

Article

# SFCA: A Scalable Formal Concepts Driven Architecture for Multi-Field Knowledge Graph Completion

Xiaochun Sun <sup>1</sup>, Chenmou Wu <sup>2</sup>  and Shuqun Yang <sup>1,\*</sup>

<sup>1</sup> School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China; sunxiaochun@sues.edu.cn

<sup>2</sup> Department of Computer Science & Engineering, Jeonbuk National University, Jeonju 54896, Republic of Korea; chenmou0410@gmail.com

\* Correspondence: shqyang@sues.edu.cn

**Abstract:** With the proliferation of Knowledge Graphs (KGs), knowledge graph completion (KGC) has attracted much attention. Previous KGC methods focus on extracting shallow structural information from KGs or in combination with external knowledge, especially in commonsense concepts (generally, commonsense concepts refer to the basic concepts in related fields that are required for various tasks and academic research, for example, in the general domain, “Country” can be considered as a commonsense concept owned by “China”), to predict missing links. However, the technology of extracting commonsense concepts from the limited database is immature, and the scarce commonsense database is also bound to specific verticals (commonsense concepts vary greatly across verticals, verticals refer to a small field subdivided vertically under a large field). Furthermore, most existing KGC models refine performance on public KGs, leading to inapplicability to actual KGs. To address these limitations, we proposed a novel Scalable Formal Concept-driven Architecture (SFCA) to automatically encode factual triples into formal concepts as a superior structural feature, to support rich information to KGE. Specifically, we generate dense formal concepts first, then yield a handful of entity-related formal concepts by sampling and delimiting the appropriate candidate entity range via the filtered formal concepts to improve the inference of KGC. Compared with commonsense concepts, KGC benefits from more valuable information from the formal concepts, and our self-supervision extraction method can be applied to any KGs. Comprehensive experiments on five public datasets demonstrate the effectiveness and scalability of SFCA. Besides, the proposed architecture also achieves the SOTA performance on the industry dataset. This method provides a new idea in the promotion and application of knowledge graphs in AI downstream tasks in general and industrial fields.

**Keywords:** formal concept; knowledge graph embedding; machine learning; knowledge graph completion



**Citation:** Sun, X.; Wu, C.; Yang, S. SFCA: A Scalable Formal Concepts Driven Architecture for Multi-Field Knowledge Graph Completion. *Appl. Sci.* **2023**, *13*, 6851. <https://doi.org/10.3390/app13116851>

Academic Editors: Sławomir Nowaczyk, Rita P. Ribeiro and Grzegorz Nalepa

Received: 20 May 2023

Revised: 2 June 2023

Accepted: 4 June 2023

Published: 5 June 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Knowledge Graphs (KGs), as a structured representation of interlinked descriptions of concepts, entities, relations and events, provide effective support for question answering [1], recommendation systems [2,3], information retrieval [4,5], and natural language processing [6]. By analyzing the public KGs (Freebase [7], YAGO [8], and DBPedia [9]), the incompleteness of KGs is the inevitable problem limited by existing KGs construction technology, requiring KGC to infer new facts. Among all available Knowledge Graph Completion (KGC) research, the knowledge graph embedding (KGE) model shows efficiency and significant performance, which embeds KGs components (entities and relations) into a latent space to learn topological structure information.

Existing KGE-based KGC methods can be divided into two streams: (1) Structure information-driven methods [10–16]. This stream focuses on learning the embedding of entities and relations from KG structural information to score the plausibility of triples for

link prediction. (2) External information fusion method [17–22]. This stream augments the training or inference process by fusing additional information with structural information to improve the link prediction performance.

### 1.1. Limitations

Judging by the results of recent methods, combining external information performed better than the structure information-driven method. Among various external information, commonsense concepts are recognized as more appropriate and effective by many researchers in enhancing the KGE model. However, there are still several challenges to catching commonsense concepts from data or a knowledge base.

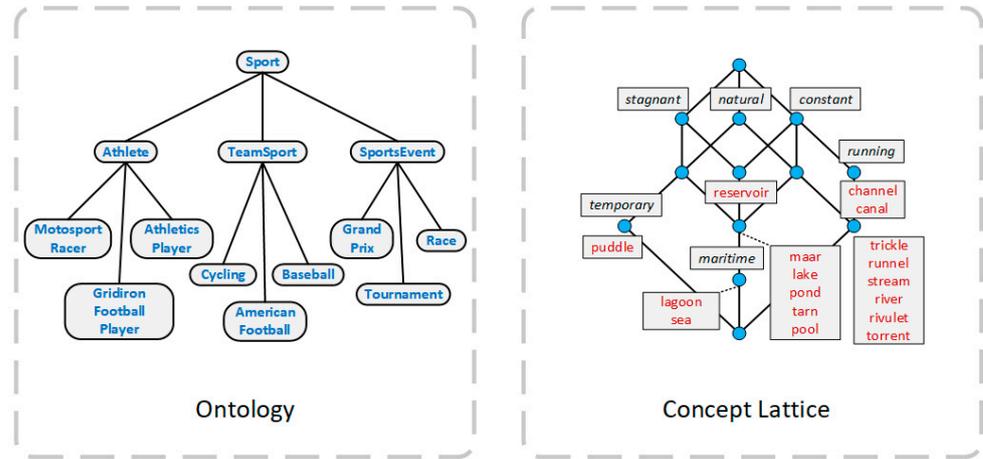
- (1) For the DBpedia-related KGs, only some famous KGs have their commonsense knowledge base, and the number of including commonsense concepts is not large. In general, the commonsense concepts from the DBpedia KGs are hard to share with other verticals KGs because commonsense concepts are mainly appropriate for the corresponding KGs (The greater the difference between common sense concepts in the more specialized fields, the more difficult it is to reuse).
- (2) For the specific KGs, such as industrial KGs, do not even have a commonsense knowledge base. Meanwhile, the commonsense concepts of specific KGs are hard to collect since they are commonly defined by corresponding researchers or experts. For the same reason, the previous models or algorithms [23,24] performed poorly on automatic commonsense extraction (The more specialized domains rely more on human-defined commonsense concepts).

### 1.2. Motivation

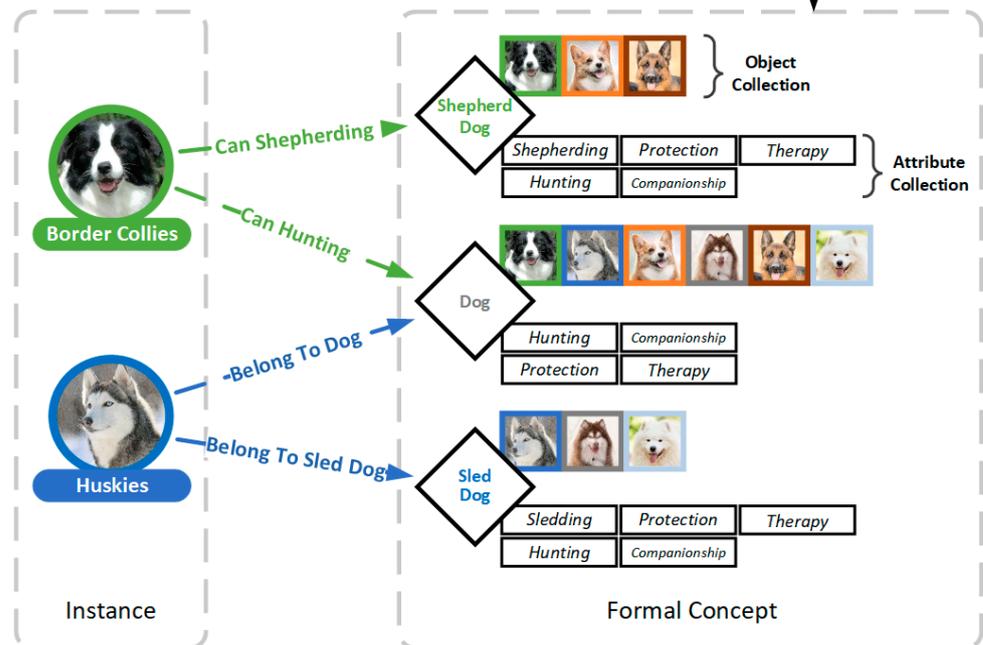
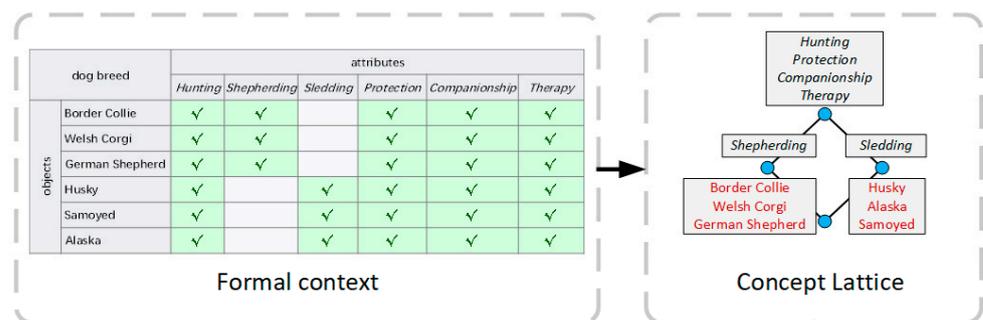
After analyzing the dilemmas involved in fetching external information, we meet three puzzles. What kind of concept can be defined as commonsense concepts? Which part of the commonsense concept provides valuable information in the KGC model? Is it possible to refine a concept from data to instead commonsense concept? The essence of the commonsense concept is summarized in many instances by human cognition. Thus, we try to explore new metaphysical ‘concepts’ from the latent space of KGs. Inspired by data mining theories, we argue that formal concepts with lattice structures are similar to ontologies with tree structures and can guide instance knowledge in KGC tasks. Thus, we use formal concepts to represent a concept, including entities and relations subset of KG. Figure 1a shows a schematic diagram of the ontology and concept lattice structure.

