Machine Intelligence Research
www.mi-research.net

21(2), April 2024, 239-256
DOI: 10.1007/s11633-023-1450-8

A Comprehensive Overview of CFN From a Commonsense Perspective

Ru Li Yunxiao Zhao Zhiqiang Wang Xuefeng Su Shaoru Guo
Yong Guan Xiaoqi Han Hongyan Zhao

Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education,
School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China

Abstract: Chinese FrameNet (CFN) is a scenario commonsense knowledge base (CKB) that plays an important role in research on Chinese language understanding. It is based on the theory of frame semantics and English FrameNet (FN). The CFN knowledge base contains a wealth of scenario commonsense knowledge, including frames, frame elements, and frame relations, as well as annotated instances with rich scenario-related labels on Chinese sentences and discourses. In this paper, we conduct a comprehensive overview of CFN from a commonsense perspective, covering topics such as scenario commonsense representation, CFN resources, and its applications. We also summarize recent breakthroughs and identify future research directions. First, we introduce the concept of scenario commonsense, including its definitions, examples, and representation methods, with a focus on the relationship between scenario commonsense and the frame concept in CFN. In addition, we provide a comprehensive overview of CFN resources and their applications, highlighting the newly proposed frame-based discourse representation and a human-machine collaboration framework for expanding the CFN corpus. Furthermore, we explore emerging topics such as expanding the CFN resource, improving the interpretability of machine reading comprehension, and using scenario CKBs for text generation.

Keywords: Chinese FrameNet (CFN), commonsense, scenario commonsense, frame, knowledge.

Citation: R. Li, Y. Zhao, Z. Wang, X. Su, S. Guo, Y. Guan, X. Han, H. Zhao. A comprehensive overview of CFN from a commonsense perspective. *Machine Intelligence Research*, vol.21, no.2, pp.239–256, 2024. http://doi.org/10.1007/s11633-023-1450-8

1 Introduction

Commonsense knowledge is an essential resource for artificial intelligence and has received significant attention in recent years. There is a growing belief that a robust intelligent system must possess rich commonsense knowledge and the ability to reason. For example, a computer vision model may incorrectly classify a bird as a bear without the commonsense that a bear on a branch is an uncommon occurrence[1], and it is difficult for a machine to accurately describe the visual relationship, such as we ski with legs, if it lacks commonsense knowledge[2]. Similarly, machine translation can be easily confused by subtle distinctions, such as quick food VS. fast food, without linguistic commonsense[3]. The well-known pretrained language models in natural language processing are also prone to confusion with modifications such as add, del, sub and swap without understanding commonsense knowledge[4]. In short, if machines can effectively integrate commonsense, they will be more robust in open and dynamic environments and perform better in many applications, such as machine translation[5], dialogue systems[6–8], and automatic driving[9].

Commonsense is prevalent in the real world and human society, but there is no strict and unified definition in the literature. Generally, commonsense is considered to be the basic level of practical knowledge and judgment that we all need to live in a reasonable and safe way[1]. While it is broad in nature, the acquisition and use of commonsense remains a challenging and unsolved problem in artificial intelligence due to its subtle and implicit nature. For instance, some commonsense frequently appears in one modality but rarely in another. Vedantam et al.[10] found that "squirrels wanting nuts" is frequently mentioned in texts, while "squirrels looking at nuts" is rarely mentioned in pictures. Additionally, Sap et al.[11] demonstrated the implicit nature of commonsense through a natural language example: We all know that the food inside the fridge might go bad if we keep the door open, but we do not explicitly say it.

Researchers have made numerous efforts to build com-

Review
Special Issue on Commonsense Knowledge and Reasoning: Representation, Acquisition and Applications
Manuscript received on December 19, 2022; accepted on April 20, 2023; published online on January 12, 2024
Recommended by Associate Editor Jun Zhao
Colored figures are available in the online version at https://link.springer.com/journal/11633

1 https://dictionary.cambridge.org/dictionary/english-chinese-simplified/common-sense

monsense knowledge bases (CKBs), such as Cyc[12], ConceptNet[13], ATOMIC[14], and HowNet[15], etc. These knowledge bases describe various types of commonsense knowledge, such as attribute commonsense[16], temporal and spatial commonsense[17, 18], social commonsense[19, 20], physical commonsense[21, 22], and causal commonsense[23], and they are represented in different forms, such as "if-then" in ATOMIC and triples $\langle head, relation, tail\rangle$ in ConceptNet. Specifically, "frame" is an important representation form for a specific type of commonsense knowledge called scenario commonsense, which describes the motivating situation occurring against a background of knowledge and experience[24]. For a specific scenario, it must include certain elements closely associated with the scenario.

For example, Fig. 1 shows a scenario (commercial event) in the mall, including two sub-scenarios inside such as "commerce money transfer" and "commerce goods transfer". The elements in these two scenarios include the buyer, who is interested in goods, the seller, who is interested in money, the goods that the buyer can acquire from the seller, and the money that the seller acquires from the buyer. It is worth noting that the "commerce money transfer" scenario places more emphasis on money, while the other scenario places more emphasis on goods.

Chinese FrameNet (CFN) is a scenario commonsense knowledge base that was developed using the frame semantics theory and modeled after the English FrameNet (FN). It provides a structured representation of scenario-based commonsense knowledge in Chinese, using the frame to represent the scenario commonsense knowledge in a way that is easy for humans to understand. More importantly, it is a valuable resource for natural language understanding tasks and applications, as it contains a wealth of scenario commonsense knowledge represented by frames, frame elements, frame relations, and lexical units (LUs), as well as annotated instances with detailed scenario-related labels on Chinese sentences and discourses.

