

融合知识表示和深度强化学习的知识推理方法

宋浩楠, 赵 刚, 王兴芬

北京信息科技大学 信息管理学院, 北京 100192

摘 要: 知识推理是解决知识图谱中知识缺失问题的重要方法, 针对大规模知识图谱中知识推理方法仍存在可解释性差、推理准确率和效率偏低的问题, 提出了一种将知识表示和深度强化学习相结合的方法 RLPTransE。利用知识表示学习方法, 将知识图谱映射到含有三元组语义信息的向量空间中, 并在该空间中建立强化学习环境。通过单步择优策略网络和多步推理策略网络的训练, 使强化学习智能体在与环境交互过程中, 高效挖掘推理规则进而完成推理。在公开数据集上的实验结果表明, 相比于其他先进方法, 该方法在大规模数据集推理任务中取得更好的表现。

关键词: 知识推理; 深度强化学习; 知识表示; 路径控制; 规则挖掘

文献标志码: A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.2104-0430

Knowledge Reasoning Method Combining Knowledge Representation with Deep Reinforcement Learning

SONG Haonan, ZHAO Gang, WANG Xingfen

School of Information Management, Beijing Information Science & Technology University, Beijing 100192, China

Abstract: Knowledge reasoning is an important method to solve the problem of lack of knowledge in the knowledge graph. The knowledge reasoning method in the large-scale knowledge graph still has the problems of poor interpretability, low reasoning accuracy and efficiency. This paper proposes a method RLPTransE that combines knowledge representation with deep reinforcement learning. Firstly, it uses the knowledge representation learning method to map the knowledge graph to the vector space containing the semantic information of the triples, and establishes a reinforcement learning environment in the space. Then, it trains through the single-step optimization strategy network and the multi-step reasoning strategy network to enable the reinforcement learning agent to efficiently mine the reasoning rules and complete the reasoning in the process interacting with the environment. According to experimental results on public datasets, compared with state-of-the-art methods, the proposed method achieves better performance in reasoning tasks for large-scale datasets.

Key words: knowledge reasoning; deep reinforcement learning; knowledge representation; path control; rule mining

随着知识图谱相关技术的快速发展, 各种大规模的知识图谱被构建出来并广泛服务于各个领域。例如 Freebase^[1]、DBpedia^[2]、NELL^[3]等知识图谱都为研究和应用提供了巨大数据支撑。但是, 无论是自动化构建或人工构建的通用知识图谱, 还是自动化构建或人工构建的领域知识图谱, 绝大多数都存在一定程度的不完备问题^[4], 知识图谱中存在大量的实体和关系缺失。知识图谱补全技术便是为了应对实体和关系缺失而出现, 该技术的最主要方法就是知识推理^[5]。

近年来, 基于知识表示学习推理和基于关系路径推理等研究方法, 已成为知识推理研究的热点。知识表示学习方法(TransE^[6]、TransH^[7]等)因其具有较好的效率和

性能, 所以被广泛应用于知识推理相关任务中, 该类模型将知识图谱中的实体和关系映射到低维向量空间, 并在此空间中进行计算完成推理; 关系路径推理方法作为一种适用于大规模知识图谱的推理方法引起关注和研究, 其主要思想是: 在知识图谱中充分挖掘和利用实体间多步关系, 组成路径信息从而完成知识推理。而知识表示学习和关系路径相融合的方法因其同时具备以上两种优势, 因此得到广泛的研究和利用。

尽管如此, 现阶段知识图谱补全研究中仍存在可解释性差、大规模知识推理效率和准确率较低的问题, 特别是推理的可解释性^[8]逐渐受到领域研究者的关注。针对这类问题, 本文提出了一种将知识表示和深度强化学

基金项目: 国家重点研发计划课题(2019YFB1405003)。

作者简介: 宋浩楠(1993—), 男, 硕士研究生, 研究方向为知识推理、强化学习, E-mail: shnbistu@sina.com; 赵刚(1965—), 男, 博士, 教授, CCF 会员, 研究方向为强化学习、风险评估; 王兴芬(1968—), 女, 博士, 教授, 研究方向为大数据分析、智能决策。

收稿日期: 2021-04-30 **修回日期:** 2021-06-15 **文章编号:** 1002-8331(2021)19-0189-09

习(Reinforcement Learning, RL)相结合的方法 RLP-TransE。将知识图谱中的所有的三元组信息通过知识表示方法映射成低维向量空间中的稠密向量,充分保留其语义信息,同时引入强化学习方法,将知识图谱中知识推理问题转化为马尔可夫序列决策问题,通过知识表示和强化学习相融合的方法充分挖掘知识图谱中的有效推理规则,并且对智能体选择的路径质量进行了控制,从而高效完成大规模知识图谱的补全任务,该融合方法也为知识推理的混合推理研究提供了新的思路。

与其他方法相比,本文方法在大规模数据集上的表现均优于知识表示学习推理和关系路径推理等方法。

本文的主要贡献:

(1)提出了一种融合知识表示和深度强化学习的知识推理方法 RLPTransE,将知识推理问题转化为序列决策问题,增强了推理的可解释性。

(2)提出了一种单步择优策略网络和多步推理策略网络的双网络结构。其目的是准确并高效地挖掘高质量推理规则。

(3)该融合推理方法充分发挥了两者的优势,在公开标准数据集上的对比实验结果显示,本方法取得了较好的性能,为知识推理的混合推理研究提供了新的思路。

1 相关工作

知识图谱的中知识缺失问题普遍存在,而知识推理是解决知识图谱补全任务的重要方法。大规模知识图谱中知识推理方法大致分为三类:基于知识表示学习的方法、基于关系路径的方法和基于知识表示学习和关系路径融合的方法。本章将按照此类别对国内外知识推理方法研究进行介绍。

