Journal of Electronic Science and Technology 20 (2022) 100159

Contents lists available at ScienceDirect

Journal of Electronic Science and Technology

journal homepage: www.keaipublishing.com/en/journals/journal-of-electronic-science-and-technology/

Invited Paper

Knowledge graph and knowledge reasoning: A systematic review

Ling Tian [a,b,c,*], Xue Zhou [c], Yan-Ping Wu [c], Wang-Tao Zhou [c], Jin-Hao Zhang [c], Tian-Shu Zhang [c]

[a] *Kashi Institute of Electronics and Information Industry, Kashi, 844099, China*
[b] *Shenzhen Institute of Information Technology, Shenzhen, 518172, China*
[c] *School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, 610054, China*

ARTICLE INFO

Publishing editor: Xuan Xie

Keywords:
Knowledge graph (KG)
Knowledge graph applications
Knowledge hypergraph
Knowledge reasoning

ABSTRACT

The knowledge graph (KG) that represents structural relations among entities has become an increasingly important research field for knowledge-driven artificial intelligence. In this survey, a comprehensive review of KG and KG reasoning is provided. It introduces an overview of KGs, including representation, storage, and essential technologies. Specifically, it summarizes several types of knowledge reasoning approaches, including logic rules-based, representation-based, and neural network-based methods. Moreover, this paper analyzes the representation methods of knowledge hypergraphs. To effectively model hyper-relational data and improve the performance of knowledge reasoning, a three-layer knowledge hypergraph model is proposed. Finally, it analyzes the advantages of three-layer knowledge hypergraphs through reasoning and update algorithms which could facilitate future research.

1. Introduction

The knowledge graph (KG) describes the objective world's concepts, entities, and their relationships in the form of graphs. It can organize, manage, and understand massive information in a way close to human cognitive thinking. In that case, KG plays an important role in a variety of downstream applications, such as semantic search, intelligent recommendation, and question answering.

KG reasoning, which is essential for the KG applications, improves the completeness of KG by inferring new knowledge. However, there is little research that systematically summarizes and analyzes KG reasoning. Furthermore, KG reasoning has attracted wide attention and a series of advanced approaches have emerged in recent years. Hence, a comprehensive review of KG reasoning is necessary and could promote the development of related research.

The main contributions of this paper include: First, it summarizes the representation, storage, and essential technologies of KG. Second, it gives a fine-grained classification of KG reasoning, and analyzes the typical methods, solutions, and advantages and shortcomings of each category in detail. Third, it summarizes the research of knowledge hypergraphs. Finally, it proposes a three-layer framework of knowledge hypergraphs, which can greatly improve the ability and efficiency of reasoning as well as

* Corresponding author. Kashi Institute of Electronics and Information Industry, Kashi, 844099, China.
E-mail addresses: lingtian@uestc.edu.cn (L. Tian), 202011090917@std.uestc.edu.cn (X. Zhou), 202021080530@std.uestc.edu.cn (Y.-P. Wu), wtzhou@std.uestc.edu.cn (W.-T. Zhou), jinhaozhang@std.uestc.edu.cn (J.-H. Zhang), tianshuzhang@std.uestc.edu.cn (T.-S. Zhang).

https://doi.org/10.1016/j.jnlest.2022.100159
Received 24 May 2021; Received in revised form 25 May 2022; Accepted 28 May 2022
Available online 31 May 2022

updating.

2. KG overview

Since Feigenbaum worked with Stanford University proposed expert systems in 1965 [1], artificial intelligence (AI) research has changed from traditional reasoning algorithms to knowledge-driven algorithms. And in 1968, Quillian proposed the knowledge expression model of semantic networks [2]. The establishment of the knowledge base (KB) and knowledge representation (KR) methods had fueled a wave of hot research topics. Then in 1977, Feigenbaum proposed the concept of knowledge engineering [3], demonstrating how to represent knowledge using the principles, methods, and technology of AI. Later, in 1989, Berners-Lee invented the World Wide Web (WWW) [4]. The Semantic Web concept, which combined traditional AI with WWW, was proposed in 1998 [5].

In 2012, Google proposed KG which was defined as KB [6]. KG uses semantic retrieval methods to collect information from multiple sources to improve the quality of searches. In essence, KG is a semantic network composed of various entities, concepts, and relationships [7]. This section mainly introduces the KG architecture, KR, storage, and essential technology. And the technology application framework of KG is shown in Fig. 1. It has 5 main components: KR learning (KRL), knowledge storage (KS), KG construction (KGC), knowledge updating (KU), and knowledge reasoning. Besides, KGC includes knowledge extraction (KE), knowledge fusion (KF), and knowledge processing (KP). The KG technology is the basis of knowledge reasoning and KG applications.

2.1. *Architecture of KG*

KG, which can be divided into a pattern layer and a data layer, needs certain constraints and norms to form a logical architecture. The pattern layer represents the data structure of knowledge, the hierarchical structure, and definitions of knowledge classes, such as the entity, relation, and attribute. It restricts the specific knowledge form of the data layer. The knowledge triples in the data layer are viewed as units to store specific data information. Thus, KG can be generally expressed in the form of triples $G = \{E,R,F\}$. Among them, E represents the set of entities $\{e_1, e_2, \cdots, e_i\}$. Entity e is the basic element in KG. It refers to things that exist objectively and can be distinguished from each other, including people, things, or abstract concepts. R represents the set of relationships $\{r_1, r_2, \cdots, r_j\}$. And the relationship r represents the edge of a certain connection between two different entities in KG. F represents the set of facts $\{f_1, f_2, \cdots, f_k\}$ and each fact is defined as a triple $(h,r,t) \in F$, where h, r, and t represent the head entity, relationship, and tail entity, respectively. For example, basic types of facts can be expressed as triples (Entity, Relation, Entity), (Entity, Attribute, Value), etc.

As shown in Fig. 2, the knowledge triple can be expressed through directed graph structures. The one-way arrow indicates the asymmetric relationship "Starring", while the two-way arrow indicates the symmetric relationship "Co-star".

2.2. *KR and KS*

KR and KS are the bases of the construction, management, and application of KG. Modern KGs are based on a massive amount of Internet data. Its growing scale poses a new challenge of the effective representation and storage of knowledge.

2.2.1. *KR*

KR refers to a method of knowledge descriptions. It could transform massive amounts of information in the real world into structured data by using information technology. Early KR methods include the first-order logic, Horn logic, semantic networks, production

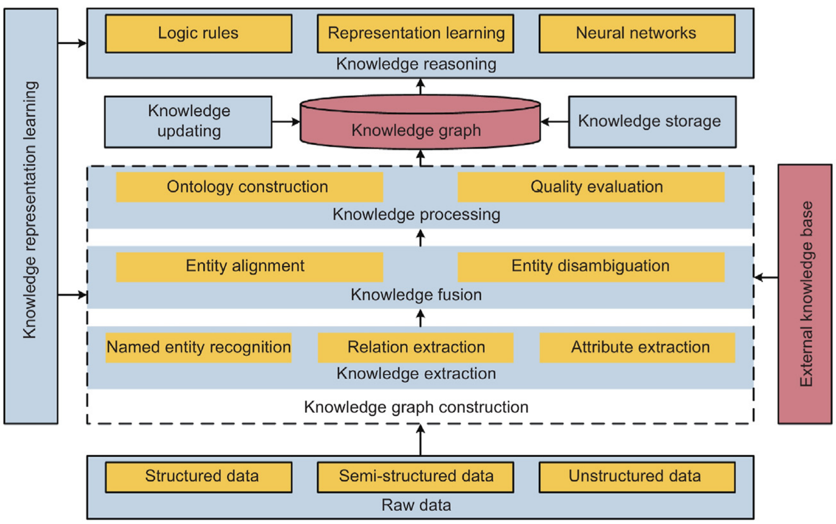


Fig. 1. Technology application framework of KG.

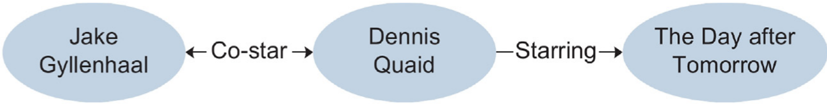


Fig. 2. Triple example.

systems, and frame systems. With the development of the Internet and the Semantic Web, WWW Consortium (W3C) proposed the resource description framework (RDF), RDF schema (RDFS), and web ontology language (OWL) description languages. However, these traditional KR methods are based on symbolic logic, which are unable to represent a large amount of the real-world data. And it is also difficult to mine and analyze the semantic relationships between knowledge entities. In recent years, KRL based on deep learning has attracted extensive attention in the fields of speech recognition, image analysis, and natural language processing (NLP). KRL can calculate complex semantic relationships among entities efficiently through projecting the semantic information (such as triples) into a dense low-dimensional vector space. And it can be easily integrated with deep learning models to realize knowledge reasoning. We will introduce the details of knowledge reasoning based on KRL in Section 3.

2.2.2. KS

The purpose of KS is to find a reasonable and efficient approach for storing KGs. In existing research, most KGs are based on graph data structures. And there are three main storage methods of KS: RDF databases, relational databases (RDBs), and graph databases (GDBs). Nowadays, the main open-source RDF databases in academia include Jena [8]、RDF4J [9], and gStore [10]. RDBs include PostgreSQL [11] and MySQL [12]. Typical GDBs include Neo4j [13], JanusGraph [14], and HugeGraph [15].

2.3. Essential technology

It requires a variety of efficient KG essential technologies to build large-scale, high-quality general KGs or domain KGs. Through the technologies, including KE, KF, KP, and KU, accurate extraction and rapid aggregation of knowledge can be achieved.

2.3.1. KE

KE is the basic and primary task of KG. Different entities, relationships, and attributes are extracted from raw data through automatic or semi-automatic KE, to support the construction of KGs. Early KE is mainly based on the expert rules. Through pre-defined KE rules, triple information can be extracted from the text. However, the traditional method relies on experts with domain knowledge to manually define the rules. When the amount of data increases, the rule-based method takes a long time and has poor portability, which is unable to support the construction of large-scale KGs.

Compared with rule-based KE, neural network-based KE can discover the entity, relationship, and attribute features automatically. And it is suitable for processing large-scale knowledge. The three main tasks of KE include named entity recognition (NER), relation extraction (RE), and attribute extraction (AE).

NER accurately extracts named entity information, such as people, locations, and organizations, from massive amounts of raw data (such as text). The main neural network-based methods for NER include the convolutional neural network (CNN) [16,17] and recurrent neural network (RNN) [18],[[19].

RE is a hot research topic in the field of KGs and also the core of KE. By obtaining semantic relationships or relationship categories among entities, RE can automatically recognize the triples formed by entity pairs and the relationship between two entities. Neural network-based RE methods include CNN [20–22], RNN [23–25], attention mechanisms [26–29], graph convolutional network (GCN) [30–32], adversarial training (AT) [33–35], and reinforcement learning (RL) [36–38]. In recent years, some researchers have proposed methods for joint entity and RE [39–42], which allows the models to incorporate the semantic relevance between entities and relationships. It can solve the problem of overlapping relationships, effectively improving KE.

AE extracts the attribute names and values of the entities from different information sources, constructs the attribute lists of entities, and achieves comprehensive descriptions of the entities. AE is generally divided into traditional supervised models (hidden Markov model, conditional random fields, etc.), unsupervised models [43,44], neural network-based models [45], and other types of AE models (such as meta-patterns [46] and multimodal [47]).

2.3.2. KF

KF involves the fusion of the same entity in different KBs, different KGs, multi-source heterogeneous external knowledge, etc. It determines equivalent instances, classes, and attributes in KGs, to facilitate the update of existing KG. The main tasks of KF consist of entity alignment (EA) and entity disambiguation (ED).

EA constitutes the majority of the work in the KF stage, aiming to discover entities that represent the same semantics in different KGs. EA methods can be classified into traditional probability models (conditional random fields, Markov logic networks, latent Dirichlet allocation, etc.), machine learning models (decision tree, support vector machine, etc.), and neural networks (embedding-based CNN, RNN, etc.).

ED eliminates the ambiguity of entities in different text according to the certain text, and maps them to actual entities that they refer to. Based on the target KB, ED includes named entity clustering disambiguation and named entity linking disambiguation.

2.3.3. KP

KP processes basic facts and forms a structured knowledge system with high-quality knowledge based on KE and KF, to realize the unification of knowledge. KP includes ontology construction (OC) and quality evaluation (QE).

OC refers to the construction of conceptual entities of knowledge in the pattern layer based on the architecture of KG. It standardizes the description of the concepts and relationship between two concepts in a specified field, including concept extraction and inter-concept relationship extraction. According to the extent of automation in the construction process, widely used OC methods include manual construction, semi-automatic construction, and automatic construction.

