



计算机科学

COMPUTER SCIENCE

QubitE:用于知识图谱补全的量子嵌入模型

林学渊, 鄂海红, 宋文字, 罗浩然, 宋美娜

引用本文

林学渊, 鄂海红, 宋文字, 罗浩然, 宋美娜. [QubitE:用于知识图谱补全的量子嵌入模型](#)[J]. 计算机科学, 2023, 50(11): 201-209.

LIN Xueyuan, E Haihong, SONG Wenyu, LUO Haoran, SONG Meina. [QubitE:Qubit Embedding for Knowledge Graph Completion](#) [J]. Computer Science, 2023, 50(11): 201-209.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于贝叶斯规则的具有层次注意力的知识补全](#)

Bayesian Rule-based Knowledge Completion with Hierarchical Attention

计算机科学, 2023, 50(11): 234-240. <https://doi.org/10.11896/jsjcx.221000056>

[路网拓扑感知的轨迹表示学习方法](#)

Road Network Topology-aware Trajectory Representation Learning

计算机科学, 2023, 50(11): 114-121. <https://doi.org/10.11896/jsjcx.221000058>

[基于对比学习的多关系属性图聚类方法](#)

Clustering Method Based on Contrastive Learning for Multi-relation Attribute Graph

计算机科学, 2023, 50(11): 62-70. <https://doi.org/10.11896/jsjcx.220900166>

[基于知识增强的企业实体关系预测模型](#)

Knowledge Enhanced Relationship Prediction Model for Enterprise Entities

计算机科学, 2023, 50(10): 146-155. <https://doi.org/10.11896/jsjcx.221000063>

[基于复合语义特征的事件图谱构建技术研究进展](#)

Overview About Composite Semantic-based Event Graph Construction

计算机科学, 2023, 50(9): 242-259. <https://doi.org/10.11896/jsjcx.230400046>

QubitE:用于知识图谱补全的量子嵌入模型

林学渊 鄂海红 宋文宇 罗浩然 宋美娜

北京邮电大学计算机学院 北京 100876

(linxy59@bupt.edu.cn)

摘要 知识图谱补全任务通过预测知识图谱中缺失的事实补全知识图谱。基于量子的知识图谱嵌入(KGE)模型利用变分量子电路,通过测量量子比特状态的概率分布对三元组进行评分,评分高的三元组即为缺失的事实。但是目前基于量子的KGE要么在优化过程中失去了量子优势,矩阵酉性被破坏,要么需要大量参数用于存储量子态,从而导致过拟合和低性能。此外,这些方法忽略了对于理解模型性能必不可少的理论分析。为了解决性能问题和弥合理论差距,提出了QubitE模型:将实体嵌入作为量子位(单位复向量),将关系嵌入作为量子门(酉复矩阵),评分过程为复矩阵乘法,利用核方法进行优化。该模型的参数化方式能在优化中保持量子优势,时空复杂度为线性,甚至可以进一步实现基于语义的量子逻辑计算。此外,从理论上可以证明该模型具有完全表达性、关系模式推理能力和包含性等,有助于理解模型性能。实验表明,QubitE在一些基准知识图谱上可以取得与最先进的经典模型相当的结果。

关键词: 知识图谱;知识图谱补全;知识图谱嵌入;表示学习;量子比特

中图法分类号 TP391

QubitE:Qubit Embedding for Knowledge Graph Completion

LIN Xueyuan, E Haihong, SONG Wenyu, LUO Haoran and SONG Meina

School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract The knowledge graph completion task completes the knowledge graph by predicting missing facts in the knowledge graph. The quantum-based knowledge graph embedding(KGE) model uses variational quantum circuits to score triples by measuring the probability distribution of qubit states, and triples with high scores are the missing facts. But the current quantum-based KGE either loses the quantum advantage in the optimization process and the matrix unitary property is destroyed, or requires a large number of parameters for storing quantum states, resulting in overfitting and low performance. Furthermore, these methods ignore the theoretical analysis that is essential for understanding model performance. In order to solve the performance problem and bridge the theoretical gap, we propose QubitE: entities are embedded as qubits(unit complex vectors), relations are embedded as quantum gates(unit unitary matrices), the scoring process is complex matrix multiplication, and kernel methods are used for optimization. The parameterization method of the model can maintain the quantum advantage in optimization, the space-time complexity is linear, and it can even further realize semantic-based quantum logic calculation. In addition, the model can be proved to be fully expressive, relational schema reasoning ability and inclusiveness, etc. theoretically, which is helpful to understand the model performance. Experiments show that QubitE can achieve results comparable to state-of-the-art classical models on some benchmark knowledge graphs.

Keywords Knowledge graph, Knowledge graph completion, Knowledge graph embedding, Representation learning, Qubit

1 引言

知识图谱(Knowledge Graph, KG)由节点(实体)和边(实体之间的关系)组成,已被广泛应用于知识驱动的任务,如问答、推荐系统和搜索引擎。然而,知识图谱是不完整的,这个问题会影响任何与知识图谱相关的算法的性能。知识图谱

嵌入(Knowledge Graph Embedding, KGE)将实体和关系嵌入到低维向量空间并对候选三元组进行评分,评分高的三元组即为缺失边。

基于量子的KGE是量子力学在知识表示学习领域的应用,但目前的研究还处于起步阶段。Ma等^[1]提出了最经典的基于量子的KGEs,包括两种类型的变分量子电路。

到稿日期:2022-11-25 返修日期:2023-02-09

基金项目:国家自然科学基金(62176026,61902034);北京市自然科学基金(M22009)

This work was supported by the National Natural Science Foundation of China(62176026,61902034) and Natural Science Foundation of Beijing, China(M22009).

通信作者:鄂海红(ehaihong@bupt.edu.cn)

(1)QCE,将实体的隐特征视为量子态的参数,将关系的隐特征视为作用于量子态的参数门。三元组的评分取决于对量子态的测量。然而,测量会导致信息丢失。量子优势,即根据量子力学概率解释的量子态和量子门的归一化约束,在优化模型时消失了。(2)F-QCE,从作用于纯态的参数化量子门生成实体嵌入,可以在保留量子优势的同时有效地训练量子嵌入。然而,它不得不面对参数爆炸的问题,因为准备多量子位进行推理的成本很高。此外,这两种类型在知识图补全(Knowledge Graph Completion, KGC)任务上的性能都很差。

为了解决性能问题,本文探索了一种基于量子的参数化方法进行嵌入,希望能够保留量子优势,实现轻量级和高性能。其中保留量子优势尤为重要。首先,借助量子优势,可以在不涉及关系的情况下通过预定义的量子门对实体的语义进行操作。这意味着模型可以感知更深层次的、关系无关的、实体特定的语义信息。其次,它允许基于语义的重新编程,利用否运算、交运算、并运算等逻辑运算创建新实体。与以前的经典KGE相比,这是新特性。最后,量子优势简化了模型,更容易从理论上对其进行研究。总而言之,探索具有物理解释的新方法是有价值的。

本文提出了一种新的基于量子的KGE,即QubitE。实体嵌入向量被视为量子态的参数,同时通过激活函数保持为单位复向量,从而保留实体的量子优势。关系被建模为参数化量子门,即酉复矩阵,保留关系的量子优势。关系作用于实体的过程等价于复矩阵乘法,其通过核方法进行参数优化,因此具有线性的计算复杂度。

此外,本文从理论上分析了包含性、完整表达性、模式推理以及空间和时间复杂度。首先证明了QubitE是完全表达的。这表明该模型可以很好地区分正负三元组。然后证明QubitE包含了TransE, RotatE, pRotatE, ComplEx和DistMult。此外,还证明了QubitE能够学习到组合、逆和对称等关系模式。最后,证明了QubitE具有线性空间复杂度和线性时间复杂度。

本文的贡献总结如下:

(1)模型设计:提出了QubitE,这是一种新的基于量子的线性KGE模型,用于知识图谱上的链接预测,轻量级、简单且表达能力强。

(2)理论贡献:全面分析了模型的包含性、完整表达性、模式推理和时空复杂度。

(3)实验:在4个标准公共数据集上进行了大量实验,其实验结果证明了模型的有效性。

2 相关工作

知识图谱嵌入模型分为4种:欧几里得嵌入、非欧几里得嵌入、张量分解嵌入和量子嵌入。

(1)欧几里得嵌入。TransE^[2]将关系建模为从头实体到尾实体的距离变换;TransR^[3]提出为每个关系设计一个投影矩阵,以便实体在不同关系下有不同的嵌入向量;RotatE^[4]将关系定义为二维复空间中头实体到尾实体的旋转变换;QuatE^[5]利用四元数的方法将旋转扩展到三维复空间;5 * E^[6]提出了一种基于射影几何的模型,为同时表示平移、

旋转、同态、反转和反射提供了统一的方法。

(2)非欧几里得嵌入。MuRP^[7]在双曲空间和欧几里得空间中建模,结合关系向量,可以处理图中存在的多种类型的关系;ATTH^[8]利用双曲空间的表现力和基于注意力的几何变换来学习低维空间中改进的KG表示。

(3)张量分解嵌入。DistMult^[9]放宽了对关系矩阵的约束,使用对角矩阵表示关系矩阵;ComplEx^[10]扩展到复空间,可以同时求解对称和非对称关系;Simple^[11]提出了一个简单的Canonical Polyadic(CP)增强,允许每个实体的两个嵌入被依赖地学习;HyperER^[12]使用超图网络为每个关系生成一维卷积滤波器,以提取关系的特定特征;TuckerER^[13]提出使用Tucker分解对KG的二进制张量表示进行链接预测。

(4)量子嵌入。Ma等^[1]提出了两种类型的变量子电路(QCE和F-QCE)用于知识图谱嵌入。Lloyd等^[14]提出了一种量子嵌入模型,该模型通过量子特征图将经典数据点表示为希尔伯特空间中的量子态。经典数据点 x 被转换为量子电路 ψ 中的一组门参数,创建一个量子状态 $|x\rangle$ 使得 $\psi: x \rightarrow |x\rangle$ 。

本文方法属于量子嵌入方法,但是与之前已有的方法大不相同。首先,本文通过嵌入向量上的距离函数而不是测量量子比特状态的概率分布来区分量子态。其次,KG中的实体被直接分配可训练参数来构建量子态,而不是使用参数量子电路。这些设计使得本文模型不依赖于量子模拟器,成本更低。

3 方法

3.1 模型公式化

KG是多关系有向图 $\mathcal{KG}=(\mathcal{E}, \mathcal{R}, \mathcal{T})$,其中 \mathcal{E} 是节点集(实体), \mathcal{R} 是边集(实体之间的关系)。集合 $\mathcal{T}=\{(h, r, t)\} \subseteq \mathcal{E} \times \mathcal{R} \times \mathcal{E}$ 包含所有形如(head, relation, tail)的三元组,如(smart-Phone, hypernym, iPhone)。为了在KG上应用嵌入学习方法,KGE学习实体(\mathcal{E})和关系(\mathcal{R})的向量表示。令 h, r, t 表示三元组 (h, r, t) 的向量表示,其中 $h, t \in \mathbb{V}^{d_e}, r \in \mathbb{V}^{d_r}$ (\mathbb{V}^d 是一个 d 维向量空间)。TransE考虑 $\mathbb{V}=\mathbb{R}$,而ComplEx和RotatE使用 $\mathbb{V}=\mathbb{C}$ (复空间)和QuatE考虑 $\mathbb{V}=\mathbb{C}^2$ (四元数空间)。本文选择二维希尔伯特空间,即 $\mathbb{V}=\mathbb{C}^2$ 。大多数KGE模型都是通过关系特定的转换函数 g_r 定义的: $\mathbb{V}^{d_e} \rightarrow \mathbb{V}^{d_e}$,它将头实体映射到尾实体,即 $g_r(h)=t$ 。在转换函数上定义评分函数 $f: \mathbb{V}^{d_e} \times \mathbb{V}^{d_e} \times \mathbb{V}^{d_r} \rightarrow \mathbb{R}$,衡量三元组的合理性: $f(h, r, t)=p(g_r(h), t)$ 。通常,评分函数的公式可以取 $p(g_r(h), t)=-\|g_r(h)-t\|$ 或 $p(g_r(h), t)=\langle g_r(h), t \rangle$ 。

本文中,给定一个三元组 (h, r, t) ,头实体和尾实体 h, t 被嵌入到一个 d 维希尔伯特空间即 $h, t \in \mathbb{C}^{2d}$,向量中每个元素都是一个二维复向量。关系 $r \in \mathcal{R}$ 嵌入为 d 维向量 r ,其中每个元素都是 2×2 复酉矩阵。 r 包含两个复向量 r_a 和 $r_b \in \mathbb{C}^d$ 。 $r_a, r_b, h_a, h_b, t_a, t_b$ 的第 i 个元素分别记成 $r_{ai}, r_{bi}, h_{ai}, h_{bi}, t_{ai}, t_{bi}$ 。

3.2 特定于实体的量子比特嵌入

量子比特:经典比特可以表示为0和1这两种状态。量子比特即量子位不仅可以是这两种离散状态之一,而且可以

是这两者所有可能的线性叠加。在数学上,量子位的量子态表示为二维希尔伯特空间 \mathbb{C}^2 中的一个状态向量,其基向量的狄拉克符号表示为:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

令向量 $|0\rangle$ 对应于经典值0,而 $|1\rangle$ 对应于1。一个量子比特状态向量可写为:

$$|\psi\rangle = \mathbf{a}|0\rangle + \mathbf{b}|1\rangle$$

其中, $\mathbf{a}, \mathbf{b} \in \mathbb{C}$, $|\mathbf{a}|^2 + |\mathbf{b}|^2 = 1$ 。复数 \mathbf{a} 和 \mathbf{b} 称为量子振幅。根据量子力学,对 $|\psi\rangle$ 进行测量,看它是在 $|0\rangle$ 还是 $|1\rangle$ 中,结果是0(1)的概率为 $|\mathbf{a}|^2$ ($|\mathbf{b}|^2$),对应状态为 $|0\rangle$ ($|1\rangle$)。状态 $|\psi\rangle$ 的密度矩阵 ρ 由下式给出:

$$\rho = |\psi\rangle\langle\psi|$$

本文使用量子比特状态的标准表示法来表示 \mathbb{C}^{2d} 中的实体。实体嵌入向量 \mathbf{h} 的第 i 个元素由下式给出:

$$\mathbf{h}_i = \mathbf{h}_{ai}|0\rangle + \mathbf{h}_{bi}|1\rangle = \begin{pmatrix} \mathbf{h}_{ai} \\ \mathbf{h}_{bi} \end{pmatrix}, i=1, 2, \dots, d$$

其中, d 是实体嵌入维度, $\mathbf{h}_{ai}, \mathbf{h}_{bi} \in \mathbb{C}$ 和 $|\mathbf{h}_{ai}|^2 + |\mathbf{h}_{bi}|^2 = 1$ 使得 $\mathbf{h} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_d]$ 。相对应实体 h 的密度矩阵为:

$$\rho_{h_i} = |\mathbf{h}_i\rangle\langle\mathbf{h}_i| = \begin{pmatrix} |\mathbf{h}_{ai}|^2 & \mathbf{h}_{ai}\mathbf{h}_{bi}^* \\ \mathbf{h}_{bi}\mathbf{h}_{ai}^* & |\mathbf{h}_{bi}|^2 \end{pmatrix}$$

3.3 特定于关系的量子门

量子门:将系统从一种状态转换为另一种状态。当不进行测量时,状态的时间演化由薛定谔方程描述。由于量子力学的概率解释,状态向量归一化为1。因此,时间演化是酉的。量子门 U 有 $U U^\dagger = U^\dagger U = \mathbf{I}$,其中 U^\dagger 是矩阵 U 的共轭转置。 2×2 酉矩阵的一般表达式是:

$$U = \begin{pmatrix} \mathbf{a} & -e^{i\psi}\mathbf{b}^* \\ \mathbf{b} & e^{i\psi}\mathbf{a}^* \end{pmatrix}$$