1.1 基于知识表示学习的推理方法

自词嵌入表示模型提出后,许多自然语言任务都证明了其重要作用。受此启发,对知识三元组的表示学习也取得了许多突破性成果。Border等人^[6]提出了TransE,将知识库中实体之间的关系当作实体间的某种平移,实现了对多元关系数据高效建模的知识补全,但该模型在处理一对多等复杂关系表示问题存在不足。针对这个问题,后续研究提出了许多衍生版本。Wang等人^[7]提出TransH,引入超平面法向量,实现了不同实体在不同关系下拥有不同的表示,但三元组知识仍处于相同的语义空间,限制了自身的表达能力;Lin等人^[9]提出了TransR,为每一种关系定义单独的语义空间,并使用不相同的映射矩阵实现从实体空间到不同关系空间的映射;Ji等人^[10]提出TransD,不仅考虑关系的多样性,而且考虑实体的多样性,提出了不同实体具有不同的映射矩阵,减少模型参数的同时使模型更加灵活。Ebisu等人^[11]提出TorusE,将TransE的思想应用在李群(Lie group)理论的环面空间中。Sun等人^[12]提出RotatE,实现知识表示学

习从实数空间到复数空间的扩展。将关系看作是从头实体到尾实体的旋转。Zhang等人^[13]引入超复数的概念,提出了QuatE,与RotatE类似,该模型将关系看作超复数平面内头实体到尾实体的旋转。相比于其他先进方法,上述方法存在以下问题:(1)将知识推理转化为单一的向量计算,带来了可解释性差的问题;(2)未能充分利用关系路径等重要信息,推理能力受限,推理准确率尚有较大提升空间。

1.2 基于关系路径的推理方法

该类方法主要是基于知识库中图的结构特性进行研究推理^[14]。Lao等人^[15]提出了路径排序算法(Path Ranking Algorithm, PRA),将知识库中连接实体的不同的关系路径作为特征,通过在知识库中统计路径来构建分类的特征向量,建立针对关系的分类器来预测两个实体之间的关系。Lao等人^[16]基于PRA算法,通过调整和组合图中不同随机游走相关权重来学习推断关系,提升路径推理质量的同时使其更适用于大规模知识图谱推理。Gardner等人^[17]提出的子图特征提取模型,将图中的节点对生成特征矩阵,通过修改PRA路径搜索的过程,提取路径之外更丰富的特征,提高推理的效率。Gardner等人^[18]通过结合文本内容对PRA算法进行修改,引入了向量空间相似性,缓解了PRA中的特征稀疏性问题。Das等人^[19]使用RNN模型,通过递归方式组合知识图谱中多跳路径的分布式语义从而构成关系路径,并且在推理过程中引入了注意力机制。Chen等人^[20]设计了概率图模型下的推理问题,在知识推理中引入变分推理框架,将路径搜索和路径推理紧密结合从而进行联合推理,大幅提升了推理效果。相比于其他先进方法,上述方法存在如下问题:(1)由于存在数据稀疏问题,知识图谱中的信息未被充分利用;(2)未考虑路径的可靠性^[5]计算问题,难以适用于大规模知识图谱。

1.3 基于知识表示学习和路径融合方法

上面的两类模型仅考虑了图谱中实体间的直接关系或者只考虑了实体间简单的路径关系,但事实上,知识图谱中实体之间的关系路径隐含着丰富的语义信息,研究知识表示和关系路径的融合方法具有重要意义。Lin等人^[21]设计了PTransE,使用语义连接算法表示路径关系,同时引入路径约束资源分配算法来衡量关系路径的可靠性,将实体和关系映射到低维空间中表示计算,从而显著提高推理能力。陈海旭等人^[22]提出了PSTransE,该模型对PTransE进行了改进,用关系和路径的向量相似度来表示路径推理关系的概率,通过互补方法计算推理概率,在综合考虑相关路径信息的同时,更注重关键路径对推理所起的决定性作用。文献[23-24]都是在PTransE的基础上的改进模型,它们的基本结构一致,仅在知识的表示方式上存在不同。

近年来,机器学习的可解释性越来越得到大家的关

注,强化学习在可解释性和性能等方面的优势,使得强化学习应用于知识推理领域成为研究热点。Xiong 等人^[25]设计了DeepPath,将知识库中知识推理过程转化为马尔可夫序列决策过程(Marcov Decision Process, MDP),以实体集合为状态空间,关系集合为动作空间,智能体通过选择最优动作以拓展其路径来实现知识库中的推理。但由于DeepPath模型简单,并且需要提供大量已知路径进行预训练,训练过程复杂,因此其推理性能存在很大提升空间。Das 等人^[26]提出了MINERVA,将起始实体到目的实体之间的路径选择问题转化为序列决策问题,通过建模以查询问题为条件引导模型在知识库中找出预测路径,解决了在已知一个实体和关系情况下的问答问题。Lin 等人^[27]提出了Multi-Hop,针对路径择优和路径多样性探索的问题分别提出了软奖励机制和随机 Action-Drop 方法。Li 等人^[28]提出了DIVINE,一种基于生成式对抗模仿学习的框架,并通过模仿从知识库中自动采样来自适应地学习推理策略和奖励函数。Wang 等人^[29]提出了AttnPath,将LSTM和图注意力机制作为记忆组件,并提出一种新的避免智能体停滞不前的强化学习机制来提高推理成功率。相比于其他先进方法,上述方法侧重于提高推理准确率,但由于引入许多新技术,模型的复杂度更高,推理效率较低。

