QUANTUM AND TRANSLATION EMBEDDING FOR KNOWLEDGE GRAPH COMPLETION

Anonymous authors
Paper under double-blind review

ABSTRACT

Knowledge Graph Completion (KGC) mainly devotes to link predicting for an entity pair in Knowledge Graph (KG) according to known facts. In this work, we present a novel model for this end. In this model, Quantum and Translation Embedding are used as components for logic and structural feature capturing in the same vector subspace, respectively. The two components have synergy with each other and achieve impressive performance at low cost which is close to the efficient model TransE. Surprisingly, the performance on challenging datasets such as `fb15k237` and `WN18RR` are up to 94.89% and 92.79% in metric Hits@1 while the dimension of embedding is only 4 in the process of training. The insight of this work enlightens the notion of dense feature model design for KGC which is a new alternative to Deep Neural networks (DNN) in this task or even a better choice.

1 INTRODUCTION

KGs are broadly applied in information retrieval, question answering, recommendation, E-commerce, etc. They can supply prior knowledge or capability of reasoning, which enables many applications to be more intelligent. Unfortunately, almost all KGs suffer from incompleteness(West et al., 2014; Krompaß et al., 2015) which harms downstream tasks. To be specific, some necessary relations among existing entities are missing, and thus applying such KGs can lead to incorrect results, which causes applications broken. For this reason, KGC has attracted considerable attention and effort. Hence, the motivation in this work is manifold. The recent proposed models including DNN based ones become more complex while the performance of them are still not so promising as expected. Furthermore, KG is a data with manifold properties such as logic, structural, graphic, etc. On the contrary, the abundant properties are usually not made full used. As an important property of this data, logic is underlying the KGs due to they are built in description logic and the things described in KG is logic. Consequently, logic rules based models are always a kind of very important research line of KGC. However, this line is somewhat declined in recent years owing to their inefficient and being weak in generalization. Although some progress is achieved (Niu et al., 2019; Meilicke et al., 2020), they also introduced DNN and greatly increasing the complexity. The statistical or probabilistic features of logic in data is well studied (Qu & Tang, 2019; Cheng et al., 2020).

Roughly, the models involve in logic rules can be attributed to four groups. The first one is logic mining system used as a reasoning engine such as AMIE(Galárraga et al., 2015), FOIL(Quinlan, 1990), ALEPH(Muggleton, 1995), etc. The second one is logic embedding and this group is usually based on first order logic, and they are used as a component to improve the performance of model such as RUGE(Guo et al., 2018). In other words, the embedding model is dominant in these models. The third group is based on other algebraic methods such as ASP(Nguyen et al., 2018) and a newly proposed model E2R(Garg et al., 2019). In this group, logic is no longer symbolic but being projected onto vector space. Furthermore, the latter models achieved promising performance for the task of KGC. The fourth group is based on neural networks, the logic feature usually mined as abstract features by increasing the number of layers or modifying the structure of their networks. Recent years, DNN achieves great success in many fields. The same thing may occur in the representation learning for KGC task, for instance, pLogicNet(Qu & Tang, 2019), pGAT(Vardhan et al., 2020), AnyBURL (Meilicke et al., 2020) are the models of this kind. Although a little improvement in performance is achieved, the time and space complexity is also increased at the meantime.

To alleviate these problems in existing models, a novel model named QLogicE is proposed. It is an embedding method based on Quantum Logic and Translation Emebedding co-work together, and then they are adjusted to the same subspace and achieve amazing performance.

2 Related Works

In this section, Quantum Embedding (QE) based model E2RB and KGE model TransE are the two components of the proposed model.

2.1 Quantum Embedding (QE)

E2R(Garg et al., 2019) is a newly proposed QE approach as a Statistical Relational Learning (SRL) model for KGC which runs on Quantum Logic (QL). And this approach is claimed with the capability of preserving underlying logic structure. It is based on a logic language QL which is proposed to explain the mechanism of quantum mechanics (Birkhoff & Von Neumann, 1936). This logic is built on the basis of quantum theory in vector space. E2R keeps the logic with distributive law holding by constraining the embedding space axis-parallel. Therefore, the vector space is more suitable for logic features characterizing and KG datasets mining.

In this case, the logic elements can be projected onto vector space. Furthermore, complex number is proved to be isomorphic to a double dimension real vector space in this case. As a result, the whole model works in real number space. This idea is coordinate with the distribution representation learning. This work is not only project the logic atoms, propositions, membership and predicates onto vector space, but also the logic conjunction, disjunction, negation and even universal type restriction. Consequently, the logic semantic and restriction are mapped into vector space and formulated as loss function. And then it obtains the solutions learned from known facts.

2.2 Knowledge Graph Embedding (KGE)

Recent years, KGE is a kind of popular models for KGC task. The typical one is TransE(Bordes et al., 2013) and it is enlightened by word2vec(Mikolov et al., 2013) under additive compositionality assumption. In this model, the entities and relations are represented as low-dimension vectors and learning from the KG data to predict new facts unobserved in it. It arouses widespread interest due to its simplicity and efficiency.

After its putting forward, a series of extensions are developed such as TransH(Wang et al., 2014), TransR(Lin et al., 2015), TransD(Ji et al., 2015), TransA(Jia et al., 2016), etc. These extensions are from different aspects to fill the gap between TransE and actual requirement. In TransH, to enable the model to deal with the relations one to many, many to one and many to many, a hyperplane is introduced. With the similar notion, TransR introduces a space to overcome the incapability of handling the three kinds of relations in TransE. TransD is a simplified version of TransR. In this model, the projection matrix is factorized into two vectors to capture the diversity of features in relation. However, they suffer from higher cost and relative lower performance.

2.3 Combination of KGE and QE

Combining KGE with logic is not a new topic but with QL is. For the former, KALE(Guo et al., 2016) and RUGE(Guo et al., 2018) combined the KGE and logic rules by enriching the training dataset to improve the performance of KGC task. Recently, there is a model UniKER(Cheng et al., 2020) devoting to combining KGE and logic rules more closely. The logic rules are restricted as Horn rules to explore the logic knowledge for better reasoning. These models are based on mining statistical features from the logic rules in the form of Horn rules. However, the gathered evidences are far from containing whole logic semantic in KG data and thus without significant performance improvements. E2R models logic elements including rules explicitly. Therefore, the whole logic language runs in a vector subspace. KGE models are also projecting all entities and relations onto a low dimension vector space. As a result, there is very important in common between them. In other words, both of them run in vector subspaces. Hence, there is a possibility that adjusts them to co-work seamlessly.