According to Fillmore′s frame semantics theory[24–26], humans understand the meaning of language by activating scenarios, known as frames, in their minds. These frames are formed through repeated experiences of real-world events and serve as conceptual structures in the brain. Frame semantics uses the concept of "frames", "schemas" and "scenarios" to connect language with the objective world, with the idea that people have a mental inventory of structures for organizing, classifying, and interpreting experiences[25]. This notion is important in linguistics, cognitive psychology, and artificial intelligence.

Over the past two decades, scenario commonsense knowledge bases such as FN and CFN have been continually expanded, and many scenario commonsense-based natural language understanding (NLU) tasks such as frame disambiguation, frame semantic role labeling, and null instantiation recognition have shown significant progress. These advances have also led to the development of various applications of natural language understanding, such as machine reading comprehension and text summarization. This paper presents a comprehensive overview of CFN from a commonsense perspective. The main contributions of this paper are as follows.

New perspectives and insights. This is the first time to introduce CFN from a commonsense perspective, and a new concept, i.e., the scenario commonsense, is presented in the paper. Moreover, we also clarify the relations between scenario commonsense and CFN.

Comprehensive introduction. We present a comprehensive overview of CFN resources and their applications, highlighting the newly proposed frame-based discourse representation and a human-machine collaboration framework for expanding the CFN corpus.

Fully comparison. This paper fully compares and analyses the differences and similarities between CFN and other existing commonsense knowledge bases, providing a comprehensive summary of CFN and its relation to these other commonsense knowledge bases.

Wide coverage on researching developments. This paper widely investigates the research progress and advancements in scenario commonsense-based natural language understanding tasks, including frame semantic parsing and relevant applications.

Challenges and future directions. We also discuss potential future challenges and outline potential research directions on the scenario CKB CFN.

The remainder of this paper is organized as follows: In Section 2, we introduce the concept of scenario commonsense and its representation. Section 3 provides a comprehensive overview of CFN resources, highlighting the frame-based discourse representation and the method for expanding the CFN corpus. In Section 4, we investigate the research developments of technologies for scenario CKB-based semantic parsing and their applications. Finally, we present our conclusion and discuss some future challenges and research directions in Section 5.

2 Scenario commonsense

2.1 What is scenario commonsense

Scenario commonsense refers to the knowledge that people have about the everyday situations and events that they encounter. It is the understanding that people have about a particular scenario, which includes certain elements that are specific to that scenario. As shown in Fig. 1, there are two scenarios depicted by blocks: the "commerce money transfer" scenario and the "commerce goods transfer" scenario. The "commerce money transfer"

Fig. 1 Examples of scenario commonsense, i.e., scenario "commerce_money_transfer" and scenario "commerce_goods_transfer". Notice that the block on the left is a money transaction, and the right is a goods transaction.

scenario represents the currency transactions that take place in people's daily lives and typically includes necessary participants such as the buyer, seller, goods, and price. These elements are all part of the scenario commonsense to "commerce money transfer". Similarly, the "commerce goods transfer" scenario includes its specific participants. Additionally, there is a commonsense relation between the two scenarios: If a goods transaction scenario ("commerce goods transfer") occurs, the currency transaction scenario ("commerce money transfer") will also occur.

According to Zang's theory[27], commonsense typically has six fundamental characteristics: shareability, fundamentality, implicitness, large scale, openness, and default. The scenario commonsense fully aligns with these characteristics as follows:

Shareability. Scenario commonsense knowledge is possessed and shared by all people. For example, everyone knows that the scenario "commerce goods transfer" typically includes elements such as a buyer, seller, goods, and price.

Fundamentality. Scenario commonsense knowledge is so fundamental to people's understanding of the world that it is taken for granted.

Implicitness. Scenario commonsense knowledge is often not explicitly discussed because it is widely understood. For example, the knowledge, the "commerce money transfer" scenario involving the participation of both the buyer and the seller, is also widely understood without specialized discussion.

Large scale. Scenario commonsense knowledge exists in large-scale and massive textual data and encompasses a wide range of aspects, including objects, events, states, behaviors, changes, people's feelings, and perceptions.

Openness. Scenario commonsense pertains to all aspects of daily life, rather than being limited to a specific domain.

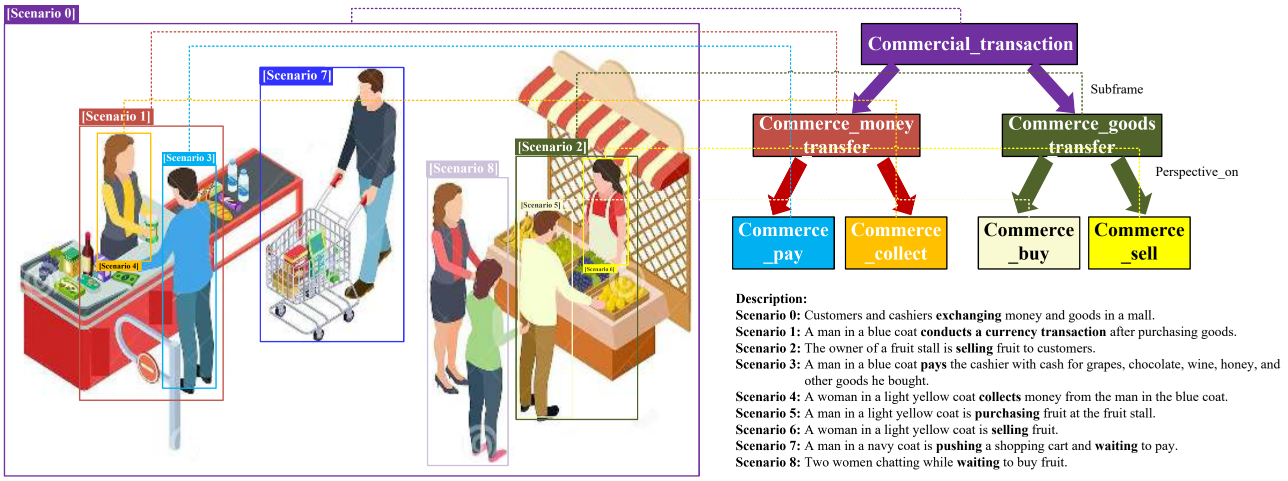
Default. In general, people assume that scenario knowledge is correct in their daily life by default.