QE is usually used in KE or the fusion stage, to improve the quality of knowledge extracted from raw data and enhance the effectiveness of KE. And it is able to obtain high-quality and high-confidence knowledge through QE.

2.3.4. KU

KU refers to the update of KG and increase of new knowledge, to ensure the validity of knowledge. As shown in Table 1, KU includes the pattern layer update, data layer update, comprehensive update, and incremental update.

2.4. Category of KG

As shown in Table 2, categories of KG include early KB, open KG, common sense KG, and domain-specific KG.

Early KB is based on expert systems, which is applied to information retrieval and question answering systems. Open KG allows anyone to access, use, and share with the rule of open-source licenses. Common sense KG is significant to many NLP tasks.

Domain-specific KG has a wide range of application scenarios. For example, medical KG can provide correct and efficient retrieval and scientific interpretation, which benefits from medical knowledge question and answer, intelligent diagnosis and treatment, medical quality control and disease risk assessment, etc. Transportation KG can execute traffic volume modeling, air traffic management, and public transport data mining. KGs in finance can save massive amounts of financial data, including entities, relationships, and attributes, allowing for the predictions of hidden risks in complex information and promoting financial upgrading and transformation.

3. Knowledge reasoning methods

Due to the limitations of KG construction and the rapid change in real-world data, KGs usually suffer from missing knowledge triples. Moreover, the performance of downstream tasks can also be severely degraded. Therefore, knowledge reasoning, to predict the missing relationships in KGs, is of great research and application value. This section introduces the details of knowledge reasoning methods based on logic rules, representation learning, and neural networks.

3.1. KG reasoning based on logic rules

Knowledge reasoning based on logic rules refers to using simple rules and features in KGs to discover new facts. These methods can make good use of symbolic representations of knowledge. In that case, they can perform with high accuracy and provide explicit explanations for reasoning results. This subsection introduces three types of KG reasoning approaches based on logic rules, specifically logic-based reasoning, statistics-based reasoning, and graph structure-based reasoning.

3.1.1. Reasoning based on logic

Logic-based knowledge reasoning refers to directly using first-order logic (FOL) and description logic to express the rules formulated by experts. According to the representation methods of rules, logic-based reasoning methods can be classified into reasoning based on FOL and reasoning based on description logic.

Knowledge reasoning based on FOL means that it adopts FOL to represent rules defined by experts, and then performs reasoning tasks by using propositions as basic units. Since it is conducted in a way that is close to natural human language, FOL-based reasoning achieves good interpretability and high accuracy for small-scale KGs.

Propositions are comprised of two parts, individuals and predicates. Individuals and predicates correspond to entities and relations in KG separately. As shown in Fig. 3, the user "Carl" likes the director "Roland Emmerich", and the movie "The Day after Tomorrow" is directed by the director "Roland Emmerich". It is possible that there is a relation "Like" between "Carl" and "The Day after Tomorrow". Therefore, a FOL rule can be obtained:

Table 1
Content of KU.

Update type	Update method
Pattern layer update	Update knowledge class, such as concepts, entities, relationships, and attributes.
Data layer update	Update specific knowledge triples.
Comprehensive update	Combine the new knowledge with the original knowledge to rebuild KG.
Incremental update	Use new knowledge as input data and add it into the existing KG.

Table 2
Categories of KG.

Category	Feature	Examples
Early KB	Constructed by experts but with a complex process and high overload	WordNet [48] and ConceptNet [49]
Open KG	Similar to the data warehouse and allowing to freely access	Freebase [50] and Wikidata [51]
Common sense KG	Diverse, huge, and important for conversational systems	Zhish.me [52] and CN-DBpedia [53]
Domain-specific KG	Complex applications in medical, traffic, finance, and so on	IBM Watson Health medical KG [54]

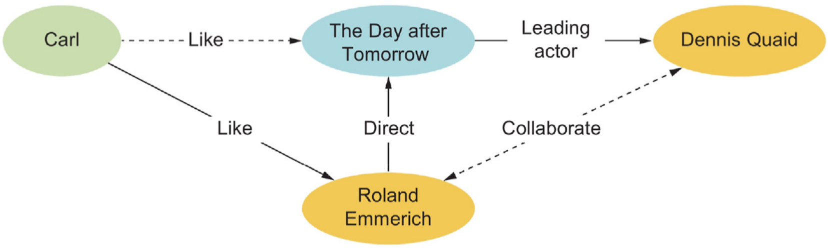


Fig. 3. Example of knowledge reasoning methods based on logic rules.

$$(Carl, Like, Rolan\ Emmerich) \Lambda (Roland\ Emmerich, Direct, The\ Day\ after\ Tomorrow)$$
$$\Rightarrow (Carl, Like, The\ Day\ after\ Tomorrow)$$

where Λ means the conjunction operator, forming a Boolean-valued function, returning true only if all propositions are true.

Reasoning based on description logic aims to transform the complex entity and relation reasoning into a consistency detection problem. This method effectively reduces the reasoning complexity over KGs and achieves the tradeoff between the expressing ability and reasoning complexity. Specifically, KG represented by description logic is composed of terminological axioms (TBoxes) and assertional sets (ABoxes) [55]. TBoxes contain a series of axioms describing concepts and relations, and ABoxes contain examples of concepts in TBoxes. This method conducts reasoning by judging whether a description satisfies logical consistency.

3.1.2. Reasoning based on statistics

Knowledge reasoning based on statistics applies machine learning approaches to automatically extract hidden logic rules from KGs, and conducts reasoning using these rules. Those methods do not depend on expert-defined rules and can interpret the results of the reasoning with automatically extracted logic rules. Statistics-based reasoning methods can be further divided into two subcategories, which are based on inductive logic programming (ILP) and association rule mining, separately.

Reasoning based on ILP refers to using the machine learning and logic programming technology to automatically summarize abstract rule sets. This method abandons the usage of manually defined rules and achieves good reasoning ability over small-scale KGs.

The key idea of reasoning based on association rule mining is to automatically extract high-confidence rules, and then employ these rules in reasoning. Compared with traditional ILP methods, reasoning based on association rule mining is faster and can handle more complex and larger scale KGs.

3.1.3. Reasoning based on graph structure

Graph structure-based reasoning refers to using the structure of the graph as the feature to conduct reasoning. The most typical structure in KG is the paths between entities, which play an important role in KG reasoning. Reasoning based on the graph structure is efficient and interpretable. For example, in Fig. 3, starting from the node "Roland Emmerich", based on the relation path "Direct→Leading actor", it can be inferred that the entity "Roland Emmerich" and the entity "Dennis Quaid" may have the relation "Collaborate". According to the granularity of features, graph structure-based reasoning methods can be divided into global structure-based and local structure-based models.

The essence of global structure-based reasoning is to extract paths of entire KG and use these paths as features to determine the existence of a target relation.

Reasoning with the local structure performs KG reasoning by using the local graph structure that is highly related to the reasoning as features. Compared with reasoning based on the global structure, this method focuses on the finer granularity of features and has a lower computational cost.

Table 3 shows the current research of knowledge reasoning methods based on logic rules.

3.2. KG reasoning based on representation learning

In the field of machine learning, representation learning is a particularly important technique, which transforms complicated data structures into vectors. The traditional one-hot embedding of relational information faces the challenge of feature sparsity under the circumstance of large-scale data. Due to the limited accuracy of knowledge reasoning, semantic information is not incorporated in the

Table 3
Comparison of knowledge reasoning methods based on logic rules.

Author (Year)	Problem	Solution	Category
Richardson and Domingos. (2006) [56]	Handles uncertainty.	Combines FOL and Markov networks to build a network in which each formula has a weight, and performs reasoning on the built networks.	Reasoning based on first-order predicate logic
Pujara et al. (2013) [57]	Incorporates the confidence value.	Develops a model based on probabilistically logic (PSL) to reason candidate facts and their confidence.	Reasoning based on first-order predicate logic
Kuželka and Davis (2020) [58]	Considers the presence of missing data.	Theoretically studies the applicability of learning Markov logic networks (MLN) weights from KB in the presence of missing data.	Reasoning based on first-order predicate logic
Halaschek-Wiener et al. (2006) [59]	Handles low reasoning efficiency.	Presents a description algorithm which completes KG by adding and removing ABox assertions.	Reasoning based on description logic
Calvanese et al. (2006) [60]	Handles incomplete information.	Proposes an epistemic first-order query language epistemic first-order query language (EQL) to handle the incomplete information during the queries over description KG.	Reasoning based on description logic
Li et al. (2006) [61]	Represents knowledge in fuzzy cases.	Proposes a discrete tableau algorithm to solve the reasoning with general TBoxes in fuzzy description logic.	Reasoning based on description logic
Stoilos et al. (2007) [62]	Represents knowledge in fuzzy cases.	Proposes fuzzy description logic SHIN which uses transitive role axioms (S), role hierarchies (H), inverse roles (I), and number restrictions (N) based on the fuzzy set theory, and then completes KR, reasoning by SHIN.	Reasoning based on description logic
Krötzsch et al. (2018) [63]	Copes with property graph reasoning.	Extends description logic by proposing attributed description logic which has sets of attribute-value pairs and can model the metaknowledge of the real world, then performs reasoning based on the attributed description logic.	Reasoning based on description logic
Bienvenu et al. (2019) [64]	Provides interpretability.	Develops a practical method for description logic: The KB query, and then proposes a framework to explain the query answer.	Reasoning based on description logic
Schoenmackers et al. (2010) [65]	Learns inference rules from ambiguous, noisy, and incomplete web extraction.	Proposes an ILP method to automatically extract first-order Horn clauses from open-domain web text.	Reasoning based on ILP
Landwehr et al. (2005) [66]	Learns inference rules from ambiguous, noisy, and incomplete web extraction.	Proposes the first-order inductive learner (FOIL) by integrating the naïve Bayes with the ILP rule-learner FOIL, and uses the naïve Bayes Criterion to guide the search.	Reasoning based on ILP
Landwehr et al. (2010) [67]	Handles low reasoning efficiency.	Develops a framework for statistic relational learning by combining ILP and kernel methods, which greatly reduces the computational complexity.	Reasoning based on ILP
Galárraga et al. (2013) [68]	Supports the open-world assumption.	Proposes an algorithm, association rule mining under incomplete evidence (AMIE), which can efficiently mine the Horn rules of high confidence on large RDF KBs, and then use the mined rules to infer new facts.	Reasoning based on association rule mining
Galárraga et al. (2015) [69]	Copes with larger KG.	Extends AMIE to AMIE + by performing pruning strategies and approximations.	Reasoning based on association rule mining
Wang and Li (2015) [70]	Supports mining multiple rules at one time.	Proposes an efficient algorithm named RDF2Rules, which generates reliable rules by the frequent predicated cycles (FPCs) that are mined from RDF KBs.	Reasoning based on association rule mining
Tanon et al. (2017) [71]	Evaluates and ranks learned rules properly.	Introduces completeness-aware scoring functions for association rules, which can better evaluate the quality of the mined rules from incomplete KGs.	Reasoning based on association rule mining
Lao et al. (2011) [72]	Considers the global structure of KG.	Uses the path connecting the entity pairs that have a target relation as features, trains a logistic regression model for each relation, and then completes KG reasoning by using the trained logistic regression model.	Reasoning based on the global structure
Lao et al. (2015) [73]	Considers the global structure of KG.	Develops an algorithm by considering relational paths with constants, bi-directional path features, and long paths, which can expand the ability of learning of relations.	Reasoning based on the global structure
Wang et al. (2016) [74]	Considers the global structure of KG.	Proposes a multi-task learning framework which considers the association between different relations. It first mines relations with high correlations, and then performs multi-task learning to couple the prediction of these relations.	Reasoning based on the global structure
Gardner and Mitchell (2015) [75]	Considers the local structure of KG.	Uses the breadth-first search method to obtain the subgraphs of the target entity, and then performs multi-feature extraction on the searched subgraphs, and finally achieves reasoning based on the extracted features.	Reasoning with the local structure
Liu et al. (2016) [76]	Considers the local structure of KG.	Proposes a hierarchical random walk algorithm, which improves the reasoning performance by extracting the topological structure of subgraphs that belong to the specific relation.	Reasoning with the local structure

representations of entities and relation types in this scenario. Thus, representations with more semantic information are needed for better reasoning quality. Through KRL, the semantic relationships between entities and relation types are well captured, and contained in the learned distributed embeddings, which guarantees a fixed semantic space. These distributed embeddings with abundant semantic information could facilitate expressive modeling of relations and boost the performance of knowledge reasoning tasks. By mapping hidden relational information in graphs into the low-dimensional space, KRL makes previously obscure correlations in graphs apparent.