其中, $\mathbf{a}, \mathbf{b} \in \mathbb{C}$, $|\mathbf{a}|^2 + |\mathbf{b}|^2 = 1$, ψ 是幅角。 \mathbf{a}^* 是 \mathbf{a} 的复共轭。

本文使用特定于关系的转换将头实体 \mathbf{h} 从源空间映射到目标希尔伯特空间。由于量子门是酉的,因此可以将关系嵌入向量 \mathbf{r} 的第 i 个元素的参数化酉矩阵写为:

$$\mathbf{r}_i = \mathbf{U}_{ri} = \begin{pmatrix} \mathbf{r}_{ai} & -e^{i\psi}\mathbf{r}_{bi}^* \\ \mathbf{r}_{bi} & e^{i\psi}\mathbf{r}_{ai}^* \end{pmatrix}, i=1, 2, \dots, d$$

其中, d 是关系嵌入维度, $\mathbf{r}_{ai}, \mathbf{r}_{bi} \in \mathbb{C}$ 和 $|\mathbf{r}_{ai}|^2 + |\mathbf{r}_{bi}|^2 = 1$ 使得 $\mathbf{r} = [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_d]$ 。这意味着 $\det(\mathbf{U}_{ri}) = e^{i\psi} \neq 0$,即 \mathbf{U}_{ri} 是可逆的。

为了将量子门作用于量子位,即将关系特定的变换 \mathbf{r} 作用于头实体 \mathbf{h} ,本文通过矩阵乘法执行逐元素变换以计算变换后的实体表示 \mathbf{h}_r :

$$\mathbf{h}_{ri} = g_{ri}(\mathbf{h}_i) = \mathbf{U}_{ri}\mathbf{h}_i = \begin{pmatrix} \mathbf{r}_{ai}\mathbf{h}_{ai} - e^{i\psi}\mathbf{r}_{bi}^*\mathbf{h}_{bi} \\ \mathbf{r}_{bi}\mathbf{h}_{ai} + e^{i\psi}\mathbf{r}_{ai}^*\mathbf{h}_{bi} \end{pmatrix}, i=1, 2, \dots, d$$

其中, $\mathbf{h}_r = [\mathbf{h}_{r1}, \mathbf{h}_{r2}, \dots, \mathbf{h}_{rd}]$ 。

3.4 评分函数

表1列出了几种流行的KGE的评分函数和参数,其中,*表示循环相关运算; \cdot 表示Hadamard(或按元素)积; \otimes 表示Hamilton积。TransE, HolE和DistMult使用欧几里德嵌入,而ComplEx和RotatE使用复数空间,QuatE使用四元数

空间。相比之下,本文模型使用超复数空间中的量子态和量子门。

表1 最先进的KGE的评分函数

Table 1 Scoring functions of state-of-the-art KGEs

模型	评分函数	参数
TransE	$\ (\mathbf{h} + \mathbf{r}) - \mathbf{t}\ $	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^d$
HolE	$\langle \mathbf{r}, \mathbf{h} * \mathbf{t} \rangle$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^d$
DistMult	$\langle \mathbf{r}, \mathbf{h} \cdot \mathbf{t} \rangle$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^d$
ComplEx	$Re(\langle \mathbf{r}, \mathbf{h}, \tilde{\mathbf{t}} \rangle)$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{C}^d$
RotatE	$\ \mathbf{h} \circ \mathbf{r} - \mathbf{t} \ $	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{C}^d$
QuatE	$\mathbf{h} \otimes \mathbf{r} \cdot \mathbf{t}$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{H}^d$
$5 * E$	$\left\ \begin{pmatrix} \mathbf{r}_1\mathbf{h} + \mathbf{r}_2 \\ \mathbf{r}_3\mathbf{h} + \mathbf{r}_4 \end{pmatrix} - \mathbf{t} \right\ $	$\mathbf{h}, \mathbf{t}, \mathbf{r}_1 \dots \mathbf{r}_4 \in \mathbb{C}^d$
QubitE	$\langle \mathbf{U}, \mathbf{h}, \mathbf{t} \rangle$	$\mathbf{h}, \mathbf{t} \in \mathbb{C}^{2d}$ $\mathbf{U}_r \in \mathbb{C}^{2 \times 2 \times d}$

传统的基于量子的方法^[1]必须在训练时测量状态,而本文通过对状态隐藏参数的相似性进行评分来区分状态。KG中三元组的得分是关系特定变换后的头 \mathbf{h}_r 和尾 \mathbf{t} 之间的相似度 $\langle \mathbf{h}_r, \mathbf{t} \rangle$ 。参数优化过程为最小化所预测实体 \mathbf{h}_r 和尾实体 \mathbf{t} 之间的距离,即最大化相似度 $\langle \mathbf{h}_r, \mathbf{t} \rangle$ 。相应地,对负三元组则最小化相似度。为了让相似度 $\langle \mathbf{h}_r, \mathbf{t} \rangle$ 符合量子解释,考虑以下定义。

(1) 迹距离

通常,迹距离测度量化两个量子态以相同方式表现的程度。迹距离越小,两个状态越相似。将相似度定义为迹距离的负值。

$$f(\mathbf{h}, \mathbf{r}, \mathbf{t}) = -\frac{1}{2} \text{tr}(\sqrt{(\rho_{h_r} - \rho_t)^\dagger (\rho_{h_r} - \rho_t)})$$

其中, ρ_{h_r}, ρ_t 分别是状态 $|\mathbf{h}_r\rangle$ 和 $|\mathbf{t}\rangle$ 的密度矩阵, $\text{tr}(\rho)$ 是密度矩阵 ρ 的迹, ρ^\dagger 是 ρ 的共轭转置。

(2) 希尔伯特-施密特距离

一般称两个状态之间的Hilbert-Schmidt距离为 l_2 距离,而 l_1 距离即是迹距离。同样,本文将相似度定义为Hilbert-Schmidt距离的负值。

$$f(\mathbf{h}, \mathbf{r}, \mathbf{t}) = -\text{tr}((\rho_{h_r} - \rho_t)^\dagger (\rho_{h_r} - \rho_t))$$

本文还探索了更多可能有助于训练过程的定义。按元素 l_1 距离和按元素内积是继承自经典KGEs的两个测度。

(3) 按元素 l_1 距离

$$f(\mathbf{h}, \mathbf{r}, \mathbf{t}) = -\|\mathbf{h}_r - \mathbf{t}\|_1 = -\sum_{i=1}^d \|\mathbf{h}_{ri} - \mathbf{t}_i\|_1$$

其中, $\|\mathbf{x}\|_1$ 是二维复向量 $\mathbf{x} \in \mathbb{C}^{2d}$ 的 l_1 模。

(4) 按元素内积

$$f(\mathbf{h}, \mathbf{r}, \mathbf{t}) = Re(\langle \mathbf{h}_r, \tilde{\mathbf{t}} \rangle)$$

其中, $Re(\mathbf{x})$ 是二维复向量 $\mathbf{x} \in \mathbb{C}^{2d}$ 的实部, $\langle \mathbf{h}_r, \tilde{\mathbf{t}} \rangle$ 是按元素内积。

3.5 损失函数

本文将链接预测任务公式化为分类问题,最小化以下损失函数:

$$\text{Loss} = -\log(\gamma - f(\mathbf{h}, \mathbf{r}, \mathbf{t})) - \sum_{i=1}^K p(h_i, r_i, t_i) \log \sigma(f(h_i, r_i, t_i) - \gamma)$$

其中, γ 是固定边距, K 是负样本的数量, (h_i, r_i, t_i) 是第 i 个负三元组, σ 是sigmoid函数。此外, $p(h_i, r_i, t_i)$ 是采样负样

本的分布,它依赖于均匀采样、伯努利采样和对抗采样等负采样策略^[4]。

3.6 初始化

本文采用特定的初始化算法来保留量子优势并提高模型效率和加快其收敛速度^[15]。实体的初始化遵循以下规则:

$$\begin{aligned} \mathbf{a}_{\text{real}} &= \cos(\theta) \\ \mathbf{a}_{\text{img}} &= \sin(\theta)\cos(\phi) \\ \mathbf{b}_{\text{real}} &= \sin(\theta)\sin(\phi)\cos(\varphi) \\ \mathbf{b}_{\text{img}} &= \sin(\theta)\sin(\phi)\sin(\varphi) \end{aligned}$$