2.2 Frame-based scenario commonsense representation

For a specific scenario, a certain number of participants are needed to express its meaning. For example, the participants (buyer, seller, goods, and money) in the scenario "commerce money transfer" are essential to understand the situation. Meanwhile, related scenarios such as "commerce_goods_transfer" and "waiting" may appear alongside "commerce_goods_transfer". As shown in Fig. 2, when people buy goods, the scenario "commerce goods transfer" occurs, and then the buyer must pay for the goods. This results in the scenario "commerce_money_transfer" occurring alongside "commerce_goods_transfer", and the scenario "waiting" may also occur because people in line must wait for the front buyers to make their payments. Additionally, a scenario may be associated with certain words, such as the word "transaction" expressing the scenario "commerce_transaction". These words are referred to as LUs in the scenarios.

For scenario representation, the concept "frame" is proposed in Fillmore's frame semantic theory, which represents the schematic scenarios stored in human experience and related to some motivating scenes occurring in daily life. In formal terms, a scenario can be understood as a "frame" in Fillmore's frame semantic theory. It consists of its definition, frame elements, frame relations, and lexical units, which can be represented by a compound 4-tuple,

$$\boldsymbol{F} = \langle d, \boldsymbol{E}, \boldsymbol{R}, \boldsymbol{U} \rangle \tag{1}$$



(a) Commercial transaction scenarios in a mall (b) Representation using CFN

Fig. 2 An example of commercial transaction scenarios in a mall and mapping representation of these scenarios

where d represents the frame definition; $\boldsymbol{E}$ represents the set of frame roles, each of which is denoted as $e_k \in \boldsymbol{E}$ with $|\boldsymbol{E}| = K$; $\boldsymbol{R}$ represents the set of frame relations between the current frame and other relevant frames, with each element $\boldsymbol{R}_l \in \boldsymbol{R}$ being a 3-tuple $\langle f_i, r_l, f_j \rangle$ that denotes the relation r_l between frames f_i and f_j, and $|\boldsymbol{R}| = L$; finally, $\boldsymbol{U}$ represents the set of lexical units.

Overall, a scenario represents a category of cognitive concepts, while a frame is a concrete and structured representation of a scenario. Based on this formalized frame representation, many scenario-based knowledge bases have been constructed for different languages, such as English FrameNet, Chinese FrameNet, Brazilian FrameNet, and Spanish FrameNet, among others.

3 CFN: An example of Chinese scenario CKB

3.1 CFN overview

CFN[2] is a Chinese scenario CKB constructed based on Fillmore′s theory of frame semantics and according to FN. Fig. 3 provides an overview of the CFN resource, which contains the frame base, sentence base, and discourse base. The frame base is what we refer to as a scenario-based knowledge base, while the sentence and discourse bases are annotated corpora based on the frame base. Their goal is to provide a valuable data foundation for frame-based natural language understanding tasks or applications in Chinese. Currently, CFN has constructed 1 327 frames, 11 567 roles, and 21 288 LUs, with an average of 8.7 frame elements and 16 LUs per frame.

The frame and sentence bases are built mainly based on FN, as shown in Figs. 4 and 5, which provide examples of a frame in the CFN frame base and a sentence annotation example based on the frame. Similar to FN, each frame in the CFN frame base is also composed of a frame definition, frame roles, lexical units, and frame relations; each sentence in the sentence base is also annotated with information related to the frame.

In addition to the frame and sentence bases, which are constructed based on the FN, the CFN has also developed a new frame-based representation method for discourse, and proposed a human-machine collaborative framework for expanding the corpus. These developments improve the knowledge-based language semantic representation and provide an important solution for constructing a large-scale, high-quality corpus. In this section, we will focus on introducing these two aspects in the following subsections.

3.2 Frame-based discourse representation

As we all know, FN has frame-based full-text annotation[28], but it remains at the sentence level and lacks effective semantic organization between frames and sentences, making it difficult to represent discourse-level semantics. To solve this problem, CFN has created a new discourse representation model called frame-based discourse tree[29], which not only takes into account the frames involved in each sentence, but also incorporates coherence relationships in Chinese to facilitate discourse organization. This allows for frame representation at both the sentence and discourse levels.