Representation learning shows a great advantage over many other approaches. Thus, KG reasoning based on representation learning has been extensively studied in recent years. This subsection introduces and compares three types of KG reasoning approaches based on representation learning, including the tensor decomposition approach, distance model, and semantic matching model.

3.2.1. Tensor decomposition approach

Tensor decomposition-based KG reasoning refers to decomposing the relation tensor of KG into multiple matrices, which can be used to construct a low-dimensional embedding of KG. Existing basic tensor decomposition algorithms are modified and applied to train a distributed representation of KG in a fast and efficient way.

RESCAL [77,78] is currently the most widely used KG reasoning approach based on tensor decomposition. This model represents KG as a three-way tensor, which models the three-way interactions of entity-relation-entity. RESCAL obtains representations of the entities and relation types by solving a simple tensor decomposition problem. The low-dimensional embedding calculated by RESCAL reflects the similarity of neighborhood structures of the entities and relation types in original KG.

A relational graph is shown in Fig. 4, which is a small part of large KG. The movies "2012" and "The Day after Tomorrow" are both directed by "Roland Emmerich" and written by "Harald Kloser", but have different leading actors, namely "John Cusack" and "Dennis Quaid", respectively. The neighborhood structure of "2012" is similar to that of "The Day after Tomorrow", causing the two movies with similar embeddings. As a result, given that the movie "The Day after Tomorrow" belongs to the type "Science fiction", RESCAL can determine that "2012" also belongs to the type "Science fiction".

RESCAL is a classic tensor decomposition model for representation learning. However, RESCAL is too simple and lacks interpretability, it cannot be directly applied in complex scenarios. Hence, numerous tensor decomposition models are developed to enhance the performance of representation learning. Nickel et al. [79], Rendle et al. [80], and Jenatton et al. [81] proposed three other KG representation learning approaches aiming at more complicated application situations. Table 4 shows the details.

3.2.2. Distance model

The distance model considers any relations in KG as a translational transformation from the subject embedding to the object embedding. By minimizing the transformation error, this type of models learns low-dimensional embeddings of all the entities and relation types in KG.

The representative distance model is TransE [82] and its derivatives. TransE, a well-known translational model, learns distributed representations of all entities and relation types by imposing $\mathbf{h}+\mathbf{r}=\mathbf{t}$ for all relation triples in KG, where $\mathbf{h}$, $\mathbf{t}$, and $\mathbf{r}$ denote the embedding of the subject entity, the object entity, and the corresponding relation type, respectively.

Fig. 5 illustrates the mechanism of TransE with the previously given example. In order to find out whether "2012" and "Science fiction" share the relationship "Type", as indicated by the dashed line, TransE projects all entities onto a low-dimensional space. "2012" and "Science fiction" are two existing entities, which are projected to Point A and Point B, respectively. Two other entities named "Mr. Bean" and "Comedy" are then projected to Point C and Point D, respectively. Since they already have the relation "Type", the relation type "Type" can then be projected onto the embedding space as $\boldsymbol{\alpha}$, that is, the translation vector between the embedding point of "Mr. Bean" (Point C) and the embedding point of "Comedy" (Point D). Once all these embeddings are obtained, the existence of relation (2012, Type, Science fiction) can be determined by checking whether the embedding of "2012" (Point A) can be translated to the embedding of "Science fiction" (Point B) via the embedding of "Type" ($\boldsymbol{\alpha}$). If and only if this condition is satisfied, the relation (2012, Type, Science fiction) holds.

TransE achieves good interpretability by formulating the problem in a straightforward way. Nevertheless, TransE has two limitations. One limitation is that the translation rule is too strict, which affects its flexibility and robustness. Thus, a number of relaxed translational models [83–87] are proposed to tackle noise in practical data. The other limitation has to do with the fact that TransE is not suitable for processing 1-to-N, N-to-1, or N-to-N relations, which largely limits its usage in practice. Some works [88–91] aim to solve

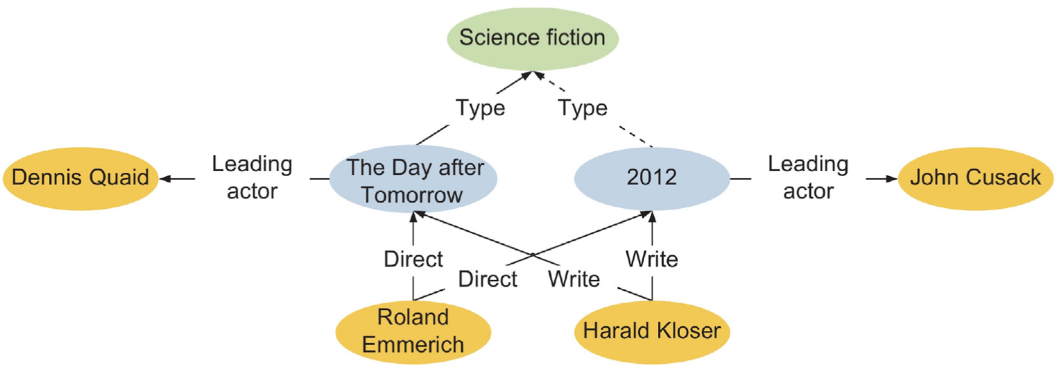


Fig. 4. Example of knowledge reasoning problem.

Table 4
Comparison of knowledge reasoning methods based on representation learning.

Author (Year)	Problem	Solution	Category
Nickel et al. (2011) [77]	Mapping non-Euclidean graph data into vector spaces	Decomposes a three-way tensor to obtain a low-dimensional representation.	Tensor decomposition
Nickel et al. (2012) [79]	Attributes of entities carrying information that may affect KR	Adds an attribute processing module into the RESCAL framework.	Tensor decomposition
Rendle et al. (2013) [80]	High complexity of large-scale KG reasoning	Scales factorization machines to learn representations in linear time.	Tensor decomposition
Jenatton et al. (2012) [81]	Large number of relation types in KG	Shares sparse latent factors across different relations using a bilinear structure.	Tensor decomposition
Bordes et al. (2011) [82]	Poor interpretability of previous KG representation learning approaches	Minimizes the sum of translation errors of all relations in KG.	Simple translations
Bordes et al. (2013) [83]	Too strict translation requirement	Defines a weight for each relation type to measure its importance in the optimization.	Relaxed translations
Xiao et al. (2016) [84]	Too strict translation requirement	Relaxes the problem by mapping the translated vector onto a certain manifold.	Relaxed translations
Feng et al. (2016) [85]	Too strict translation requirement	Applies a scoring function to facilitate flexible translations.	Relaxed translations
Qian et al. (2018) [86]	Too strict translation requirement	Attaches an attention module to TransE, in order to focus on more important relations and ignore the others.	Relaxed translations
Xiao et al. (2015) [87]	Too strict translation requirement	Replaces the Euclidean distance with the Mahalanobis distance, which is more flexible.	Relaxed translations
Wang et al. (2014) [88]	Difficulty in modeling 1-to-*N*, *N*-to-1, or *N*-to-*N* relations	Models relations as hyperplanes rather than vectors, which allows one entity to play different roles in different relations.	Projection space translations
Lin et al. (2015) [89]	Difficulty in modeling 1-to-*N*, *N*-to-1, or *N*-to-*N* relations	Introduces the idea of the entity space and relation space, and models entities and relation types separately in their own spaces.	Projection space translations
Ji et al. (2015) [90]	Time-consuming matrix multiplication of projection operations	Each entity or relation type is represented by two vectors.	Projection space translations
Ji et al. (2016) [91]	A large number of parameters in projection matrices	Simplifies projection matrices by using sparse matrices to reduce training complexity.	Projection space translations
He et al. (2015) [92]	Uncertainty of relations is not modeled by previous models.	Maps entities and relation types onto the multi-dimensional Gaussian distribution space.	Stochastic distance models
Xiao et al. (2016) [93]	Uncertainty of relations is not modeled by previous models.	Represents entities and relation types as stochastic vectors which follow the Gaussian distribution.	Stochastic distance models
Sun et al. (2019) [94]	Subjects and objects are considered to have multiple semantics in different contexts.	Maps entities and relation types onto a complex space and considers a relation as a rotation from the subject to the object.	Rotation models
Ebisu and Ichise (2018) [95]	Regularization in previous translational models may warp the embeddings.	Maps entities and relation types onto a torus to avoid regularization.	Other translational models
García-Durán et al. (2014) [96]	Difficulty in measuring the validity of relations	Evaluates the validity of the relations using a linear scoring function.	Linear semantic matching
Yang et al. (2015) [97]	A large number of parameters in TATEC	Restricts the projection matrix in TATEC to a symmetric matrix.	Linear semantic matching
Nickel et al. (2016) [98]	Balance of efficiency and accuracy of semantic matching	Applies vector convolution and the Fourier transform to fuse the information of subjects and objects.	Linear semantic matching
Trouillon et al. (2016) [99]	High complexity of state-of-the-art semantic matching	Expands the embedding space to a complex space.	Linear semantic matching
Zhang et al. (2019) [100]	High complexity of state-of-the-art semantic matching	Represents entities using hypercomplex-valued embeddings with three imaginary components, and models relations as rotations in the quaternion space.	Linear semantic matching
Liu et al. (2017) [101]	Analogical relations among entities are not considered.	Optimizes the latent representations with aspect to the analogical properties of entities and relation types.	Linear semantic matching
Bordes et al. (2014) [102]	Non-linear relations are hard to capture while representation learning.	Trains a neural network which assigns low energy to plausible relations.	Neural network-based semantic matching
Socher et al. (2013) [103]	Non-linear relations are hard to capture while representation learning.	Trains a neural tensor network, where entities are represented as their constituting word vectors.	Neural network-based semantic matching
Dong et al. (2014) [104]	Non-linear relations are hard to capture while representation learning.	Adopts a simple multi-layer perceptron approach to fuse the information of subjects, relation types, and objects.	Neural network-based semantic matching
Liu et al. (2016) [105]	Non-linear relations are hard to capture while representation learning.	Trains neural networks which take one event as input and compute a conditional probability of the other event to model how likely these two events are associated.	Neural network-based semantic matching

this problem by separating the entity space and relation space to model their interactions in a projection space. Apart from these two types of derivatives, there are works focusing on other problems, including stochastic distance models [92,93], rotation models [94], and other distance models [95]. See Table 4 for details of distance models.

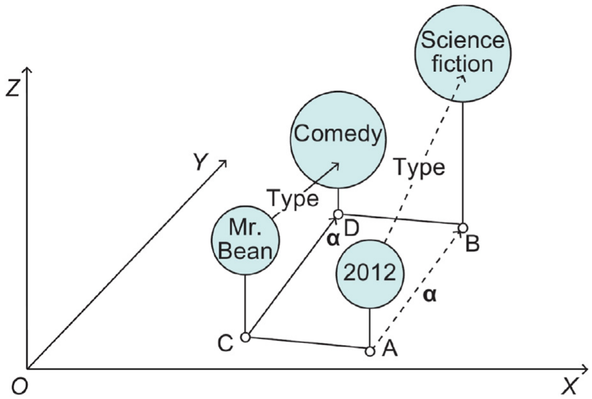


Fig. 5. Example of entity and relation mapping of TransE.

3.2.3. *Semantic matching model*

The semantic matching model measures the validity of relation triples by matching hidden semantics among different entities and relation types in the low-dimensional space, using a scoring function. These models consider relations presented in KG similar, and those not in KG dissimilar.

Currently, the most widely used semantic matching models, e.g., two- and three-way embedding combination (TATEC) [96], mainly focus on matching two-way and three-way semantics in KG. Specifically, the validity of relations is evaluated using a linear scoring function.

Given a relational graph as shown in Fig. 4, TATEC first defines a scoring function for three-way relations, such as (Roland Emmerich, Direct, The Day after Tomorrow). For two-way relations, such as (Roland Emmerich, Directs), (Direct, The Day after Tomorrow), and (Roland Emmerich, The Day after Tomorrow), TATEC defines scoring functions in a similar way to evaluate their validity. For example, the three-way score of (Roland Emmerich, Direct, The Day after Tomorrow) is 0.35, and the two-way scores of (Roland Emmerich, Directs), (Direct, The Day after Tomorrow), and (Roland Emmerich, The Day after Tomorrow) are 0.25, 0.12, and 0.18, respectively. Then the total score for the triple (Roland Emmerich, Direct, The Day after Tomorrow) should be 0.90. The training process maximizes scores of the three-way and two-way relations presented in KG. When determining whether “2012” and “Science fiction” have the relation “Type”, TATEC simply calculates the score of the triple (2012, Type, Science fiction), and establishes the relation if the score is greater than an empirical threshold, such as 0.75.