In light of the definition of formal concept analysis in data mining, a formal concept is an idea or category defined by a concrete or specific set of rules, guidelines, or properties. Extending to KGs, formal concepts can be treated as an ensemble of two sets. One is its extension to denote entity sets. The other is its intention to represent latent relations of entities. For the first set, our identification of instance membership in formal concepts relies on the formal concept’s instance set. By comparing the source instances and instances in the formal concept instances collection to judge whether this instance belongs to the formal concept. As shown in Figure 1b, In the lower part of the figure, diamonds represent formal concepts, and the squares and the rectangles to the right of the diamond, respectively, represent the objects and properties that make up the formal concept. The source of formal concepts is shown in the upper part of the figure. Specifically, huskies can be recognized as an instance in the formal concept of the “Dog” and an instance in the formal concept of the “Sled Dog” since it is contained in the instances collection of “Dog” and “Sled Dog” in formal concepts (The three names of “Dog”, “Herding Dog” and “Sled Dog” are matched by ourselves according to the instance set an attribute set of formal concept, formal concepts themselves have no names, but the individual elements of instance sets and property sets have their names). For the second set, formal concepts involve potential information of instances concepts, including possible properties (According to the different attributes of the instance, we can find the different formal concepts corresponding to the instance that are more focused on a certain attribute in meaning). For example, Border

Collies, in this instance, hide the attribute of “Hunting” and the attribute of “Herding”, which are respectively contained in the formal concept of “Dog” and the formal concept of “Herding Dog”.



(a)



(b)

Figure 1. (a) A schematic diagram of ontology structure and concept lattice structure. (b) Examples of Correspondence between instances and formal concepts and sources of formal concepts.

This sparked our interest in exploring the role of formal concepts in KGC research. Formal concepts can be found from the binary relations between known instances and attributes and exactly corresponds to the entity and relation in KG. For example, in a triplet SPO, P can be regarded as an attribute of S, and S can be regarded as an instance with the P property, from which we can naturally mine the formal concept in KG. In this work, we apply formal concepts, such as metaphysical structural information, to enhance the KGE model for the KGC task. Compared with commonsense concepts, formal concepts have the following advantages:

- (1) Formal concepts can be efficient and automatically generated, while commonsense concepts require expensive manual annotation. The formal concept is derived from the KG itself and belongs to the information of the KGs. On the contrary, the commonsense concept must be manually annotated for information outside the KG.
- (2) Formal concepts are not subject to KGs, while commonsense concepts are limited to the corresponding KGs. To be detailed, the formal concept can be applied to any KG, including a commonsense KG, while the commonsense concept is only applicable to the KG that has corresponding commonsense in theory and has artificially annotated commonsense concepts.

### 1.3. Architecture for KGC

Based on this, we proposed a Scalable Formal Concepts driven Architecture (SFCA) to extract formal concepts from the KGs to improve the performance of KGE. The SFCA consists of three modules:

- (1) Formal concept extraction module (FCE) extracts the formal concepts in triples and links the formal concepts with the entities.
- (2) Formal concept sampling module (FCS) sifting streamlined formal concepts from redundant formal concepts.
- (3) Formal concept-driven link prediction module (FCLP) leverages sampled formal concepts to improve KGC performance.

This framework applies to all knowledge graphs and fully benefits from the self-supervision of the data itself. The contributions of the proposed method are summarized as follows:

- (1) We propose a scalable KGC architecture based on formal concept analysis to generate formal concepts from KG. In comparison, existing methods either perform poorly or require commonsense concepts of manual extraction. To our knowledge, we are the first to apply formal concepts to KGC.
- (2) We design a coarser-to-fine formal concept extraction strategy to choose streamlined formal concepts of entities in the KG that can be used to improve the computational efficiency of the model.
- (3) Comprehensive experiments on four public datasets demonstrate the effectiveness and robustness of the proposed architecture with each module. Furthermore, the results show that SFCA helps generate KG's formal concepts, which greatly supplements the existing methods.
- (4) Extensive experiments on real data of industrial proving the practicality of the proposed model. The results show that SFCA is effective and robust in various fields of KGC tasks.

## 2. Related Works

### 2.1. KGE Models

According to KGE models' input data, the KGE models currently in use can be broadly split into two main streams:

- (1) The KG structural information-based methods include translation-based and semantic matching models. Translation-based models [10,12,15] are models that utilize entity and relation embeddings to compute translation scores, where relations represent

the translation operations between entities. Translating embeddings for modeling multi-relational data (TransE) [10] is the pioneer of the translation-based model, which embeds entities and relations into the space with the same dimension, and regards relations as translation operations between entity vectors. The advantage of TransE is that it is simple and efficient, while it cannot model various relation patterns. Knowledge graph embedding by relational rotation in complex space (RotatE) [15] embeds entities and relations in a complex vector space, treats relations as rotation operations between entity vectors, effectively models and infers various relational patterns, and refreshes the best results of the KGC task. Learning hierarchy-aware knowledge graph embeddings for link prediction (HAKE) [12] effectively embeds the semantic hierarchy by mapping the entity to the polar coordinate system, achieving the STOA result of the KGC task. Semantic matching models [11,13,14,16] compute semantic matching scores for entity and relation embeddings in the latent space. A three-way model for collective learning on multi-relational data (RESCAL) [16] treats entities as vectors and relations as matrices and calculates scores with bilinear functions. Embedding entities and relations for learning and inference in knowledge bases (DistMult) [13] simplifies RESCAL by restricting the relation matrix to be a diagonal matrix. Complex embeddings for simple link prediction (ComplEx) [14] extend DistMult to embed entities and relations into complex space. Quaternion knowledge graph embeddings (QuatE) [11] embed a hypercomplex value with three imaginary components to represent entities and model the relation as a rotation on a 4-dimensional space (hypercomplex space), thus unifying ComplEx [14] and RotatE [15].

- (2) External information-based methods focus on adding extra information to enrich the KGE models. Most models that add external information [18,20,21] utilize the logic rules mined from the knowledge graph to improve the link prediction results. Fast rule mining in ontological knowledge bases with AMIE++ (AMIE+) [20], End-to-end differentiable rule mining on knowledge graphs (DRUM) [21], and Knowledge graph embedding with iterative guidance from soft rules (RUGE) [18] automatically mine the logic rules in KG and apply the logic rules to KGC tasks. A considerable part of the models that add internal information [17,19,22] use the known concept information corresponding to the entity to improve the link prediction outcomes. Representation learning of knowledge graphs with hierarchical types (TKRL) [17] uses the type information of the entity to design a scoring function with a hierarchical projection matrix for the entity, which improves the performance of the KGC task, Type-based multiple embedding representations for knowledge graph completion (TransT) [22] adopts entity types to construct relation types and takes the similarity between relative entities and relations as prior knowledge, and utilizes prior knowledge to improve the KGC task results. A scalable commonsense-aware framework for multi-view knowledge graph completion (CAKE) [19] leverages commonsense concepts of entities to improve the quality of negative sampling and the accuracy of link prediction candidate entities. Ontology-guided Entity Alignment via Joint Knowledge Graph Embedding (OntoEA) [25] embeds ontology and knowledge graphs together to get better entity embedding.