3 METHODOLOGY

Our model is composed of binary relation version of quantum embedding(Garg et al., 2019) (E2RB) and a canonical KGE model TransE(Bordes et al., 2013).

3.1 QUANTUM LOGIC COMPONENT (E2RB)

Representation. QL is an algebraic logic and all its elements are represented in vectors space with some constraints. On the contrary to Boolean Logic, it is not only binary but also multi-valued. It runs in the continuous vector space. Furthermore, this model is devised not only for binary relations but also unary relations. In this work, only binary relations are considered due to the datasets of KGC are binary only. As a consequence, it is tailored into a binary version and named `E2RB`.

The relation is modeled as a $2d$-dimension complex vector. For instance, a relation `isFather(Bob, Ann)` is represented as $\mathbf{r}_{isFather} = \mathbf{v}_{Bob} + \imath\mathbf{v}_{Ann}$, where entity vectors $\mathbf{v}_{Bob}, \mathbf{v}_{Ann} \in \mathbb{R}^d$ and relation vectors $\mathbf{r}_{isFather} \in \mathbb{C}^d$. For entity alone, taking `Bob` for instance, the vector is $\mathbf{e}_{Bob} = \mathbf{v}_{Bob} + \imath\mathbf{v}_0$, where $\mathbf{v}_0$ is a d-dimension 0 vector. Similarly, $\mathbf{e}_{Ann} = \mathbf{v}_0 + \imath\mathbf{v}_{Ann}$. Due to the $2d$-dimension complex apace is isomorphic to vector space $\mathbb{C}^{2d}$. In other words, the relation vector assumed to be isomorphic to a vector subspace $\mathbb{R}^{2d}$, consequently, there exists a bijection map F: $\mathbb{C}^d \mapsto \mathbb{R}^{2d}$. As a result, the entity pair can be represented as $\mathbf{e}_{BA} = \binom{\mathbf{v}_{Bob}}{\mathbf{v}_{Ann}}$. Formally, for a fact, the head entity, relation and tail entity vector are represented as $\mathbf{e}_r = \binom{\mathbf{v}_h}{\mathbf{v}_t}$, $\mathbf{e}_h = \binom{\mathbf{v}_h}{\mathbf{0}}$, $\mathbf{e}_t = \binom{\mathbf{0}}{\mathbf{v}_t}$, where $\mathbf{e}_r, \mathbf{e}_h, \mathbf{e}_t \in \mathbb{R}^{2d}$ and $\mathbf{v}_h, \mathbf{v}_t \in \mathbb{R}^d$.

Score for Logic As mentioned above, the datasets only involve in binary relations. Therefore, only these relations are modeled, and the score function is as follows according to the vectors of every proposition and its negation are orthogonal with each other.

$$f_{E2RB} = \|(\mathbf{1} - \mathbf{e}_r) \cdot \mathbf{e}_r\|^p \tag{1}$$

where p is the kind of vector norm, that is $p \in \{\ell_1, \ell_2\}$, $p = \ell_1$ here.
Loss for Logic To keep the projection running properly, there are some constrains to be satisfied.

$$\mathcal{L}_h = \left\|\mathbf{e}_r \odot \begin{pmatrix}\mathbf{1}_d \\ \mathbf{0}_d\end{pmatrix}\right\|^2, \mathcal{L}_t = \left\|\mathbf{e}_r \odot \begin{pmatrix}\mathbf{0}_d \\ \mathbf{1}_d\end{pmatrix}\right\|^2 \tag{2}$$

The Equation 2 ensures the entity vectors learning into expecting form and the similar reason for relation vector as follows.

$$\mathcal{L}_{r_i} = \left(\min\left\{0, \mathbf{I}_{r_i}^\top \begin{pmatrix}\mathbf{0}_d \\ \mathbf{1}_d\end{pmatrix} - 1\right\}\right)^2 + \left(\min\left\{0, \mathbf{I}_{r_i}^\top \begin{pmatrix}\mathbf{1}_d \\ \mathbf{0}_d\end{pmatrix} - 1\right\}\right)^2 \tag{3}$$

where $\mathbf{I}_{r_i}$ is an indicator. Due to the dimension of it is $2d$, $\mathbf{I}_{r_i} = \mathbf{1}_{2d}$ is true. To ensure it is a binary relation and $\mathbf{I}_{r_i}$ to be as close $\mathbf{1}_{2d}$ as possible, the loss term has to meet with

$$\mathcal{L}_{I_i} = \left\|\mathbf{I}_{r_i} \odot \bar{\mathbf{I}}_{r_i}\right\|^2 \tag{4}$$

where $\bar{\mathbf{I}}_{r_i}$ is the bit flipped version of $\mathbf{I}_{r_i}$, that is, $\mathbf{I}_{r_i} + \bar{\mathbf{I}}_{r_i} = \mathbf{1}_{2d}$. For membership (i.e. $\mathcal{L}_{r_i \in \mathcal{R}}$, this means r_i is a relation included in the relation set $\mathcal{R}$ of KG.), the loss term constructed via the residual length of the projection. This relation is multi-hop one, that means this model contains multi-hop relations via projection of the entity and relation onto vector space and learns it via the loss term Equation 5.

$$\begin{aligned}\mathcal{L}_{r_i(h,t)} = {} & \left\|\bar{\mathbf{I}}_{r_i} \odot \mathbf{e}_r\right\|^2 + \left(\mathbf{1}_{2d}^\top \left(\left(\begin{pmatrix}\mathbf{1}_d & \mathbf{0}_d \\ \mathbf{0}_d & \mathbf{1}_d\end{pmatrix}\bar{\mathbf{I}}_{r_i}\right) \odot \mathbf{e}_r\right)\right)^2 \\ & + \left(1 - \mathbf{e}_h^\top \mathbf{e}_t\right)^2 + \left\|\begin{pmatrix}\mathbf{0}_d \\ \mathbf{1}_d\end{pmatrix} \odot \mathbf{e}_r - \mathbf{e}_h\right\|^2 + \left\|\begin{pmatrix}\mathbf{1}_d \\ \mathbf{0}_d\end{pmatrix} \odot \mathbf{e}_r - \mathbf{e}_t\right\|^2\end{aligned} \tag{5}$$

