Pintscher。2016年。从Freebase到维基数据：大迁移。载于第25届万维网国际会议论文集。国际万维网会议指导委员会，M

1419–1428。[47] Pengda Qin、Xin Wang、Wenhu Chen、Chunyun Zhang、Weiran Xu 和 William Yang Wang。2020。面向知识图谱的生成对抗零样本关系学习。载于 AAAI。[48] Alec Radford、Karthik Narasimhan、Tim Salimans、Ilya Sutskever 等。2018。通过生成式预训练提升语言理解。（2018）。[49] Sachin Ravi 和 Hugo Larochelle。2016。将优化作为少样本学习的模型。载于 ICLR。[50] Maarten Sap、Ronan Le Bras、Emily Allaway、Chandra Bhagavatula、Nicholas Lourie、Hannah Rashkin、Brendan Roof、Noah A Smith 和 Yejin Choi。2019。ATOMIC：用于如果—那么推理的机器常识图谱。载于《AAAI人工智能会议论文集》，第 33 卷。3027–3035。[51] Michael Schlichtkrull、Thomas N Kipf、Peter Bloem、Rianne van den Berg、Ivan Titov 和 Max Welling。2018。使用图卷积网络对关系数据建模。载于欧洲语义网会议。Springer，593–607。[52] Haseeb Shah、Johannes Villmow、Adrian Ulges、Ulrich Schwanecke 和 Faisal Shafait。2019。知识图谱补全模型的开放世界扩展。载于《AAAI人工智能会议论文集》，第 33 卷。3044–3051。[53] Chao Shang、Yun Tang、Jing Huang、Jinbo Bi、Xiaodong He 和 Bowen Zhou。2019。用于知识库补全的端到端结构感知卷积网络。载于《AAAI人工智能会议论文集》，第 33 卷。3060–3067。[54] Jiawei Sheng、Shu Guo、Zhenyu Chen、Juwei Yue、Lihong Wang、Tingwen Liu 和 Hongbo Xu。2020。用于少样本知识图谱补全的自适应注意力网络。载于 EMNLP。[55] Baoxu Shi 和 Tim Weninger。2018。开放世界知识图谱补全。载于《AAAI人工智能会议论文集》，第 32 卷。[56] Jake Snell、Kevin Swersky 和 Richard Zemel。2017。用于少样本学习的原型网络。《神经信息处理系统进展》30（2017）。[57] Robyn Speer 和 Catherine Havasi。2013。ConceptNet 5：用于关系知识的大型语义网络。载于《大众网络与自然语言处理相遇》。Springer，161–176。[58] Fabian M Suchanek、Gjergji Kasneci 和 Gerhard Weikum。2008。YAGO：一个源自维基百科和 WordNet 的大型本体。《网络语义学杂志》6，3（2008），203–217。[59] Zhiqing Sun、Zhi-Hong Deng、Jian-Yun Nie 和 Jian Tang。2018。RotatE：通过复数空间中的关系旋转实现知识图谱嵌入。载于国际学习表征会议。[60] Dídac Surís、Dave Epstein、Heng Ji、Shih-Fu Chang 和 Carl Vondrick。2020。从视觉场景中学习词语。载于欧洲计算机视觉会议。Springer，434–452。[61] Zhen Tan、Kaize Ding、Ruocheng Guo 和 Huan Liu。2021。图少样本类增量学习。载于 WSDM。[62] Jie Tang、Jing Zhang、Limin Yao、Juanzi Li、Li Zhang 和 Zhong Su。2008。Arnetminer：学术社交网络的抽取与挖掘。载于第 14 届 ACM SIGKDD 国际知识发现与数据挖掘会议论文集。990–998。[63] Komal Teru、Etienne Denis 和 Will Hamilton。2020。通过子图推理进行归纳式关系预测。载于国际机器学习会议。PMLR，9448–9457。[64] Théo Trouillon、Johannes Welbl、Sebastian Riedel、Éric Gaussier 和 Guillaume Bouchard。