Explainable Knowledge Reasoning Framework Using Multiple Knowledge Graph Embedding

Mori Kurokawa
mo-kurokawa@kddi-research.jp
KDDI Research, Inc.
Fujimino-shi, Saitama, Japan

ABSTRACT

Knowledge reasoning using knowledge graphs has attracted much attention. However, there is difficulty in integrating various related works to realize complex reasoning with explanation using multiple knowledge graphs. To do this, I propose a reasoning framework which combines multiple knowledge graph embedding techniques with corresponding explainable AI techniques. Experiments using the third knowledge graph reasoning challenge dataset demonstrate the effectiveness of the framework.

CCS CONCEPTS

• **Computing methodologies** → **Knowledge representation and reasoning**; • **Information systems** → *Information integration.*

KEYWORDS

knowledge graph; knowledge reasoning; knowledge graph embedding; explainable AI

ACM Reference Format:

Mori Kurokawa. 2021. Explainable Knowledge Reasoning Framework Using Multiple Knowledge Graph Embedding. In *The 10th International Joint Conference on Knowledge Graphs (IJCKG'21), December 6–8, 2021, Virtual Event, Thailand*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3502223.3502248

1 INTRODUCTION

Incorporating human knowledge into Artificial Intelligence (AI) has been seen as a promising direction to enhance AI ability. Although AI has acquired powerful inductive ability by machine learning - oriented technologies, AI has still limitations in adapting to unseen or rare cases. Knowledge reasoning with human knowledge bases has been expected to overcome the limitations by completing reasoning in such difficult cases. Sherlock Holmes said *"I say now, as I said then, that a man should keep his little brain-attic stocked with all the furniture that he is likely to use, and the rest he can put away in the lumber-room of his library, where he can get it if he wants it."* (Sir Arthur Conan Doyle, THE FIVE ORANGE PIPS). Here, "the lumber-room of his library" is I think a metaphor of a human knowledge base, and he mentioned that a human can adapt to difficult cases by referring to it accordingly.

Knowledge Graph (KG) is a famous structure of human knowledge bases. KG consists of triples which represent relations between entities, e.g., (Tokyo, capital of, Japan) that represents two entities, "Tokyo" and "Japan", are linked with a relation of "capital of". Typical knowledge reasoning over KGs enables to reveal missing entities or relations in KGs. In recent years, KG embedding has been used for this problem[1–5]. KG embedding represents the entities and relations in a low-dimensional vector space and the representations are used to predict possible missing entities or relations.

KGs have also attracted much attention as an important information source to explain AI behavior. Typical methods of explainable AI (XAI) is to explain AI behavior by deriving important features or feature subspaces which have relatively high influence on target prediction. Some of recent XAI technologies employ human knowledge bases such as KGs and specify the important subgraph included in the knowledge bases which has relatively high relevance to target prediction[6].

Under conditions that sufficient knowledge has been accumulated in KGs, knowledge representation and explanation using KGs can be powerful tools to infer unseen or rare cases with reasons. However, it is difficult to implement such tools because it is necessary to combine two different types of technologies, knowledge representation and XAI. In this study, I propose a unified reasoning framework which combines KG embedding techniques with corresponding XAI techniques. The contributions of this study are the unified framework and the practical experimental results using the third knowledge graph reasoning challenge dataset.

The rest of the paper is organized as follows. I first discuss the related work, and then propose our framework. Then I present the experimental results. Finally, I conclude the paper in the last section.

2 RELATED WORK

First, let me introduce the notations used in this paper. A i-th knowledge graph KG_i is formalized as $KG_i = (E_i, R_i, T_i)$, where E_i, R_i, T_i are the set of entities, relations, and triples respectively. I describe triples (h, r, t), in which $h \in E_i$ and $t \in E_i$ denote head and tail entities respectively and $r \in R_i$ denotes the relation between h and t. Let me denote a set of KGs, $KG = \{KG_i\}$. Knowledge reasoning or completion on KGs is formalized as a task for predicting a missing part in a triple, e.g., a missing head entity denoted by h' in (h', r, t).