2 融合知识表示和深度强化学习的知识推理方法

上述推理方法,各有其优点和不足。因此本文考虑

结合这两类方法的优势来提高模型推理的可解释性、准确性和推理效率,于是本文提出了一种融合知识表示和深度强化学习的推理方法RLPTransE。如图1所示,使用知识表示学习方法,将知识图谱映射到含有三元组语义信息的向量空间中,然后在该空间中建立深度强化学习的环境。通过基于有监督的单步择优策略网络的训练,降低RL智能体单步错误动作的选择率,再通过基于奖励函数的多步推理策略网络的训练,提升RL智能体搜索正确路径的成功率。最终,实现RL智能体在与知识图谱环境交互过程中,成功挖掘推理规则进而完成推理任务。

2.1 知识表示学习模块

为了解决大规模知识推理面临的复杂数据高效利用的问题,在词嵌入表示模型的启发下,研究人员基于分布式思想提出知识表示学习(Knowledge Representation Learning, KRL)的方法^[6],将实体和关系的语义信息映射到低维向量空间,使得语义相近对象的向量表示距离也相近。

在TransE模型中,知识库中的关系当作实体间的某种平移。对于知识库中三元组,用 l_h 、 l_r 和 l_t 依次表示头向量、关系向量和尾向量。该模型的核心思想如公式所示:

$$l_h + l_r \approx l_t \quad (1)$$

得分函数定义如下:

$$d_r(h, t) = \|l_h + l_r - l_t\|_{L1/L2} \quad (2)$$

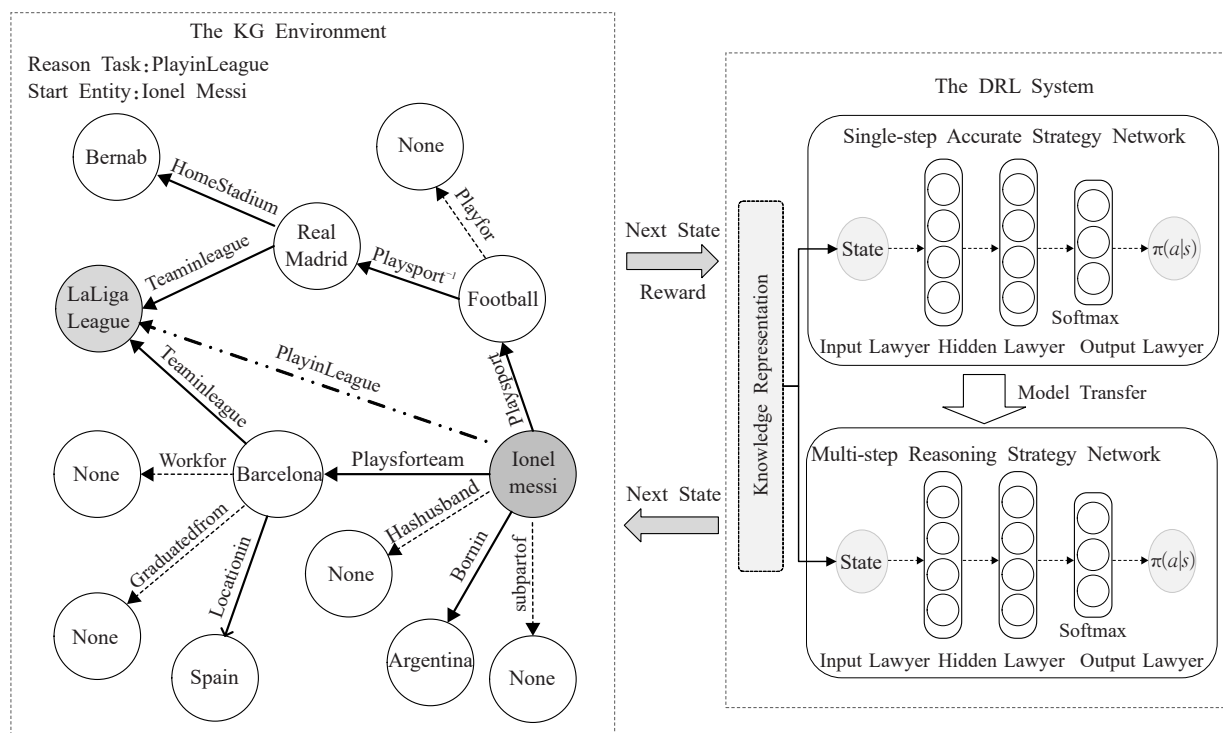


图1 融合知识表示和深度强化学习的知识推理模型框架图

Fig.1 Overall framework of integrating knowledge representation and deep reinforcement learning model for knowledge reasoning

使用基于边界的方法,定义了如下优化目标函数:

$$L = \sum_{(h,r,t) \in S} \sum_{(h',r',t') \in S'} [\gamma + d_r(h,t) - d_r(h',t')]_+ \quad (3)$$

其中, S 是正样本三元组集合; S' 是负样本三元组集合,该集合是通过随机替换正样本三元组的头实体或者尾实体得到; $[x]_+$ 表示 $\max(0, x)$; γ 表示一个边界参数,是一个需要设置为大于零的超参。

当前知识表示学习模型多样。本文选用了TransE模型,因其参数数量较少,在大规模稀疏数据集上效果明显,在与深度强化学习融合过程中,解决了知识稀疏性问题,提高了模型整体效果,实验结果证明该模型对本文方法具有支撑作用。