## 2.2. KG Embedding with Ontology

At present, many scholars are studying KG embedding with ontology. Semantically smooth embedding for knowledge graphs (SSE) [26] models the intrinsic geometry of KG based on the assumption that entities belonging to the same semantic category are close to each other in the embedding space (semantically smooth). SSE [26] uses two manifold learning algorithms, Laplacian Eigenmap and Local Linear Embedding, as regularization terms to model the smoothness hypothesis. Differentiating concepts and instances for knowledge graph embedding (TransC) [27] models concept embeddings as a sphere and assumes that the embedding vectors corresponding to instances of categories belonging to concepts should lie in this sphere. TKRL [17] focuses on the Type hierarchy of KG and

believes that entities have different representations under different categories. And use the layered type as a mapping matrix and use type encoders to design. A knowledge-driven representation learning method with ontology information constraints (TransO) [28] model considers the limitations of the three types of ontology information, type information, relation constraint information, and hierarchical structure information, and maps entities and relationships to the ontology information-limited space. Based on the TransE model, loss functions are defined in the basic space and ontology information-limited space and combine the two for representation learning. Knowledge graph embedding with hierarchical relation structure (HRS) [29] learns representations of Relation clusters, Relations, and Sub-relations separately. It sums the three as the embedding vector of the relationship, thus modeling the hierarchical structure of the relationship. A representation learning method for knowledge graphs with relation hierarchical structure (TransRHS) [30] builds on HRS [29]. It encodes each relation as a vector and relation-specific spheres in the same space. TransRHS [30] uses the relative positions between vectors and spheres to model sub-relationships, which embodies the inherent generalization relationship between relations. The Universal representation learning of knowledge bases by jointly embedding instances and ontological concepts (JOIE) [31] model proposes a method that considers that KG and ontology use different embedding spaces and enables cross-space interaction between the two embeddings. A Modified Joint Knowledge Graph Embedding Model for Concepts and Instances (JECI++) [32] simplifies hierarchical concepts and links instances to them, making identifying cases based on neighboring instances and simplified concepts easier. It uses circular convolution to locate instances in the embedding space and employs CBOW and Skip-Gram strategies to embed simplified concepts and instances jointly.

### 2.3. Formal Concept Analysis

Formal Concept Analysis (FCA) [33] is a powerful and widely used method in information science that enables the creation of a concept hierarchy or formal ontology from a given set of objects and their properties. This approach is based on the mathematical theory of lattices and ordered sets, which allows for the identification of shared properties and relationships between objects in a structured and systematic way.

The resulting hierarchy of concepts represents a logical and intuitive organization of the objects and their properties, with each concept capturing a group of objects that share a set of common attributes. Moreover, the sub-concepts in the hierarchy represent a more specific grouping of the objects and a superset of the attributes from the abovementioned concepts.

Introduced by Rudolf Wille in 1981, Formal Concept Analysis (FCA) [33] has become a fundamental tool in various fields, including data mining, text mining, machine learning, knowledge management, semantic web, software development, chemistry, and biology. Its practical applications are diverse and numerous, ranging from discovering hidden patterns and relationships in data sets to developing more effective search algorithms and enhancing the quality of knowledge representation in various domains.

## 3. Methodology

We will introduce the SFCA framework in this section. As shown in Figure 2, the whole architecture includes three modules. The FCE module realizes the conversion of KG to formal context and formal context to formal concepts. The FCS module realizes the conversion of the whole formal concepts to the partial formal concepts. Finally, the FCLP module implements partial formal concepts supervision for link prediction.

As shown in Figure 2, In the Vertical field KG, points of different colors correspond to different types of real-world things (type information is not included in this KG). In the formal context of the FCE module, the various entities and relations are denoted by name abbreviations and the color of the type to which they belong. The FCE module's concept lattice's formal concept is depicted as a diamond, and the colored rectangle on the right

side of the diamond identifies the different kinds of entities and relations that make up the formal concept’s entity set and relation set, respectively.

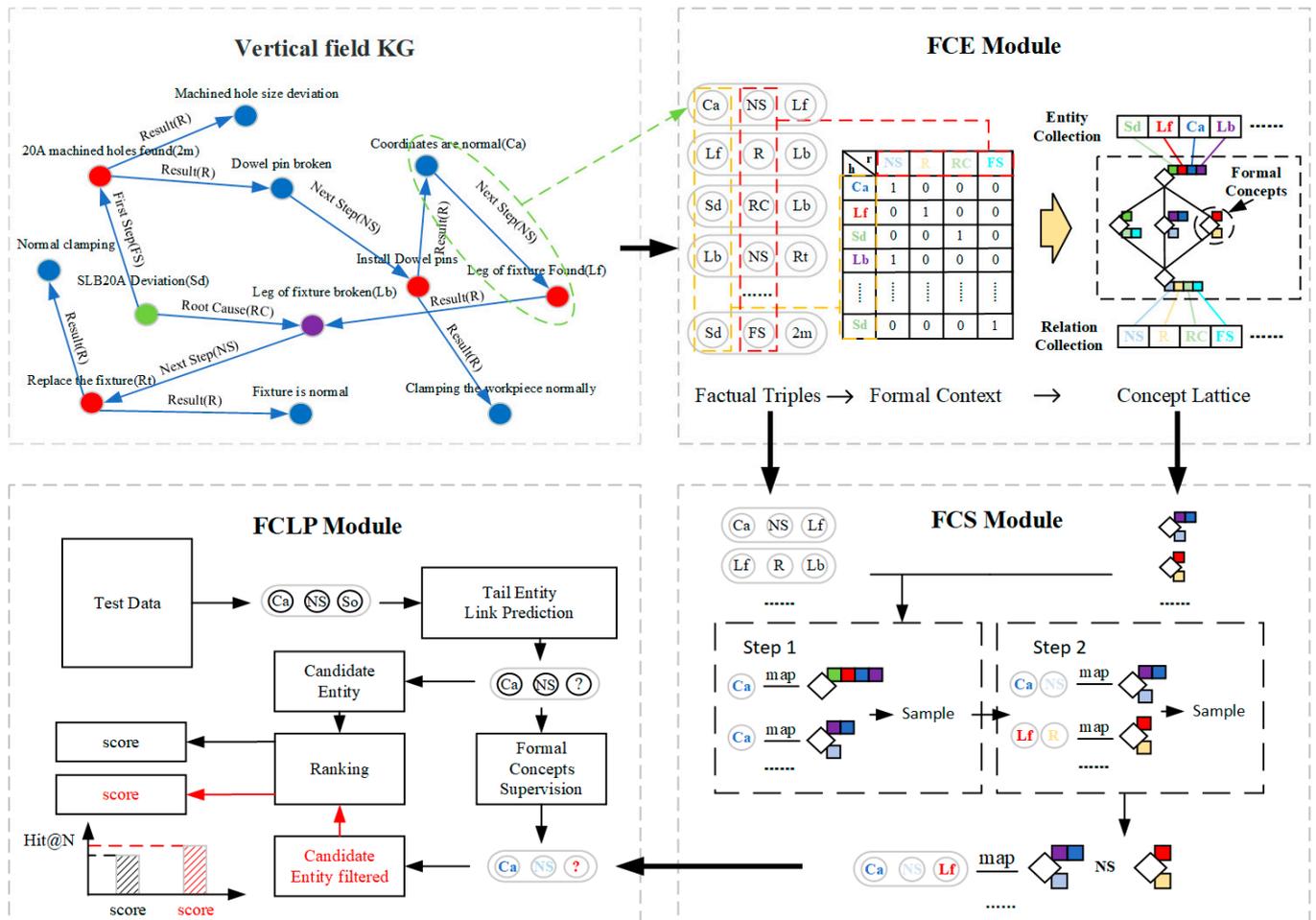


Figure 2. An overview of the SFCA framework.