KG embedding represents the entities and relations of KGs in a low-dimensional vector space. The representations are used to predict possible missing entities or relations. Translation-based methods preserve triple information in the embedding space so

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
IJCKG'21, December 6–8, 2021, Virtual Event, Thailand
© 2021 Association for Computing Machinery.
ACM ISBN 978-1-4503-9565-6/21/12…$15.00
https://doi.org/10.1145/3502223.3502248

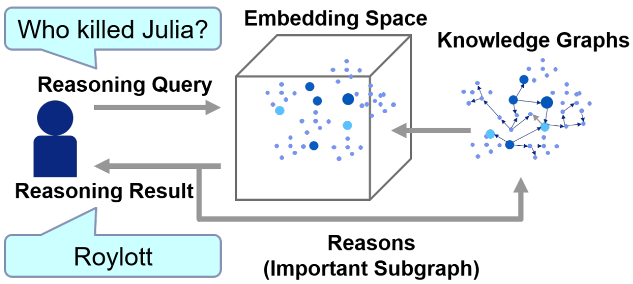

Figure 1: Proposed Reasoning Framework using KGs

that translation operation over the space characterizes each triple. TransE[1] is an initial work of this family which assumes $\boldsymbol{h}+\boldsymbol{r} \simeq \boldsymbol{t}$ in which $\boldsymbol{h}, \boldsymbol{r}, \boldsymbol{t}$ is the corresponding embeddings of h, r, t. Based on this assumption, the scoring function of each triple is designed as $-\|\boldsymbol{h}+\boldsymbol{r}-\boldsymbol{t}\|$. TransE has limitations in representing complex relations. Later works such as PTransE[2] have improved relational representations. Recently, Graph Neural Networks (GNNs) have been extended to represent both entities and relations for KGs. R-GCN[3] utilizes this extension of GNNs as a encoder, and realizes knowledge reasoning using DistMult[4] as a decoder.

Multiple KG embedding methods have been utilized to obtain better embedding using multiple KGs. Some of these methods are extensions of translation-based methods. ITransE and IPTransE[5] apply parameter sharing which shares same entity embeddings with respect to common entities between KGs, and iteratively adds likely aligned entity pairs as a soft alignment to augment a limited number of prior entity alignments. ITransE utilizes TransE as a base model, whereas IPTransE utilizes PTransE as a base model.

Explainable AI (XAI) technologies have been developed to make AI to explain its behavior. Early works of XAI are based on visualization[7]. Recently, LIME[8] realizes XAI by deriving important features and Grad-CAM[9] realizes XAI by highlighting input image regions. Some recent works have focused on graphs or KGs. GNNExplainer[10] specifies an important subgraph related to target prediction. XKE[6] specifies important paths related to target relation from KGs. [11] realizes XAI by linking target prediction with external KGs.

3 PROPOSED FRAMEWORK

Figure 1 illustrates our reasoning framework which combines multiple KG embedding methods and XAI technologies. As a pre-process, the framework learns an embedding function $F(\cdot;KG)$ of entities and relations from given KGs. By this embedding function, we can build the embedding space which consists of embedded representations, $\boldsymbol{h}=F(h;KG)$, $\boldsymbol{r}=F(r;KG)$, and $\boldsymbol{t}=F(t;KG)$. As a main process, a human enters a query for inference of a missing fact in KGs, e.g., a missing head entity h' in a triple (h', r, t). Then, a human receives the most plausible query result from the framework which calculates the plausibility of candidate answers by a scoring function of a triple $G(\cdot;KG)$ correspondent to the embedding method. By this scoring function, a completed triple (h', r, t) with a candidate missing head entity h' can be evaluated as $G((h', r, t);KG)$. At the same time, the framework specifies an important subgraph of KGs as reasons which support the query result by a subgraph selection function $H(\cdot;(h', r, t), KG)$.

3.1 Implementation Details

Here, I incorporated three KG embedding methods for learning the embedding function $F(\cdot;KG)$: two major translation-based methods (TransE, PTransE) and one major GNNs-based method (R-GCN with DistMult). To integrate multiple KGs, I applied parameter sharing proposed in ITransE, which shares same embeddings for common entities, e.g., "Holmes" and "Watson", and for common relations, e.g., "kill". By this, for instance, entity embeddings of "Holmes" are regularized to be same for all KGs in which "Holmes" appears. On the other hand, I did not apply soft entity alignment which is also utilized in ITransE, for there are too little common entities among KGs. The scoring function for knowledge reasoning $G(\cdot;KG)$ is designed consistent to the objective function utilized in each KG embedding method. For example,in case of TransE, this function is represented as $G((h', r, t);KG) = -\|\boldsymbol{h'}+\boldsymbol{r}-\boldsymbol{t}\|$ using embedded representations $\boldsymbol{h'}$, $\boldsymbol{r}$, and $\boldsymbol{t}$.