TATEC suffers from high computational complexity due to a large number of parameters. Hence, a series of linear and bilinear models [97–101] try to balance its performance and complexity.

In order to better capture non-linear patterns of relations, numerous neural network-based representation learning models [102–105] based on semantic matching have been proposed. These models calculate scores of relations through deep neural networks.

The representation learning-based methods of knowledge reasoning are summarized and compared in Table 4.

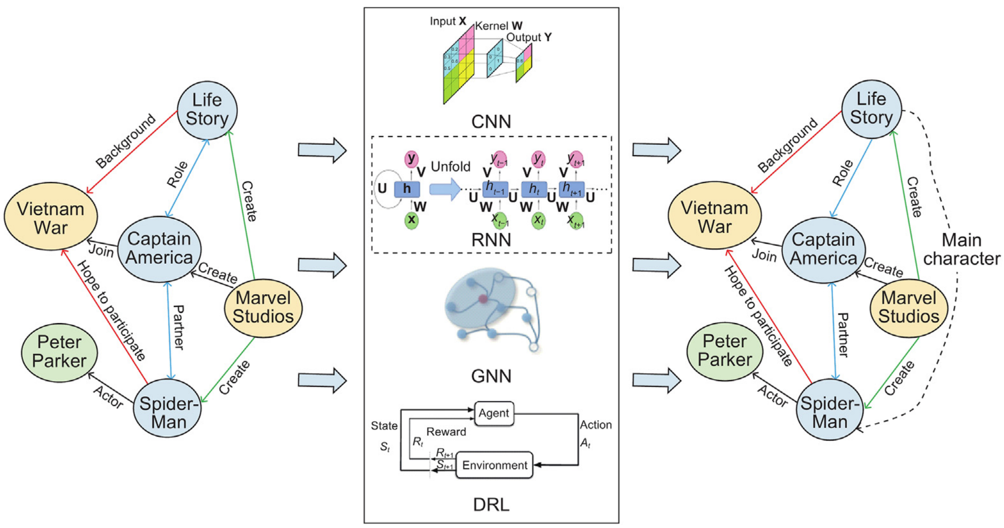


Fig. 6. Example of KG reasoning methods based on RNN.

Table 5
Comparison of knowledge reasoning methods based on neural network models.

Author (Year)	Problem	Solution	Category
Xie et al. (2016) [107]	Ignoring the entity text descriptions	Combines CBOW and CNN to learn features of entity text descriptions.	Reasoning based on entity text descriptions
Shi and Weninger (2018) [108]	The closed world hypothesis	Utilizes CNN to fuse text description features and entity semantic information.	Reasoning based on entity text descriptions
Zhao et al. (2021) [119]	Ignoring entity text descriptions	Uses two convolution operations successively to extract text description features.	Reasoning based on entity text descriptions
Dettmers et al. (2018) [106]	Poor feature learning ability of the traditional model	Performs simple operations, such as two-dimensional convolution, embedding projection, and inner product, on a two-dimensional entity relation embedding matrix.	Reasoning based on entity relation interactions
Vashishth et al. (2020) [120]	Restriction of entity relation interaction features learning	Utilizes some simple techniques like feature replacement, square feature reshaping, and circular convolution to increase the number of interactions between entities and relations.	Reasoning based on entity relation interactions
Zhang et al. (2022) [121]	Poor ability of handling complex relations	Regards alternate entities and relations as the input and learns the sequence characteristics of the interactions among knowledge triples.	Reasoning based on entity relation interactions
Annervaz et al. (2018) [122]	Difficulty in learning knowledge on demand	Integrates the attention mechanism to achieve the deep analysis of the aggregation structure among entities and relations.	Reasoning based on entity relation interactions
Neelakantan et al. (2015) [109]	Using only a single relation to predict new relations	Iteratively employs RNN and learns a combination representation of knowledge paths.	Reasoning based on knowledge path semantics
Das et al. (2017) [110]	Insufficient accuracy of multi-step inference	Introduces the neural attention mechanism and synthesizes the semantic information of multiple relation paths.	Reasoning based on knowledge path semantics
Jagvaral et al. (2020) [111]	Focus only on a single relation.	Models the bidirectional semantics of paths simultaneously.	Reasoning based on knowledge path semantics
Guo et al. (2018) [112]	Indiscriminate learning of entities and relations	Utilizes two independent RNN units to model entity and relations separately.	Reasoning based on knowledge path semantics
Chen et al. (2020) [113]	Word order information loss by CNN	Conducts LSTM to encode related text descriptions, and conducts triples to encode entity descriptions.	Reasoning based on entity text descriptions
Zhao et al. (2019) [123]	The same representation of the same entity in different triples	Integrates the complete attention mechanism to encode entity descriptions and learn the specific semantics of entities in different triples.	Reasoning based on entity text descriptions
An et al. (2018) [124]	The same representation of the same entity in different triples	Synthesizes the mutual attention mechanism to integrate multiple text corpus information.	Reasoning based on entity text descriptions
Shang et al. (2019) [115]	Ignoring the knowledge structure	Utilizes weighted GCN as the encoder and a convolution network Conv-TransE as a decoder.	Reasoning based on GCN
Schlichtkrull et al. (2018) [125]	Ignoring the knowledge structure	Introduces a transformation matrix specific to the relation types in the message transmission process.	Reasoning based on GCN
Zhang et al. (2020) [126]	Ignoring the structural information of neighbors around the entity	Combines the hierarchical attention mechanism and analyzes entity neighborhood information more effectively.	Reasoning based on GCN
Xu et al. (2020) [127]	Lack of reasoning interpretability	Simultaneously encodes a general full graph representation and a local representation of the input information.	Reasoning based on GAT
Zhang et al. (2020) [128]	Cannot incorporate the priori logic rules.	Unites the Markov logic network with GNN and executes probabilistic logic reasoning.	Reasoning based on GAT
Xie et al. (2020) [116]	Ignoring the knowledge structure	Integrates ConvE with KBGAT and learns structural information adaptively.	Reasoning based on GAT
Xiong et al. (2017) [118]	Insufficient accuracy and diversity	Defines behavior as selecting a certain edge connected by the current node and achieves path expansion based on the extension of the behavior sequence.	Reasoning based on relation path exploration
Das et al. (2017) [129]	Lack of reasoning ability of unknown triples	Treats the relation as a query with a known answer to an entity, and guides the map to discover new paths based on the input query.	Reasoning based on relation path exploration
Wang et al. (2020) [130]	Lack of reasoning interpretability	Integrates LSTM with the attention mechanism, and effectively learns the entity neighbor structure.	Reasoning based on relation path exploration
Tiwari et al. (2021) [131]	The same reward in different states	Integrates the memory mechanism of GRU to capture more complete entity information in the path neighbor structure.	Reasoning based on relation path exploration
Li et al. (2018) [132]	Ignoring the entity selection in the inference process	Alternately selects entities and relations and realizes the knowledge reasoning of joint entity relation semantics.	Reasoning based on relation path exploration
Wan et al. (2020) [133]	Ignoring the multiple semantics of entities and relations on multi-hop paths	Encodes the historical information and action space successively and effectively deals with multi-semantic problems.	Reasoning based on relation path exploration
Wang et al. (2020) [134]	Ignoring the domain rules	Integrates the generative adversarial network (GAN) with LSTM to generate trajectories and utilizes domain rules to achieve knowledge reasoning.	Reasoning based on relation path exploration

3.3. KG reasoning based on neural network

Though more and more KG reasoning methods have been proposed, it still remains an open issue on complex and multi-hop relation reasoning. To analyze such relations, the neural network could be a more powerful tool than other reasoning methods based on logic rules or representation learning. After representation learning based on neural networks like CNN or RNN, the accurate understanding of knowledge semantics could benefit the subsequent reasoning process based on the fully connected layer and softmax layer. Moreover, such methods could realize automatical reasoning on KG without logical inference or theoretical modeling. Due to the generality of structures and the convenience of reasoning, KG reasoning methods based on neural networks emerge endlessly. Therefore, this subsection introduces the knowledge reasoning methods based on CNN, RNN, the graph neural network (GNN), and deep reinforcement learning (DRL) in detail.

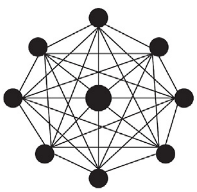
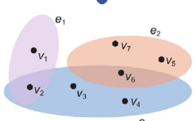
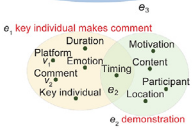
Considering the knowledge subgraph shown in Fig. 6, RNN chooses the red, blue, and green paths as the input to reasoning the target relation between "Spider-Man" and "Life Story". For the red path, RNN pays attention to the words like "hope" and "background" and keeps them in memory cells to increase the probability of the "main character" relation. For the blue path, since the participation of "Captain America" in "Vietnam War", RNN notices the role of "Captain America" in "Life Story" and reasonings the friendship between "Spider-Man" and "Captain America". For the green path, RNN forgets the "create" relation to some extent because all the relations are the same. At last, RNN synthesizes all the information in the three paths and obtains a high probability of the "main character" relation between the "Life Story" entity and the "Spider-Man" entity. Based on the final reasoning result of the friendship of "Spider-Man" and "Captain America" in "Life Story", RNN decides to add a new triple (Life Story, main character, Spider-Man) to the knowledge subgraph.

CNN, which is capable of capturing the local characteristics of KG, is one of the earliest neural networks applied for knowledge reasoning. First, the knowledge triples could be an ideal learning object. With a two-dimensional convolution on a two-dimensional entity relation embedding matrix, ConvE achieves the effective capture of interactive features between entities and relations [106]. Besides, entity text descriptions provide more information for the entities in the knowledge triples. The most typical example is description-embodied knowledge representation learning (DKRL) [107,108], which utilizes continuous bag-of-words (CBOW) and CNN to learn the unordered features and word-order features of text descriptions separately. After that, the fused feature achieves effective discovery of new entities.

Additionally, only focusing on single knowledge triples would limit the reasoning scope. Many studies seek to explore a wider reasoning scope, focusing on using RNN to analyze the knowledge path that is alternately composed of entities and relations [109–112]. This research could effectively combine different representations of knowledge paths. Moreover, RNNs could also be used to analyze entity text descriptions. As a typical method, the learning knowledge graph embedding with entity descriptions based on long short-term memory (LSTM) networks (KGDL) [113] utilizes LSTM to encode related text descriptions, then uses triples to encode entity descriptions, and realizes the prediction of missing knowledge.

Since KG is based on the graph structure, GNN [114] has been widely explored for knowledge reasoning in recent years. It enlarges the learning scope from a single triple in CNN or knowledge path in RNN to knowledge subgraphs in GNN. For instance, the structure aware convolutional network (SACN) [115] utilizes one weighted GCN as an encoder and a convolution network Conv-TransE as a

Table 6
Definitions of knowledge hypergraph and related graphs.

Class	Graph	Definition	Characteristic	Task
Isomorphic graph		$H = (V, E)$, where V represents the node collection and E represents the relation collection.	$\|T_V\| = 1$ and $\|T_E\| = 1$, where $\|T_V\|$ represents the number of node types and $\|T_E\|$ represents the number of relationship types. This graph contains only one kind of nodes and relations, and can only represent binary relationship data.	Link prediction; node classification; clustering
Heterogeneous graph		$H = (V, E)$, where V represents the node collection and E represents the relation collection.	$\|T_V\| + \|T_E\| > 2$. This graph contains more than one kind of nodes and relations, and can only represent binary relationship data.	Link prediction; node classification; clustering
Heterogeneous hypergraph		$H = (V, E)$, where V represents the node collection and E represents the relation collection. For any $e \in E$ and $e(v_1, v_2, \cdots, v_n) \subseteq V$.	$\|T_V\| > 1$. This graph contains more than one kind of nodes and can represent hyper-relational data.	Link prediction; node classification; clustering
Knowledge hypergraph		$H = (V, E)$, where V represents the entity collection and E represents the relation collection. Hyper relational fact $= (e, v_1, v_2, \cdots, v_n)$.	$\|T_V\| > 1$ and $\|TE\| > 1$. This graph contains more than one kind of nodes and relations and can represent hyper-relational data.	Link prediction

decoder. As a result, SACN realizes the adaptive learning of semantic information in the node's neighborhood structure. While GCN has the powerful subgraph learning ability, it simply models a one-way relation as a two-way relation and suffers from the modeling error of relation direction. However, the graph attention network (GAT) explicitly distinguishes the bidirectionality of edges. As an improvement, ReInceptionE [116] combines ConvE and KBGAT to achieve deep understanding of KG structural information.