2016。用于简单链接预测的复嵌入。载于国际机器学习会议。PMLR，2071–2080。[65] Oriol Vinyals、Charles Blundell、Timothy Lillicrap、Daan Wierstra 等。2016。用于单样本学习的匹配网络。《神经信息处理系统进展》29（2016）。[66] Denny Vrandečić 和 Markus Krötzsch。2014。维基数据：一个免费的协作知识库。（2014）。[67] Bin Wang、Guangtao Wang、Jing Huang、Jiaxuan You、Jure Leskovec 和 C-C Jay Kuo。2021。常识知识图谱补全的归纳学习。载于 2021 年国际神经网络联合会议（IJCNN）。IEEE，1–8。[68] Ning Wang、Minnan Luo、Kaize Ding、Lingling Zhang、Jundong Li 和 Qinghua Zheng。2020。基于属性匹配的图少样本学习。载于 CIKM。[69] Quan Wang、Pingping Huang、Haifeng Wang、Songtai Dai、Wenbin Jiang、Jing Liu、Yajuan Lyu、Yong Zhu 和 Hua Wu。2019。Coke：上下文化知识图谱嵌入。arXiv 预印本 arXiv:1911.02168（2019）。[70] Song Wang、Xiao Huang、Chen Chen、Liang Wu 和 Jundong Li。2021。REFORM：误差感知少样本知识图谱补全。载于第 30 届 ACM 国际信息与知识管理会议论文集。1979–1988。[71] Shen Wang、Xiaokai Wei、Cicero Nogueira Nogueira dos Santos、Zhiguo Wang、Ramesh Nallapati、Andrew Arnold、Bing Xiang、Philip S Yu 和 Isabel F Cruz。2021。用于知识图谱补全的混合曲率多关系图神经网络。载于 2021 年网络会议论文集。1761–1771。[72] Xiaozhi Wang、Tianyu Gao、Zhaocheng Zhu、Zhengyan Zhang、Zhiyuan Liu、Juanzi Li 和 Jian Tang。2021。KEPLER：知识嵌入与预训练语言表征的统一模型。《计算语言学协会汇刊》9（2021），176–194。[73] Yaqing Wang、Abulikemu Abuduweili、Quanming Yao 和 Dejing Dou。2021。用于少样本分子属性预测的属性感知关系网络。载于 NeurIPS。[74] Yaqing Wang、Quanming Yao、James T Kwok 和 Lionel M Ni。2020。从少数示例中泛化：少样本学习综述。《计算调查》（2020）。[75] Bo Wu、Shoubin Yu、Zhenfang Chen、Joshua B Tenenbaum 和 Chuang Gan。2021。STAR：现实世界视频情境推理的基准数据集。载于第三十五届神经信息处理系统会议数据集与基准赛道（第二轮）。已提交 ACM 的稿件

[76] Han Wu、Jie Yin、Bala Rajaratnam 和 Jianyuan Guo。2022。面向少样本知识图谱补全的层次关系学习。arXiv 预印本 arXiv:2209.01205（2022）。[77] Wenhan Xiong、Thien Hoang 和 William Yang Wang。2017。DeepPath：一种用于知识图谱推理的强化学习方法。载于《2017年自然语言处理实证方法会议论文集》。564–573。[78] Wenhan Xiong、Mo Yu、Shiyu Chang、Xiaoxiao Guo 和 William Yang Wang。2018。面向知识图谱的单样本关系学习。载于《2018年自然语言处理实证方法会议论文集》。1980–1990。[79] Jingwen Xu、Jing Zhang、Xirui Ke、Yuxiao Dong、Hong Chen、Cuiping Li 和 Yongbin Liu。2021。P-INT：用于少样本知识图谱补全的基于路径的交互模型。载于 EMNLP 研究成果。[80] Cheng Yang、Chunchen Wang、Yuanfu Lu、Xumeng Gong、Chuan Shi、Wei Wang 和 Xu Zhang。2022。动态网络中的少样本链接预测。