For learning the subgraph selection function $H(\cdot;(h', r, t), KG)$, I incorporated one method specific to KGs (XKE) and one method for general graphs (GNNExplainer). The original XKE evaluates a path as a coefficient value of a surrogate linear predictor model which predicts a binary output indicating relation existence (1: exist, 0: not exist) from multiple existing paths for given head and tail entities. In implementing XKE, I changed the surrogate predictor a little by making $softmax(-\|\boldsymbol{h}+\boldsymbol{r}-\boldsymbol{t}\|)$ as the output variable so that the predictor can more easily reflect the characteristics of embedding space. GNNExplainer employs a different principle based on information theory and evaluates a subgraph of KGs. I utilized XKE in case of using TransE or PTransE for KG embedding, and utilized GNNExplainer in case of using R-GCN with DistMult for KG embedding.

4 EXPERIMENTS

In this section, I illustrate experiments using the third knowledge graph reasoning challenge dataset. I first show experimental settings including KGs and query design, and then show experimental results including query result accuracy, embedding space visualization, and example reasons derived by XAI methods. There are two research questions (RQs): **RQ1 (Accuracy)** : How accurate can the learned scoring function $G(\cdot;KG)$ infer missing facts?, and **RQ2 (Explainability)**: How many relevant reasons can the learned subgraph selection function $H(\cdot;(h', r, t), KG)$ extract from KGs?.

4.1 Experimental settings

Table 1 shows statistics of KG datasets. Novel-related KGs, namely novel KGs, are obtained using a SPARQL query applied to the third knowledge graph reasoning challenge (KGRC) dataset[1], which are made from seven Sherlock Holmes short novels manually by KGRC committee. The SPARQL query extracts a subject, a predicate, and the other elements from an entry. I map the subject to a head entity, the predicate to a relation, and the other elements to tail entities. Multiple triples can be extracted from an entry when there

[1]https://github.com/KnowledgeGraphJapan/KGRC-RDF/tree/master/2020v2

Table 1: Statistics of Datasets

KGs	#entity	#relation	#triple
Speckled Band (SB)	277	107	596
Devil's Foot (DF)	597	161	794
A Case of Identity (CI)	655	186	798
Crooked Man (CM)	389	149	563
Abbey Grange (AG)	253	74	390
Resident Patient (RP)	347	194	497
Silver Blaze (SBz)	446	151	561
ConceptNet	2,272	750	4,847

Table 2: Query Design

Novel	Query	Answer
SB	?, kill, julia	roylott
DF(Case1)	?, kill, brenda	mortimer
DF(Case2)	?, kill, mortimer	standale
CI	?, hide, hozma	windybank
AG	?, kill, sir eustace brackenstall	jack crocker
RP	?, kill, blessington	elderly man
		young man
		helper boy
SBz	?, take, silver blaze	john straker

exist multiple tail entities for the same head entity and relation. I regard entities with the same name in different novel KGs as distinct entities, while relations with the same name in different novel KGs as an identical relation. Hereafter, I abbreviate the novel name like "SB" for "Speckled Band". In test, I eliminate latter 10% triples of novel KGs which are deeply related to the true answer for reasoning.

I introduce an external commonsense KG, ConceptNet[12], by relating novel entities and commonsense entities (E-E) and novel relations and commonsense entities (E-R). I collect ConceptNet triples within one-hop apart from direct E-E or E-R links.

Table 2 shows query design for each novel. The query is designed to answer the same conclusion as Sherlock Holmes solved in each novel. For example, in the novel titled "Speckled Band (SB)", "Who killed Julia?" is what to infer and I convert it to incomplete triple, (?, kill, julia), as a query. For this, "Roylott" is the true answer. The true answer for the query is not included in the corresponding KG except a KG of "Abbey Grange (AG)" in cases where I eliminate latter 10% triples from each KG.

4.2 Results and discussions for RQ1 (Accuracy)

To answer RQ1 (Accuracy), Table 3 shows the reasoning accuracy by illustrating the ranks of true answer in query results. The ranks in braces indicate the rank after post-process in which I filter only entities related to the characters in each novel from query results. The lower the rank, the better: "1" means that the best plausible query result is the true answer, "2" means that the second best plausible query result is the true answer, and so on. For "Resident Patient (RP)" which has three true answers, I evaluate the rank by the top rank at which either of the three appears in the query results. Each row indicates the combination of a KG embedding method and input KGs, and each column indicates the target novel.