Xu′s discourse theory[30] posits that discourse-level semantic structures are formed through the continual organization of smaller semantic units into larger ones. Drawing inspiration from this theory, CFN proposes a frame-based approach to organizing the semantics of Chinese discourse: Frames are first taken as the most basic semantic units, and are then combined according to discourse relationships to form larger semantic scenarios. These scenarios are then further combined with adjacent ones to create a hierarchical tree of discourse semantic

[2] Some resource is available through http://www.cfn-lab.com

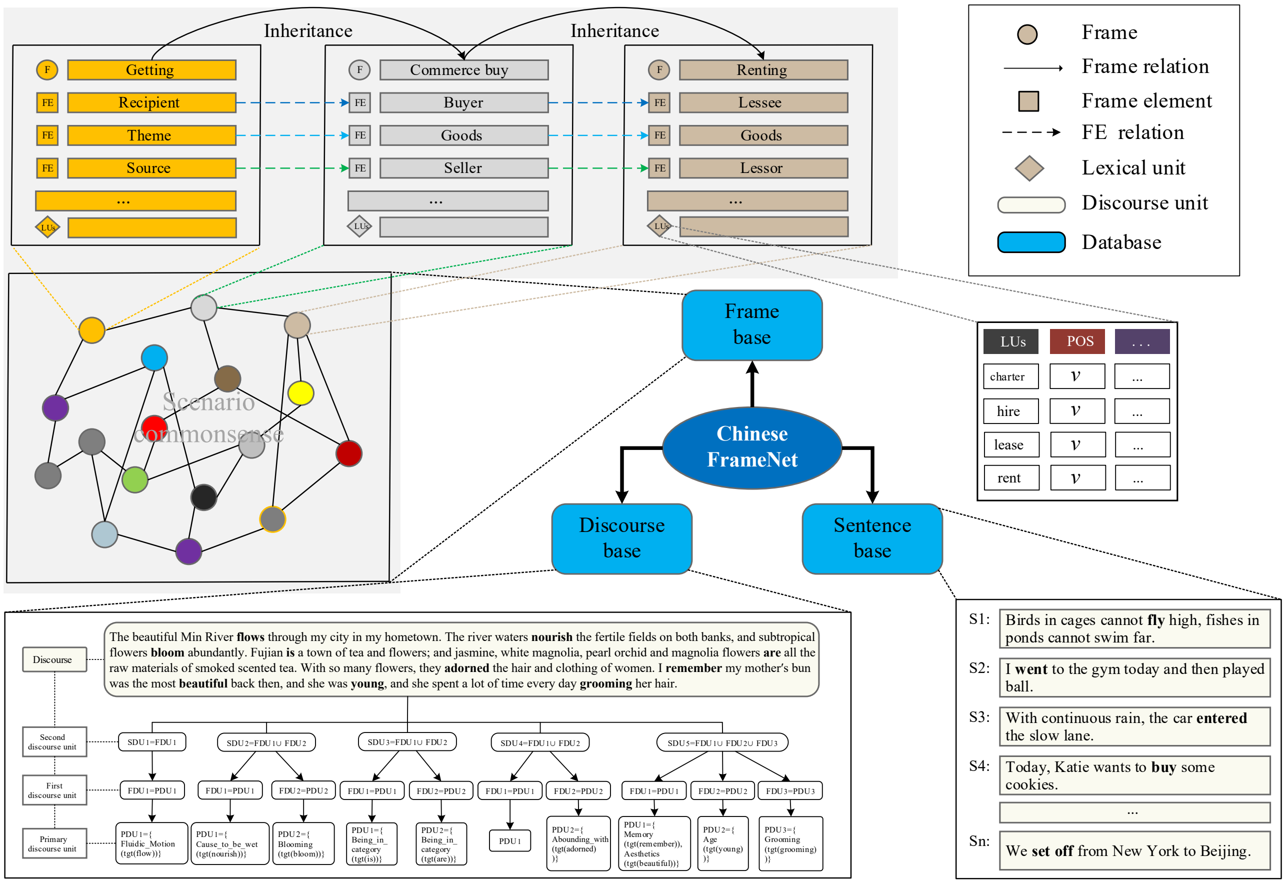


Fig. 3 Chinese FrameNet resources consist of three resource bases: the frame base, sentence base, and discourse base. Frame base is a scenario commonsense knowledge base that contains various frame-based scenarios. The sentence and discourse bases are annotated corpora that are based on the frame base.

Frame: Commercial_transaction

Definition: It describes basic commercial transactions involving a buyer and a seller who exchange money and goods.

Frame Elements:
- Buyer. The buyer wants the goods and offers money to a seller in exchange for them.
- Money. The money is the thing given in exchange for goods in a transaction.
- …

LUs: Transaction. ***n***

F-to-F Relations:
- Inherits from: Reciprocality (frame R for short)
- Subframe of: Commerce_scenario (frame C for short)
- …

This indicates that there is an inheritance relation between current frame and frame R, and it is a subframe of frame C.

Fig. 4 Frame “Commercial_transaction” contains frame definition, frame roles, LUs, and frame relations.

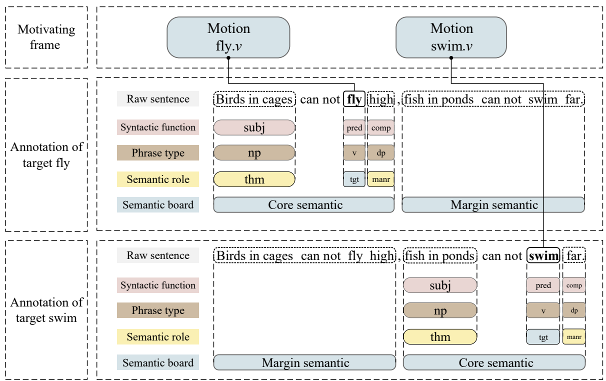


Fig. 5 A frame-based representation of a Chinese sentence. Note that due to different core semantics of concern, a sentence may need to annotate repeatedly.

structures that describes a complete discourse semantic scenario. Specifically, the frame-based discourse tree includes primary discourse unit (PDU), first discourse unit (FDU), and second discourse unit (SDU).

1) Primary discourse unit is a basic semantic unit, which can be separated by a comma, semicolon, colon, etc.