For logic inclusion, it is one relation between other two relations. In this sense, it may be also multi-hop relation and the loss term can be formulated as follows.

$$\mathcal{L}_{r_i \sqsubseteq r_j} = \left\|\mathbf{I}_{r_i} \odot \bar{\mathbf{I}}_{r_j}\right\|^2 \tag{6}$$

For logic conjunction and disjunction, these are very important relations in logic, and so they are latent in KG data. The loss terms can be devices as

$$\mathcal{L}_{(r_i=r_j\sqcap r_k)} = \left\|\mathbf{I}_{r_i} - \mathbf{I}_{r_j} \odot \mathbf{I}_{r_k}\right\|^2 \qquad \mathcal{L}_{(r_i=r_j\sqcup r_k)} = \left\|\mathbf{I}_{r_i} - \max(\mathbf{I}_{r_j}, \mathbf{I}_{r_k})\right\|^2 \tag{7}$$

and logic negation is also a category of relation between relations in KG and the loss term is

$$\mathcal{L}_{(r_i=\neg r_j)} = \left(\mathbf{I}_{r_i}^\top \mathbf{I}_{r_j}\right)^2 + \left(\bar{\mathbf{I}}_{r_i}^\top \bar{\mathbf{I}}_{r_j}\right)^2 \tag{8}$$

then the loss terms are added together to learn the logic features latent behind the symbolic KG data via vector operations.

$$\mathcal{L}_{pE2RB} = \mathcal{L}_h + \mathcal{L}_t + \mathcal{L}_{r_i} + \mathcal{L}_{r_i\in\mathcal{R}} + \mathcal{L}_{I_r} + \mathcal{L}_{r_i\sqsubseteq r_j} + \mathcal{L}_{(r_i=r_j\sqcap r_k)} + \mathcal{L}_{(r_i=r_j\sqcup r_k)} + \mathcal{L}_{(r_i=\neg r_j)} \tag{9}$$

Although the paper(Garg et al., 2019) also pointed out that the universal type restriction could be expressed. It is ignored due to it involves in unary relations. To avoid subspace collapsing, two regular terms are formulated. Owing to there exists both logic negation and universal type restriction in the datasets, they are also ignored accordingly. Furthermore, the negative sample loss $\mathcal{L}_{nE2RB}$ is also computed similar $\mathcal{L}_{E2RB}$. Finally, the loss function of logic semantic is

$$\mathcal{L}_{E2RB} = \max\{0, \gamma + \mathcal{L}_{pE2RB} + \mathcal{L}_{nE2RB}\} \tag{10}$$

3.2 Knowledge Graph Embedding Component (TransE)

TransE is not the only choice for KGE component, and any other KGE model are candidates and the representations, score functions and loss functions should be adapted, accordingly.
Presentation. In this model, for a fact, the head entity, relation and tail entity vector represented as $\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^{2d}$, respectively.
Score for KGE Every entity and relation are project to a vector space, according the additive compositionality assumption for word embedding(Mikolov et al., 2013), the positive sample score is devised

$$f_{TransE} = \|\mathbf{h} + \mathbf{r} - \mathbf{t}\|^p \tag{11}$$

It is known that, KG data is not including negative samples.
Loss for KGE (TransE) Due to the KGE model TransE is chosen for simplicity, the loss function of it can be formulated as

$$\mathcal{L}_{TransE} = \max\{0, \gamma + f_{TransE} + f'_{TransE}\} \tag{12}$$

3.3 QLogicE

Representation. Because of the E2RB and TransE model representing entity and relation according to different theories, every fact is represented in the corresponding form. Thus, they have to be handled and coordinated with each other. This causes space complexity increasing in reasonable scale.
Score for QLogicE The final score is the result of the weighted sum of the two score functions.

$$f_{QLogicE} = f_{E2RB} + \lambda_s f_{TransE} \tag{13}$$

where λ_s is the weight of f_{TransE}, it is used for adjusting the score value to suitable scale for f_{E2RB}.
Loss for QLogicE. Loss function is composed of several loss terms according to conditions they have to satisfy. They are the components of objective function. This function drives the training process convergence to reasonable value. They are also the reflection of the notions behind the whole model. It is known that, there is no negative sample in KG. To improve the generalizing capacity of the model, usually corrupting the positive samples to obtain negative ones which follows the paradigm of TransE(Bordes et al., 2013).

The loss function of QLogicE is also a weighted sum of the loss functions of E2RB and TransE.

$$\mathcal{L}_{QLogicE} = \mathcal{L}_{E2RB} + \lambda_l \mathcal{L}_{TransE} \tag{14}$$

where λ_l is the weight of $\mathcal{L}_{TransE}$. The results of the model are sensitive to this value. The objective of training is to minimize the loss function $\mathcal{L}_{QLogicE}$.

4 EXPERIMENTS

To empirically evaluate the model proposed for the KGC task, link prediction experiments on the widely used datasets are conducted. Furthermore, the proposed model shows good complexity performance, and this point is also verified by further experiments.

4.1 DATASETS

There are seven widely used datasets to be tested. However, three results of them thought to be more challenging shows in the paper, and the left are in appendices. **FB15K237** is a subset of `FB15K`. In paper(Toutanova & Chen, 2015), the author believed that `FB15K` was test leakage for some models which get negative samples by swapping head entity with tail entity. Therefore, the inverse relations in the dataset were deleted to avoid testing leakage. It is usually considered to be more challenging than `FB15K`. **WN18RR** is the subset of `WN18` for the same reason by similar operating(Dettmers et al., 2018). And then this dataset becomes widely used dataset for evaluating the capacity of new proposed models. Together with **FB15K237**, they become more popular for evaluating new devised models. **YAGO3-10** is the subset of real world KG YAGO [1] and it was selected in work (Mahdisoltani et al., 2015). This is a multiple language KG and based on Wikipedias. It is used for KGC evaluation task in(Dettmers et al., 2018).

Table 1: The statics of datasets and parameters for the best performance on them.