From this table, I can find, for each column except "Resident Patient (RP)", at least one row in which the result is "1", i.e. the true answer can be accurately inferred. I can also find that ConceptNet is effective in some cases, e.g., "SB", "DF(Case1)", and "RP" in case of PTransE. The performance of R-GCN with DistMult is lower than those of TransE and PTransE but it can be found that the performance becomes much better after post-process. It might imply that R-GCN with DistMult can order entities relative to the queries while it has weakness in distinguishing persons from entities. The reason is still unclear but oversmoothing by the graph convolution operator might be one cause.

4.3 Results and discussions for RQ2 (Explainability)

To inspect potential explainability, Figure 2 visualizes KG embedding space of PTransE using Novel KGs + ConceptNet (E-E) by principal component analysis in 2D. In this figure, I can find comprehensive characteristics: i) Holmes and Watson are set close to the criminals, ii) criminals and victims are set close to each other, and iii) weapons used by criminals are set close to criminals but motivations are set relatively apart from criminals.

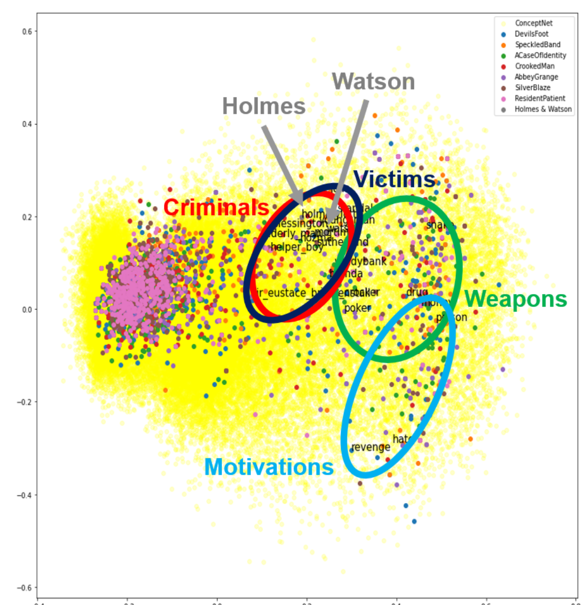

Figure 2: KG embedding visualization

To answer RQ2 (Explainability), I investigate the reasons for query results of "Speckled Band (SB)", "Devil's Foot (DF)", and "Abbey Grange (AG)" in case of PTransE using Novel KGs + ConceptNet (E-E). The most important paths using XKE related to the query results are as follows:

SB: roylott -(come)- house_of_holmes -(_come)- helen -(return)- julia

Table 3: Reasoning Results: Rank of true answer in query results

Methods	Input KGs	Novel						
		SB	DF		CI	AG	RP	SBz
			Case1	Case2				
TransE	Novel KGs	5 (2)	6 (5)	2 (2)	12 (5)	**1** (**1**)	5 (5)	3 (3)
	Novel KGs + ConceptNet (E-E)	6 (3)	4 (4)	2 (2)	3 (3)	2 (2)	6 (6)	4 (4)
	Novel KGs + ConceptNet (E-E + E-R)	2 (2)	2 (2)	4 (3)	2 (2)	**1** (**1**)	8 (8)	4 (3)
PTransE	Novel KGs	4 (2)	4 (4)	2 (2)	3 (3)	**1** (**1**)	7 (7)	**1** (**1**)
	Novel KGs + ConceptNet (E-E)	**1** (**1**)	**1** (**1**)	**1** (**1**)	4 (4)	**1** (**1**)	6 (6)	2 (2)
	Novel KGs + ConceptNet (E-E + E-R)	**1** (**1**)	3 (3)	5 (4)	34 (4)	2 (2)	6 (5)	**1** (**1**)
R-GCN with DistMult	Novel KGs	264 (7)	311 (7)	516 (8)	4 (**1**)	2 (2)	415 (16)	371 (3)
	Novel KGs + ConceptNet (E-E)	267 (7)	116 (3)	576 (10)	607 (8)	113 (3)	228 (10)	213 (5)
	Novel KGs + ConceptNet (E-E + E-R)	20 (**1**)	164 (6)	20 (2)	56 (2)	261 (7)	228 (13)	224 (7)

DF(Case1): mortimer -(equalto)- victim_of_case_2 -(_equalto)- mortimer -(_meet)- standale -(love)- brenda

DF(Case2): standale -(say)- 422 -(_say)- standale -(meet)- mortimer

AG: windybank -(equalsto)- hozma

Here, terms in braces like "(equalto)" mean relations, terms in braces with underscore like "(_equalto)" mean reversed relations, and the others mean entities. In SB, the reason can be seen related to the circumstances of the criminal. In DF, the reason for Case1 can be seen related to the relationship with Case2, and the reason for Case2 can be seen related to the relationship between the victim and the criminal. In AG, the reason can be seen related to the key point of the case, i.e. Windybank and Hozma are the identical person.