In terms of interactive modeling ideas, DRL [117] provides a new perspective for knowledge reasoning. As the most typical model, Deeppath [118] treats the knowledge entities as the state space and wanders between entities by choosing relations. If it reaches the correct answer entity, Deeppath will receive a reward. In essence, this kind of method establishes a reasoning plan based on relation path exploration. As a result, DRL methods could significantly improve the effectiveness and diversity of knowledge reasoning.

Table 5 lists the details and comparison of reasoning methods based on neural network models.

4. Knowledge hypergraph theory

Despite KG being universally adopted, representation methods based on triples often oversimplify the complexity of the data stored in KG, especially for hyper-relational data. In hyper-relational data, each fact contains multiple relations and entities. To understand the characteristics of knowledge hypergraphs more clearly, Table 6 shows the definitions of knowledge hypergraphs and related graphs.

Table 7 gives the comparison of knowledge hypergraph representation methods which mainly include soft rules-based, translation-based, tensor decomposition-based, and neural networks-based methods. To realize reasoning, the soft rules-based method considers relation and node as predicate and variable separately. Then it sets the logic rules and constraints for relational reasoning. The translation-based method aims to model the relation as a transformation operation between entities. The tensor decomposition-based method represents a hyper-relational fact as an n-order tensor (n is the number of entities in a hyper-relational fact). Then it learns the embedding of entities through tensor decomposition. The neural networks-based method can learn the interaction information between entities, the topological structure information of the graph, etc.

The knowledge hypergraph adopts a flat structure to organize knowledge and lacks the ability to explicitly leverage temporal and spatial knowledge, it may lead to vague spatio-temporal relations, slow update, and reasoning speed. As shown in Fig. 7, a three-layer knowledge hypergraph (TLKH) is proposed to combine the event layer, concept layer, and instance layer. Compared with the existing knowledge hypergraph, TLKH has the following three advantages: Explicit spatio-temporal relations, fast knowledge update, and fast knowledge reasoning.

TLKH is represented as $G = \{L_1, L_2, L_3, R\}$, where $R = \{R_{(L_1,L_2)}, R_{(L_2,L_3)}\}$ represents the collection of cross-layer relations. The first layer is the event layer, defined as $L_1 = \{E_{L_1}, R_{L_1}\}$, where E_{L_1} and R_{L_1} are the sets of event entities and logical relations between entities separately. The second layer is the concept layer, defined as $L_2 = \{E_{L_2}, R_{L_2}\}$, where E_{L_2} and R_{L_2} are the sets of concept entities and hyperedges separately. The third layer is the instance layer, defined as $L_3 = \{E_{L_3}, R_{L_3}\}$, where E_{L_3} and R_{L_3} are the sets of instance entities and hyperedges separately.

Table 7
Comparison of knowledge hypergraph representation methods.

Author (Year)	Problem	Solution	Category
De Raedt et al. (2007) [135]	Variable hyper-relations	Mines the soft rules in the knowledge hypergraph.	Representation learning based on soft rules
Kazemi (2014) [136]	Variable hyper-relations	Extends logistic regression to hyper-relational data.	Representation learning based on soft rules
Wen et al. (2016) [137]	Limited learning ability	Maps the entity to the hyperplane related to the hyper-relation.	Representation learning based on a translation
Zhang et al. (2018) [138]	Complexity behaves as a high-degree polynomial	Utilizes schema-based filtering as well as relatedness filtering for complexity reduction.	Representation learning based on a translation
Fatemi et al. (2020) [139]	Tensor decomposition models only for binary relations.	Constructs convolution filters based on position.	Representation learning based on tensor decomposition
Liu et al. (2020) [140]	Cannot fully express the knowledge hypergraph.	Combines the Tucker decomposition and tensor ring decomposition and decomposes the core tensor into three-order tensors.	Representation learning based on tensor decomposition
Guan et al. (2019) [141]	Decomposing hyper-relations into triples increases complexity	Uses convolution and fully connected networks to learn hyper-relational facts.	Representation learning based on the fully connected network
Guan et al. (2020) [142]	Ignores knowledge inference on partial facts.	Uses fully connected networks to learn hyper-relational facts.	Representation learning based on the fully connected network
Rosso et al. (2020) [143]	Ignores that the triplet structure is the fundamental data structure, and poor performance.	Adopts CNN to learn hyper-relational facts.	Representation learning based on CNN
Galkin et al. (2020) [144]	Lacks structural information.	Adopts GCN to learn the structure information of knowledge hypergraph.	Representation learning based on GCN

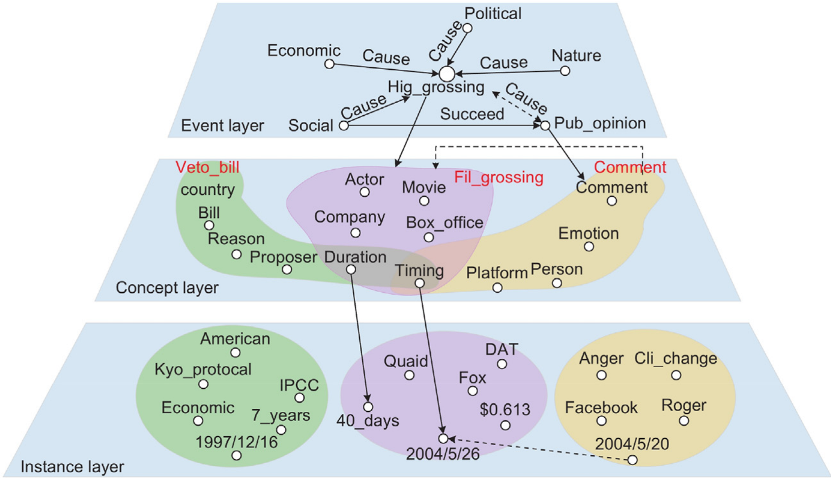


Fig. 7. Structure of three-layer knowledge hypergraph.

Fig. 7 shows the structure of TLKH which organizes the instance, concept, and event knowledge hierarchically. It includes the related entities and relations in the scenario of the highest-grossing movies. Entities in the event layer are abstracted events. For example, "Hig_grossing" can represent multiple specific Hig_grossing events, such as the highest-grossing Hollywood film and the highest-grossing Chinese film. The edges between event entities are logical relations which have various types in different application scenarios. This paper adopts the four main relations which include "Condition", "Compose", "Succeed", and "Cause". For example, "Political", "Nature", "Pub_opinion", "Economic", and "Social" cause "Hig_grossing".

The entity in the concept layer represents a group of different instances sharing common properties, such as "Country". The edge between concept entities is a hyperedge, such as the hyperedge "Fil_grossing" which connects (Company, Actor, Movie, Timing, Duration, Box_office). This hyperedge can clearly express the relevance between concept entities.

The entity in the instance layer is concrete entities that describe the real world, such as "American". The edge between instance entities is hyperedges, such as the hyperedge "Fil_grossing" connecting (Fox, Quaid, DAT, 2004/5/28, 40_days, $0.6B).

The cross-layer relation between the event layer and the concept layer is the relation between the event entity and the concept hyperedge. For example, the event entity "Hig_grossing" corresponds to the concept hyperedge "Fil_grossing". The relation between the concept layer and the instance layer is the mapping relation between the concept entity and the instance entity. For example, the concept entity "Country" corresponds to the instance entity "American".

Timing and duration are proposed to represent the spatio-temporal characteristics of knowledge hypergraphs. Timing is an attribute entity, which represents the entity or hyperedge generated or occurred at a specific time. For example, "Timing" is in the concept layer and "2004/5/20" is in the instance layer. Duration is another attribute entity, which means that the entity or hyperedge is valid within a period, such as "Duration" at the concept layer and "140_days" at the instance layer.

According to TLKH in Fig. 7, the reasoning process is shown in Algorithm 1 (Table 8). According to the entities s_1 (2004/5/20) and s_2 (2004/5/28), there is a "Cause" relation between "Social" and "Hig_grossing" in the event layer. Algorithm 2 (Table 9) shows the update process of TLKH. Inputting a new event entity e_1 and its related concept and instance entities, these entities can be easily added to the right places quickly.

To reason the implicit relations between event entities, the normal knowledge hypergraph needs to find the implicit relation between any entity pairs by reasoning $(n_e+n_c+n_s)^2$ times, where n_e, n_c, and n_s are the numbers of entities in the event, concept, and instance layers. According to Algorithm 1, TLKH needs to find the implicit relation between any instance entity pairs by reasoning n_s^2 times. Implicit relations in the concept and instance layers can be found through cross-layer relations. Through the knowledge complementarity between layers and the spatio-temporal characteristics, the query space of reasoning can be reduced, and the speed of knowledge reasoning can be accelerated dramatically.

To update an event $\{e_1,\{c_1,c_2,\cdots,c_i\},\{s_1,s_2,\cdots,s_j\}\}$, the normal knowledge hypergraph needs to calculate the similarity between $\{e_1, \{c_1,c_2,\cdots,c_i\},\{s_1,s_2,\cdots,s_j\}\}$ and all existing entities with $n_e+in_c+jn_s$ times, where i and j are the numbers of concept and instance entities related to event e_1. According to Algorithm 2, TLKH needs to calculate the similarity between e_1 and other event entities by n_e times. Related entities $\{\{c_1,c_2,\cdots,c_i\},\{s_1,s_2,\cdots,s_j\}\}$ can be located through cross-layer relations. It shows that TLKH is highly efficient, which can greatly accelerate the update process.

Based on the above analysis, the advantages of TLKH are obvious. The hierarchical knowledge hypergraph is powerful when it comes to organizing the domain-specific knowledge. It can learn implicit spatio-temporal relations and greatly accelerate the speed of knowledge update and reasoning. Moreover, TLKH can facilitate the future research of large-scale KG by utilizing hierarchical organization and spatio-temporal characteristics.

Table 8
Algorithm 1.

Reasoning in TLKH
Input: TLKH and a parameter δ Output: A set of new relations R' in event layer
$s_1, s_2 \in E_{L_3}$, $r_{c_1}, r_{c_2} \in R_{L_2}$, $e_1, e_2 \in E_{L_1}$
If time$(s_2) <$ time(s_1) then
$s_1 \rightarrow s_2 \leftarrow$ find explicit relation based on timeline
End if
$r_{c_1} := R_{(L_2,L_3)}(s_1) \leftarrow$ find concept hypergraph through cross-layer relation
$r_{c_2} := R_{(L_2,L_3)}(s_2)$
If similarity$(r_{c_1}, r_{c_2}) > \delta$ then $\leftarrow$ basedon$s_1 \rightarrow s_2$, caculate the similarity of two hypergraphs
$r_{c_1} \overset{\text{cause}}{\rightarrow} r_{c_2}$
End if
$e_1 := R_{(L_1,L_2)}(r_{c_1}) \leftarrow$ find event through cross-layer relation
$e_2 := R_{(L_1,L_2)}(r_{c_2})$
If similarity$(e_1, e_2) > \delta$ then $\leftarrow$ basedon$e_1 \overset{\text{cause}}{\rightarrow} e_2$, caculate the similarity of two events
$e_1 \overset{\text{cause}}{\rightarrow} e_2$
End if
$R' := R' + \{e_1 \overset{\text{cause}}{\rightarrow} e_2\}$

Table 9
Algorithm 2.

Update in TLKH
Input: new entities $\{e_1, \{c_1, c_2, \cdots, c_i\}, \{s_1, s_2, \cdots, s_j\}\}$
Output: The updated knowledge hypergraph
$s_1, (s_1, s_2, \cdots, s_j \in E_{L_3})$; $c_1, (c_1, c_2, \cdots, c_i \in E_{L_2})$; $e_1 \in E_{L_1}$
If $e_1 \notin E_{L_1}$ then
Locate$(e_1) \leftarrow$ find the right place based on entity linking method
Add(e_1)
End if
$\{c_1, c_2, \cdots, c_i\} := R_{(L_1,L_2)}(e_1) \leftarrow$ find concept hypergraph through cross-layer relation
For $c_n \in \{c_1, c_2, \cdots, c_i\}$ do
If $c_n \notin E_{L_2}$ then
Locate$(c_n \setminus e_1) \leftarrow$ find the right place based on event e_1
Add(c_n)
End if
$\{s_1, s_2, \cdots, s_j\} := R_{(L_2,L_3)}(\{c_1, c_2, \cdots, c_i\})$
For $s_m \in \{s_1, s_2, \cdots, s_j\}$ do
If $s_m \notin E_{L_3}$ then
Locate$(s_m \setminus \{c_1, c_2, \cdots, c_i\}) \leftarrow$ find the right place based on event $\{c_1, c_2, \cdots, c_i\}$
Add(s_m)
End if

5. Conclusion

This paper introduces KG and summarizes knowledge reasoning and knowledge hypergraph. Especially, this paper focuses on the reasoning methods, because of the important role of knowledge reasoning in the practical application of KG, based on logic rules, representation learning, and neural networks. And it is beneficial to downstream tasks, such as the link prediction. For effective and efficient reasoning and updating of KG, this paper proposes TLKH, which could express the spatio-temporal relations and make the knowledge reasoning and updating faster.