As shown in Figure 2, In the formal concept sampling module in the lower right corner, the initial input is the natural mapping relationship between the formal concept and the entity. First, the natural mapping between the formal concept and the relationship is considered to find the relationship between each triple “head entity-relationship” and the formal concept. And then consider the natural mapping of formal concepts and entities again to find the correspondence between each triplet and the “formal concept triplet”. Finally, in the link prediction module driven by the formal concept in the lower left corner, whether the new triplet combined with the triplet of the missing head entity or the missing tail entity and the candidate entity can correspond to the known “formal concept triplet” is initially carried out a filter, and then enter the ranking stage to output scores.

### 3.1. Notations and Problem Formalization

#### 3.1.1. Preliminary Knowledge of FCA

For better understanding, we first provide a brief introduction to Formal Concept Analysis [19] (FCA). FCA is a method for knowledge representation, information management and data analysis. Generally, we regarded FCA as a conceptual clustering method used to determine implicit associations between objects and attributes. Formal context, formal concept, and concept lattice are three central notions in FCA. The following are the key definitions:

**Definition 1.** Let  $G$  be the set of objects,  $M$  be the set of attributes, and  $I$  be the binary relationship between object set  $G$  and attribute set  $M$ . Then, the triple  $(G, M, I)$  is a formal context.

**Definition 2.** Given a formal context  $K = (G, M, I)$ , for  $A \subseteq G, B \subseteq M$ , the following operations are defined:

$$A' = \{m \in M | gIm, \forall g \in A\} \tag{1}$$

$$B' = \{g \in G | gIm, \forall m \in B\} \tag{2}$$

$A'$  is the set of attributes shared by all objects in  $A$ ;  $B'$  is the set of objects corresponding to all attributes in  $B$ .

If  $A' = B, B' = A$ , then the binary  $(A, B)$  is a concept in the formal context, where  $A$  is the intent of the concept  $(A, B)$  and  $B$  is the extent of the concept  $(A, B)$ .

**Example 1.** A formal context  $K = (G, M, I)$  is shown in Table 1, where the object set  $G = \{1, 2, 3, 4, 5\}$  and the attribute set  $M = \{a, b, c, d, e\}$ .

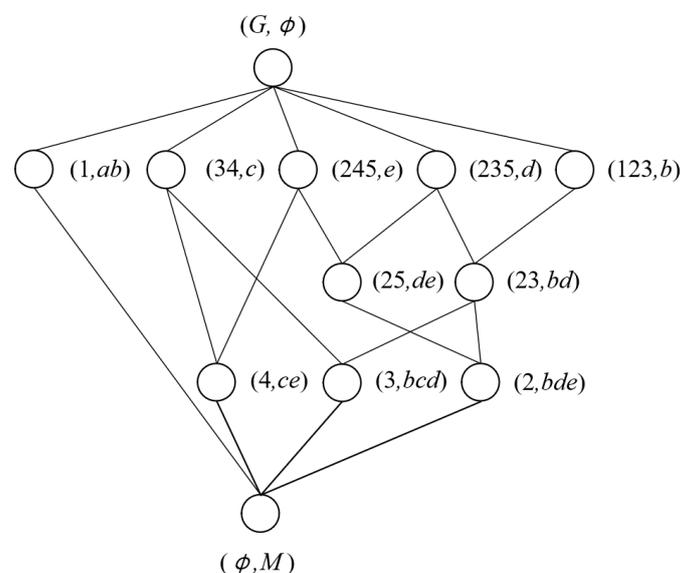
**Table 1.** A formal context.

	a	b	c	d	e
1	1	1	0	0	0
2	0	1	0	1	1
3	0	1	1	1	0
4	0	0	1	0	1
5	0	0	0	1	1

According to Definition 2, We can get all the concepts in the formal context  $K: (G, \emptyset), (1, ab), (2, bde), (3, bcd), (4, ce), (23, bd), (245, e), (235, d), (135, b), (\emptyset, M)$ .

**Definition 3.** Let  $(A_1, B_1)$  and  $(A_2, B_2)$  be two concepts on the formal context  $K = (G, M, I)$ , and  $A_1 \subseteq A_2$  (equivalent to  $B_1 \supseteq B_2$ ), then we call  $(A_1, B_1)$  the subconcept of  $(A_2, B_2)$  and  $(A_2, B_2)$  the parent concept of  $(A_1, B_1)$ , denoted as  $(A_1, B_1) \leq (A_2, B_2)$ .

The concept lattice can be visualized by a Hasse diagram. For example, the corresponding concept lattice in Example 1 is shown in Figure 3.



**Figure 3.** The lattice of formal concepts for Example 1.

### 3.1.2. KGE Score Function

Inspired by the CAKE [12], we adopt a scalable architecture design that takes any KGE model as a plugin module for SFCA direct usage. The KGE model plays the role of extracting the entity and relation embeddings in architecture. Here, we give a uniform symbol  $E(h, r, t)$  to clearly describe the score function of any KGE model for assessing the plausibility of a triple  $(h, r, t)$ . Table 2 shows the definitions of various KGE models.

**Table 2.** Details of several knowledge graph embedding models.

Model	Score Function	Parameters
TransE	$-\ h + r - t\ _{1/2}$	$h, r, t \in \mathbb{R}^k$
DistMult	$h^\top \text{diag}(M_r)t$	$h, r, t \in \mathbb{R}^k$
Complex	$\text{Re}(h^\top \text{diag}(M_r)\bar{t})$	$h, r, t \in \mathbb{C}^k$
RotatE	$-\ h \circ r - t\ _2$	$h, r, t \in \mathbb{C}^k,  r^i  = 1$
HAKE	$-\ h_m \circ r_m - t_m\ _2 - \lambda \ \sin((h_p + r_p - t_p)/2)\ _1$	$h_m, t_m \in \mathbb{R}^k,$ $r_m \in \mathbb{R}_+^k,$ $h_p, r_p, t_p \in [0, 2\pi)^k, \lambda \in \mathbb{R}$

### 3.1.3. KGC

KGC is commonly divided into three subtasks: triple classification, link prediction and relation prediction. In this work, we only concern about link prediction, not involving other KGC tasks. The link prediction task refers to finding the missing entity when the head or tail entity in the triple is missing. Specifically, we treat link prediction as an entity prediction task that searches for the reasonable entity when the triple's a head entity or tail entity is missing. Every entity in KGs will be considered a candidate when encountering a missing entity or in a triple query. We choose the top  $n$  ( $n = 1, 3, 10$ ) hits of correct entities as predicted results by ranking the scores of the candidate entities.

### 3.2. Formal Concept Extraction Module

According to the formal concepts definition (see in Section 3.1.1), SFCA automatically generate formal concepts from arbitrary KGs without external annotated knowledge. To get high-quality formal concepts, we developed a FCE module to generate massive corresponding formal concepts from KGs in mining valuable information. All generated formal concepts contain corresponding entity set and relation set.

Each fact triple set will first be encoded into a two-dimensional tabular as the formal context. Then, we produce formal concepts by the entity and relation binary relations in these formal contexts. The formal representation of KG:

$$KG = \{(e_i, r_j, e_k) | e_i \in E, e_k \in E, r_j \in R\} \tag{3}$$

where  $E$  is the set of entities,  $R$  is the set of relations from  $E$  to  $E$ .

In this paper, entity  $e_i$  is regarded as the object and relation  $r_j$  as the attribute of  $e_i$ . The formal context can be obtained by a KG and then the concept lattice  $K$  can be induced by the formal context.

For every  $e_i \in E$ , let  $C_{e_i}^{temp} = \{(A_l, B_l) | e_i \in A_l\}$ . Where  $(A_l, B_l)$  is the concept in the concept lattice  $K$ .

The mapping from the formal concepts to the entity is formulated as:

$$f : C_{e_i}^{temp} \rightarrow \{e_i\} \tag{4}$$

for every  $(A_l, B_l) \in C_{e_i}^{temp}, f((A_l, B_l)) = e_i$ .

### 3.3. Formal Concept Sampling Module

An obvious problem with dense formal concepts produced by the FCE module: Not all formal concepts are requisite. Similar semantic information and negative gain information consist of dense formal concepts (Not all formal concepts can have corresponding real concepts in the real world, so formal concepts need to be screened before using them). Thus, we proposed an FCS module to reduce complexity and improve the quality of formal concepts.