Using GNNExplainer, I investigate the reasons for query results of "A Case of Identity (CI)" in case of R-GCN with DistMult using Novel KGs. Table 4 shows top-5 important triples related to the query results. In this table, it can be found that the important triples are related to confidential relationship between Sutherland and Hozma.

Table 4: Important Triples for A Case of Identity (CI) by GNNExplainer

Triple	Score
(sutherland, notworry, windybank)	0.883
(sutherland, getengaged, hozma)	0.874
(sutherland, believe, hozma)	0.841
(sutherland, unforgettable, hozma)	0.834
(sutherland, love, hozma)	0.834

I search the plausible weapons and motivations from the commonsense knowledge in ConceptNet triples for "Devil's Foot (DF) (Case1)" using GNNExplainer. Here, I utilize R-GCN with DistMult using Novel KGs + ConceptNet (E-E + E-R) for KG embedding. Table 5 and 6 show the results. In ConceptNet triples, there are 19 candidate weapons, i.e., head entities in triples like (?, used for, kill), and 38 candidate motivations, i.e., tail entities in triples like (kill, motivated by, ?). In the top-5 candidates, I can find the true weapon (rank 5: poison) and motivation (rank 4: money).

Table 5: Candidate Weapons for Devil's Foot (DF) (Case1) by GNNExplainer

Triple	Score
(sword, used for, kill)	0.383
(bomb, used for, kill)	0.378
(gun, used for, kill)	0.375
(bullet, used for, kill)	0.375
(**poison**, used for, kill)	0.374

Table 6: Candidate Motivations for Devil's Foot (DF) (Case1) by GNNExplainer

Triple	Score
(kill, motivated by goal, evil)	0.384
(kill, motivated by goal, lunch)	0.381
(kill, motivated by goal, die)	0.379
(kill, motivated by goal, **money**)	0.378
(kill, motivated by goal, fool)	0.378

5 CONCLUSION

I proposed a reasoning framework which combines multiple KG embedding techniques with corresponding XAI techniques. Experiments using the third knowledge graph reasoning challenge dataset demonstrated the effectiveness of our framework. For further study, I will incorporate various related methods into our framework and evaluate XAI results in detail.

ACKNOWLEDGMENTS

This research was partially supported by JST CREST Grant Number JPMJCR21F2, Japan.

REFERENCES

[1] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. In Advances in neural information processing systems.

[2] Y. Lin, Z. Liu, H. Luan, M. Sun, S. Rao, and S. Liu. 2015. Modeling Relation Paths for Representation Learning of Knowledge Bases. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 705-714.

[3] M. Schlichtkrull, T. N. Kipf, P. Bloem, R. van den Berg, I. Titov, and M. Welling. 2018. Modeling relational data with graph convolutional networks. In European

Semantic Web Conference, 593-607.
[4] B. Yang, W.-t. Yih, X. He, J. Gao, and L. Deng. 2015. Embedding entities and relations for learning and inference in knowledge bases. In Proceedings of the 3rd International Conference on Learning Representations.
[5] H. Zhu, R. Xie, Z. Liu, and M. Sun. 2017. Iterative entity alignment via knowledge embeddings. In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI).
[6] A. C. Gusmao, A. H. C. Correia, G. De Bona, and F. G. Cozman. 2018. Interpreting Embedding Models of Knowledge Bases: A Pedagogical Approach. In ICML Workshop on Human Interpretability in Machine Learning (WHI 2018).
[7] O. Biran, and C. Cotton. 2017. Explanation and justification in machine learning: A survey. In IJCAI-17 workshop on explainable AI (XAI).
[8] M. T. Ribeiro, S. Singh, and C. Guestrin. 2016. Why should I trust you? Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 1135-1144.
[9] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. 2017. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE international conference on computer vision, 618-626.
[10] Z. Ying, D. Bourgeois, J. You, M. Zitnik, and J. Leskovec. 2019. GNNExplainer: Generating explanations for graph neural networks. In Advances in neural information processing systems.
[11] M. Fuji, H. Morita, K. Goto, K. Maruhashi, H. Anai, and N. Igata. 2019. Explainable AI through combination of deep tensor and knowledge graph. Fujitsu Scientific and Technical Journal, 55(2), 58-64.
[12] R. Speer, J. Chin, and C. Havasi. 2017. ConceptNet 5.5: An Open Multilingual Graph of General Knowledge. In Proceedings of the 31th AAAI.