其中, $\mathbf{a}_{\text{real}}, \mathbf{a}_{\text{img}}, \mathbf{b}_{\text{real}}, \mathbf{b}_{\text{img}}$ 分别表示 \mathbf{a} 和 \mathbf{b} 的实部和虚部。 θ, ϕ, φ 随机采样于区间 $[-\pi, \pi]$ 。关系的初始化也遵循类似的规则。 \mathbf{a} 和 \mathbf{b} 的参数按上述规则初始化,而幅角 ψ 随机采样于区间 $[-\pi, \pi]$ 。这种初始化方法是可选的。

4 理论分析

为了探究本文方法和经典 KGE 方法的关系,有以下命题。

命题 1 量子表示等价于单位四元数表示。此时,某些特定的量子门可被视为四元数空间中的旋转。

对于每个量子比特表示,有 4 个自由变量归一化为 1。存在自然的一对一映射 ϕ :

$$\begin{aligned} \phi: \mathbb{C}^2 &\rightarrow \mathbb{H} \\ (a+bi)|0\rangle + (c+di)|1\rangle &\rightarrow a+bi+cj+dk \\ a^2+b^2+c^2+d^2 &= 1 \end{aligned}$$

将每个量子位映射到单位四元数。类似地,如果限制酉矩阵中的幅角 $\psi=0$,则关系表示也映射到单位四元数。

$$\begin{aligned} \varphi: \mathbb{C}^{2 \times 2} &\rightarrow \mathbb{H} \\ \begin{pmatrix} a+bi & -c+di \\ c+di & a-bi \end{pmatrix} &\rightarrow a+bi+cj+dk \\ a^2+b^2+c^2+d^2 &= 1 \end{aligned}$$

因此,作用于量子位状态的特殊量子门等于两个单位四元数的哈密顿积。取 $\psi=0$,这可以导出 QubitE 的一个变体,称为 QubitE₂。显然,将实体表示为四元数并将关系表示为四元数空间中的旋转的 QuatE^[5] 包含 QubitE₂。但是,QuatE 不包含 QubitE,原因是 $\psi \neq 0$,并且一般的量子门不等于单位四元数。QubitE 中西矩阵的行列式是 $e^{i\psi}$ 而不是 1。

4.1 包含性

本节展示了 QubitE 包含其他模型并继承了它们在学习各种推理模式方面的有利特性。第 7 章中提供了完整的数学证明。

定义 1^[16] 一个模型 M_1 包含另一个模型 M_2 ,当且仅当一个 KG 中任意三元组的存在能由 M_2 判定,也能由 M_1 判定。

命题 2 QubitE 包含 DistMult, pRotatE, RotatE, TransE 和 ComplEx。

4.2 完全表达性

定义 2^[11] 一个模型 M 称为完全表达的(具有完全表达性),仅当存在对实体和关系的嵌入的赋值,可以精确地区分任何符合给定基本事实的正确三元组。

命题 3 QubitE 具有完全表达性。

4.3 推理模式

知识图谱中基于关系的推理模式普遍存在。这些推理模型包括逆模式、对称(反对称)模式、可交换(不可交换)组合模式。这些推理模式 QubitE 都能处理。更正式地,有以下命题。

命题 4 设 $r_2 \in \mathcal{R}$ 是 $r_1 \in \mathcal{R}$ 的逆。QubitE 通过 $\mathbf{u}_{r_2, i} = \mathbf{u}_{r_1, i}^{-1}$ 推理逆模式,其中 $i=1, 2, \dots, d, d$ 是嵌入维度。

命题 5 设 $r \in \mathcal{R}$ 是对称的(反对称的)。QubitE 通过判断 $\mathbf{u}_{r, i} = \mathbf{u}_{r, i}^{-1}$ 成立(或不成立)来推理对称模式(反对称模式),其中 $i=1, 2, \dots, d, d$ 是嵌入维度。

命题 6 设 $r_1, r_2, r_3 \in \mathcal{R}$ 是关系且满足 r_3 是 r_1 和 r_2 的组合。QubitE 通过 $\mathbf{u}_{r_3, i} \mathbf{u}_{r_1, i} = \mathbf{u}_{r_2, i}$ 推理组合模式。当 r_1 和 r_2 可交换,则 $\mathbf{u}_{r_2, i} \mathbf{u}_{r_1, i} = \mathbf{u}_{r_1, i} \mathbf{u}_{r_2, i}$ 。当 r_1 和 r_2 不可交换,则 $\mathbf{u}_{r_2, i} \mathbf{u}_{r_1, i} \neq \mathbf{u}_{r_1, i} \mathbf{u}_{r_2, i}$ 。其中 $i=1, 2, \dots, d, d$ 是嵌入维度。

上述命题的证明在第 7 章中给出。命题 4—命题 6 可以合并为一个结论,即定理 1。

定理 1 QubitE 可以推断对称/反对称、逆和交换/非交换组合模式。

4.4 复杂度分析

表 2 列出了 QubitE 与几种流行模型的空间和时间复杂度。可以看出, QubitE 与经典的 KGE 具有相似的复杂度,例如 TransE, RotatE 和 QuatE 等。

表 2 不同 KGE 模型的空间和时间复杂度对比

Table 2 Comparison of space and time complexity of different

KGE models		
方法	空间复杂度	时间复杂度
TransE	$O(\mathcal{E} n + \mathcal{R} n)$	$O(n)$
TransH	$O(\mathcal{E} n + \mathcal{R} n)$	$O(n)$
TransR	$O(\mathcal{E} n + \mathcal{R} n^2)$	$O(n^2)$
RESCAL	$O(\mathcal{E} n + \mathcal{R} n^2)$	$O(n^2)$
DistMult	$O(\mathcal{E} n + \mathcal{R} n)$	$O(n)$
ComplEx	$O(\mathcal{E} n + \mathcal{R} n)$	$O(n)$
RotatE	$O(\mathcal{E} n + \mathcal{R} n)$	$O(n)$
QuatE	$O(\mathcal{E} n + \mathcal{R} n)$	$O(n)$
$5 * E$	$O(\mathcal{E} n + \mathcal{R} n)$	$O(n)$
QubitE	$O(\mathcal{E} n + \mathcal{R} n)$	$O(n)$

5 实验

5.1 实验设置

5.1.1 数据集

本文在 4 个广泛使用的基准数据集上进行评估,即 FB15k^[17], FB15k-237^[18], WN18^[19] 和 WN18RR^[18]。表 3 列出了这 4 个数据集的三元组数量。FB15k 是从原始 FreeBase^[15] 知识图谱中创建的标准基准。WN18^[17] 是一个从原始 WordNet 数据集派生出来的具有英语分层集合的词汇数据库。根据文献[20], FB15k 和 WN18 存在测试泄漏问题。训练集包含许多反向测试三元组。为解决该问题, FB15k-237 和 WN18RR 被提出,分别作为 FB15k 和 WN18 的子版本,去除了逆关系。FB15k-237 和 WN18RR 数据集均包含多种关系模式,例如组合(awardnominee/.../nominatedfor)、对称(derivationally_related_form)和反对称(has_part)。

表3 数据集统计信息

Table 3 Dataset statistics

数据集	训练集	验证集	测试集
FB15k	483 142	50 000	59 071
WN18	141 442	5 000	5 000
FB15k-237	272 115	17 535	20 466
WN18RR	86 835	3 034	3 134

5.1.2 评估方法

为了加快评估速度,本文一次对所有实体进行评分。具体地,首先,对于每个测试三元组,将尾实体替换为 KG 中的所有实体以获得候选三元组;然后,计算所有候选三元组的分数,并按分数升序排序;最后,计算正确三元组的排名。遵循嵌入模型评估的最佳实践,本文使用最常用的指标(MR)、倒数排名(MRR)和 Hits@ n ($n=1,3,10$)。所有指标的值越大表示模型性能越好。