An entity can be mapped into many formal concepts in a sizable KG. Formal concepts, including the same entity, can be considered a hierarchy. Among all formal concepts, we argue that those located in the top-most node formal concepts involve the most valuable information, and relation in factual triples is sufficient to filter out them (Here, the topmost formal concept refers to the formal concept that has a non-zero and minimum number of elements in the relationship set, and the number of elements in the entity set owned by this formal concept is closest to all entities). As shown in Figure 4, our FCS module is designed by a rough-to-fine sampling strategy in the usage of instance relations and partial order relations. Specifically, the formal concepts mapped by Entity 2 are marked by the blue dots in the concept lattice on the left top. The formal concepts mapped by Entity 2 and relation e after the first sampling are marked by the blue dots in the left middle concept lattice. The formal concept mapped by entity 2 and relation e after the second sampling is marked by the blue dots in the concept lattice on the lower left. The blue dots in the green parallelogram on top of the figure are the sampled formal concepts.

(1) Sampling with instance relations: In the first stage, we only use the instance relations for sampling by inferring that formal concepts mapped by the same entity under different relations should differ. Given an instance triplet, if a formal concept in the set is mapped to the head entity of the instance triplet and the relation of the instance triplet belongs to the relation set of the formal concept, the formal concept will be mapped to the entity-relation pair composed of the head entity and the relation of the instance triple:

For every  $(e_i, r_j, e_k) \in KG$ , let  $C_{e_i r_j}^{temp} = \{(A_l, B_l) \mid (A_l, B_l) \in C_{e_i}^{temp}, r_j \in B_l\}$ . Where  $(A_l, B_l)$  is the concept in the concept lattice  $K$ .

We can get the following mapping:

$$g : C_{e_i r_j}^{temp} \rightarrow \{(e_i, r_j)\},$$

$$\text{for every } (A_l, B_l) \in C_{e_i r_j}^{temp}, g((A_l, B_l)) = (e_i, r_j). \tag{5}$$

(2) Sampling with partial order relations: After the first sampling stage, the selected formal concept set mapped to the head entity-relation pair still contains some formal concepts, and the formal concepts in this set can form a partial order relation between formal concepts. Thus, to simplify the mapping, we select the formal concept with the same link position for all formal concept links. Here, we select the most valuable formal concept in the formal concept set. Substitute formal concept set as a mapping formal concept of head entity-relation pair.

Given an instance triplet, if a formal concept in the set mapped to the head entity-relation pair of the instance triplet and the relation set of the formal concept is included in the relation set of all formal concepts in the formal concept set, then the formal concept is selected to be mapped to the entity-relation pair composed of the instance triple head entity and relation, according to Definition 3, there exists a maximum concept in the  $C_{e_i r_j}^{temp}$ .

For every  $(e_i, r_j, e_k) \in KG$ , let  $C_{e_i r_j}^{final} = \{(A_l, B_l) \mid (A_l, B_l) \in C_{e_i r_j}^{temp}, \forall (A_m, B_m) \in C_{e_i r_j}^{temp}, (A_m, B_m) \leq (A_l, B_l)\}$ .

We can get the following mapping:

$$h : C_{e_i r_j}^{final} \rightarrow \{(e_i, r_j)\},$$

$$\text{for every } (A_l, B_l) \in C_{e_i r_j}^{final}, h((A_l, B_l)) = (e_i, r_j). \tag{6}$$

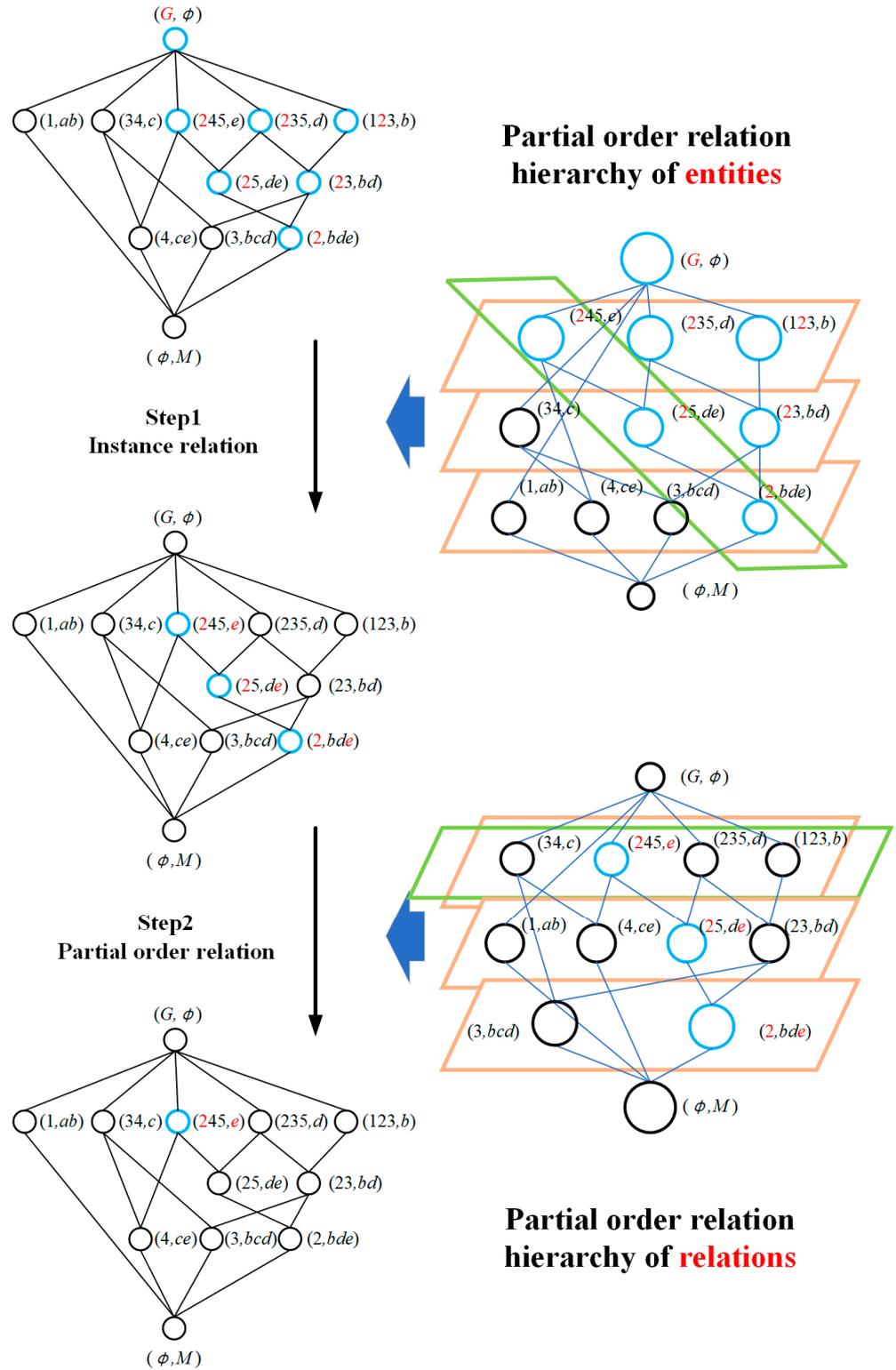


Figure 4. Formal Concept Sampling Module.

After sampling twice, the mapping from formal concepts to instance entities and the formal concept triples can be obtained—the mapping of entity-relation pairs from formal concepts to instance triples after sampling twice.

For every  $e_i \in E$ , let  $C_{e_i}^{final} = \{(A_l, B_l) \mid (A_l, B_l) \in C_{e_i r_j}^{final}, (e_i, r_j, e_k) \in KG\}$ .