2) First discourse unit is composed of PDUs with frame evoked and adjacent PDUs without frame evoked.

3) Second discourse unit is a high-level semantic unit, which can be separated by a period, exclamation and

question mark, etc.

As shown in the bottom left of Fig. 3, a discourse in the knowledge base is composed of multiple second-level discourse units, which are made up of multiple first-level discourse units. These first-level units are composed of several basic frame units. For a more in-depth understanding of the organizational structure of discourse, refer to literature [29].

In addition, to depict the relationships between semantic units in the frame-based discourse tree, CFN introduces the relationships used to describe the relationships between sentences in Modern Chinese[31], and establishes a three-level representation structure of discourse relationships. The first level adopts the relationships from Modern Chinese that describes the semantic equality between discourse units: the conjunction relationship and the partial positive relationship. In the second level of discourse relationships, the conjunction relationship and the partial positive relationship are further refined into 5 and 6 types, respectively. In the third level of discourse relationships, some of the discourse relationships in the second level are further divided. See Fig. 6 for details.

3.3 Human-machine collaborative corpus expansion

CFN has a variety of frame types with varying semantic labels, making it difficult to build large-scale resources efficiently and with high quality. To address this challenge, we propose a corpus annotation framework that leverages human-machine collaboration. This framework utilizes active learning techniques to identify high-quality samples, and combines human and machine efforts to improve the efficiency of the annotation process at each stage. Using this framework, we are able to more efficiently build large-scale CFN resources and ensure a higher quality annotated corpus.

As illustrated in Fig. 7, CFN expands the corpus using a three-step semi-automatic annotation process consisting of retrieval-based target identification, knowledge-guided frame identification, and active learning-based semantics role labeling. This forms a pipeline method for Chinese corpus expansion. Additionally, before starting the annotation work, we prepare an annotation guideline that includes the annotation standard, training tutorial system, annotated examples, frequently asked questions (FAQ), and other relevant information. We then invite graduate students from our group to participate in the annotation work. To ensure high-quality and consistent annotations, we use the results of consistent annotation as the final annotated sentence instances and require annotators to discuss and record any differences in their annotations.

Retrieval-based target identification. To improve our corpus annotation framework, we have invited linguistic experts to construct a large new lexicon for Chinese that maps to existing frames. For example, the frame "Commercial_transaction" can be evoked by the word "transaction" (see Fig. 4). We use these words from the lexicon as a retrieval base to identify target words that can evoke frames in sentences. To expand the lexicon, we collect corpora from various sources, such as the CCL corpus[3], China Gaokao Chinese, and China People's Daily. Additionally, we update the lexicon with lexical units mapped to frames from FrameNet 1.5 and 1.7. By building a comprehensive and up-to-date lexicon, we can improve the effectiveness of our corpus annotation framework.

Knowledge-guided frame identification. To address the issue that traditional methods often lack sufficient knowledge to accurately identify frames, we have developed a knowledge-guided frame identification method[32]. This method introduces three types of frame knowledge, i.e., frame definitions, frame elements and frame relations. For example, the key information to distinguish "Activity_stop" and "Process_stop" exists in their frame definitions. Besides, frame elements and frame relation are essential to frame identification since frame knowledge defined by linguists can rich frame semantic information potentially. We incorporate these frame knowledge into frame representation which guides models to learn the representation of target words and frames in the same embedding space.

Based on the results of target and frame identification, our annotators select the correct result from multiple identification options. Additionally, the data that machines identify incorrectly will be retained to improve the training of our models. This allows us to ensure high-quality annotations and improve the performance of our models over time.

Active learning based semantics role labeling. Frame semantics role labeling can be performed using either manual or automated annotation methods. To generate high-quality data, sentences with consistent annotations are stored in the annotated sentence base, while those with inconsistent annotations are returned to the unannotated sentence base for further review. To improve the effectiveness of the annotation process, we combine active learning[33] techniques with our frame semantics role labeling approach. This allows us to select and prioritize sentences with inconsistent annotations for further review, improving the overall performance of the model. We train our initial model using few-shot learning, and set a confidence threshold to monitor its performance.

3.4 CFN VS. related CKBs

Researchers have made many efforts in building CKBs, such as ConceptNet, OpenCyc, ATOMIC, and FrameNet, which describe various types of commonsense,