Dataset	Entity	Relation	Train	Valid	Test	Triple	Parameters					
							d	b	μ	γ	λ_s	λ_l
Kinships	104	25	8,544	1,068	1,074	10,686	4	2,848	0.1	2	0.1	3
UMLS	135	46	5,216	652	661	6,529	4	2,500	0.2	2	0.2	2
FB15k	14,951	1,345	483,142	50,000	59,071	592,213	4	219,610	0.05	2	0.2	2
WN18	40,943	18	141,442	5,000	5,000	151,442	4	141,442	0.1	2	0.1	2
FB15k237	14,505	237	272,115	17,535	20,466	310,116	4	90,705	0.1	2	0.1	2
WN18RR	40,943	11	86,835	3,034	3,134	93,003	4	86,835	0.2	2	0.1	2
YAGO3-10	123,182	37	1,079,040	5,000	5,000	1,089,040	4	359,680	0.1	2	0.1	2

4.2 EXPERIMENTAL SETUP

4.2.1 RUNNING ENVIRONMENT AND EXPERIMENTAL SETUP

The algorithm QLogicE is implemented by Python 3.7.6 with PyTorch 1.1.0 on operating system of Ubuntu 18.04. The code is trained with Stochastic Gradient Descent (SGD) with Adam algorithm. The hardware is a computer with an Intel Core i5-3350P CPU that has 8 cores with a 2.30 GHz main frequency and 12 GB RAM. GPU is Nvidia GeForce GTX 1080Ti. The code can be found at site [2].

4.2.2 HYPERPARAMETERS

There are mainly six hyperparameters in the proposed model when it trains. d is the dimension of entity or relation vectors. b is the batch-size in every epoch of training. γ is the margin of ranking criterion and composed of two parts from the both components. μ is learning rate, which is also sensitive for the model. These parameters are inherent in the models. Our model introduces the other two ones. λ_s is the coefficient of additional score term from KGE model and λ_l is for loss term. For the best performance, the parameters are as the Table 1. The running epochs is 1000 for all datasets in the experiments. Besides, the negative sample ratio is also a parameter that can be set before running the model. In all experiments, any positive sample corresponding to a negative one.

[1] http://yago-knowledge.org

[2] https://github.com/PandaCoding2020/QLogicE

4.2.3 EVALUATION PROTOCOL

This work mainly report the metrics `MRR` and `Hits@N`. The former is short for Mean Reciprocal Rank. In this paper, `Hits@N` are including `Hits@1` and `Hits@10`. It is short for the proportion of correct entities ranked in the top N. To better adapting to the table, it is written as `H@N`.

4.3 RESULTS AND ANALYSIS

4.3.1 EXPERIMENTAL RESULTS

In this section, the performance of link prediction and its complexity demonstrates in two tables and a figure, respectively. We only focus on the three challenging datasets and more results refer to appendices.

4.3.2 LINK PREDICTION

Link prediction is the key task of KGC. The results are analyzed in this section.

Table 2: Results of QLogicE comparing with current state of the art (SOTA). This table demonstrates the SOTA results of the three main categories. They are neural network, translation and factorization based models. These results are adapted from the survey(Rossi et al., 2020).

Model	FB15k			WN18			FB15k237			WN18RR			YAGO3-10		
	M	H1	H10	M	H1	H10	M	H1	H10	M	H1	H10	M	H1	H10
ConvE	0.688	59.46	84.94	0.945	93.89	95.68	0.305	21.90	47.62	0.427	38.99	50.75	0.488	39.93	65.75
ConvKB	0.211	11.44	40.83	0.709	52.89	94.89	0.230	13.98	41.46	0.249	5.63	52.50	0.420	32.16	60.47
ConvR	0.773	70.57	88.55	0.950	94.56	95.85	0.346	25.56	52.63	0.467	43.73	52.68	0.527	44.62	67.33
CapsE	0.087	1.934	21.78	0.890	84.55	95.08	0.160	7.34	35.60	0.415	33.69	55.98	0.000	0.00	0.00
RSN	0.777	72.34	87.01	0.928	91.23	95.10	0.280	19.84	44.44	0.395	34.59	48.34	0.511	42.65	66.43
TransE	0.628	49.36	84.73	0.646	40.56	94.87	0.310	21.72	49.65	0.206	2.79	49.52	0.501	40.57	67.39
STransE	0.543	39.77	79.60	0.656	43.12	93.45	0.315	22.48	49.56	0.226	10.13	42.21	0.049	3.28	7.35
CrossE	0.702	60.08	86.23	0.834	73.28	95.03	0.298	21.21	47.05	0.405	38.07	44.99	0.446	33.09	65.45
TorusE	0.746	68.85	83.98	0.947	94.33	95.44	0.281	19.62	44.71	0.463	42.68	53.35	0.342	27.43	47.44
RotatE	0.791	73.93	88.10	0.949	94.43	**96.02**	0.336	23.83	53.06	0.475	42.60	57.35	0.498	40.52	67.07
DistMult	0.784	73.68	86.32	0.824	72.60	94.61	0.313	22.44	49.01	0.433	39.68	50.22	0.501	41.26	66.12
ComplEx	0.848	81.56	90.53	0.949	94.53	95.50	0.349	25.72	52.97	0.458	42.55	52.12	0.576	50.48	70.35
Analogy	0.726	65.59	83.74	0.934	92.61	94.42	0.202	12.59	35.38	0.366	35.82	38.00	0.283	19.21	45.65
SimplE	0.726	66.13	83.63	0.938	93.25	94.58	0.179	10.03	34.35	0.398	38.27	42.65	0.453	35.76	63.16
HolE	0.800	75.85	86.78	0.938	93.11	94.94	0.303	21.37	47.64	0.432	40.28	48.79	0.502	41.84	65.19
TuckER	0.788	72.89	88.88	**0.951**	**94.64**	95.80	0.352	25.90	53.61	0.459	42.95	51.40	0.544	46.56	68.09
QLogicE	**0.969**	**96.93**	**96.93**	0.914	91.42	91.42	**0.949**	**94.89**	**94.89**	**0.928**	**92.79**	**92.79**	**0.937**	**93.74**	**93.74**

Table 2 displays the results on all five datasets. Our method QLogicE outperforms on all of them except `WN18`. Especially, on the datasets `FB15k237`, `WN18RR` and `YAGO3-10` with large margin. To be specific, on the dataset `FB15k237`, our model is better than the state of the art 169.60% in MRR, 266.37% in Hits@1 and 77.00% in Hits@10, respectively. Similarly, on the dataset `WN18RR`, the counterpart values are 11.43%, 112.19% and 61.80% and dataset `YAGO3-10` are 62.67%, 85.70% and 33.25%.