2.2 深度强化学习模块

本文将知识图谱中推理问题转化为马尔可夫序列决策问题,RL智能体的动作选择和状态转移都是在该框架中进行的,故本部分将介绍对知识图谱的强化学习建模过程。

该过程主要由 $\langle S, A, P, \gamma \rangle$ 四部分组成,下面将其进行详细介绍。

(1) 状态空间 S

本模型的状态空间 S 是由知识图谱中的有效实体集合 E 组成的。针对本文的双策略网络结构,设计了两种状态。在单步择优策略网络中,将每个实体作为RL智能体的具体状态 s ;在多步推理策略网络中,将当前实体和目标实体作为RL智能体的具体状态 s ,其中状态 $s \in S$ 。状态的两种表示如下:

$$s1_t = \text{TransE}(e_c) \quad (4)$$

$$s2_t = \text{TransE}(e_c, e_t) \quad (5)$$

其中, e_c 表示知识图谱中的当前实体, e_t 表示知识图谱中的目的实体, $s1_t, s2_t$ 是当前实体在状态空间中的向量化表示。为了规范化表达,如无特殊说明,本文状态均用 s 表示。

(2) 动作空间 A

RL智能体在当前状态下,经过动作选择后,基于环境的交互反馈实现状态转移。本模型的动作空间 A 是由智能体可能选择的动作集合 A_s 组成的,它是由状态 s 的当前位置实体在知识图谱 G 中所有可能的输出边组成的,动作集合表示如下:

$$A_s = \{(e_c, a, e_n) \in E: S = \{s_t\}; a \in R; e_c, e_n \in E\} \quad (6)$$

其中, e_c, e_n 分别表示当前位置实体和下一个可能的位置实体, a 表示采取的动作, S 表示状态空间, R 表示关系集合, E 表示实体集合。

特别地,关系集合 R 由知识图谱中已存在的关系 r 和新添加关系 r^{-1} 两部分组成, r^{-1} 是关系 r 的逆关系。大量实验研究表明,将 r^{-1} 关系添加到动作空间中,不仅可以使智能体自动撤销错误的决策,而且还可能发现一些隐含的推理信息。

(3) 状态转移 P

RL智能体通过选择当前状态下的最优动作实现状态转移,具体而言,智能体以知识图谱中的某个实体为当前位置,选择一个与当前实体相连的某个具体关系作为下一步执行的动作,然后执行该动作,实现智能体状态转移。状态转移 P 表示如下:

$$P = \{s' | (s, a); s \in S, a \in A\} \quad (7)$$

(4) 奖励函数 γ

RL智能体完成一次完整的任务过程后,环境都会给予智能体一定的奖励,包括正向奖励和负向奖励,智能体根据这些反馈的奖励值来更新自己的策略,以实现最大化的奖励。本文采取了多样化的奖励方式,下面将对奖励函数进行详细介绍。

全局奖励 在同环境交互过程中,智能体会有大量的可选动作,这就意味着智能体很有可能选择错误的动作,从而导致无法到达目标状态。为了解决该问题,强化学习方法添加了一个全局奖励函数。若智能体在与环境的交互过程中,从起始状态 e_{cur} 成功达到目标状态 e_{tar} ,则给予智能体一个正向奖励,否则无奖励。其定义如下:

$$\gamma_{\text{Global}}(e_{\text{cur}}) = \begin{cases} 1 & \text{if } e_{\text{cur}} = e_{\text{tar}} \\ 0 & \text{if } e_{\text{cur}} \neq e_{\text{tar}} \end{cases} \quad (8)$$

单步负向奖励 在同环境交互过程中,智能体可能选择大量错误动作,为了降低智能体的错误动作选择率,定义了单步负向奖励函数,当智能体选择的动作不能推理出目标实体时,给予负向奖励。

$$\gamma_{\text{Step_wrong}}(e_{\text{cur}}) = \begin{cases} -0.1 & \text{if } e_{\text{cur}} \neq \text{path}_{\text{tar}} \\ 0 & \text{if otherwise} \end{cases} \quad (9)$$

路径长度奖励 对于知识图谱中的推理任务,大量的研究表明:短关系路径 p 比长关系路径 p 更能提供有价值的推理关系。为了限制推理路径的长度,提高推理效率。本文定义了如下的路径长度奖励函数:

$$\gamma_{\text{path_length}} = \frac{1}{\text{length}(p)} \quad (10)$$

路径多样性奖励 为了使智能体推理出不同的关系路径,本文定义了如下的路径多样性奖励函数:

$$\gamma_{\text{DIVERSITY}} = -\frac{1}{|F|} \sum_{i=1}^{|F|} (p, p_i) \quad (11)$$

其中 $|F|$ 表示已发现的路径的数量, p 和 p_i 表示关系路径组成的表示向量。

单步择优策略网络仅使用全局奖励进行训练,多步推理策略网络则综合使用4种奖励函数进行训练,训练过程中保证正向奖励总和大于负向奖励总和。

(5) 策略神经网络

由于大规模知识图谱的关系数量众多,建模出来的强化学习方法的动作空间规模庞大,因此本文直接选择基于策略梯度的深度强化学习来完成该任务。本

文使用三层全连接神经网络设计策略函数,其中在每个隐藏层之后添加非线性层(ReLU),并使用softmax函数对输出层进行归一化处理。该策略网络实现了将状态向量 s 映射到所有选择动作的概率分布中,本方法采用REINFORCE^[30]策略进行参数优化,如下面公式所示:

$$\nabla_{\theta} J(\theta) = \nabla_{\theta} \sum_t |r_t| s_t; \theta) \gamma \quad (12)$$

其中 θ 是策略网络的参数, $\pi(a=r_t|s_t; \theta)$ 是基于状态 s_t 时策略网络输出动作为 r_t 的概率, γ 是选择该动作获得的奖励值。