5.1.3 实现细节

本模型使用 PyTorch^[21]实现,在一张 GTX1080 显卡上进行训练和测试,使用 Adam 作为梯度优化器。此外,本文采用与 QuatE^[5]相同的类型约束。具体地,类型约束是对评估中目标实体的类型(头或尾)进行约束。最后根据 MRR 在验证集上执行网格搜索以获得最佳超参数。超参数选择如下:嵌入维度 $n \in \{100, 200, 300, 400, 500, 600, 800, 1000\}$, dropout 率 $drop \in \{0.1, 0.2, 0.3\}$, 批次大小 $B \in \{256, 512,$

$1024\}$ 。本文采用 1-N 评分^[20]来加速训练。表 4 列出了 QubitE 在所有数据集上的超参数值,其中, lr 表示学习率, $drop$ 表示丢失率, d_e 表示实体嵌入维度, d_r 表示关系嵌入维度, bs 表示批次大小。

表4 QubitE 在所有数据集上的超参数值

Table 4 Hyper-parameter values of QubitE on all datasets

数据集	lr	$drop$	d_e	d_r	bs
FB15k	0.00005	0.1	600	600	512
FB15k-237	0.00050	0.2	200	200	512
WN18	0.00010	0.1	400	400	512
WN18RR	0.00005	0.2	400	400	512

5.1.4 基线

作为对比模型的 4 个强基线包括:欧几里得 KGEs (TransE, RotatE, QuatE, $5 * E$ 和 HopfE)、非欧几里得 KGE (MuRP 和 ATTH)、张量分解 KGE (DistMult, ComplEx, SimpleE, HyperE 和 TuckER), 以及量子 KGEs (QCE 及其变体 F-QCE)。

5.2 主要结果

本文研究了 QubitE 在链接预测任务上的表现。表 5 列出了 FB15k-237 和 WN18RR 上的结果,表 6 列出了 FB15k 和 WN18 上的结果。其中最佳结果用粗体显示,次优结果用下划线表示,排名第三的结果用斜体表示。

表5 FB15k-237 和 WN18RR 上的链接预测结果

Table 5 Link prediction results on FB15k-237 and WN18RR

模型	FB15k-237				WN18RR			
	MRR	Hits@10	Hits@3	Hits@1	MRR	Hits@10	Hits@3	Hits@1
TransE	0.294	0.465	—	—	0.226	0.501	—	—
RotatE	0.338	0.533	0.375	0.241	0.476	0.571	0.492	0.428
QuatE	0.348	0.550	0.382	0.248	0.488	<u>0.582</u>	0.508	0.438
NagE	0.340	0.530	0.378	0.244	0.477	0.574	0.493	0.432
$5 * E$	0.350	0.530	0.380	0.260	0.470	<u>0.580</u>	<u>0.500</u>	0.410
HopfE	0.343	0.534	0.379	0.247	0.472	0.586	<u>0.500</u>	0.413
MuRP	0.340	0.520	0.370	0.240	<u>0.480</u>	0.570	<u>0.500</u>	0.440
ATTH	0.311	0.488	0.339	0.223	0.456	0.526	0.471	0.419
DistMult	0.241	0.419	0.263	0.155	0.430	0.490	0.440	0.390
ComplEx	0.247	0.428	0.275	0.158	0.440	0.510	0.460	0.410
HyperE	0.341	0.520	0.376	0.252	0.465	0.522	0.477	0.436
TuckER	<u>0.358</u>	0.544	<u>0.394</u>	<u>0.266</u>	0.470	0.526	0.482	0.443
QCE	—	0.350	0.225	—	—	0.323	0.195	—
F-QCE	—	0.337	0.198	—	—	0.378	0.274	—
QubitE(ours)	0.366	<u>0.554</u>	<u>0.400</u>	0.273	0.467	0.525	0.478	0.437
QubitE ₂ (ours)	0.366	0.555	0.401	0.273	0.471	0.531	0.482	<u>0.441</u>

表6 在 FB15k 和 WN18 上链接预测结果

Table 6 Link prediction results on FB15k and WN18

模型	FB15k				WN18			
	MRR	Hits@10	Hits@3	Hits@1	MRR	Hits@10	Hits@3	Hits@1
TransE	0.463	0.749	0.578	0.297	0.495	0.943	0.888	0.113
RotatE	0.797	0.884	0.830	0.746	0.949	<u>0.959</u>	0.952	0.944
QuatE	0.782	0.900	0.835	0.711	0.950	<u>0.959</u>	<u>0.954</u>	0.945
NagE	—	—	—	—	0.950	0.960	0.953	0.944
$5 * E$	0.730	0.860	0.780	0.660	0.950	0.960	0.950	0.950
HopfE	—	—	—	—	0.949	0.960	<u>0.954</u>	0.938
DistMult	0.798	0.893	—	—	0.797	0.893	—	—
ComplEx	0.692	0.840	0.759	0.599	0.941	0.947	0.936	0.936
SimpleE	0.727	0.838	0.773	0.660	0.942	0.947	0.944	0.939
HyperE	0.790	0.885	0.829	0.734	<u>0.951</u>	0.958	0.955	0.947
TuckER	0.795	0.892	0.833	0.741	0.953	0.958	0.955	<u>0.949</u>
QubitE(ours)	<u>0.807</u>	<u>0.894</u>	<u>0.838</u>	0.758	0.950	0.957	0.952	0.945
QubitE ₂ (ours)	0.818	0.897	0.846	<u>0.753</u>	0.950	<u>0.959</u>	<u>0.954</u>	0.946

总体而言,与最先进的(SOTA)经典模型相比,QubitE取得了有竞争力的结果。在FB15k-237和FB15k上,QubitE在所有指标上都优于所有SOTA模型。但是,在WN18和WN18RR上,所有方法的表现都相近,最优和次优结果仅差0.001,KGE没有力压群雄、在所有指标上都优于其他方法。在WN18上,QubitE和QubitE₂的性能与SOTA模型(QuatE,NagE,HopfE等)在所有指标上相差在0.003以内,在Hits@1和Hits@3上也取得了次优结果。因此在0.005的误差允许范围内,可以认为WN18数据集不能显著区分这些方法的优劣。

FB15k-237和WN18RR主要包含对称/反对称和组合的推理模式。在欧几里得KGEs方法中,TransE的性能最差,因为它无法推断出反对称或反转模式,而RotatE的推理能力更好。从RotatE,QuatE到HopfE,其MRR和Hits@10随着在复空间、四元数空间等方面的推广而稳步提高。在张量分解和KGEs方法中,CompLex和DistMult表现不佳,因为它们无法推断组合模式。TuckER的表现则要好得多,因为它具有完全表达性。但与线性模型QubitE相比,TuckER需要更多的参数。整体而言,具有竞争性的性能表现展示了QubitE较高的表示能力。FB15k和WN18主要包含对称/

反对称和反演的推理模式。在欧几里得KGE方法中,TransE在这两个数据集上表现不佳,因为TransE无法处理对称模式。RotatE将关系转换为复空间中的旋转,而四元数空间中的QuatE则表现更好。正如QuatE所观察到的,与单位四元数关系的归一化是嵌入性能的关键步骤。确切地说,由于量子机制,QubitE自然满足归一化约束以保持量子优势,因此其性能比TransE和RotatE好得多。

作为一种基于量子的方法,QubitE的性能明显优于两种代表性的基于量子的模型QCE和F-QCE。与QCE和F-QCE相比,QubitE在FB15k和WN18上的所有指标上平均提高了50%,因为QCE无法在训练中保持量子优势,而F-QCE则面临着参数爆炸和过度拟合的问题。因此可以认为QubitE的提升还源于其模式推理能力、完全表达性、包含性以及量子机制在链路预测任务上的正确应用。

5.3 主要结果的误差限

为了评估QubitE的链接预测性能,本文使用不同的随机种子1,10,100,1000,10000分别运行模型5次。本节介绍了这些结果的误差限。表7列出了QubitE在4个数据集上的结果的误差。总体而言,方差很小,说明QubitE的性能是稳定的。