We can get the following mapping:

$$l : C_{e_i}^{final} \rightarrow \{e_i\},$$

$$\text{for every } (A_l, B_l) \in C_{e_i}^{final}, l((A_l, B_l)) = e_i. \tag{7}$$

A collection of formal concept triples is denoted as  $FC$ , where each triple consists of a head entity’s formal concept set  $C_{hr}^{final}$  and a tail entity’s formal concept set  $C_t^{final}$  associated with their instance-level relation  $r$ , defined as:

$$FC = \{(C_{hr}^{final}, r, C_t^{final}) \mid (h, r, t) \in KG\} \tag{8}$$

### 3.4. Formal Concept-Driven Link Prediction Module

To find better candidate entities and improve prediction outcomes, we propose a novel two-stage formal concept supervised link prediction mechanism. In the first stage, candidate entities are selected from the perspective of formal concepts; specifically, accepting a query  $(h, r, ?)$  to filter plausible formal concepts of tail entities using the set of formal concept triples  $FC$ , the set of candidate formal concepts of tail entity  $t$  is  $C_t^{final}$ , and then determine the entity belonging to the formal concept set as the candidate entity.

In the second stage, for each candidate entity  $e_i$  that has been screened, the score is calculated by the scoring function. The candidate triplet that finally calculates the score is:

$$score(e_i) = E(h, r, e_i) \tag{9}$$

Among them,  $E(h, r, e_i)$  is the scoring function used to train the KGE model, and then the prediction results will arrange the scores of the candidate entities in ascending order and output the top  $n$  hits of the correct entities.

## 4. Experiments

For a comprehensive comparison, we evaluate our SFCA on five real-world datasets and one industry dataset. In this section, the detail of the setting of the experiment will be introduced. First, the performance of SFCA on four public datasets will be shown. Second, the comparison with common sense concepts will be discussed third, and the effectiveness of real data will be proved finally.

### 4.1. Experiment Settings

#### 4.1.1. Datasets

Our evaluation is based on five public datasets (FB15K237 [34], YAGO3-10 [35], WN18RR [36], NELL-995 [37], DBpedia-242 [19]) and an industrial KG datasets collecting from the workshop of actual factory. Table 3 shows the statistics of public and industry datasets.

**Table 3.** Statistics of datasets.

Dataset	#Rel	#Ent	#Train	#Valid	#Test
FB15K237	237	14,541	272,115	17,535	20,466
YAGO3-10	37	123,182	1,079,040	5000	5000
WN18RR	11	40,943	86,835	3034	3134
NELL-995	200	75,492	123,370	15,000	15,843
Dbpedia-242	298	99,744	529,654	35,850	30,000
Industrial KG	15	57,373	78,787	3000	3000

FB15K237 is a link prediction dataset created from FB15k. FB15k-237 was created by Toutanova and Chen to ensure that the test and evaluation datasets do not have an inverse relation with test leakage. YAGO3-10 is a benchmark dataset for knowledge base completion. It is a subset of YAGO3 (YAGO3 is an extension of YAGO) and contains entities associated with at least ten different relations. WN18RR is a link prediction dataset created by WN18, a subset of WordNet. However, many text triples are obtained by inversely finding triples from the training set. The WN18RR dataset was therefore created to ensure that the evaluation dataset does not have an inverse relation with test leaks. NELL-995 is a subset of NELL suitable for multi-hop inference proposed from the 995th iteration of the NELL system. Useless triples are first removed using relations that occur more than 2 M times in the NELL dataset. After this step, the triples with Top-200 relations are selected, and the dataset is obtained after adding the inverse triples. DBpedia-242 is extracted from DBpedia [6], which contains 242 concepts. It is worth mentioning that the entities in FB15K237, YAGO3-10 and NELL-995 have a corresponding ontology, while in WN18RR, The entity does not have a corresponding ontology.

#### 4.1.2. Baselines

We compare our SFCA model with five baseline models, including TransE [10], Dist-Mult [13], ComplEx [14], RotatE [15], and HAKE [12], and we also integrate these baseline models into our framework. All baseline models are KG-structured models. As a result, our framework does not require the input of external data, and we show through experiments that our model can be applied to most models without the input of external expert data.

#### 4.1.3. Implementation Details

We use the Adam optimizer for training, and all models adopt the self-adversarial negative sampling method. In terms of the parameters, under the same dataset, we use the same parameters of different baseline models, including embedding size, batch size, negative sampling size, learning rate, margin, and sampling temperature. All experiments are performed on Pytorch and NVIDIA Quadro RTX 5000 GPU.

#### 4.1.4. Evaluation Protocol

We choose three recognized evaluation metrics for comparison, including mean rank (MR), mean reciprocal rank (MRR), and the proportion of top-N rankings of correct entities (Hits@N). Notably, we filtered out all candidate triples in the datasets. The detailed computing formulas and notation definition as shown in Table 4.

**Table 4.** Detailed computing formulas of evaluation metrics for KGC.

Metrics	Computing Formula	Notation Definition
MRR	$MRR = \frac{1}{ Q } \sum_{ Q }^{i=1} \frac{1}{rank_i}$	Q : query sets;  Q  : queries numbers; $rank_i$ : the rank of the first correct answer for the $i$ th query
MR	$MR = \frac{1}{ Q } \sum_{ Q }^{i=1} rank_i$	Q : query sets;  Q  : queries numbers; $rank_i$ : the rank of the first correct answer for the $i$ th query
Hits@n	$Hits@n = \frac{1}{ Q } Count(rank_i \leq n),$ $0 < i \leq  Q $	Count() : the hit test number in the top n rankings among test examples; Q : query sets;  Q  : queries numbers; $rank_i$ : the rank of the first correct answer for the $i$ th query

#### 4.2. Evaluation of Public KGs

Table 5 shows the link prediction performance of SFCA on the four public datasets. The formal concepts-driven KGE module improved significantly, which has an average increase of 11.83% (3.78 points), 16.53% (4.12 points), and 19.13% (5.82 points) in the MRR indicators of different baselines on the three datasets of FB15K237, YAGO3-10 and NELL-995, respectively. On WN18RR, the Hit@10 indicator of different baselines improved

by an average of 2.63% (1.34 points). These results prove formal concepts, as metaphysical features of KG-structured information, are more splendid and effective for link prediction.

**Table 5.** Link prediction results on four datasets. Bold numbers are the best results for each type of model. \* denotes SFCA is used.

Model	FB15K237					YAGO3-10				
	MR	MRR	Hits@1	Hits@3	Hits@10	MR	MRR	Hits@1	Hits@3	Hits@10
TransE	175	0.331	0.234	0.369	0.526	1124	0.504	0.411	0.563	0.672
TransE *	<b>110</b>	<b>0.371</b>	<b>0.275</b>	<b>0.410</b>	<b>0.563</b>	<b>648</b>	<b>0.535</b>	<b>0.452</b>	<b>0.583</b>	<b>0.688</b>
DistMult	216	0.284	0.223	0.310	0.447	1723	0.130	0.084	0.135	0.219
DistMult *	<b>145</b>	<b>0.320</b>	<b>0.240</b>	<b>0.346</b>	<b>0.482</b>	<b>1110</b>	<b>0.174</b>	<b>0.116</b>	<b>0.185</b>	<b>0.287</b>
ComplEx	181	0.308	0.220	0.337	0.486	1118	0.198	0.131	0.214	0.325
ComplEx *	<b>119</b>	<b>0.344</b>	<b>0.258</b>	<b>0.372</b>	<b>0.518</b>	<b>660</b>	<b>0.253</b>	<b>0.181</b>	<b>0.276</b>	<b>0.393</b>
RotatE	177	0.336	0.240	0.374	0.530	1859	0.497	0.406	0.553	0.665
RotatE *	<b>108</b>	<b>0.376</b>	<b>0.280</b>	<b>0.413</b>	<b>0.567</b>	<b>795</b>	<b>0.542</b>	<b>0.469</b>	<b>0.591</b>	<b>0.694</b>
HAKE	184	0.343	0.246	0.380	0.537	1384	0.531	0.444	0.586	0.687
HAKE *	<b>112</b>	<b>0.380</b>	<b>0.285</b>	<b>0.417</b>	<b>0.572</b>	<b>613</b>	<b>0.562</b>	<b>0.485</b>	<b>0.605</b>	<b>0.703</b>