本文方法中单步择优策略网络和多步推理策略网络使用了相同的神经网络结构。

2.3 训练过程

训练过程由单步择优策略网络训练和多步推理策略网络训练两部分组成。如图1所示,首先使用有监督策略学习方法对单步择优策略网络进行训练,提高RL智能体在推理过程中单步择优能力。将训练后的参数作为多步推理策略网络的初始化参数,基于奖励函数对多步推理策略网络进行再训练,提高智能体在推理任务中的多步路径择优能力。

2.3.1 基于有监督的单步择优策略网络训练

本文首先对RL智能体进行有监督策略学习的单步择优策略网络训练任务,目的是让智能体尽可能地在第一步就选择正确动作。单步择优策略网络在知识图谱推理任务中只需训练一次,极大提高了推理效率。

(1) 训练集

知识图谱中的三元组集合 $F = \{e_h, r, e_t\}$ 是RL智能体知识推理的环境。首先,为三元组集合 F 中元素添加一个反向关系,生成一个新的三元组集合 $F' = \{(e_h, r, e_t), (e_t, r', e_h)\}$ 。取出 F' 中三元组的前两部分并将相同部分合并组成二元组训练集合 $D_{\text{train}} = \{(e_h, r), (e_t, r')\}$ 。

(2) 训练流程及算法

RL智能体依次将训练集合 D_{train} 中的实体作为的起始状态 e_{cur} ,并输入到深度强化学习的策略网络中,根据策略网络的输出结果,选择一个关系 r 作为下一步的执行动作,此时判断起始状态 e_{cur} 和单步动作选择的关系 r 组成的新二元组是否属于 D_{train} ,若是,给予+1奖励并更新策略网络。预训练算法如算法1所示。

算法1 单步择优策略网络训练算法

输入: D_{train}

输出:强化学习智能体的策略网络参数

1. $\text{Entity}_{\text{embedding}}, \text{Action}_{\text{embedding}} = \text{TransE}(D_{\text{train}})$

2. for D_{train} 中的二元组 (e_{cur}, r) :

3. $s_t = \text{getEntity}_{\text{embedding}}(e_{\text{cur}})$

4. $r_{\text{cur}} = \text{policy_NN}(s_t)$

5. If $\langle s_t, r_{\text{cur}} \rangle$ in D_{train} :

6. $\gamma(e_t, r_{\text{cur}}) = 1$

7. 更新策略网络

8. $\nabla_{\theta} \log \pi(a = r_{\text{cur}}, s_t; \theta) \gamma(e_t, r_{\text{cur}})$

9. end for

2.3.2 基于奖励函数的多步推理策略网络训练

经过单步择优策略网络训练后,RL智能体对单步动作选择具有很高的成功率,但它在知识图谱环境交互过程中多步动作选择的成功率却很低。而对于知识推理任务而言,实现多步推理才是任务目标。本部分的主要目的是通过基于奖励函数的再训练,提高智能体在推理任务中的多步路径选择能力。

(1) 训练集

为了提高本模型的整体性能,对于知识图谱中的三元组集合 $F = \{e_h, r, e_t\}$,将 r 作为推理任务。针对特定的推理任务 r' ($r' \in r$),将三元组集合 F 中含关系 r' 的三元组分离出来,组成推理任务三元组集合 $T = \{e_h, r', e_t\}$,按照比例7:3分为训练集 Trainset 和测试集 Testset 。

(2) 训练流程及算法

与单步择优策略网络训练任务不同,多步推理策略网络训练的的目的是使RL智能体高效的完成多步关系路径推理任务。对于推理任务三元组集合 $T = \{e_h, r', e_t\}$ 中的三元组 (e_h, r', e_t) ,RL智能体从起始状态 e_h 出发,通过与知识图谱环境不断地交互中发现有效路径,寻找除了关系 r' 外到达目的状态 e_t 的路径。在与知识图谱环境不断地交互过程中,使用多种奖励函数对深度强化学习进行多步推理策略网络训练。训练的算法2如下所示。

算法2 多步推理策略网络训练算法

输入:训练集 Trainset

输出:强化学习智能体的策略网络参数

1. 初始化RL环境 env ,加载单步择优策略网络参数

2. for 训练集 Trainset 中的三元组 (e_h, r', e_t) :

3. $\text{cur_}s_t, \text{tar_}s_t = \text{getEntity}_{\text{embedding}}(e_h, e_t)$

4. $\text{succ} = \text{Qtrain_step} = \text{Q}$

5. while $\text{train_step} < \text{max_step}$ and $\text{succ} == 1$:

6. $a_{\text{next}} = \text{policy_NN}(\text{cur_}s_t)$

7. $\text{Reward}, \text{new_}s_t, \text{done} = \text{env.interact}((\text{cur_}s_t, \text{tar_}s_t), a_{\text{next}})$

8. $\text{train_step} += 1$

9. if $\text{new_}s_t == e_t$:

10. $\text{succ} = 1$

11. $\text{new_}s_t = \text{cur_}s_t$

12. end while

13. 更新策略网络

14. $\nabla_{\theta} \log \pi(a = r_{\text{cur}}, s_t; \theta) \gamma$

15. end for

3 实验与分析

通过公开数据验证方法有效性,并通过对比和消融

实验来进一步分析说明。代码使用Python编写,基于TensorFlow框架实现。运行环境为Ubuntu 18.04.5操作系统,Intel® Xeon Silver4210 2.20 GHz CPU和NVIDIA Tesla V100S GPU。

3.1 数据集及参数介绍

本文实验中,采用知识推理领域通用的两个基准数据集FB15K-237^[31]和NELL-995^[25]作为对比实验的实验数据集,两者都是较大数据集的子集,其中FB15K-237中的三元组是从FB15K中去除了冗余关系后得到的数据集。NELL-995是基于NELL系统的第995次迭代产生的数据集整理后的数据集。数据的统计信息如表1所示。