[3] http://ccl.pku.edu.cn:8080/ccl_corpus

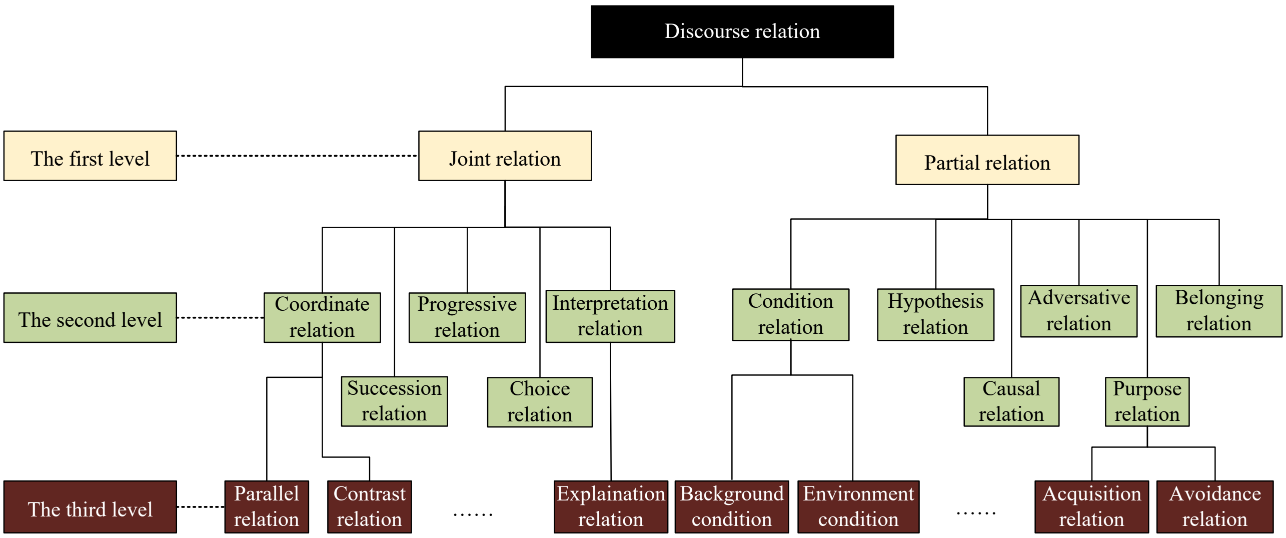


Fig. 6 Structure diagram of discourse relationships[29]. Notice that, the third level contains a total of 22 detailed relationships, and only partial details are shown here.

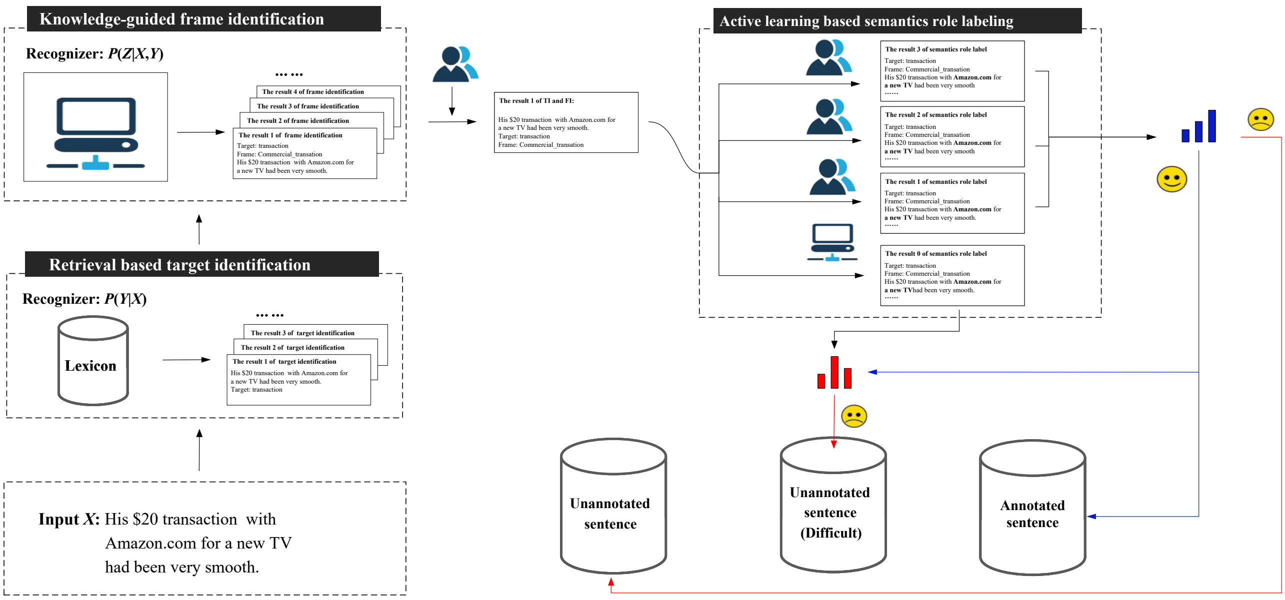


Fig. 7 Chinese sentence instances annotation process in detail

such as concepts, entity relationships, objective laws, and scenarios. It is worth noting that these CKBs use different types of representation formats, design approaches, and acquisition methods. As shown in Table 1, we compare the scenario CKB (i.e., CFN) with several typical CKBs, including OpenCyc[34], HowNet[15], ConceptNet[13], and ATOMIC[14], in terms of content, representation, creation, topic, design approach, and data size. In the following, we will provide a brief summary of each of the CKBs and then conduct a comparison and analysis.

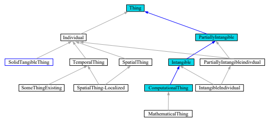
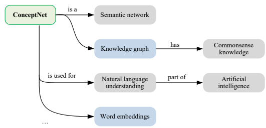
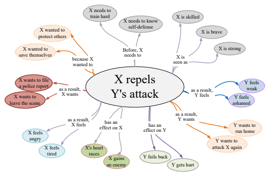
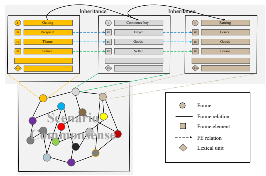
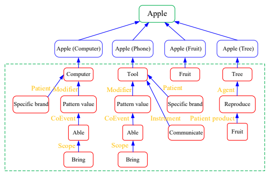
OpenCyc is derived from the Cyc commonsense knowledge base, which represents commonsense knowledge using predicate logic[13]. Cyc aims to create a massive knowledge base containing human commonsense[12, 35, 36]. It was founded in 1984 by Lenat et al.[34] and aims to create a comprehensive ontology and knowledge base of commonsense knowledge.

ConceptNet is an open and multi-lingual knowledge graph and is designed to help computers understand the meanings of words that people use in a freely-available semantic network[13]; it is created by the open mind common sense (OMCS)[37], and the source of ConceptNet comes from WordNet, DBpedia, OpenCyc, etc.