Model	WN18RR					NELL-995				
	MR	MRR	Hits@1	Hits@3	Hits@10	MR	MRR	Hits@1	Hits@3	Hits@10
TransE	5081	0.191	0.004	0.343	0.478	2238	0.239	0.079	0.348	0.508
TransE *	<b>4209</b>	<b>0.221</b>	<b>0.029</b>	<b>0.382</b>	<b>0.500</b>	<b>353</b>	<b>0.325</b>	<b>0.202</b>	<b>0.394</b>	<b>0.548</b>
DistMult	6043	0.406	0.358	0.437	0.490	2523	0.348	0.268	0.385	0.500
DistMult *	<b>4935</b>	<b>0.414</b>	<b>0.364</b>	<b>0.444</b>	<b>0.503</b>	<b>472</b>	<b>0.398</b>	<b>0.311</b>	<b>0.435</b>	<b>0.564</b>
ComplEx	6394	0.457	0.416	0.481	0.526	3664	0.369	0.289	0.405	0.518
ComplEx *	<b>5274</b>	<b>0.463</b>	<b>0.421</b>	<b>0.488</b>	<b>0.535</b>	<b>727</b>	<b>0.417</b>	<b>0.331</b>	<b>0.453</b>	<b>0.582</b>
RotatE	5648	0.466	0.435	0.480	0.527	2327	0.358	0.269	0.404	0.526
RotatE *	<b>4597</b>	<b>0.473</b>	<b>0.440</b>	<b>0.488</b>	<b>0.539</b>	<b>337</b>	<b>0.405</b>	<b>0.313</b>	<b>0.450</b>	<b>0.580</b>
HAKE	4132	0.492	0.450	0.509	0.574	2419	0.313	0.205	0.374	0.514
HAKE *	<b>3325</b>	<b>0.504</b>	<b>0.463</b>	<b>0.519</b>	<b>0.585</b>	<b>348</b>	<b>0.373</b>	<b>0.270</b>	<b>0.427</b>	<b>0.567</b>

#### 4.3. Common Sense Concepts vs. Formal Concepts

We also compare our SFCA with an external information-based method: Commonsense-Aware Knowledge Embedding (CAKE) [19] framework. The comparison results are obtained by combining uniform sampling [10] and self-adversarial sampling [15] with the KGE model TransE [10] and RotatE [15]. Table 6 presents the link prediction evaluation results on the three datasets. By comparison, our SFCA has an average increase of 147.54% (45.13 points) and 9.34% (3.33 points) higher than CAKE in the MRR indicators of different baselines on the two datasets of FB15K237 and NELL-995, respectively. On DBpedia-242, in the best effect on each indicator of different baselines, the highest is 14.47% (2.3 points) higher than CAKE, and the lowest is 3.36% (1.5 points) lower than CAKE. These results prove formal concepts are more effective in most cases than common sense concepts for the KGC task.

#### 4.4. Evaluation of Actual Industrial KGs

Table 7 shows the link prediction evaluation results on the industrial KG dataset. By comparison, we can see that our SFCA framework has improved by more than 13.11% (7.3 points) in the MRR indicators of different baselines on the fault diagnosis industrial dataset. Our SFCA achieves remarkable performance in industrial filed KGs. With this result, we believe our SFCA can play well in multi-field KGs.

**Table 6.** Link prediction results on three datasets. Bold numbers are the best results for each type of model. \* Denotes SFCA is used. +U denotes uniform sampling strategies are used. +S denotes self-adversarial negative sampling strategies used. +C denotes Commonsense-Aware Knowledge Embedding framework used.

Model	FB15K237					NELL-995					Dbpedia-242				
	MR	MRR	Hits@1	Hits@3	Hits@10	MR	MRR	Hits@1	Hits@3	Hits@10	MR	MRR	Hits@1	Hits@3	Hits@10
TransE + U	167	0.309	0.215	0.340	0.496	1261	0.249	0.153	0.288	0.428	1332	0.281	0.123	0.393	0.538
TransE + S	172	0.303	0.208	0.337	0.494	1212	0.255	0.156	0.295	0.436	1302	0.302	0.132	0.427	0.573
TransE + C	177	0.295	0.199	0.330	0.484	485	0.310	0.187	0.374	0.532	901	0.331	0.159	0.461	0.596
TransE + U *	29	<b>0.756</b>	<b>0.709</b>	0.779	<b>0.850</b>	232	0.344	0.248	0.386	0.527	635	0.316	0.170	<b>0.415</b>	<b>0.557</b>
TransE + S *	<b>28</b>	0.754	0.704	<b>0.780</b>	0.849	<b>226</b>	<b>0.351</b>	<b>0.253</b>	<b>0.392</b>	<b>0.536</b>	<b>606</b>	<b>0.336</b>	<b>0.182</b>	0.446	0.586
RotatE + U	188	0.316	0.221	0.351	0.507	1389	0.342	0.249	0.379	0.522	1935	0.337	0.197	0.434	0.560
RotatE + S	190	0.324	0.230	0.357	0.514	1357	0.351	0.260	0.385	0.533	1788	0.372	0.239	0.464	0.592
RotatE + C	192	0.318	0.226	0.348	0.503	503	0.404	0.305	0.455	0.593	1049	0.419	0.310	0.486	0.605
RotatE + U *	34	<b>0.761</b>	<b>0.712</b>	<b>0.784</b>	0.853	223	<b>0.433</b>	<b>0.343</b>	0.469	<b>0.611</b>	772	<b>0.366</b>	<b>0.236</b>	<b>0.454</b>	0.576
RotatE + S *	<b>32</b>	0.760	0.711	<b>0.784</b>	<b>0.855</b>	<b>220</b>	0.433	0.341	<b>0.470</b>	<b>0.611</b>	<b>718</b>	0.413	0.299	0.485	<b>0.607</b>

**Table 7.** Link prediction results on Industrial datasets. Bold numbers are the best results for each type of model. \* denotes SFCA is used.

Model	Industrial KG				
	MR	MRR	Hits@1	Hits@3	Hits@10
TransE	5949	0.353	0.143	0.505	0.758
TransE *	<b>1000</b>	<b>0.681</b>	<b>0.609</b>	<b>0.741</b>	<b>0.796</b>
DistMult	6811	0.545	0.492	0.572	0.671
DistMult *	<b>847</b>	<b>0.636</b>	<b>0.574</b>	<b>0.680</b>	<b>0.744</b>
Complex	6246	0.557	0.498	0.591	0.685
Complex *	<b>835</b>	<b>0.630</b>	<b>0.562</b>	<b>0.678</b>	<b>0.752</b>
RotatE	7086	0.588	0.510	0.634	0.760
RotatE *	<b>1061</b>	<b>0.672</b>	<b>0.602</b>	<b>0.726</b>	<b>0.788</b>
HAKE	3538	0.584	0.521	0.610	0.728
HAKE *	<b>448</b>	<b>0.699</b>	<b>0.629</b>	<b>0.756</b>	<b>0.819</b>

## 5. Conclusions

Motivated by the formal concept analysis theory, we propose a novel scalable formal concept-driven knowledge graph completion framework (SFCA) applying to multiple verticals. SFCA can automatically generate formal concepts from KG with a coarse-to-fine extraction strategy and a formal concept-supervised link prediction module to filter candidate entities from the perspective of formal concepts. Experiments on five public datasets demonstrate the effectiveness and scalability of SFCA. In addition, our model is also experimentally performed on real industrial datasets to demonstrate that the model has high performance in both general and industrial domains.

Despite the performance of SFCA, it is still some areas for improvement. First, the application of our method on the KGC task is based on the closed-world hypothesis, which considers any triplet not explicitly present in the graph as a negative triplet. Second, our method only applies to the KGC task and does not extend to other knowledge-related tasks. Thus, our further work is to study the application of formal concept analysis on KGC tasks under the open-world hypothesis and explore the application of formal concept analysis on various concept-related KG tasks, such as life-long learning. Third, this paper's knowledge graph embedding-related tasks only involve knowledge graph completion. Future work can consider other knowledge map-related tasks such as named entity recognition and relationship extraction, that is, using formal concept analysis to assist tasks such as named entity recognition and relationship extraction.