表7 QubitE在所有数据集上结果的均值和方差

Table 7 Mean values and variances of QubitE's results on all datasets

数据集	MRR	Hits@10	Hits@3	Hits@1
FB15k-237	$0.366 \pm 3 \times 10^{-7}$	$0.554 \pm 2 \times 10^{-6}$	$0.400 \pm 2 \times 10^{-6}$	$0.273 \pm 3 \times 10^{-7}$
WN18RR	$0.467 \pm 9 \times 10^{-7}$	$0.525 \pm 2 \times 10^{-2}$	$0.478 \pm 3 \times 10^{-2}$	$0.437 \pm 1 \times 10^{-2}$
FB15k	$0.807 \pm 2 \times 10^{-6}$	$0.894 \pm 1 \times 10^{-3}$	$0.838 \pm 3 \times 10^{-2}$	$0.758 \pm 2 \times 10^{-2}$
WN18	$0.950 \pm 3 \times 10^{-6}$	$0.957 \pm 3 \times 10^{-6}$	$0.952 \pm 5 \times 10^{-6}$	$0.945 \pm 8 \times 10^{-7}$

5.4 模型分析

消融研究 ψ 。限制 $\psi=0$ 构造变体QubitE₂,以探究参数 ψ 对模型性能的影响。从表5和表6可以观察到,在所有数据集上,QubitE₂都略好于标准QubitE。结果表明, ψ 不是提升性能的核心参数,其他参数对于提升性能才更为关键。根据量子力学解释, ψ 不改变酉矩阵的模,即它不会影响量子比特的物理测量结果,因此不会显著影响实验结果。

嵌入维度的影响。实验表明,嵌入维度的选择对QubitE的有效性和效率都有重大影响。设置嵌入维度 $d \in \{100, 200, 400, 800, 1000, 1200\}$,训练QubitE,并根据验证集绘制结果,如图1所示。

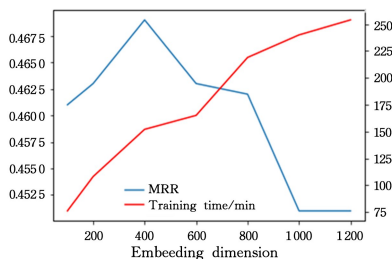


图1 QubitE在WN18RR上MRR和训练时间随嵌入维度的变化

Fig. 1 Convergence MRR and training time of QubitE on WN18RR

随着 d 的增加,训练时间增加,而模型性能(由MRR表示)在 $d=100$ 到 $d=400$ 之间缓慢提升,但在 $d=400$ 之后

急剧下降。因此,本文选择 $d=400$ 作为嵌入维度的最佳设置。

语义逻辑计算。语义逻辑计算是不同于以往所有经典KGEs的特征。借助量子力学,可以对学习到的量子嵌入的语义进行量子逻辑计算。例如,给定实体A,NOT量子电路将计算实体A的语义否定。此外,AND量子门将计算给定实体A和B的语义交集。非门和与门是非参数的,表明逻辑计算对于实体量子嵌入是关系无关的。当然,QubitE支持所有的量子逻辑运算算子。现在考虑非门。在数学上,非门可以写成如下形式:

$$\text{NOT} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

于是可以通过以下等式创建新实体 $\text{NOT}(\mathbf{h})$,即实体 \mathbf{h} 的语义否定:

$$\text{NOT}(\mathbf{h}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{h}_a \\ \mathbf{h}_b \end{pmatrix} = \begin{pmatrix} \mathbf{h}_b \\ \mathbf{h}_a \end{pmatrix}$$

对 $\text{NOT}(\mathbf{h})$ 和所有实体进行评分,得分最高的实体被认为是 $\text{NOT}(\mathbf{h})$ 的最佳解释。在FB15k中随机选择5个实体,并将它们的语义否定的解释在表8中列出。从表中可以观察到,否定在“艺术家 artist”和“电视演员 tv_actor”、“获奖者 award_winner”和“获奖作品 award_winning_work”、“电视制片人 tv_producer”和“电影演员 film_actor”之间建立了联系。总的来说,从实体的类型来看,目标实体是源实体的否定是有道理的。

表 8 由量子门 NOT 生成源实体的否定实体

Table 8 Negative entities of source entities generated by quantum gate NOT

源实体	源实体类型	否定实体	否定实体类型	分数
Hermann Hesse	/music/artist	Dannii Minogue	/tv/tv_actor	0.8221
Norman Stiles	/award/award_winner	The Verdict	/award/award_winning_work	0.9402
Edward G. Robinson	/award/award_winner	Snow White and the Huntsman	/award/award_winning_work	0.8526
Martin Scorsese	/tv/tv_producer	Liza Minnelli	/film/actor	0.8513
Ellie Kemper	/tv/tv_actor	Amy Winehouse	/music/artist	0.6913

6 理论证明

6.1 包含性

这里证明命题 2: QubitE 包含 DistMult, pRotatE, RotatE, TransE 和 ComplEx, 并继承了它们在学习各种推理模式方面的有利特性。

在证明之前, 先给出以下命题:

命题 7 \forall 单位四元数 q , 存在双射 $\phi: \mathbb{H} \rightarrow \mathbb{C}$ 使得 $\phi(q)$ 是复数。进一步, $\phi(q)$ 可写成四元数的形式 $\phi(q) = a + 0i + bj + 0k$, $a, b \in \mathbb{R}$, 且四元数空间中的 Hamilton 积退化为复数积。

Proof: 对任意给定的单位四元数 $q = a + bi + cj + dk$, 有:

$$a = \cos(\theta)$$

$$b = \sin(\theta) \cos(\varphi)$$

$$c = \sin(\theta) \sin(\varphi) \cos(\varphi)$$

$$d = \sin(\theta) \sin(\varphi) \sin(\varphi)$$

其中, $\theta, \phi, \varphi \in [-\pi, \pi]$ 。考虑 $\phi(q) = a' + 0i + b'j + 0k$ 的形式, 其中 $a', b' \in \mathbb{R}$ 。

首先用 a 表示 a' :

$$a' = \frac{a}{1-a^2} \in \mathbb{R}$$

其次有:

$$\frac{c}{b} = \tan(\phi) \cos(\varphi)$$

$$\frac{d}{b} = \tan(\phi) \sin(\varphi)$$

$$\frac{c^2}{b^2} + \frac{d^2}{b^2} = \tan^2(\phi)$$

$$\frac{c^2}{b} + \frac{d^2}{b} = b \left(\frac{c^2}{b^2} + \frac{d^2}{b^2} \right) = \sin(\theta) \cos(\phi) \tan^2(\phi) \in \mathbb{R}$$

所以可以用 b, c, d 表示 b' :

$$b' = \frac{c^2}{b} + \frac{d^2}{b} \in \mathbb{R}$$

所以映射可以是:

$$\phi: \mathbb{H} \rightarrow \mathbb{C}$$

$$a + bi + cj + dk \rightarrow a' + 0i + b'j + 0k$$

$$a' = \frac{a}{1-a^2}$$

$$b' = \frac{c^2}{b} + \frac{d^2}{b}$$

此时四元数空间中的 Hamilton 积退化为复数积。

然后开始证明包含性(命题 2)。

Proof: 对任意给定实体 h 和关系 r , 前文已经证明它们可以映射为单位四元数(命题 1)。对任意单位四元数, 前文也证明了存在它到复数的映射(命题 7)。令 $z_e = a_e' + 0i + b_e'j + 0k$, 其中 e 表示量子态, z_e 是 e 的四元数表示。

因此有以下推导:

$$\begin{aligned} f(h, r, t) &= \text{Re}(\langle \mathbf{h}_r, \overline{\mathbf{t}} \rangle) \\ &= \text{Re}(\langle \mathbf{z}_{h_r}, \overline{\mathbf{z}_t} \rangle) \\ &= \sum_{i=1}^d \text{Re}(\langle \mathbf{z}_{h_{r_i}}, \overline{\mathbf{z}_{t_i}} \rangle) \\ &= \sum_{i=1}^d \text{Re}(\langle \mathbf{z}_{h_i}, \mathbf{z}_{r_i}, \overline{\mathbf{z}_{t_i}} \rangle) \\ &= f_{\text{ComplEx}}(h, r, t) \end{aligned}$$