表1 数据集统计信息

Table 1 Statistics of the datasets

数据集	实体量	关系量	事实量	任务数
FB15K-237	14 505	237	310 116	20
NELL-995	75 492	200	154 213	12

本文方法知识表示模型TransE用Fast-TransX(<https://github.com/thunlp/Fast-TransX>)中的方法训练,嵌入维度设置为100维;策略网络的隐藏层是由两个全连接网络和ReLU激活函数构成,神经元分别设置为512和1 024,输出层节点数等于RL环境中所有关系数量;FB15K-237是474,NELL-995是400。单步择优策略网络训练任务上batchsize设置为1 000,训练1 000个epochs。

3.2 评价方式和评价指标

对于知识推理任务的评价方式,通常是链接预测(Link Prediction, LP)和事实预测(Fact Prediction, FP)。链接预测是预测三元组中缺失的部分。事实预测是在判断三元组的正确与否。数据集按7:3的比例分为训练集和初始测试集,而测试集是由初始测试集和其生成的负样本组合而成,正负样本比例约为1:10,其中负样本是由正样本被替换尾实体生成。此次实验使用平均精度均值(Mean Average Precision, MAP),定义如下:

$$MAP = \frac{1}{m} \sum_{i=1}^m \frac{1}{|t_i|} \sum_{y \in t_i} \frac{|\{rank(y') \leq rank(y); y' \in t_i\}|}{rank(y)} \quad (13)$$

其中 $rank(y)$ 和 $rank(y')$ 分别为正样本和负样本的排名, t_i 为测试集, m 为样本总数。

本文实验结果使用所有任务MAP的平均值 $\frac{\sum_T M}{|T|}$ 作为评价指标,其中 T 为任务集合。

3.3 实验结果及分析

为了验证RLPTransE方法的有效性,将本文方法RLPTransE与基于知识表示学习算法(TransE^[6]、TransH^[7]、TransR^[9]、TransD^[10])、基于关系路径算法(PRA^[16]、DIVA^[20])和基于融合算法(DeepPath^[25]、AttnPath^[29]等)在两个公开数据集上进行对比实验,实验结果如表2和表3所示。

表2 链接预测结果

Table 2 Link prediction results

算法	模型	FB15K-237	NELL-995
基于知识表示	TransE	0.532	0.737
学习算法	TransR	0.540	0.789
基于关系路径	PRA	0.541	0.675
算法	DIVA	0.598	0.886
	DeepPath	0.572	0.796
	MINERVA ^[26]	0.553	0.885
基于融合算法	DININE(MINERVA)	0.568	0.893
	AttnPath	0.585	0.834
本文算法	RLPTransE	0.616	0.801

表3 事实预测结果

Table 3 Fact prediction results

算法	模型	FB15K-237	NELL-995
	TransE	0.277	0.383
基于知识表示	TransH	0.309	0.389
学习算法	TransR	0.302	0.406
	TransD	0.303	0.413
	DeepPath	0.311	0.493
基于融合算法	DeepPath TransD	0.313	0.535
	AttnPath	0.315	0.598
本文算法	RLPTransE	0.347	0.589

3.3.1 链接预测

链接预测是指三元组在给定头实体和关系二元组 (e_h, r) 条件下预测三元组的尾实体 e_o 。本文采用DeepPath中测试方法,通过对候选尾实体打分来进行排名。实验结果由表2所示。

由表2所示的实验结果可知,本文方法在FB15K-237数据集上取得了最优的链接预测结果,性能比AttnPath高出0.031,而在NELL-995数据集上取得了优于知识表示学习方法、关系路径方法中PRA和融合方法中DeepPath的性能,但略逊色于MINERVA等其他模型的性能。本文方法在规模较大的FB15K-237数据集上性能提升更明显,主要原因是:相比于关系稀疏的NELL-995数据集,FB15K-237数据集中实体之间平均路径长度较长,动作选择的错误率更高,导致大量正确路径难以被挖掘,模型性能降低。而本文的优势在于降低错误动作的选择率,提高正确路径的挖掘成功率,因此在FB15K-237数据集上效果提升更明显。

3.3.2 事实预测

事实预测旨在判断未知事实是否为真,对于给定的三元组 (e_h, r, e_o) ,模型通过判断符合路径的个数作为分数,从而判断其正确与否。本文延用DeepPath中的评价方法,对测试集所有输出采取全排名方式计算结果,由于PRA未提供所有三元组全排名结果,因此这里不考虑PRA作为基线。表3显示了所有方法的结果。

如表3所示,本文方法同样在FB15K-237数据集上的事实预测结果达到了最优性能,性能比AttnPath高出

0.032,而在NELL-995上取得优于其他模型,但略逊色于AttnPath的结果。对于FB15K-237数据集,本文方法比知识表示学习方法中性能最好的TransH高出0.038。对于NELL-995数据集,本文方法比知识表示学习方法中性能最好的TransD高出0.176。相比于融合模型中的DeepPath及其衍生模型,本文方法性能均有较大提升。

3.4 消融实验

为了进一步分析单步择优策略网络训练对本文模型推理效果的影响,本节对RLPTransE做了消融实验分析。将RL智能体直接从多步推理策略网络开始推理任务,即在原方法的基础上去掉单步择优策略网络训练任务,得到模型RLPTransE-part,使用和RLPTransE相同的训练集和测试集对该模型进行训练和测试。消融实验的结果如表4所示。两个模型的测试次数统一设为1。