KG changes the traditional KS and usage approaches, providing a solid knowledge foundation for the development of knowledge-driven artificial intelligence. In the future, KG could influence cognitive intelligence, and knowledge reasoning could be applied and extended to other aspects.

Funding

This work was supported by the Key Science and Technology R&D Project of Sichuan Province under Grants No. 2022YFG0038 and No. 2021YFG0018.

Disclosures

The authors declare no conflicts of interest.

Declaration of competing interest

We declare that we have no financial and personal relationships with other people or organizations that could inappropriately influence or bias our work.

References

[1] Kenyon, October, AI: Timeline. [Online]. Available: https://biology.kenyon.edu/slonc/bio3/AI/TIMELINE/timeline.html, 2018.
[2] M.R. Quillian, Semantic memory, in: M. Minsky (Ed.), Semantic Information Processing, MIT Press, Cambridge, 1968, pp. 227–270.
[3] E.A. Feigenbaum, The art of artificial intelligence: themes and case studies of knowledge engineering, in: Proc. of the 5th Intl. Joint Conf, . on Artificial Intelligence, Cambridge, 1977, pp. 1014–1029.
[4] T. Berners-Lee, R. Cailliau, J.F. Groff, B. Pollermann, World-Wide Web: the information universe, Internet Res. 2 (1) (1992) 52–58.
[5] T. Berners-Lee, J. Hendler, O. Lassila, May, The semantic web [Online]. Available: https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf, 2001.
[6] A. Singhal, May, Introducing the Knowledge Graph: Things, Not Strings, 2012 [Online]. Available: https://www.blog.google/products/search/introducing-knowledge-graph-things-not/.
[7] Wikipedia, March, Google knowledge graph [Online]. Available: https://en.wikipedia.org/wiki/Google_Knowledge_Graph, 2021.
[8] The Apache Software Foundation. (March, Apache jena. [Online]. Available: https://jena.apache.org/, 2021.
[9] Eclipse, RDF4J. [Online]. Available, https://rdf4j.org/, November 2021.
[10] L. Zou, J. Mo, L. Chen, M.T. Özsu, D. Zhao, gStore: answering SPARQL queries via subgraph matching, Proc. VLDB Endowment. 4 (8) (2011) 482–493.
[11] The Postgresql Global Development Group. (March, PostgreSQL: the world's most advanced open source relational database [Online]. Available: https://www.postgresql.org/, 2021.
[12] Oracle, MySQL. [Online]. Available: https://www.mysql.com/, March 2021.
[13] Neo4j Company, Neo4j. [Online]. Available: https://neo4j.com/, February 2021.
[14] The Linux Foundation, March, JanusGraph. [Online]. Available: https://janusgraph.org/, 2021.
[15] Baidu (March, HugeGraph. [Online]. Available: https://github.com/hugegraph/hugegraph, 2021.
[16] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, P. Kuksa, Natural language processing (almost) from scratch, J. Mach. Learn. Res. 12 (2011) 2493–2537, Feb.
[17] E. Strubell, P. Verga, D. Belanger, A. McCallum, Fast and accurate entity recognition with iterated dilated convolutions, in: Proc. of Conf. on Empirical Methods in Natural Language Processing, 2017, pp. 2670–2680. Copenhagen.
[18] Z.-H. Huang, W. Xu, K. Yu (August, Bidirectional LSTM-CRF models for sequence tagging [Online]. Available: https://arxiv.org/abs/1508.01991, 2015.
[19] A. Žukov-Gregorič, Y. Bachrach, S. Coope, Named entity recognition with parallel recurrent neural networks, in: Proc. of the 56th Annual Meeting of the, Association for Computational Linguistics, Melbourne, 2018, pp. 69–74.
[20] D.-J. Zeng, K. Liu, S.-W. Lai, G.-Y. Zhou, J. Zhao, Relation classification via convolutional deep neural network, in: Proc. of the 25th Intl. Conf. on Computational Linguistics: Technical Papers, 2014, pp. 2335–2344. Dublin.
[21] T.H. Nguyen, R. Grishman, Relation extraction: perspective from convolutional neural networks, in: Proc. of the 1st Workshop on Vector Space Modeling for Natural Language Processing, 2015, pp. 39–48. Denver.
[22] J. Kuang, Y.-X. Cao, J.-B. Zheng, X.-N. He, M. Gao, A.-Y. Zhou, Improving neural relation extraction with implicit mutual relations, in: Proc. of IEEE the 36th Intl, . Conf. on Data Engineering, Dallas, 2020, pp. 1021–1032.
[23] R. Socher, B. Huval, C.D. Manning, A.Y. Ng, Semantic compositionality through recursive matrix-vector spaces, in: Proc. of Joint Conf. on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Jeju Island, 2012, pp. 1201–1211.
[24] Y. Xu, L.-L. Mou, G. Li, Y.-C. Chen, H. Peng, Z. Jin, Classifying relations via long short term memory networks along shortest dependency paths, in: Proc. of Conf. on Empirical Methods in Natural Language Processing, 2015, pp. 1785–1794. Lisbon.
[25] Z.-Q. Geng, G.-F. Chen, Y.-M. Han, G. Lu, F. Li, Semantic Relation Extraction Using Sequential and Tree-Structured LSTM with Attention, 509, Information Sciences, Jan. 2020, pp. 183–192.
[26] Y.-T. Shen, X.-J. Huang, Attention-based convolutional neural network for semantic relation extraction, in: Proc. of COLING, the 26th Intl. Conf. on Computational Linguistics: Technical Papers, Osaka, 2016, pp. 2526–2536.
[27] Y.-K. Lin, S.-Q. Shen, Z.-Y. Liu, H.-B. Luan, M.-S. Sun, Neural relation extraction with selective attention over instances, in: Proc. of the 54th Annual Meeting of the, Association for Computational Linguistics, Berlin, 2016, pp. 2124–2133.
[28] G.-L. Ji, K. Liu, S.-Z. He, J. Zhao, Distant supervision for relation extraction with sentence-level attention and entity descriptions, in: Proc. of the 31st AAAI Conf, Artificial Intelligence, San Francisco, 2017, pp. 3060–3066.
[29] Y. Li, G.-D. Long, T. Shen, et al., Self-attention enhanced selective gate with entity-aware embedding for distantly supervised relation extraction, in: Proc. of the 34th AAAI Conf, Artificial Intelligence, New York, 2020, pp. 8269–8276.
[30] Y.-H. Zhang, P. Qi, C.D. Manning, Graph convolution over pruned dependency trees improves relation extraction, in: Proc. of Conf. on Empirical Methods in Natural Language Processing, 2018, pp. 2205–2215. Brussels.
[31] Z.-J. Guo, Y. Zhang, W. Lu, Attention guided graph convolutional networks for relation extraction, in: Proc. of the 57th Annual Meeting of the, Association for Computational Linguistics, Florence, 2019, pp. 241–251.
[32] S.K. Sahu, D. Thomas, B. Chiu, N. Sengupta, M. Mahdy, Relation extraction with self-determined graph convolutional network, in: Proc. of the 29th ACM Intl. Conf. on Information & Knowledge Management, 2020, pp. 2205–2208. Virtual Event.
[33] Y. Wu, D. Bamman, S. Russell, Adversarial training for relation extraction, in: Proc. of Conf. on Empirical Methods in Natural Language Processing, 2017, pp. 1778–1783. Copenhagen.
[34] X.-Z. Wang, X. Han, Y.-K. Lin, Z.-Y. Liu, M.-S. Sun, Adversarial multi-lingual neural relation extraction, in: Proc. of the 27th Intl. Conf, . on Computational Linguistics, Santa Fe, 2018, pp. 1156–1166.
[35] P.-D. Qin, W.-R. Xu, W.-Y. Wang, DSGAN: generative adversarial training for distant supervision relation extraction, in: Proc. of the 56th Annual Meeting of the, Association for Computational Linguistics, Melbourne, 2018, pp. 496–505.
[36] P.-D. Qin, W.-R. Xu, W.-Y. Wang, Robust distant supervision relation extraction via deep reinforcement learning, in: Proc. of the 56th Annual Meeting of the, Association for Computational Linguistics, Melbourne, 2018, pp. 2137–2147.
[37] R. Takanobu, T.-Y. Zhang, J.-X. Liu, M. Huang, A hierarchical framework for relation extraction with reinforcement learning, in: Proc. of the 33rd AAAI Conf, Artificial Intelligence, Honolulu, 2019, pp. 7072–7079.
[38] Y. Xiao, C.-X. Tan, Z.-J. Fan, Q. Xu, W.-Y. Zhu, Joint entity and relation extraction with a hybrid transformer and reinforcement learning based model, in: Proc. of the 34th AAAI Conf, Artificial Intelligence, New York, 2020, pp. 9314–9321.