使得 QubitE 包含 ComplEx。通过去除 z_e 的虚部, 评分函数变为 $f(h, r, t) = \sum_{i=1}^d \langle \text{Re}(\mathbf{z}_{h_i}), \text{Re}(\mathbf{z}_{r_i}), \text{Re}(\mathbf{z}_{t_i}) \rangle$, 在这种情况下退化为 DistMult。另一方面, 有:

$$\begin{aligned} f(h, r, t) &= -\|\mathbf{h}_r - \mathbf{t}\| \\ &= -\|\mathbf{z}_{h_r} - \mathbf{z}_t\| \\ &= -\|\mathbf{z}_h \circ \mathbf{z}_r - \mathbf{z}_t\| \\ &= f_{\text{RotatE}}(h, r, t) \end{aligned}$$

使得 QubitE 包含 RotatE。根据 RotatE^[4] 的结论, RotatE 包含 pRotatE 和 TransE。根据传递性, QubitE 也包含 pRotatE 和 TransE。证毕。

6.2 完全表达性

本节证明命题 3: QubitE 具有完全表达性。

Proof: 证明包含两个步骤, 首先证明 QubitE 具有表达性, 其次证明表达性是完全的。

首先证明 QubitE 可以表达任何排序张量 $\mathcal{A} \in \mathbb{R}^{n_e \times n_r \times n_t}$, 其中 n_e 是实体的数量, n_r 是 KG 中的关系数量。 \mathcal{A} 的第 ikj 元素表示为 α_{ikj} , 对应于三元组 (h_i, r_k, t_j) 。如果模型对三元组 (h_i, r_k, t_j) 的评分高于 (h_i', r_k', t_j') , 则排名张量中三元组 (h_i, r_k, t_j) 的排名低于 (h_i', r_k', t_j') 的排名。其次, 对于任意布尔张量 $\mathcal{B} \in \{0, 1\}^{n_e \times n_r \times n_t}$, QubitE 得到一个与 \mathcal{B} 一致的排序张量。也就是说, 对于 $\beta_{ikj} = 1$, 三元组 (h_i, r_k, t_j) 为正例, 对于 $\beta_{i'k'j'} = 0$, 三元组 (h_i', r_k', t_j') 为负例, 有 $\alpha_{ikj} > \alpha_{i'k'j'}$, 以正确区分三元组是正例还是负例。

对于第一步, Wang 等^[16]证明了 ComplEx 模型可以获得满足排序规则的分数张量 $\mathcal{M}^{n_e \times n_r \times n_t}$ 。该模型给三元组 (h_i, r_k, t_j) 评分为 $\mu_{ikj} = f(h_i, r_k, t_j)$, 使得 $\mu_{ikj} < \mu_{i'k'j'}$ 满足排序张量 \mathcal{A} 的定义。命题 2(包含性)证明了 QubitE 包含 ComplEx。因此, 一定存在实体和关系的嵌入, 使得 QubitE 获得一个排序张量。

对于第二步, Wang 等^[16]表明对于给定的布尔矩阵 \mathcal{B} , 存在一个与 \mathcal{B} 一致的排序矩阵。因此, QubitE 得到与 \mathcal{B} 一致的排序矩阵也成立。

综合第一步和第二步, 得出结论: 存在这样的实体和关系嵌入, 对于任何事实三元组, QubitE 都可以正确区分其为正例还是负例。这意味着 QubitE 具有完全表达性。证毕。

6.3 模式推理

定义 3 关系 r 是对称的(反对称的), 仅当 $\forall x, y \in \mathcal{E}$, $(x, r, y) \in \mathcal{T} \Rightarrow (y, r, x) \in \mathcal{T} (x, r, y) \in \mathcal{T} \Rightarrow (y, r, x) \notin \mathcal{T}$ 。

命题 8 设 $r \in \mathcal{R}$ 是对称的(反对称的)。QubitE 通过判断 $\mathbb{U}_{r,i} = \mathbb{U}_{r,i}^{-1}$ 成立(或不成立)来推理对称模式(反对称模式), 其中 $i=1, 2, \dots, d, d$ 是嵌入维度。

Proof: 首先考虑关系 r 是对称的情况。

根据定义 3, 当对于所有给定实体 x, y , 如果 (x, r, y) 为正例, 模型推断出对称模式, 则 (y, r, x) 也为正例。则

$$g_{r,i}(\mathbf{x}_i) = \mathbf{y}_i$$

然后 $g_{r,i}(\mathbf{y}_i) = \mathbf{x}_i$ 。而 $\mathbf{y}_i = g_{r,i}(\mathbf{x}_i) = \mathbb{U}_{r,i} \mathbf{x}_i$ 。由于 $g_{r,i}$ 是其矩阵表示, $\mathbb{U}_{r,i}$ 是酉和可逆的量子门, 可以限制 $\mathbb{U}_{r,i} = \mathbb{U}_{r,i}^{-1}$ (命题 5)。然后有

$$\mathbf{y}_i = g_{r,i}^{-1}(\mathbf{x}_i)$$

等价于 $\mathbf{x}_i = g_{r,i}(\mathbf{y}_i)$ 。这意味着三元组 (y, r, x) 必须是正例, 即推断出对称模式。

其次, 如果关系 r 是反对称的, 只需限制 $\mathbb{U}_{r,i} \neq \mathbb{U}_{r,i}^{-1}$ 即可得到 $\mathbf{x}_i \neq g_{r,i}(\mathbf{y}_i)$, 这意味着三元组 (y, r, x) 被推断为负例。

定义 4 关系 r_2 (e. g. StudentOf) 是关系 r_1 (e. g. SupervisorOf) 的逆, 仅当 $\forall x, y \in \mathcal{E}$, $(x, r_1, y) \in \mathcal{T} \Rightarrow (y, r_2, x) \in \mathcal{T}$ 。

命题 9 设 $r_2 \in \mathcal{R}$ 是 $r_1 \in \mathcal{R}$ 的逆。QubitE 通过 $\mathbb{U}_{r_2,i} = \mathbb{U}_{r_1,i}^{-1}$ 推理逆模式, 其中 $i=1, 2, \dots, d, d$ 是嵌入维度。

Proof: 根据定义 4, 当对于所有给定实体 x, y , 如果 (x, r_1, y) 为正例, 模型推断出逆模式, 则 (y, r_2, x) 也为正例。则

$$g_{r_1,i}(\mathbf{x}_i) = \mathbf{y}_i$$

那么 $g_{r_2,i}(\mathbf{y}_i) = \mathbf{x}_i$ 。而 $\mathbf{y}_i = g_{r_1,i}(\mathbf{x}_i) = \mathbb{U}_{r_1,i} \mathbf{x}_i$ 。由于 r_1 的矩阵表示 $\mathbb{U}_{r_1,i}$ 是酉和可逆的量子门, 可以约束 $\mathbb{U}_{r_2,i} = \mathbb{U}_{r_1,i}^{-1}$ 满足逆。然后有

$$\mathbf{y}_i = g_{r_2,i}^{-1}(\mathbf{x}_i)$$

等价于 $\mathbf{x}_i = g_{r_2,i}(\mathbf{y}_i)$ 。这意味着三元组 (y, r_2, x) 必须是正例, 即推断出逆模式。证毕。

定义 5 关系 r_1 和关系 r_2 可交换(非交换), 仅当

$$\forall x, y \in \mathcal{E}, (x, r_1 \circ r_2, y) \in \mathcal{T} \Rightarrow (x, r_2 \circ r_1, y) \in \mathcal{T}$$

$$(\exists x, y \in \mathcal{E}, (x, r_1 \circ r_2, y) \in \mathcal{T} \Rightarrow (x, r_2 \circ r_1, y) \notin \mathcal{T})$$

其中, \circ 是组合运算。

定义 6 关系 r_3 (如 UncleOf) 是关系 r_1 (如 FatherOf) 和关系 r_2 (如 BrotherOf) 的组合, 仅当

$$\forall x, y, z \in \mathcal{E}, (x, r_1, y) \in \mathcal{T} \wedge (y, r_2, z) \in \mathcal{T} \\ \Rightarrow (x, r_3, z) \in \mathcal{T}$$