载于第十五届 ACM 国际网络搜索与数据挖掘会议论文集。1245–1255。[81] Huaxiu Yao、Chuxu Zhang、Ying Wei、Meng Jiang、Suhang Wang、Junzhou Huang、Nitesh Chawla 和 Zhenhui Li。2020。通过知识迁移实现图少样本学习。载于 AAAI。[82] Liang Yao、Chengsheng Mao 和 Yuan Luo。2019。KG-BERT：用于知识图谱补全的 BERT。arXiv 预印本 arXiv:1909.03193（2019）。[83] Hongbin Ye、Ningyu Zhang、Shumin Deng、Xiang Chen、Hui Chen、Feiyu Xiong、Xi Chen 和 Huajun Chen。2022。用于少样本学习的本体增强提示调优。载于《2022年 ACM 网络会议论文集》。778–787。[84] Donghan Yu、Yiming Yang、Ruohong Zhang 和 Yuexin Wu。2021。基于知识嵌入的图卷积网络。载于 2021 年网络会议论文集。1619–1628。[85] Hanwen Zha、Zhiyu Chen 和 Xifeng Yan。2022。通过 BERT 进行归纳式关系预测。载于《AAAI人工智能会议论文集》，第 36 卷。5923–5931。[86] Chuxu Zhang、Huaxiu Yao、Chao Huang、Meng Jiang、Zhenhui Li 和 Nitesh V Chawla。2020。少样本知识图谱补全。载于 AAAI。[87] Chuxu Zhang、Lu Yu、Mandana Saebi、Meng Jiang 和 Nitesh Chawla。2020。知识库上的少样本多跳关系推理。载于 EMNLP 研究成果。[88] Ningyu Zhang、Shumin Deng、Zhanlin Sun、Jiaoyan Chen、Wei Zhang 和 Huajun Chen。2020。面向低资源知识图谱补全的关系对抗网络。载于《2020年网络会议论文集》。1–12。[89] Yiming Zhang、Yiyue Qian、Yanfang Ye 和 Chuxu Zhang。2022。为知识图谱上的少样本关系推理调整蒸馏知识。载于 SDM。[90] Yongqi Zhang 和 Quanming Yao。2022。使用关系有向图进行知识图谱推理。载于《2022年 ACM 网络会议论文集》。912–924。[91] Zhanqiu Zhang、Jianyu Cai、Yongdong Zhang 和 Jie Wang。2020。学习层次感知知识图谱嵌入以进行链接预测。载于《AAAI人工智能会议论文集》，第 34 卷。3065–3072。[92] Zhanqiu Zhang 、 Jie Wang 、 Jieping Ye 和 Feng Wu。2022。重新思考用于知识图谱补全的图卷积网络。载于《2022年 ACM 网络会议论文集》。798–807。[93] Huan Zhao、Quanming Yao、Jianda Li、Yangqiu Song 和 Dik Lun Lee。2017。基于元图的异构信息网络推荐融合。载于第 23 届 ACM SIGKDD 国际知识发现与数据挖掘会议论文集。635–644。[94] Shangfei Zheng、Wei Chen、Pengpeng Zhao、An Liu、Junhua Fang 和 Lei Zhao。2021。难度何以重要：少样本关系上的多跳知识图谱推理。载于第 30 届 ACM 国际信息与知识管理会议论文集。2688–2697。[95] Fan Zhou、Chengtai Cao、Kunpeng Zhang、Goce Trajcevski、Ting Zhong 和 Ji Geng。2019。Meta-GNN：图元学习中的少样本节点分类。载于 CIKM。[96] Yuke Zhu、Alireza Fathi 和 Li Fei-Fei。2014。基于知识库表征对物体可供性进行推理。载于欧洲计算机视觉会议。Springer，408–424。[97] Zhaocheng Zhu、Zuobai Zhang、Louis-Pascal Xhonneux 和 Jian Tang。2021。神经 Bellman-Ford 网络：用于链接预测的通用图神经网络框架。《神经信息处理系统进展》34（2021），29476–29490。[98] Marinka Zitnik 和 Jure Leskovec。2017。通过多层组织网络预测多细胞功能。《生物信息学》33，14（2017），i190–i198。