Table 3 displays the results of performance on the newly proposed logic rules based models. They are attributed to two groups according the key technology they based on. The first group comes from the logic rule based research line and most of them are published recently or to be published in near future. The second group is the newly proposed models. On the dataset `FB15k237`, our model is also outstanding, and they outperform the existing model 33.29% in MRR, 221.77% in Hits@1 and 19.21% in Hits@10. On the dataset `WN18RR`, the counterpart values are 31.82%, 35.86% and 28.34%. On the large dataset `YAGO3-10`, our model outperforms the best reported metrics 67.32% in MRR, 89.83% in Hits@1 and 35.07% in Hits@10.

Table 3: Results of QLogicE comparing with existing logic or rule based models. RPJE(Niu et al., 2019), pGAT(Vardhan et al., 2020). In this table, the models are newly proposed, ADRL(Wang et al., 2020), CoPER-MINERVA(CoPER-ConvE)(Stoica et al., 2020), ParamE-CNN(ParamE-MLP, ParamE-Gate)(Che et al., 2020). The results are adapted from the original papers.

Model	FB15k237			WN18RR			YAGO3-10		
	M	H@1	H@10	M	H@1	H@10	M	H@1	H@10
IterE	0.247	17.90	39.20	0.274	25.40	31.40			
pLogicNet	0.330	23.10	52.80	0.230	1.50	53.10			
pLogicNet*	0.332	23.70	52.40	0.441	39.80	53.70			
RARL	0.247	17.90	39.20	0.274	25.40	31.40	0.560	48.20	69.30
AnyBURL	0.346	27.34	52.25	0.555	49.24	68.94	0.556	49.38	69.10
RPJE	0.470		62.50						
pGAT	0.457	37.70	60.90	0.459	39.50	57.80			
UniKER-TransE	0.522	46.30	63.00	0.307	4.00	56.10			
UniKER-DistMult	0.533	50.70	58.70	0.485	43.20	53.80			
ADRL	0.712	57.40	79.60	0.704	68.30	72.30			
CoPER-MINERVA	0.365	29.49	50.39	0.465	42.66	50.99			
CoPER-ConvE	0.426	32.18	62.92	0.483	44.05	56.12			
ParamE-MLP	0.314	24.00	45.90	0.407	38.40	44.50			
ParamE-CNN	0.393	30.40	57.60	0.461	43.40	51.30			
ParamE-Gate	0.399	31.00	57.30	0.489	46.20	53.80			
InteractE	0.354	26.30	53.50	0.463	43.00	52.80	0.541	46.20	68.70
HAKE	0.346	25.00	54.20	0.497	45.20	58.20	0.545	46.20	69.40
DPMPN	0.369	28.60	53.30	0.485	44.40	55.80	0.553	48.40	67.90
QLogicE	**0.949**	**94.89**	**94.89**	**0.928**	**92.79**	**92.79**	**0.937**	**93.74**	**93.74**

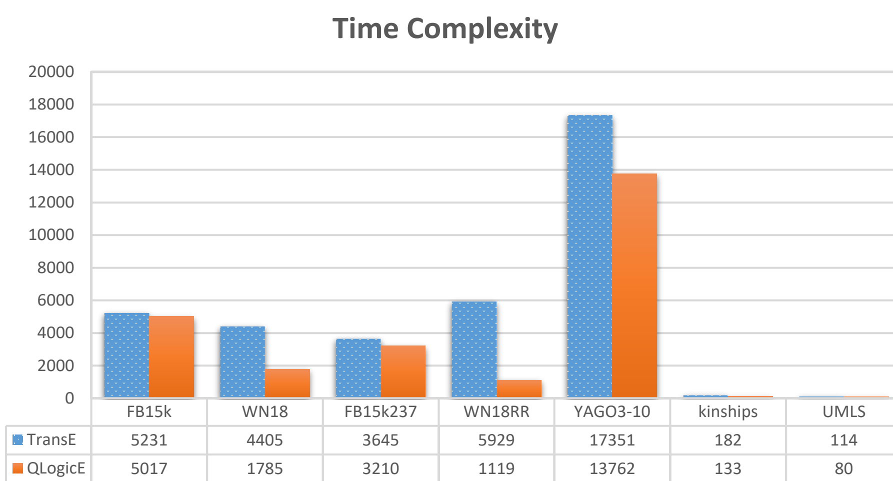

Figure 1: Illustration the comparison of time complexity, and the time unit is second.

4.3.3 Complexity

Figures 1 demonstrates the comparison of time complexity. The baseline TransE is considered to be one of simplest and the most efficient in the existing KGC models. For the sake that the best performance reported in existing papers usually set the dimension of the embedding as 200 recently and the best of QLogicE is 4, and thus the dimensions of embedding for TransE and QLogicE are set as 200 and 4, respectively. The results of TransE are from running the code[3], ℓ_1 norm for datasets `FB15k`, `FB15k237`, `YAGO3-10` and the other ones with ℓ_2 norm. For QLogicE, the space number

[3]https://github.com/ttrouill/complex

is the highest one in the training or testing process. Concerning time-consuming, the proposed model is obviously less than TransE. Notably, when the score and loss functions are computed in ℓ_1 norm, the time is very close to our model and when they are computed in ℓ_2 norm for the sake of better performance, the time is obviously increased and more expensive than our model. The time contains training and testing two terms. On the contrary, the space consumption of TransE is 33% to 76% relative to the QLogicE. There exists a trade-off between time and space in this sense.

4.3.4 Dense Feature Model and Embedding Dimension

Our model runs on the basis of quantum and translation embedding in a vector subspace. It achieves promising results both in performance of link prediction and time complexity. Notably, it breaks the lower bound of the E2R(Garg et al., 2019). We believe that the explicit logic feature modeling is the key reason for the breakthrough in the KGC task. And the framework is capable of capturing dense features in unit dimension of embedding which is named dense feature model. The phenomenon illustrates in Table 1 and Figure 1.