表4 预训练消融实验结果

Table 4 Pre-training ablation experimental results

模型	单步择优策略网络	多步推理策略网络	MAP(LP/FP)
RLPTransE-part	×	√	0.727/0.471
RLPTransE	√	√	0.801/0.589

如表4所示,RLPTransE-part模型在链接预测和事实预测的实验结果都不及RLPTransE模型,主要原因是缺乏单步择优策略网络训练的方法,单步选择有效路径能力不足,路径的搜索能力弱,直接影响多步推理效果。

表6 RL智能体发现的推理规则

Table 6 Inference formulas found by RL Agent

关系	推理规则
athlethomestadium	playsforteam→teamhomestadium
	belongstoorganization→leaguestadiums
	ledsportsteam→teamplayssport→sportusesstadium
athleteplaysforteam	athlethomestadium→teamhomestadium_inv
	athleteplayssport→teamplayssport_inv
	athleteplaysinleague→belongstoleague_inv→belongstoorganization
athleteplaysinleague	playsforteam→teamplaysinleague
	athlethomestadium→leaguestadiums_inv
	playssport→teamplayssport_inv→teamplaysinleague
	playsforteam→teamplaysagainstteam→teamplaysinleague
athleteplayssport	playsforteam→teamplayssport
	homestadium→sportusesstadium_inv
	playsinleague→teamplaysinleague_inv→teamplayssport
personborninlocation	haswife→personbornincity
	hashusband→personbornincity
	personbornincity→mutualproxyfor_inv
	graduatedfromuniversity→graduatedfromuniversity_inv→personbornincity
teamplaysinleague	teamhomestadium→leaguestadiums_inv
	subpartoforganization→agentactsinlocation→leaguestadiums_inv
	teamplaysincity→proxyfor→leaguestadiums_inv
teamplayssport	teamhomestadium→sportusesstadium_inv
	teamplaysincity→stadiumlocatedincity_inv→sportusesstadium_inv
	teamplaysinleague→athleteplaysinleague_inv→athleteplayssport

因此,实验结果表明,引入单步择优策略网络对推理任务的完成具有明显的提升效果。

为了进一步分析单步择优策略网络在本方法中的重要性。本文对该网络的单步择优能力作了对比实验,使用不同训练集对单步择优策略网络进行训练,然后对网络的单步选择成功率进行统计。定义如下三种模型:All-train,使用 D_{train} 的样本集合 $\{(e_h, r), (e_t, r)\}$ 训练;Pos-train,使用 D_{train} 中正向样本集合 $\{(e_h, r)\}$ 训练;No-train,不使用样本集合训练,即使用初始化参数。不同epoch次数下的实验结果如表5所示。

表5 对比实验结果

Table 5 Comparative experimental results

模型	10	20	30	40	50	平均值
All-train	0.940 8	0.940 6	0.940 6	0.940 1	0.941 2	0.940 7
Pos-train	0.786 0	0.786 7	0.786 2	0.786 4	0.786 1	0.786 3
No-train	0.009 3	0.009 5	0.009 2	0.009 5	0.009 3	0.009 3

如表5所示,使用 D_{train} 训练的All-train在单步选择成功率上的均值达到了94.07%,Pos-train结果达到了78.65%,而No-train的选择成功率几乎为0。该实验充分验证单步择优策略网络对本方法具有重要支撑作用。

3.5 推理规则的分析

为了分析本文方法对知识推理可解释性的增强作用,从RL智能体在NELL-995数据集上挖掘出的推理规则中选择部分任务结果,如表6所示。

如表6所示,对于任务关系“athletheplaysinleague”,对应的推理规则为“playsforteam→teampaysinleague”和“athlethehomestadium→leaguestadiums_inv”,即运动员效力的球队所属的联赛就是运动员效力的联赛和运动员主场所注册的联赛就是运动员效力的联赛。其他任务的说明类似,不再逐一展开分析。因此,分析结果表明,本文方法对于增强知识推理的可解释性具有重要支撑。

4 结束语

本文提出了一种融合知识表示和深度强化学习方法RLPTransE。该模型通过知识表示方法,将知识图谱映射到含有三元组语义信息的向量空间中,然后在该空间建立强化学习环境,将知识推理成功转化为马尔可夫序列决策过程。基于有监督的单步择优策略网络的训练和基于奖励函数的多步推理策略网络的训练,使得RL智能体在推理过程中挖掘出高质量推理规则,从而完成大规模知识图谱推理任务。在公开数据集上的对比实验表明,本文方法提升了推理性能,特别是大规模知识图谱推理任务。本文还通过消融实验对单步择优策略网络对本文方法的影响做了进一步分析。通过对挖掘出来的推理规则分析,验证了本文方法对知识推理可解释性具有增强作用。

参考文献:

- [1] BOLLACKER K, EVANS C, PARITOSH P, et al. Freebase: A collaboratively created graph database for structuring human knowledge[C]//Proceedings of 2008 ACM SIGMOD International Conference on Management of Data, 2008: 1247-1250.
- [2] AUER S, BIZER C, KOBILAROV G, et al. DBpedia: A nucleus for a Web of open data[C]//Proceedings of the 6th Int'l Semantic Web Conference, 2007: 722-735.
- [3] CARLSON A, BETTERIDGE J, KISIEL B, et al. Toward an architecture for never-ending language learning[C]//Proceedings of the 24th AAAI Conference on Artificial Intelligence, 2010: 1306-1313.
- [4] 张仲伟, 曹雷, 陈希亮, 等. 基于神经网络的知识推理研究综述[J]. 计算机工程与应用, 2019, 55(12): 8-19.
ZHANG Z W, CAO L, CHEN X L, et al. Survey of knowledge reasoning based on neural networks[J]. Computer Engineering and Applications, 2019, 55(12): 8-19.
- [5] 官赛萍, 靳小龙, 贾岩涛, 等. 面向知识图谱的知识推理研究进展[J]. 软件学报, 2018, 29(10): 2966-2994.
GUAN S P, JIN X L, JIA Y T, et al. Knowledge reasoning over knowledge graph: A survey[J]. Journal of Software, 2018, 29(10): 2966-2994.
- [6] BORDES A, USUNIER N, GARCIA-DURAN A, et al. Translating embeddings for modeling multi-relational data[C]//Advances in Neural Information Processing Systems, 2013: 2787-2795.
- [7] WANG Z, ZHANG J, FENG J, et al. Knowledge graph embedding by translating on hyperplanes[C]//Proceedings of the 28th Conference on Artificial Intelligence(AAAI), 2014: 1112-1119.
- [8] 阮小芸, 廖健斌, 李祥, 等. 基于人才知识图谱推理的强化学习可解释推荐研究[J]. 数据分析与知识发现, 2020, 5(6): 36-50.
RUAN X Y, LIAO J B, LI X, et al. Interpretable recommendation of reinforcement learning based on talent knowledge graph reasoning[J]. Data Analysis and Knowledge Discovery, 2020, 5(6): 36-50.
- [9] LIN Y, LIU Z, SUN M, et al. Learning entity and relation embeddings for knowledge graph completion[C]//Proceedings of the 29th Conference on Artificial Intelligence, 2015: 2181-2187.
- [10] JI G, HE S, XU L, et al. Knowledge graph embedding via dynamic mapping matrix[C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, 2015: 687-696.
- [11] EBISU T, ICHISE R. Toruse: Knowledge graph embedding on a lie group[C]//Proceedings of the 32nd AAAI Conference on Artificial Intelligence, 2018: 1819.
- [12] SUN Z Q, DENG Z H, NIE J Y, et al. Rotate: Knowledge graph embedding by relational rotation in complex space[J]. arXiv: 1902.10197, 2019.
- [13] ZHANG S, TAY Y, YAO L N, et al. Quaternion knowledge graph embeddings[C]//Proceedings of the 33rd Conference on Neural Information Processing Systems, 2019: 2735.
- [14] 刘峤, 李杨, 段宏, 等. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(3): 582-600.
LIU Q, LI Y, DUAN H, et al. Knowledge graph construction techniques[J]. Journal of Computer Research and Development, 2016, 53(3): 582-600.
- [15] LAO N, COHEN W W. Relational retrieval using a combination of path-constrained random walks[J]. Machine Learning, 2010, 81(1): 53-67.
- [16] LAO N, MITCHELL T, COHEN W W. Random walk inference and learning in a large scale knowledge base[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 2011: 529-539.
- [17] GARDNER M, MITCHELL T. Efficient and expressive knowledge base completion using subgraph feature extraction[C]//Proceedings of the Conference on Empirical

- Methods in Natural Language Processing,2015;1488-1498.
- [18] GARDNER M, TALUKDAR P, KRISHNAMURTHY J, et al. Incorporating vector space similarity in random walk inference over knowledge bases[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing(EMNLP), 2014;397-406.
- [19] DAS R, NEELAKANTAN A, BELANGER D, et al. Chains of reasoning over entities, relations, and text using recurrent neural networks[J].arXiv:1607.01426, 2016.
- [20] CHEN W, XIONG W, YAN X, et al. Variational knowledge graph reasoning[J].arXiv:1803.06581, 2018.
- [21] LIN Y, LIU Z, LUAN H, et al. Modeling relation paths for representation learning of knowledge bases[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015;705-714.
- [22] 陈海旭, 周强, 刘学军. 一种结合路径信息和嵌入模型的知识推理方法[J]. 小型微型计算机系统, 2020, 41(6): 1147-1151.
- CHEN H X, ZHOU Q, LIU X J. Knowledge graph reasoning combining path information and embedding model[J]. Journal of Chinese Computer Systems, 2020, 41(6):1147-1151.
- [23] JIA Y, WANG Y, JIN X, et al. Path-specific knowledge graph embedding[J]. Knowledge-Based Systems, 2018, 151:37-44.
- [24] WU Y B, ZHU D H, LIAO X W, et al. Knowledge graph reasoning based on paths of tensor factorization[J]. Pattern Recognition and Artificial Intelligence, 2017, 30(5):473-480.
- [25] XIONG W, HOANG T, WANG W Y. DeepPath: A reinforcement learning method for knowledge graph reasoning[J].arXiv:1707.06690, 2017.
- [26] DAS R, DHULIAWALA S, ZAHEER M, et al. Go for a walk and arrive at the answer: Reasoning over paths in knowledge bases using reinforcement learning[J].arXiv:1711.05851, 2017.
- [27] LIN X V, SOCHER R, XIONG C. Multi-hop knowledge graph reasoning with reward shaping[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018.
- [28] LI R, CHENG X. DIVINE: A generative adversarial imitation learning framework for knowledge graph reasoning[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing(EMNLP-IJCNLP), 2019.
- [29] WANG H, LI S, PAN R, et al. Incorporating graph attention mechanism into knowledge graph reasoning based on deep reinforcement learning[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing(EMNLP-IJCNLP), 2019.
- [30] WILLIAMS R J. Simple statistical gradient-following algorithms for connectionist reinforcement learning[J]. Machine Learning, 1992, 8(3/4):229-256.
- [31] TOUTANOVA K, CHEN D, PANTEL P, et al. Representing text for joint embedding of text and knowledge bases[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015;1499-1509.