**Author Contributions:** Conceptualization, X.S., C.W. and S.Y.; methodology, X.S.; validation, X.S.; investigation, X.S. and C.W.; resources, C.W. and S.Y.; data curation, X.S.; writing—original draft preparation, X.S.; writing—review and editing, X.S. and C.W.; supervision, C.W. and S.Y.; project administration, S.Y.; funding acquisition, S.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key R&D Program of China, grant number: 2020AAA0109300.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** Thanks to teachers and classmates for their suggestions on this paper and the program.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Wang, X.; Luo, F.; Wu, Q.; Bao, Z. How Context or Knowledge Can Benefit Healthcare Question Answering? *IEEE Trans. Knowl. Data Eng.* **2021**, *35*, 575–588. [[CrossRef](#)]
2. Liu, Y.; Yang, S.; Xu, Y.; Miao, C.; Wu, M.; Zhang, J. Contextualized graph attention network for recommendation with item knowledge graph. *IEEE Trans. Knowl. Data Eng.* **2021**, *35*, 181–195. [[CrossRef](#)]
3. Kang, S.; Shi, L.; Zhang, Z. Knowledge Graph Double Interaction Graph Neural Network for Recommendation Algorithm. *Appl. Sci.* **2022**, *12*, 12701. [[CrossRef](#)]
4. Zhong, M.; Zheng, Y.; Xue, G.; Liu, M. Reliable keyword query interpretation on summary graphs. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 5187–5202. [[CrossRef](#)]
5. Sun, Y.; Chun, S.-J.; Lee, Y. Learned semantic index structure using knowledge graph embedding and density-based spatial clustering techniques. *Appl. Sci.* **2022**, *12*, 6713. [[CrossRef](#)]
6. Gu, R.; Wang, T.; Deng, J.; Cheng, L. Improving Chinese Named Entity Recognition by Interactive Fusion of Contextual Representation and Glyph Representation. *Appl. Sci.* **2023**, *13*, 4299. [[CrossRef](#)]
7. Bollacker, K.; Evans, C.; Paritosh, P.; Sturge, T.; Taylor, J. Freebase: A collaboratively created graph database for structuring human knowledge. In Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data, Vancouver, BC, Canada, 10–12 June 2008; pp. 1247–1250.
8. Suchanek, F.M.; Kasneci, G.; Weikum, G. Yago: A core of semantic knowledge. In Proceedings of the 16th international conference on World Wide Web, Banff, AB, Canada, 8–12 May 2007; pp. 697–706.
9. Lehmann, J.; Isele, R.; Jakob, M.; Jentzsch, A.; Kontokostas, D.; Mendes, P.N.; Hellmann, S.; Morse, M.; Van Kleef, P.; Auer, S. Dbpedia—a large-scale, multilingual knowledge base extracted from wikipedia. *Semant. Web* **2015**, *6*, 167–195. [[CrossRef](#)]
10. Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; Yakhnenko, O. Translating embeddings for modeling multi-relational data. *Adv. Neural Inf. Process. Syst.* **2013**, *26*, 2787–2795.
11. Zhang, S.; Tay, Y.; Yao, L.; Liu, Q. Quaternion knowledge graph embeddings. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 2735–2745.
12. Zhang, Z.; Cai, J.; Zhang, Y.; Wang, J. Learning hierarchy-aware knowledge graph embeddings for link prediction. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 3065–3072.
13. Yang, B.; Yih, W.-t.; He, X.; Gao, J.; Deng, L. Embedding entities and relations for learning and inference in knowledge bases. *arXiv* **2014**, arXiv:1412.6575.
14. Trouillon, T.; Welbl, J.; Riedel, S.; Gaussier, É.; Bouchard, G. Complex embeddings for simple link prediction. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; pp. 2071–2080.
15. Sun, Z.; Deng, Z.-H.; Nie, J.-Y.; Tang, J. Rotate: Knowledge graph embedding by relational rotation in complex space. *arXiv* **2019**, arXiv:1902.10197.
16. Nickel, M.; Tresp, V.; Krieger, H.-P. A three-way model for collective learning on multi-relational data. In Proceedings of the IJCAI, Bellevue, WA, USA, 28 June–2 July 2011; pp. 3104482–3104584.
17. Xie, R.; Liu, Z.; Sun, M. Representation learning of knowledge graphs with hierarchical types. In Proceedings of the IJCAI, New York, NY, USA, 9–15 July 2016; pp. 2965–2971.
18. Guo, S.; Wang, Q.; Wang, L.; Wang, B.; Guo, L. Knowledge graph embedding with iterative guidance from soft rules. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018.
19. Niu, G.; Li, B.; Zhang, Y.; Pu, S. CAKE: A scalable commonsense-aware framework for multi-view knowledge graph completion. *arXiv* **2022**, arXiv:2202.13785.
20. Galárraga, L.; Teflioudi, C.; Hose, K.; Suchanek, F.M. Fast rule mining in ontological knowledge bases with AMIE+. *VLDB J.* **2015**, *24*, 707–730. [[CrossRef](#)]

21. Sadeghian, A.; Armandpour, M.; Ding, P.; Wang, D.Z. Drum: End-to-end differentiable rule mining on knowledge graphs. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 1–13.
22. Ma, S.; Ding, J.; Jia, W.; Wang, K.; Guo, M. Transt: Type-based multiple embedding representations for knowledge graph completion. In Proceedings of the Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2017, Proceedings, Part I 10. Skopje, North Macedonia, 18–22 September 2017; pp. 717–733.
23. Liu, J.; Chen, T.; Wang, C.; Liang, J.; Chen, L.; Xiao, Y.; Chen, Y.; Jin, K. VoCSK: Verb-oriented commonsense knowledge mining with taxonomy-guided induction. *Artif. Intell.* **2022**, *310*, 103744. [[CrossRef](#)]
24. Li, C.; Liang, J.; Xiao, Y.; Jiang, H. Towards Fine-Grained Concept Generation. *IEEE Trans. Knowl. Data Eng.* **2021**, *35*, 986–997. [[CrossRef](#)]
25. Xiang, Y.; Zhang, Z.; Chen, J.; Chen, X.; Lin, Z.; Zheng, Y. OntoEA: Ontology-guided entity alignment via joint knowledge graph embedding. *arXiv* **2021**, arXiv:2105.07688.
26. Guo, S.; Wang, Q.; Wang, B.; Wang, L.; Guo, L. SSE: Semantically smooth embedding for knowledge graphs. *IEEE Trans. Knowl. Data Eng.* **2016**, *29*, 884–897. [[CrossRef](#)]
27. Lv, X.; Hou, L.; Li, J.; Liu, Z. Differentiating concepts and instances for knowledge graph embedding. *arXiv* **2018**, arXiv:1811.04588.
28. Li, Z.; Liu, X.; Wang, X.; Liu, P.; Shen, Y. Transo: A knowledge-driven representation learning method with ontology information constraints. *World Wide Web* **2023**, *26*, 297–319. [[CrossRef](#)]
29. Zhang, Z.; Zhuang, F.; Qu, M.; Lin, F.; He, Q. Knowledge graph embedding with hierarchical relation structure. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, 31 October–4 November 2018; pp. 3198–3207.
30. Zhang, F.; Wang, X.; Li, Z.; Li, J. TransRHS: A representation learning method for knowledge graphs with relation hierarchical structure. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, Yokohama, Japan, 7–15 January 2021.
31. Hao, J.; Chen, M.; Yu, W.; Sun, Y.; Wang, W. Universal representation learning of knowledge bases by jointly embedding instances and ontological concepts. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 1709–1719.
32. Wang, P.; Zhou, J. JECI++: A Modified Joint Knowledge Graph Embedding Model for Concepts and Instances. *Big Data Res.* **2021**, *24*, 100160. [[CrossRef](#)]
33. Ganter, B.; Wille, R. *Formal Concept Analysis: Mathematical Foundations*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
34. Toutanova, K.; Chen, D. Observed versus latent features for knowledge base and text inference. In Proceedings of the 3rd Workshop on Continuous Vector Space Models and Their Compositionality, Beijing, China, 31 July 2015; pp. 57–66.
35. Mahdisoltani, F.; Biega, J.; Suchanek, F. Yago3: A knowledge base from multilingual wikipedias. In Proceedings of the 7th Biennial Conference on Innovative Data Systems Research, Asilomar, CA, USA, 4–7 January 2015.
36. Shang, C.; Tang, Y.; Huang, J.; Bi, J.; He, X.; Zhou, B. End-to-end structure-aware convolutional networks for knowledge base completion. In Proceedings of the AAAI Conference on Artificial Intelligence, Hilton, HI, USA, 27 January–1 February 2019; pp. 3060–3067.
37. Xiong, W.; Hoang, T.; Wang, W.Y. Deeppath: A reinforcement learning method for knowledge graph reasoning. *arXiv* **2017**, arXiv:1707.06690.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.