5 Conclusions and Future Works

In this section, some conclusions can be drawn from the experiments in this section, and further research works worth doing in the future.

5.1 Conclusions and Discussion

From the process of model formulation and experimental results, some conclusions can be drawn, accordingly. (1) The proposed model improves the performance of KGC task on the widely used and challenging benchmark datasets, significantly. (2) The proposed model improves performance with low time complexity and competitive space complexity. (3) The result of `FB15k237` and `WN18RR` now get close to the datasets of `FB15k` and `WN18` and all performance on the metrics of MRR, Hits@1 and Hits@10 are over 90%. On the large dataset `YAGO3-10`, the performance is also up to 93.74%. These promising results are not only better than all existing DNN based models, but also newly hot spot reinforce learning based ones.

As mentioned above, our model based on E2RB and TransE, these two models are on the basis of logic and translation, respectively. Embedding couples them together and enable the two components synergy with each other by objective function in the process of training. The outstanding results lead us to insight into the dense feature model can be used as an alternative to DNN for KGC task. We don't know whether it can be extended to other type of datasets except KG. Besides, the logic feature is a very important one lies in KG, deeper mining this kind of feature may also a very important reason for promising results. The breakthrough of lower bound (Garg et al., 2019) also needs more work to clarify whether it is caused by the dense feature model.

5.2 Future Works

As preciously mentioned, our model is a dense feature one. We believe that it is a very dense feature model and this is the key reason for the high performance with low cost. It is known that neural network is universal function for approximating and deep model can capture more abstract features. However, it increases complexity seriously without expected performance improvements. As a result, there are some works worth doing: (1) How to measure the terminology of dense feature model? (2) Why the dimension reduce so much? In fact, it is fewer than the lower bound of embedding in its original paper acclaimed(Garg et al., 2019). (3) How the two model interact or co-work with each other and improving the performance so much. (4) How to apply the model to question answering, recommendation, etc.

References

Garrett Birkhoff and John Von Neumann. The logic of quantum mechanics. *Annals of mathematics*, pp. 823–843, 1936.

Antoine Bordes, Nicolas Usunier, Alberto García-Durán, Jason Weston, and Oksana Yakhnenko. Translating embeddings for modeling multi-relational data. In Christopher J. C. Burges, Léon Bottou, Zoubin Ghahramani, and Kilian Q. Weinberger (eds.), *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States*, pp. 2787–2795, 2013. URL http://papers.nips.cc/paper/5071-translating-embeddings-for-modeling-multi-relational-data.

Feihu Che, Dawei Zhang, Jianhua Tao, Mingyue Niu, and Bocheng Zhao. Parame: Regarding neural network parameters as relation embeddings for knowledge graph completion. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pp. 2774–2781. AAAI Press, 2020. URL https://aaai.org/ojs/index.php/AAAI/article/view/5665.

Kewei Cheng, Ziqing Yang, Ming Zhang, and Yizhou Sun. Uniker: A unified framework for combining embedding and horn rules for knowledge graph inference. 2020.

Tim Dettmers, Pasquale Minervini, Pontus Stenetorp, and Sebastian Riedel. Convolutional 2d knowledge graph embeddings. In Sheila A. McIlraith and Kilian Q. Weinberger (eds.), *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, February 2-7, 2018*, pp. 1811–1818. AAAI Press, 2018. URL https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17366.

Luis Galárraga, Christina Teflioudi, Katja Hose, and Fabian M. Suchanek. Fast rule mining in ontological knowledge bases with AMIE+. *VLDB J.*, 24(6):707–730, 2015. doi: 10.1007/s00778-015-0394-1. URL https://doi.org/10.1007/s00778-015-0394-1.

Dinesh Garg, Shajith Ikbal, Santosh K. Srivastava, Harit Vishwakarma, Hima P. Karanam, and L. Venkata Subramaniam. Quantum embedding of knowledge for reasoning. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, 8-14 December 2019, Vancouver, BC, Canada*, pp. 5595–5605, 2019. URL http://papers.nips.cc/paper/8797-quantum-embedding-of-knowledge-for-reasoning.

Shu Guo, Quan Wang, Lihong Wang, Bin Wang, and Li Guo. Jointly embedding knowledge graphs and logical rules. In Jian Su, Xavier Carreras, and Kevin Duh (eds.), *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, EMNLP 2016, Austin, Texas, USA, November 1-4, 2016*, pp. 192–202. The Association for Computational Linguistics, 2016. doi: 10.18653/v1/d16-1019. URL https://doi.org/10.18653/v1/d16-1019.

Shu Guo, Quan Wang, Lihong Wang, Bin Wang, and Li Guo. Knowledge graph embedding with iterative guidance from soft rules. In Sheila A. McIlraith and Kilian Q. Weinberger (eds.), *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, February 2-7, 2018*, pp. 4816–4823. AAAI Press, 2018. URL https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16369.

Guoliang Ji, Shizhu He, Liheng Xu, Kang Liu, and Jun Zhao. Knowledge graph embedding via dynamic mapping matrix. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 1: Long Papers*, pp. 687–696. The Association for Computer Linguistics, 2015. doi: 10.3115/v1/p15-1067. URL https://doi.org/10.3115/v1/p15-1067.

Yantao Jia, Yuanzhuo Wang, Hailun Lin, Xiaolong Jin, and Xueqi Cheng. Locally adaptive translation for knowledge graph embedding. In Dale Schuurmans and Michael P. Wellman (eds.), *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA*, pp. 992–998. AAAI Press, 2016. URL http://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12018.

Charles Kemp, Joshua B. Tenenbaum, Thomas L. Griffiths, Takeshi Yamada, and Naonori Ueda. Learning systems of concepts with an infinite relational model. In *Proceedings, The Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference, July 16-20, 2006, Boston, Massachusetts, USA*, pp. 381–388. AAAI Press, 2006. URL http://www.aaai.org/Library/AAAI/2006/aaai06-061.php.