[39] M. Miwa, M. Bansal, End-to-end relation extraction using LSTMs on sequences and tree structures, in: Proc. of the 54th Annual Meeting of the, Association for Computational Linguistics, Berlin, 2016, pp. 1105–1116.
[40] S.-C. Zheng, F. Wang, H.-Y. Bao, Y.-X. Hao, P. Zhou, B. Xu, Joint extraction of entities and relations based on a novel tagging scheme, in: Proc. of the 55th Annual Meeting of the, Association for Computational Linguistics, Vancouver, 2017, pp. 1227–1236.
[41] S.-C. Zheng, Y.-X. Hao, D.-Y. Lu, et al., Joint entity and relation extraction based on a hybrid neural network, Neurocomputing 257 (Sept. 2017) 59–66.
[42] T. Nayak, H.T. Ng, Effective modeling of encoder-decoder architecture for joint entity and relation extraction, in: Proc. of the 34th AAAI Conf, Artificial Intelligence, New York, 2020, pp. 8528–8535.
[43] S. Raju, P. Pingali, V. Varma, An unsupervised approach to product attribute extraction, in: Proc. of the 31st European Conf. on IR Research: Advances in Information Retrieval, 2009, pp. 796–800. Toulouse.
[44] K. Shinzato, S. Sekine, Unsupervised extraction of attributes and their values from product description, in: Proc. of the 6th Intl. Joint Conf. on Natural Language Processing, 2013, pp. 1339–1347. Nagoya.
[45] G. Zhao, T. Zhang, C.-X. Wang, P. Lyu, J. Wu, Applications of BERT based sequence tagging models on Chinese medical text attributes extraction [Online]. Available: https://arxiv.org/abs/2008.09740v1.
[46] M. Jiang, J.-B. Shang, T. Cassidy, et al., MetaPAD: meta pattern discovery from massive text corpora, in: Proc. of the 23rd ACM Intl. Conf, . *on Knowledge Discovery and Data Mining*, Halifax, 2017, pp. 877–886.
[47] R.L. Logan IV, S. Humeau, S. Singh, Multimodal attribute extraction, in: Proc. of the 6th Workshop on Automated Knowledge Base Construction, Long Beach, 2017, pp. 1–7.
[48] G.A. Miller, WordNet: a lexical database for English, Commun. ACM 38 (11) (Nov. 1995) 39–41.
[49] Oschina (January, ConceptNet. [Online]. Available: http://www.oschina.net/p/conceptnet, 2016.
[50] K. Bollacker, R. Cook, P. Tufts, Freebase: a shared database of structured general human knowledge, in: Proc. of the 22nd AAAI Conf, . on Artificial Intelligence, Vancouver, 2007, pp. 1962–1963.
[51] Wikimedia Foundation, Wikidata. [Online]. Available: https://www.wikidata.org/wiki/Wikidata:Main_Page, November 2015.
[52] X. Niu, X.-R. Sun, H.-F. Wang, S. Rong, G.-L. Qi, Y. Yu, Zhishi.me-weaving Chinese linking open data, in: Proc. of the 10th Intl. Semantic, Web Conf., Bonn, 2011, pp. 205–220.
[53] B. Xu, Y. Xu, J.-Q. Liang, , et al.Cn-DBpedia, A never-ending Chinese knowledge extraction system, in: Proc. of the 30th Intl. Conf. on Industrial Engineering and Other Applications of Applied Intelligent Systems Advances in Artificial Intelligence: From Theory to Practice, Arras, 2017, pp. 428–438.
[54] I.B.M.(March, Watson health. [Online]. Available: https://www.ibm.com/cn-zh/watson-health/learn/healthcare-data-analytics, 2021.
[55] T.W. Lee, M.S. Lewicki, M. Girolami, T.J. Sejnowski, Blind source separation of more sources than mixtures using overcomplete representations, IEEE Signal Process. Lett. 6 (4) (Apr. 1999) 87–90.
[56] M. Richardson, P. Domingos, Markov logic networks, Mach. Learn. 62 (1–2) (Jan. 2006) 107–136.
[57] J. Pujara, H. Miao, L. Getoor, W. Cohen, Knowledge graph identification, in: Proc. of the 12th Intl. Semantic, Web Conf., Sydney, 2013, pp. 542–557.
[58] O. Kuželka, J. Davis, Markov logic networks for knowledge base completion: a theoretical analysis under the MCAR assumption, in: Proc. of the 35th Conf. on Uncertainty in Artificial Intelligence, Tel Aviv, 2019, pp. 1138–1148.
[59] C. Halaschek-Wiener, B. Parsia, E. Sirin, A. Kalyanpur, Description logic reasoning for dynamic aboxes, in: Proc. of Intl. Workshop on Description Logics, Windermere, 2006, pp. 1–8.
[60] D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, R. Rosati, Epistemic first-order queries over description logic knowledge bases, in: Proc. of Intl. Workshop on Description Logics, Windermere, 2006, pp. 1–11.
[61] Y.-H. Li, B.-W. Xu, J.-J. Lu, D.-Z. Kang, Discrete tableau algorithms for FSHI, in: Proc. of Intl. Workshop on Description Logics, Windermere, 2006, pp. 1–12.
[62] G. Stoilos, G. Stamou, J.-Z. Pan, V. Tzouvaras, I. Horrocks, Reasoning with very expressive fuzzy description logics, J. Artif. Intell. Res. 30 (Sept. 2007) 273–320.
[63] M. Krötzsch, M. Marx, A. Ozaki, V. Thost, Attributed description logics: reasoning on knowledge graphs, in: Proc. of the 27th Intl. Joint Conf, . on Artificial Intelligence, Stockholm, 2018, pp. 5309–5313.
[64] M. Bienvenu, C. Bourgaux, F. Goasdoué, Computing and explaining query answers over inconsistent DL-Lite knowledge bases, J. Artif. Intell. Res. 64 (Mar. 2019) 563–644.
[65] S. Schoenmackers, O. Etzioni, D. Weld, J. Davis, Learning first-order horn clauses from web text, in: Proc. of Conf. on Empirical Methods in Natural Language Processing, 2010, pp. 1088–1098. Cambridge.
[66] N. Landwehr, K. Kersting, L. De Raedt, nFOIL: integrating naïve Bayes and FOIL, in: Proc. of the 20th National Conf, . on Artificial Intelligence, Pittsburgh, 2005, pp. 795–800.
[67] N. Landwehr, A. Passerini, L. De Raedt, P. Frasconi, Fast learning of relational kernels, Mach. Learn. 78 (3) (Jan. 2010) 305–342.
[68] L.A. Galárraga, C. Teflioudi, K. Hose, F. Suchanek, AMIE: association rule mining under incomplete evidence in ontological knowledge bases, in: Proc. of the 22nd Intl. Conf. on World Wide Web, 2013, pp. 413–422. Rio de Janeiro.
[69] L. Galárraga, C. Teflioudi, K. Hose, F.M. Suchanek, Fast rule mining in ontological knowledge bases with AMIE+, The VLDB J. 24 (6) (Jul. 2015) 707–730.
[70] Z.-C. Wang, J.-Z. Li, December, RDF2Rules: Learning Rules from RDF Knowledge Bases by Mining Frequent Predicate Cycles, 2015 [Online]. Available: https://arxiv.org/abs/1512.07734v1.
[71] T.P. Tanon, D. Stepanova, S. Razniewski, P. Mirza, G. Weikum, Completeness-aware rule learning from knowledge graphs, in: Proc. of the 16th Intl. Semantic, Web Conf., Vienna, 2017, pp. 507–525.
[72] N. Lao, T. Mitchell, W.W. Cohen, Random walk inference and learning in a large scale knowledge base, in: Proc. of Conf. on Empirical Methods in Natural Language Processing, 2011, pp. 529–539. Edinburgh.
[73] N. Lao, E. Minkov, W.W. Cohen, Learning relational features with backward random walks, in: Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Intl. Joint Conf, . on Natural Language Processing, Beijing, 2015, pp. 666–675.
[74] Q. Wang, J. Liu, Y.-F. Luo, B. Wang, C.-Y. Lin, Knowledge base completion via coupled path ranking, in: Proc. of the 54th Annual Meeting of the, Association for Computational Linguistics, Berlin, 2016, pp. 1308–1318.
[75] M. Gardner, T. Mitchell, Efficient and expressive knowledge base completion using subgraph feature extraction, in: Proc. of Conf. on Empirical Methods in Natural Language Processing, 2015, pp. 1488–1498. Lisbon.
[76] Q. Liu, L.-Y. Jiang, M.-H. Han, Y. Liu, Z.-G. Qin, Hierarchical random walk inference in knowledge graphs, in: Proc. of the 39th Intl. ACM SIGIR Conf. on, Research and Development in Information Retrieval, Pisa, 2016, pp. 445–454.
[77] M. Nickel, V. Tresp, H.P. Kriegel, A three-way model for collective learning on multi-relational data, in: Proc. of the 28th Intl. Conf. on Machine Learning, 2011, pp. 809–816. Bellevue.
[78] M. Nickel, V. Tresp, Tensor factorization for multi-relational learning, in: Proc. of European Conf. on Machine Learning and Knowledge Discovery in Databases, 2013, pp. 617–621. Prague.
[79] M. Nickel, V. Tresp, H.P. Kriegel, Y.A.G.O. Factorizing, Scalable machine learning for linked data, in: Proc. of the 21st Intl. Conf, . on World Wide Web, Lyon, 2012, pp. 271–280.
[80] S. Rendle, Scaling factorization machines to relational data, Proc. of the VLDB Endowment 6 (5) (Mar. 2013) 337–348.
[81] R. Jenatton, N. Le Roux, A. Bordes, G. Obozinski, A latent factor model for highly multi-relational data, in: Proc. of the 25th Intl. Conf. on Neural Information Processing Systems, Lake Tahoe, 2012, pp. 3167–3175.

[82] A. Bordes, J. Weston, R. Collobert, Y. Bengio, Learning structured embeddings of knowledge bases, in: Proc. of the 25th AAAI Conf, Artificial Intelligence, San Francisco, 2011, pp. 301–306.
[83] A. Bordes, N. Usunier, A. Garcia-Durán, J. Weston, O. Yakhnenko, Translating embeddings for modeling multi-relational data, in: Proc. of the 26th Intl. Conf. on Neural Information Processing Systems, Lake Tahoe, 2013, pp. 2787–2795.
[84] H. Xiao, M.-L. Huang, X.-Y. Zhu, From one point to a manifold: knowledge graph embedding for precise link prediction, in: Proc. of the 25th Intl. Joint Conf, . on Artificial Intelligence, New York, 2016, pp. 1315–1321.
[85] J. Feng, M.-L. Huang, M.-D. Wang, M.-T. Zhou, Y. Hao, X.-Y. Zhu, Knowledge graph embedding by flexible translation, in: Proc. of the 15th Intl. Conf. on Principles of Knowledge Representation and Reasoning, 2016, pp. 557–560. Cape Town.
[86] W. Qian, C. Fu, Y. Zhu, D. Cai, X.-F. He, Translating embeddings for knowledge graph completion with relation attention mechanism, in: Proc. of the 27th Intl. Joint Conf, . on Artificial Intelligence, Stockholm, 2018, pp. 4286–4292.
[87] H. Xiao, M.-L. Huang, Y. Hao, X.-Y. Zhu, September, TransA: an adaptive approach for knowledge graph embedding [Online]. Available: https://arxiv.org/abs/1509.05490v1, 2015.
[88] Z. Wang, J.-W. Zhang, J.-L. Feng, Z. Chen, Knowledge graph embedding by translating on hyperplanes, in: Proc. of 28th AAAI Conf. on Artificial Intelligence, Québec City, 2014, pp. 1112–1119.
[89] Y.-K. Lin, Z.-Y. Liu, M.-S. Sun, Y. Liu, X. Zhu, Learning entity and relation embeddings for knowledge graph completion, in: Proc. of the 29th AAAI Conf. on Artificial Intelligence, 2015, pp. 2181–2187. Austin.
[90] G.-L. Ji, S.-Z. He, L. Xu, K. Liu, J. Zhao, Knowledge graph embedding via dynamic mapping matrix, in: Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Intl. Joint, Conf. on Natural Language Processing, Beijing, 2015, pp. 687–696.
[91] G.-L. Ji, K. Liu, S.-Z. He, J. Zhao, Knowledge graph completion with adaptive sparse transfer matrix, in: Proc. of the 30th AAAI Conf, Artificial Intelligence, Phoenix, 2016, pp. 985–991.
[92] S.-Z. He, K. Liu, G.-L. Ji, J. Zhao, Learning to represent knowledge graphs with Gaussian embedding, in: Proc. of the 24th ACM Intl. on Conf, . on Information and Knowledge Management, Melbourne, 2015, pp. 623–632.
[93] H. Xiao, M.-L. Huang, X.-Y. Zhu, TransG: a generative model for knowledge graph embedding, in: Proc. of the 54th Annual Meeting of the, Association for Computational Linguistics, Berlin, 2016, pp. 2316–2325.
[94] Z.-Q. Sun, Z.-H. Deng, J.-Y. Nie, J. Tang, RotatE: knowledge graph embedding by relational rotation in complex space, in: Proc. of the 7th Intl. Conf. on Learning Representations, 2019, pp. 1–18. New Orleans.
[95] T. Ebisu, R. Ichise, TorusE: knowledge graph embedding on a lie group, in: Proc. of the 32nd AAAI Conf. on Artificial Intelligence, 2018, pp. 1819–1826. New Orleans.
[96] A. García-Durán, A. Bordes, N. Usunier, Effective blending of two and three-way interactions for modeling multi-relational data, in: Proc. of European Conf. on Machine Learning and Knowledge Discovery in Databases, Nancy, 2014, pp. 434–449.
[97] B.-S. Yang, W.-T. Yih, X.-D. He, J.-F. Gao, L. Deng, August, Embedding Entities and Relations for Learning and Inference in Knowledge Bases, 2015 [Online]. Available: https://arxiv.org/abs/1412.6575.
[98] M. Nickel, L. Rosasco, T. Poggio, Holographic embeddings of knowledge graphs, in: Proc. of the 30th AAAI Conf. on Artificial Intelligence, 2016, pp. 1955–1961. Phoenix.
[99] T. Trouillon, J. Welbl, S. Riedel, É. Gaussier, G. Bouchard, Complex embeddings for simple link prediction, in: Proc. of the 33rd Intl. Conf, . on Machine Learning, New York, 2016, pp. 2071–2080.
[100] S. Zhang, Y. Tay, L. Yao, Q. Liu, Quaternion knowledge graph embeddings, in: Proc. of the 33rd Conf. on Neural Information Processing Systems, 2019, pp. 2731–2741. Vancouver.
[101] H.-X. Liu, Y.-X. Wu, Y.-M. Yang, Analogical inference for multi-relational embeddings, in: Proc. of the 34th Intl. Conf, . on Machine Learning, Sydney, 2017, pp. 2168–2178.
[102] A. Bordes, X. Glorot, J. Weston, Y. Bengio, A semantic matching energy function for learning with multi-relational data, Mach. Learn. 94 (2) (May 2014) 233–259.
[103] R. Socher, D.-Q. Chen, C.D. Manning, A.-Y. Ng, Reasoning with neural tensor networks for knowledge base completion, in: Proc. of the 26th Intl. Conf. on Neural Information Processing Systems, Lake Tahoe, 2013, pp. 926–934.
[104] X. Dong, E. Gabrilovich, G. Heitz, et al., Knowledge vault: a web-scale approach to probabilistic knowledge fusion, in: Proc. of the 20th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, 2014, pp. 601–610. New York.
[105] Q. Liu, H. Jiang, A. Evdokimov, , et al.August, Probabilistic Reasoning via Deep Learning: Neural Association Models, 2016 [Online]. Available: https://arxiv.org/abs/1603.07704.
[106] T. Dettmers, P. Minervini, P. Stenetorp, S. Riedel, Convolutional 2D knowledge graph embeddings, in: Proc. of the 32nd AAAI Conf. on Artificial Intelligence, 2018, pp. 1811–1818. New Orleans.
[107] R.-B. Xie, Z.-Y. Liu, J. Jia, H.-B. Luan, M.-S. Sun, Representation learning of knowledge graphs with entity descriptions, in: Proc. of the 30th AAAI Conf, Artificial Intelligence, Phoenix, 2016, pp. 2659–2665.
[108] B.-X. Shi, T. Weninger, Open-world knowledge graph completion, in: Proc. of the 32nd AAAI Conf. on Artificial Intelligence, 2018, pp. 1957–1964. New Orleans.
[109] A. Neelakantan, B. Roth, A. McCallum, Compositional vector space models for knowledge base completion, in: Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Intl. Joint Conf, . on Natural Language Processing, Beijing, 2015, pp. 156–166.
[110] R. Das, A. Neelakantan, D. Belanger, A. McCallum, Chains of reasoning over entities, relations, and text using recurrent neural networks, in: Proc. of the 15th Conf. of the, European Chapter of the Association for Computational Linguistics, Valencia, 2017, pp. 132–141.
[111] B. Jagvaral, W.K. Lee, J.S. Roh, M.S. Kim, Y.T. Park, Path-based reasoning approach for knowledge graph completion using CNN-BiLSTM with attention mechanism, Expert Syst. Appl. 142 (Mar. 2020) 1–10.
[112] L.-B. Guo, Q.-H. Zhang, W.-Y. Ge, W. Hu, Y.-Z. Qu, DSKG: a deep sequential model for knowledge graph completion, in: Proc. of the 3rd China Conf. on Knowledge Graph and Semantic Computing, Tianjin, 2018, pp. 65–77.
[113] W.-R. Chen, D.-P. Hong, C. Zheng, Learning knowledge graph embedding with entity descriptions based on LSTM networks, in: Proc. of IEEE Intl. Symposium on Product Compliance Engineering-Asia, 2020, pp. 1–7. Chongqing.
[114] D.Q. Nguyen, V. Tong, D. Phung, D.Q. Nguyen, Node co-occurrence based graph neural networks for knowledge graph link prediction, in: Proc. of the 15th ACM Intl. Conf. on, Web Search and Data Mining, Atlanta, 2022, pp. 1589–1592.
[115] C. Shang, Y. Tang, J. Huang, J.-B. Bi, X.-D. He, B.-W. Zhou, End-to-end structure-aware convolutional networks for knowledge base completion, in: Proc. of the 33rd AAAI Conf, Artificial Intelligence, Honolulu, 2019, pp. 3060–3067.
[116] Z.-W. Xie, G.-Y. Zhou, J. Liu, J.-X. Huang, ReInceptionE: relation-aware inception network with joint local-global structural information for knowledge graph embedding, in: Proc. of the 58th Annual Meeting of the, Association for Computational Linguistics, 2020, pp. 5929–5939, https://doi.org/10.18653/v1/2020.acl-main.526.
[117] Z.-J. Yang, J.-B. Ye, L.-L. Wang, X. Lin, L. He, Inferring substitutable and complementary products with knowledge-aware path reasoning based on dynamic policy network, Knowl. Base Syst. 235 (Jan. 2022) 1–12.
[118] W.-H. Xiong, T. Hoang, W.-Y. Wang, DeepPath: a reinforcement learning method for knowledge graph reasoning, in: Proc. of Conf. on Empirical Methods in Natural Language Processing, 2017, pp. 564–573. Copenhagen.
[119] F. Zhao, T. Xu, L.-J.-Q. Jin, H. Jin, Convolutional network embedding of text-enhanced representation for knowledge graph completion, IEEE Internet Things J. 8 (23) (2021) 16758–16769. Dec.