ATOMIC is an atlas of everyday commonsense reasoning that is organized through 877 K textual descriptions of inferential knowledge[14]. It focuses on commonsense knowledge of causes and effects using the form of "if-then". The ATOMIC creators believe that humans have a theory of mind, which allows us to make inferences about people's mental states and understand likely events that precede and follow certain actions.

CFN[38] is a scenario CKB for Chinese language that

Table 1 Comparison of existing commonsense knowledge sources

Content		Representation		Creation	Topic	Design	Size
		Examples	Describes				
OpenCyc	Properties restrictions instances relations		It is described by CycL, which enables differentiation between knowledge involving a concept, i.e., (*#$genls #$Collie #$Dog*) means that a collie is a dog.	Expert manual	General	Top-down	239 000 concepts 2 039 000 facts
ConceptNet	Objects actions states relations		It is represented by triples, i.e., ⟨node_1, relation, node_2⟩, wherein the node is the word or phrase of natural language,the relation, e.g., IsA, isUsedFor, etc.	Crowdsourcing	General	Bottom-up	8 million nodes 21 million edges 36 relations
ATOMIC	Events pre-post conditions		It is represented by triples, i.e., If-Event-Then-*, which focuses on inferential knowledge, e.g., A to B.	Crowdsourcing	Social	Bottom-up	309 515 nodes 877 108 triples 9 relations
CFN	Objects events states behaviors changes feelings perceptions		It is represented by 4-tuple, which pays more attention to the knowledge in a scenario, e.g., “commercial transaction”.	Expert manual curated automatic	General	Bottom-up	1 327 frames 11 567 roles
HowNet	Words senses sememes relations		It is represented by sememe tree, which focuses on the relationship between concepts and their attributes.	Expert manual	General	Bottom-up	229 767 senses 2 187 sememes

is based on frame semantic theory and the FN. It focuses on the commonsense knowledge contained in scenarios, which are schematic scenarios stored in human experience and involve objects, events, states, behaviors, changes, memories, feelings, and perceptions in the real world.

HowNet[15] is a CKB that reveals the relationship between concepts and their attributes. It was initially designed and constructed by Dong et al. and follows the reductionism idea, aiming to study smaller semantic units called sememes. The latest version of HowNet, published in 2019, contains 229 767 senses and 2 187 sememes[39].

As shown in Table 1, we can see that the CKBs OpenCyc, ConceptNet, CFN, and HowNet all contain descriptions of entities such as concepts, objects, or words. In contrast, ATOMIC focuses more on events due to its emphasis on inferential facts. Additionally, we can see that all of the CKBs include descriptions of relations. OpenCyc, ConceptNet, and HowNet are more concerned with the association of concepts, while ATOMIC and CFN focus on the reasoning through relations and describing the connections between scenarios and roles through relations, respectively. It is worth noting that CFN covers a wider range of aspects, including objects, events, states, behaviors, changes, feelings, and perceptions, among others.

Table 1 also provides examples of the different CKBs. We can see that OpenCyc contains ontology knowledge and can be described using the CycL language, which enables differentiation between knowledge involving a concept[40, 41]. For example, a collie is a dog = (*#$genls #$Collie #Dog*) means that “collie” is a collection and “dog” is one generalization of what a collie is[42]. As a knowledge graph, ConceptNet connects words and phrases of natural language with labeled edges[13] and is represented by triples. Because ATOMIC is more focused on inferential knowledge organized as typed if-then relations with variables[14], it is presented in the form of knowledge inference (e.g., If-Event-Then-*). In contrast to the other CKBs, HowNet and CFN contain more detailed knowledge, such as senses, sememes, scenarios, and roles. HowNet uses a sememe tree structure to represent words, senses, sememes, relations, etc., while CFN uses compound 4-tuples to represent scenario commonsense and organizes all scenarios into a graph structure based on scenario relations. Fig. 8 illustrates the differences between CFN, ConceptNet, and ATOMIC in a discourse example. We can see that ATOMIC is more focused on

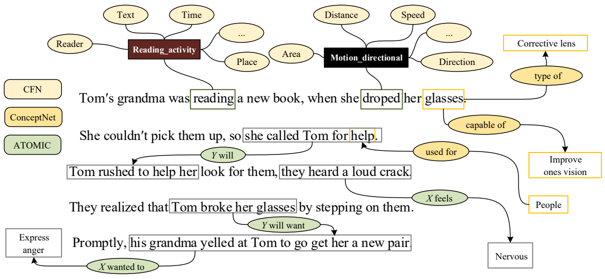


Fig. 8 An example comparison[11] with scenario commonsense. Note that the rectangles in the figure represent frames, while the ellipses represent the common sense elements within the scenario.

inferential knowledge in pre-event and post-events, while ConceptNet and CFN pay more attention to knowledge related to concepts.

As for the creative methods of these CKBs, OpenCyc and HowNet are created manually by experts, while ConceptNet and ATOMIC are created through crowdsourcing. ConceptNet, in particular, is collected from various sources including expert-created resources, crowdsourcing, and games with a purpose[13]. As for CFN, its construction process includes both expert creation and automated curation to improve the quality and efficiency of resource building. Because frame semantics does not provide an explicit upper architecture, the resources for FrameNet (including FN and CFN) adopt a bottom-up approach to build, with the aim of covering as many domains as possible. As a result, the topic of CFN resources focuses on the general domain. In terms of resource size, CFN has reached a significant level of development, containing 1 327 frames and 11 567 roles.