Denis Krompaß, Stephan Baier, and Volker Tresp. Type-constrained representation learning in knowledge graphs. In *The Semantic Web - ISWC 2015 - 14th International Semantic Web Conference, Bethlehem, PA, USA, October 11-15, 2015, Proceedings, Part I*, pp. 640–655, 2015. doi: 10.1007/978-3-319-25007-6_37. URL https://doi.org/10.1007/978-3-319-25007-6_37.

Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. Learning entity and relation embeddings for knowledge graph completion. In *Twenty-ninth AAAI conference on artificial intelligence*, 2015.

Farzaneh Mahdisoltani, Joanna Biega, and Fabian M. Suchanek. YAGO3: A knowledge base from multilingual wikipedias. In *CIDR 2015, Seventh Biennial Conference on Innovative Data Systems Research, Asilomar, CA, USA, January 4-7, 2015, Online Proceedings*. www.cidrdb.org, 2015. URL http://cidrdb.org/cidr2015/Papers/CIDR15_Paper1.pdf.

Alexa T McCray. An upper-level ontology for the biomedical domain. *International Journal of Genomics*, 4(1):80–84, 2003.

Christian Meilicke, Manuel Fink, Yanjie Wang, Daniel Ruffinelli, Rainer Gemulla, and Heiner Stuckenschmidt. Fine-grained evaluation of rule- and embedding-based systems for knowledge graph completion. In Denny Vrandecic, Kalina Bontcheva, Mari Carmen Suárez-Figueroa, Valentina Presutti, Irene Celino, Marta Sabou, Lucie-Aimée Kaffee, and Elena Simperl (eds.), *The Semantic Web - ISWC 2018 - 17th International Semantic Web Conference, Monterey, CA, USA, October 8-12, 2018, Proceedings, Part I*, volume 11136 of *Lecture Notes in Computer Science*, pp. 3–20. Springer, 2018. doi: 10.1007/978-3-030-00671-6_1. URL https://doi.org/10.1007/978-3-030-00671-6_1.

Christian Meilicke, Melisachew Wudage Chekol, Manuel Fink, and Heiner Stuckenschmidt. Reinforced anytime bottom up rule learning for knowledge graph completion. *CoRR*, abs/2004.04412, 2020. URL https://arxiv.org/abs/2004.04412.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. In Christopher J. C. Burges, Léon Bottou, Zoubin Ghahramani, and Kilian Q. Weinberger (eds.), *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States*, pp. 3111–3119, 2013.

Stephen Muggleton. Inverse entailment and progol. *New generation computing*, 13(3-4):245–286, 1995.

Hien D. Nguyen, Chiaki Sakama, Taisuke Sato, and Katsumi Inoue. Computing logic programming semantics in linear algebra. In Manasawee Kaenampornpan, Rainer Malaka, Duc Dung Nguyen, and Nicolas Schwind (eds.), *Multi-disciplinary Trends in Artificial Intelligence - 12th International Conference, MIWAI 2018, Hanoi, Vietnam, November 18-20, 2018, Proceedings*, volume 11248 of *Lecture Notes in Computer Science*, pp. 32–48. Springer, 2018. doi: 10.1007/978-3-030-03014-8_3. URL https://doi.org/10.1007/978-3-030-03014-8_3.

Guanglin Niu, Yongfei Zhang, Bo Li, Peng Cui, Si Liu, Jingyang Li, and Xiaowei Zhang. Rule-guided compositional representation learning on knowledge graphs. *arXiv preprint arXiv:1911.08935*, 2019.

Giuseppe Pirrò. Relatedness and tbox-driven rule learning in large knowledge bases. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pp. 2975–2982. AAAI Press, 2020. URL `https://aaai.org/ojs/index.php/AAAI/article/view/5690`.

Meng Qu and Jian Tang. Probabilistic logic neural networks for reasoning. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, 8-14 December 2019, Vancouver, BC, Canada*, pp. 7710–7720, 2019. URL `http://papers.nips.cc/paper/8987-probabilistic-logic-neural-networks-for-reasoning`.

J. Ross Quinlan. Learning logical definitions from relations. *Machine learning*, 5(3):239–266, 1990.

Andrea Rossi, Donatella Firmani, Antonio Matinata, Paolo Merialdo, and Denilson Barbosa. Knowledge graph embedding for link prediction: A comparative analysis. *CoRR*, abs/2002.00819, 2020. URL `https://arxiv.org/abs/2002.00819`.

George Stoica, Otilia Stretcu, Emmanouil Antonios Platanios, Tom M. Mitchell, and Barnabás Póczos. Contextual parameter generation for knowledge graph link prediction. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pp. 3000–3008. AAAI Press, 2020. URL `https://aaai.org/ojs/index.php/AAAI/article/view/5693`.

Kristina Toutanova and Danqi Chen. Observed versus latent features for knowledge base and text inference. In *Proceedings of the 3rd Workshop on Continuous Vector Space Models and their Compositionality*, pp. 57–66, 2015.

L. Vivek Harsha Vardhan, Guo Jia, and Stanley Kok. Probabilistic logic graph attention networks for reasoning. In Amal El Fallah Seghrouchni, Gita Sukthankar, Tie-Yan Liu, and Maarten van Steen (eds.), *Companion of The 2020 Web Conference 2020, Taipei, Taiwan, April 20-24, 2020*, pp. 669–673. ACM / IW3C2, 2020. doi: 10.1145/3366424.3391265. URL `https://doi.org/10.1145/3366424.3391265`.

Qi Wang, Yongsheng Hao, and Jie Cao. ADRL: an attention-based deep reinforcement learning framework for knowledge graph reasoning. *Knowl. Based Syst.*, 197:105910, 2020. doi: 10.1016/j.knosys.2020.105910. URL `https://doi.org/10.1016/j.knosys.2020.105910`.

Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. Knowledge graph embedding by translating on hyperplanes. In Carla E. Brodley and Peter Stone (eds.), *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada*, pp. 1112–1119. AAAI Press, 2014. URL `http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8531`.

Robert West, Evgeniy Gabrilovich, Kevin Murphy, Shaohua Sun, Rahul Gupta, and Dekang Lin. Knowledge base completion via search-based question answering. In *23rd International World Wide Web Conference, WWW '14, Seoul, Republic of Korea, April 7-11, 2014*, pp. 515–526, 2014. doi: 10.1145/2566486.2568032. URL `https://doi.org/10.1145/2566486.2568032`.