命题 10 设 $r_1, r_2, r_3 \in \mathcal{R}$ 是关系且满足 r_3 是 r_1 和 r_2 的组合。QubitE 通过 $\mathbb{U}_{r_2,i} \mathbb{U}_{r_1,i} = \mathbb{U}_{r_3,i}$ 推理组合模式。当 r_1 和 r_2 可交换, 则 $\mathbb{U}_{r_2,i} \mathbb{U}_{r_1,i} = \mathbb{U}_{r_1,i} \mathbb{U}_{r_2,i}$ 。当 r_1 和 r_2 不可交换, 则 $\mathbb{U}_{r_2,i} \mathbb{U}_{r_1,i} \neq \mathbb{U}_{r_1,i} \mathbb{U}_{r_2,i}$ 。其中 $i=1, 2, \dots, d, d$ 是嵌入维度。

Proof: 根据定义 6, 当对于所有给定实体 x, y, z , 如果 (x, r_1, y) 和 (y, r_2, z) 为正例, 模型推断出组合模式, 则 (x, r_3, z) 也为正例。因此

$$g_{r_1,i}(\mathbf{x}_i) = \mathbf{y}_i$$

$$g_{r_2,i}(\mathbf{y}_i) = \mathbf{z}_i$$

那么有 $g_{r_3,i}(\mathbf{x}_i) = \mathbf{z}_i, i=1, 2, \dots, d$ 以及 $g_{r_j,i}(\mathbf{h}_i) = \mathbb{U}_{r_j,i} \mathbf{h}_i, j=1, 2, 3, i=1, 2, \dots, d$ 。代入 $\mathbf{y}_i = g_{r_1,i}(\mathbf{x}_i)$ 到 $g_{r_2,i}(\mathbf{y}_i) = \mathbf{z}_i$, 得到 $g_{r_2,i}(g_{r_1,i}(\mathbf{x}_i)) = \mathbf{z}_i$ 。所以, 有

$$g_{r_2,i} \circ g_{r_1,i}(\mathbf{x}_i) = \mathbb{U}_{r_2,i} \mathbb{U}_{r_1,i} \mathbf{x}_i = \mathbf{z}_i$$

考虑组合性, 约束 $\mathbb{U}_{r_2,i} \mathbb{U}_{r_1,i} = \mathbb{U}_{r_3,i}$, 有 $g_{r_2,i} \circ g_{r_1,i}(\mathbf{x}_i) = g_{r_3,i}(\mathbf{x}_i) = \mathbf{z}_i$ 。这意味着三元组 (x, r_3, z) 必须是正例, 即推断出组合模式。如果 r_1 和 r_2 可交换, 则 $\mathbb{U}_{r_2,i} \mathbb{U}_{r_1,i} = \mathbb{U}_{r_1,i} \mathbb{U}_{r_2,i}$ 。如果 r_1 和 r_2 不可交换, 则 $\mathbb{U}_{r_2,i} \mathbb{U}_{r_1,i} \neq \mathbb{U}_{r_1,i} \mathbb{U}_{r_2,i}$ 。证毕。

结束语 本文提出了一种名为 QubitE 的新型 KGE, 以应用量子力学进行知识图谱补全。QubitE 将实体建模为量子位状态, 并将关系表示为量子门。凭借细粒度的初始化和评分函数, QubitE 可以保留量子优势并正确区分正负三元组。通过详细的理论分析, QubitE 具有完全表达性、包含性、模式推理能力和线性时空复杂度等优势。在 4 个基准数据集上的实验表明, QubitE 实现了整体上有竞争性的性能。

参考文献

- [1] MA Y P, TRESP V, ZHAO L M, et al. Variational Quantum Circuit Model for Knowledge Graph Embedding [J]. Advanced Quantum Technologies, 2019, 2(7/8): 1-13.
- [2] BORDES A, USUNIER N, GARCIA-DURAN A et al. Translating Embeddings for Modeling Multi-Relational Data [C] // NIPS. 2013.
- [3] LIN Y K, LIU Z Y, SUN M S, et al. Learning Entity and Relation Embeddings for Knowledge Graph Completion [C] // AAAI. 2015.
- [4] SUN Z Q, DENG Z H, NIE J Y, et al. RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space [C] // International Conference on Learning Representations. 2019.
- [5] ZHANG S, TAY Y, YAO L N, et al. Quaternion Knowledge Graph Embeddings [J]. arXiv:1904.10281, 2019.
- [6] NAYYERI M, VAHDATI S, AYKUL C, et al. 5 * Knowledge Graph Embeddings with Projective Transformations [J]. arXiv: 2006.04986, 2020.
- [7] IVANA B, ALLEN C, TIMOTHY M. Multi-Relational Poincaré Graph Embeddings [C] // Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, Vancouver, BC, Canada; NeurIPS 2019, 2019: 4465-4475.
- [8] INES C, WOLF A, DA-CHENG JUAN D C, et al. Low-Dimensional Hyperbolic Knowledge Graph Embeddings [C] // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. online; ACL 2020, 2020: 6901-6914.
- [9] YANG B S, YIH W T, HE X D, et al. Embedding Entities and Relations for Learning and Inference in Knowledge Bases [C] // 3rd International Conference on Learning Representations, ICLR 2015. San Diego; Conference Track Proceedings, 2015.
- [10] THÉO T, WELBL J, RIEDEL S, et al. Complex Embeddings for Simple Link Prediction [C] // Proceedings of the 33rd International Conference on Machine Learning, ICML 2016. New York:

JMLR.org,2016;2071-2080.

- [11] KAZEMI S M, POOLE D. Simple Embedding for Link Prediction in Knowledge Graphs[C]// Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018. Canada: NeurIPS 2018, 2018: 4289-4300.
- [12] IVANA B,CARL ALLEN C,TIMOTHY M. Hypernetwork Knowledge Graph Embeddings [C] // Artificial Neural Networks and Machine Learning — ICANN 2019 — 28th International Conference on Artificial Neural Networks. Germany: Springer,2019;553-565.
- [13] IVANA B,CARL ALLEN C,TIMOTHY M. Tucker: Tensor Factorization for Knowledge Graph Completion [C]// Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Hong Kong: EMNLP-IJCNLP,2019;5184-5193.
- [14] SETH L,SCHULD M,IJAZ A,et al. Quantum Embeddings for Machine Learning[J]. arXiv:2001.03622,2020.
- [15] GLOROT X,BENGIO Y. Understanding the Difficulty of Training Deep Feedforward Neural Networks [C]// Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. JMLR Workshop,2010;249-256.
- [16] WANG Y J,GEMULLA R,LI H. On Multi-Relational Link Prediction with Bilinear Models [C]// AAAI. 2018.
- [17] BOLLACKER K D,EVANS C,PARITOSH P,et al. Freebase: A Collaboratively Created Graph Database for Structuring Human Knowledge [C]// Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data. New York: SIGMOD 2008;1247-1250.
- [18] TOUTANOVA K,CHEN D Q. Observed Versus Latent Fea-

tures for Knowledge Base and Text Inference[C] // The 3rd Workshop on Continuous Vector Space Models and their Compositionality. 2015.

- [19] MILLER G A. WordNet: A Lexical Database for English [J]. Communications of the ACM,1992,38(1):39-41.
- [20] DETTMERS T,MINERVINI P,STENETORP P,et al. Convolutional 2d Knowledge Graph Embeddings [C]// Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th Innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18). New Orleans, Louisiana, USA. AAAI Press,2018;1811-1818.
- [21] ADAM P,GROSS S,CHINTALA S,et al. Automatic Differentiation in PyTorch [C]// NIPS. 2017.



LIN Xueyuan, born in 1998, postgraduate. His main research interests include deep learning, knowledge graph, natural language processing, big data and artificial intelligence.



E Haihong, born in 1982, Ph.D, associate professor, is a member of China Computer Federation. Her main research interests include deep learning, knowledge graph, natural language processing, big data and artificial intelligence.

(责任编辑:杨雪敏)