[120] S. Vashishth, S. Sanyal, V. Nitin, N. Agrawal, P.P. Talukdar, InteractE: improving convolution-based knowledge graph embeddings by increasing feature interactions, in: Proc. of the 34th AAAI Conf, Artificial Intelligence, New York, 2020, pp. 3009–3016.
[121] Z.-L. Zhang, Z.-F. Li, H. Liu, N.-N. Xiong, Multi-scale dynamic convolutional network for knowledge graph embedding, IEEE Trans. Knowl. Data Eng. 34 (5) (May 2022) 2335–2347.
[122] K.M. Annervaz, S.B.R. Chowdhury, A. Dukkipati, Learning beyond datasets: knowledge graph augmented neural networks for natural language processing, in: Proc. of Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2018, pp. 313–322. New Orleans.
[123] M.-J. Zhao, Y.-W. Zhao, B. Xu, Knowledge graph completion via complete attention between knowledge graph and entity descriptions, in: Proc. Of the 3rd Intl. *Conf. on Computer Science and Application Engineering*, Sanya, 47, 2019, pp. 1–6.
[124] B. An, B. Chen, X.-P. Han, L. Sun, Accurate text-enhanced knowledge graph representation learning, in: Proc. of Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2018, pp. 745–755. New Orleans.
[125] M.S. Schlichtkrull, T.N. Kipf, P. Bloem, R. van den Berg, I. Titov, M. Welling, Modeling relational data with graph convolutional networks, in: Proc. of the 15th Intl. Conf. on the Semantic Web, 2018, pp. 593–607. Heraklion.
[126] Z. Zhang, F.-Z. Zhuang, H.-S. Zhu, Z.-P. Shi, H. Xiong, Q. He, Relational graph neural network with hierarchical attention for knowledge graph completion, in: Proc. of the 34th AAAI Conf, Artificial Intelligence, New York, 2020, pp. 9612–9619.
[127] X.-R. Xu, Y. Feng, Y.-S. Jiang, X.-H. Xie, Z.-Q. Sun, Z.-H. Deng, Dynamically pruned message passing networks for large-scale knowledge graph reasoning, in: Proc. of the 8th Intl. Conf. on Learning Representations, Addis Ababa, 2020, pp. 1–30, https://doi.org/10.1109/ISPCE-CN51288.2020.9321857.
[128] Y.-Y. Zhang, X.-S. Chen, A. Yang, et al., Efficient probabilistic logic reasoning with graph neural networks, in: Proc. of the 8th Intl. Conf. on Learning Representations, Addis Ababa, 2020, pp. 1–20, https://doi.org/10.48550/arXiv.2001.11850.
[129] R. Das, S. Dhuliawala, M. Zaheer, et al., Go for a walk and arrive at the answer: reasoning over knowledge bases with reinforcement learning, in: Proc. of the 6th Workshop on Automated Knowledge Base Construction, Long Beach, 2017, pp. 1–18.
[130] Q. Wang, Y.-S. Hao, J. Cao, ADRL: an attention-based deep reinforcement learning framework for knowledge graph reasoning, Knowl. Base Syst. 197 (Jun. 2020) 1–9, 105910.
[131] P. Tiwari, H.-Y. Zhu, H.M. Pandey, DAPath: distance-aware knowledge graph reasoning based on deep reinforcement learning, Neural Network. 135 (Mar. 2021) 1–12.
[132] Z.-X. Li, X.-L. Jin, S.-P. Guan, Y.-Z. Wang, X.-Q. Cheng, Path reasoning over knowledge graph: a multi-agent and reinforcement learning based method, in: Proc. of IEEE Intl. Conf. on Data Mining Workshops, 2018, pp. 929–936. Singapore.
[133] G.-J. Wan, S.-R. Pan, C. Gong, C. Zhou, G. Haffari, Reasoning like human: hierarchical reinforcement learning for knowledge graph reasoning, in: Proc. of the 29th Intl. Joint Conf. on Artificial Intelligence, 2020, pp. 1926–1932. Yokohama.
[134] Q. Wang, Y.-D. Ji, Y.-S. Hao, J. Cao, GRL: knowledge graph completion with GAN-based reinforcement learning, Knowl. Base Syst. 209 (Dec. 2020) 1–8, 10642.
[135] L. De Raedt, A. Kimmig, H. Toivonen, ProbLog: a probabilistic prolog and its application in link discovery, in: Proc. of the 20th Intl. Joint Conf. on Artificial Intelligence, 2007, pp. 2462–2467. Hyderabad.
[136] S.M. Kazemi, M.S. thesis, Relational Logistic Regression, Department of Computer Science, University of British Columbia, Vancouver, 2014.
[137] J.-F. Wen, J.-X. Li, Y.-Y. Mao, S.-N. Chen, R.-C. Zhang, On the representation and embedding of knowledge bases beyond binary relations, in: Proc. of the 25th Intl. Joint Conf, . on Artificial Intelligence, New York, 2016, pp. 1300–1307.
[138] R.-C. Zhang, J.-P. Li, J.-J. Mei, Y.-Y. Mao, Scalable instance reconstruction in knowledge bases via relatedness affiliated embedding, in: Proc, of World Wide Web Conf., Lyon, 2018, pp. 1185–1194.
[139] B. Fatemi, P. Taslakian, D. Vazquez, D. Poole, Knowledge hypergraphs: prediction beyond binary relations, in: Proc. of the 29th Intl. Joint Conf. on Artificial Intelligence, 2020, pp. 2191–2197. Yokohama.
[140] Y. Liu, Q.-M. Yao, Y. Li, Generalizing Tensor Decomposition for N-Ary Relational Knowledge Bases, Proc. of the Web Conf., Taipei, 2020, pp. 1104–1114.
[141] S.-P. Guan, X.-L. Jin, Y.-Z. Wang, X.-Q. Cheng, Link prediction on N-ary relational data, in: Proc. of the World Wide Web Conf., 2019, pp. 583–593. San Francisco.
[142] S.-P. Guan, X.-L. Jin, J.-F. Guo, Y.-Z. Wang, X.-Q. Cheng, Neuinfer: knowledge inference on N-ary facts, in: Proc. of the 58th Annual Meeting of the, Association for Computational Linguistics, 2020, pp. 6141–6151, https://doi.org/10.18653/v1/2020.acl-main.546.
[143] P. Rosso, D.-Q. Yang, P. Cudré-Mauroux, Beyond Triplets: Hyper-Relational Knowledge Graph Embedding for Link Prediction, Proc. of the Web Conf., Taipei, 2020, pp. 1885–1896.
[144] M. Galkin, P. Trivedi, G. Maheshwari, R. Usbeck, J. Lehmann, Message passing for hyper-relational knowledge graphs, in: Proc. of Conf. on Empirical Meth. Nat. Lang. Proces., 2020, pp. 7346–7359, https://doi.org/10.18653/v1/2020.emnlp-main.596.

Ling Tian received the B.S., M.S., and Ph.D. degrees from the School of Computer Science and Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu in 2003, 2006, and 2010, respectively. She is currently a professor with the School of Computer Science and Engineering, UESTC. And she also works with Kashi Institute of Electronics and Information Industry, Kashi; Shenzhen Institute of Information Technology, Shenzhen. She has edited 2 books and holds over 30 China patents. She has contributed over 10 technology proposals to the standardizations, such as China Audio and Video Standard (AVS) and China Cloud Computing Standard. Her research interests include image/video coding, knowledge graphs, and knowledge reasoning.

Xue Zhou was born in Sichuan in 1995. She received the B.S. degree in software engineering from Shandong University, Jinan in 2018. She is currently pursuing the Ph.D. degree with the School of Information and Software Engineering, UESTC. Her research interests include knowledge graph, knowledge hypergraph, and presentation learning

Yan-Ping Wu was born in Sichuan in 1998. She received the B.S. degree in software engineering from UESTC in 2020. She is currently pursuing the M.S. degree with the School of Computer Science and Engineering, UESTC. Her interests include knowledge graph and few-shot learning

Wang-Tao Zhou was born in Sichuan in 1997. He received his B.E. degree in computer science and technology from UESTC in 2020. He is currently pursuing his Ph.D. degree with the School of Computer Science and Engineering, UESTC. His research interest includes event forecasting, social media analysis, and graph neural networks

Jin-Hao Zhang was born in Chongqing in 1998. He received the B.S. degree from UESTC in 2020. He is currently pursuing the Ph.D. degree with the School of Information and Software Engineering, UESTC. His research interests include knowledge graph and knowledge base question answering

Tian-Shu Zhang was born in Inner Mongolia in 1996. She received the B.S. degree in computer science and technology from Shandong University in 2019. She is currently pursuing the M.S. degree with the School of Computer Science and Engineering, UESTC. Her research interests include linking prediction, knowledge graph, and deep learning