Fan Yang, Zhilin Yang, and William W. Cohen. Differentiable learning of logical rules for knowledge base reasoning. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, pp. 2319–2328, 2017.

Wen Zhang, Bibek Paudel, Liang Wang, Jiaoyan Chen, Hai Zhu, Wei Zhang, Abraham Bernstein, and Huajun Chen. Iteratively learning embeddings and rules for knowledge graph reasoning. In

Ling Liu, Ryen W. White, Amin Mantrach, Fabrizio Silvestri, Julian J. McAuley, Ricardo Baeza-Yates, and Leila Zia (eds.), *The World Wide Web Conference, WWW 2019, San Francisco, CA, USA, May 13-17, 2019*, pp. 2366–2377. ACM, 2019. doi: 10.1145/3308558.3313612. URL https://doi.org/10.1145/3308558.3313612.

A Appendix

A.1 Small datasets

Kinships were created by Denham(Kemp et al., 2006). It includes anthropological data of relations from the Central Australia tribe named Alyawarra. It is a typical relational dataset but not a subset of a real world KG. **UMLS** was gathered by McCray(McCray, 2003). It is from a special medical ontology called the Unified Medical Language System[4].

Table 4: Results of QLogicE comparing with current state of the art on small datasets. This table demonstrates the experimental results on the two widely used small datasets. The results are adapted from the paper(Stoica et al., 2020).

Model	Kinship			UMLS		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
DistMult	0.614	48.70	90.40	0.868	82.10	96.70
ComplEx	0.838	75.40	98.00	0.894	82.30	99.50
Neural LP	0.619	47.50	91.20	0.778	64.30	96.20
NTP-λ	0.793	75.90	87.80	0.912	84.30	**100.00**
MINERVA	0.720	60.50	92.40	0.841	75.30	96.70
MultiHop-KG	0.865	78.90	98.20	0.940	90.20	99.20
ConvE	0.830	74.21	97.86	0.954	92.89	99.70
CoPER-MINERVA	0.760	66.20	94.23	0.854	77.76	97.43
CoPER-CovE	0.895	83.62	98.42	0.971	95.46	99.70
QLogicE	**0.999**	**99.91**	**99.91**	**0.998**	**99.77**	99.77

Table 4 displays the performance on widely used small datasets. The results show that our model is also outstanding on them. On the dataset `Kinship`, our model is better than the best metrics 11.51% in MRR, 19.31% in Hits@1 and 1.46% and on the `UMLS` the counterpart data is 2.68%, 4.44% and -0.3%. In the case of very close to 100%, our model also better than almost all the metrics expect Hits@10 on `UMLS`.

A.2 Large datasets

FB15K and **WN18** were first used in paper (Bordes et al., 2013) to evaluate their model TransE. It is a subset of the real KG of Freebase [5] and WordNet[6], respectively. They are widely used datasets for evaluating new KGC model.

Table 5 records the data of results in two widely used datasets `FB15k` and `WN18`. On the dataset `FB15k`, our model outperforms the best reported metrics 0.52% in MRR, 0.55% in Hits@1 and 0.55% and on the `WN18` the counterpart data is -3.38%, -3.26% and -4.57%.

[4]https://www.nlm.nih.gov/research/umls/index.html

[5]http://www.freebase.be/

[6]https://wordnet.princeton.edu/

Table 5: Results of QLogicE comparing with existing logic rules based models. In this table, all data are from their originally paper. They are KALE(Guo et al., 2016), Neural LP(Yang et al., 2017), RUGE(Guo et al., 2018), RuleN(Meilicke et al., 2018), E2R(Garg et al., 2019), IterE(Zhang et al., 2019), pLogicNet(pLogicNet*) (Qu & Tang, 2019), RARL(Pirrò, 2020), AnyBURL(Meilicke et al., 2020). The results are adapted from the original papers.

Model	FB15k			WN18		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
KALE	0.523	38.30	76.20			
Neural LP	0.760		83.70	0.940		94.50
RUGE	0.768	70.30	86.50			
RuleN		77.20	87.00		**94.50**	**95.80**
E2R	0.964	96.40	96.40	0.710	71.10	71.10
IterE	0.628	55.10	77.10	0.913	89.10	94.80
pLogicNet	0.792	71.40	90.10	0.832	71.60	95.70
pLogicNet*	0.844	81.20	90.20	0.945	93.90	**95.80**
RARL	0.628	55.10	77.10	0.913	89.10	94.80
RPJE	0.816		90.30	**0.946**		95.10
QLogicE	**0.969**	**96.93**	**96.93**	0.914	91.42	91.42

B Appendix

B.1 Ablation Study

In **Table** 6 The proposed model captures dense features from data via the binary relation version of E2RB co-work with TransE.

Table 6: Results of ablation study. The data of TransE is from (Rossi et al., 2020) and E2R is from (Garg et al., 2019).

Model	FB15k			WN18			FB15k237			WN18RR			YAGO3-10		
	M	H1	H10	M	H1	H10	M	H1	H10	M	H1	H10	M	H1	H10
TransE	0.628	49.36	84.73	0.646	40.56	**94.87**	0.310	21.72	49.65	0.206	2.79	49.52	0.501	40.57	67.39
E2R	0.964	96.40	96.40	0.710	71.10	71.10									
QLogicE	**0.968**	**96.84**	**96.84**	**0.915**	**91.48**	91.48	**0.949**	**94.94**	**94.94**	**0.928**	**92.79**	**92.79**	**0.937**	**93.74**	**93.74**

Table 6 demonstrates the ablation of the models. The results on datasets `FB15k` our model slightly better than the baseline E2R but much better than TransE. For the dataset `WN18`, only in Hits@10, the performance of TransE is better than our model QLogicE while the 41.64% in MRR and 125.54% worse than our model, which is promising progress. The paperGarg et al. (2019) didn't provide results on the datasets `FB15k237` and `WN18RR`. For them, our model is better than TransE with 206.13% in MRR,337.11% in Hits@1 and 91.22% in Hits@10 on the former and 350.49% in MRR,3225.81% in Hits@1 and 87.38% in Hits@10. On the dataset `YAGO3-10`, the result in MRR, Hits@1 and Hits@10 is better than TransE with margin of 87.03%, 131.06% and 39.10%, respectively.