4 CFN/FN-based natural language understanding

Scenario commonsense offers a new perspective for natural language understanding and promotes the study of CFN/FN-based NLU tasks, such as frame semantic parsing, which is a more detailed form of semantic parsing than traditional semantic role labeling tasks[43]. Additionally, some exploratory studies have been conducted on improving NLU applications using CFN/FN, including machine reading comprehension, text summarization, and information extraction, among others. In the following sections, we will provide a brief review of the progress made in frame semantic parsing and several typical NLU applications based on CFN/FN.

4.1 Frame semantic parsing

Frame semantic parsing (FSP) is a task that aims to automatically extract frame-based scenario commonsense from raw text. Formally, given a sentence X with n words $[w_1, w_2, \cdots, w_n]$, FSP involves three subtasks: target identification, frame identification, and frame semantic role labeling.

Target identification, which is to identify all valid frame-evoking predicates from X, outputting a set $\boldsymbol{U} = (u_i, \cdots, u_k)$.

Frame identification, which aims to predict the frame $F_i \in \boldsymbol{F}$ evoked by a given target $u_i \in \boldsymbol{U}$ in sentence X.

Frame semantic role labeling, which aims to identify all the $\langle argument, role\rangle$ pairs R_i given a certain $\langle target, frame\rangle$ pair (u_i, F_i) in sentence X.

In addition to sentence-level FSP tasks, null instantiation identification and resolution is a task that crosses sentences and belongs to the category of discourse-level FSP tasks.

Null instantiation identification and resolution, which involves identifying missing core frame elements in the present content and filling them in using context information.

Target identification involves predicting the words or phrases that evoke frames in a given sentence. Previous studies on this topic have included rule-based methods, sequence labeling methods, and span-based classification methods. Where the rule-based methods[44, 45] use a small set of manually-designed rules to identify targets, and it relies heavily on the results of morphological and syntactic analysis tools. Sequence labeling methods treat target identification as a sequence labeling task using the "BIO" labeling scheme, but this scheme can be inconvenient for handling discontinuous targets like "set$\cdots$down". Bastianelli et al.[46] proposed a "BIOC" labeling scheme to address this issue, where the label "C" denotes the discontinuous token of a target. Span-based methods, on the other hand, treat target identification as a span-level binary classification problem, where a span can consist of one or multiple continuous words. For discontinuous targets, Lin et al.[47] proposed a two-stage method involving partial target span identification and partial target span merging based on the identification of relations between these spans. Unlike sequence labeling methods, span-based methods need to enumerate all possible spans to identify targets in a sentence and often use parallel computing for this purpose.

Frame identification (FI) aims to infer the correct frame for a target word in a sentence, and is typically treated as a classification task. Previous methods for FI have included feature engineering-based methods, distributed representation-based methods, and pre-trained-based methods. Feature engineering-based methods use manually designed features and statistical machine learning methods to learn FI models, such as LTH[45] and SEMAFOR[44]. However, the effectiveness of these methods depends heavily on the quality of the manually designed features. In contrast, distributed representation-based methods use distributed vectors to represent features and neural networks to construct classification models for the FI task[48, 49]. While these methods have some improvements over feature engineering-based methods, they still rely on dependency and part-of-speech (POS) features[50]. Pre-trained-based methods[32, 51, 52] have achieved state-of-the-art performance on the FI task due to their powerful representation ability, which is learned from large-scale unlabeled text corpora. These methods are also convenient for incorporating knowledge and able to represent a variety of features, such as morphology, dependency, and semantics, without relying on manual features.

Frame semantic role labeling (FSRL) involves identifying the frame roles relative to the frame evoked by a target. Previous methods for FSRL can be divided into three categories: chart-based methods, sequence labeling-based methods, and graph-based methods. Chart-based methods use local scoring functions to model each span (candidate frame role) of every segmentation of the sentence given a target, and then search for the globally optimal solution of segmentation scores using an optimization algorithm subject to certain structural constraints. Early studies[44, 53−55] used manually designed discrete features of spans as input for the local scoring functions, while more recent studies[56−58] have used pre-trained static word embeddings and contextualized embeddings of spans as input features for the local scoring functions. The goal of chart-based methods is to find the globally optimal solution under the structural constraints of frame semantics, which can lead to semantically consistent frame role structures. However, these methods typically have high computational complexity. Sequence labeling-based methods can be divided into two categories: token-level sequence labeling and span-label sequence labeling. Token-level sequence labeling methods[59, 60] use traditional BIO tagging schemes, where the role label of each argument span is denoted as "B-rolename, I-rolename, ⋯". Span-label sequence labeling methods[46, 61] typically consist of two subtasks: argument identification and argument classification. Compared to chart-based methods, sequence labeling-based methods can implicitly capture some semantic constraints. However, the abundance of roles and a relatively small amount of labeled data make it difficult to model the precise semantic dependency of the role sequence. In addition, graph-based methods view the FSRL task as a process of graph construction, in which the model incrementally identifies the nodes and their relations[47, 62].

Null instantiation identification and resolution (NIIR) aims to identify missing core frame elements in the present content and fill them in using context information. This task includes identifying null instantiations (NII) and resolving null instantiations (NIR). The SemEval-2010 Task 10 was the first shared task to address the NIIR problem and provided a dataset annotated with implicit roles. Previous methods for NIIR can be divided into two categories: linguistically-motivated methods and neural network-based methods. Linguistically-motivated methods[63−66] try to improve performance using linguistic knowledge and feature engineering, while neural network-based methods use unsupervised or semi-supervised deep learning models[67−69] to solve the NIIR task and mitigate resource deficiencies. Currently, a lack of annotations remains a challenge in training implicit argument models for NIR resolution.

4.2 Applications

Machine reading comprehension (MRC) is a fundamental application in natural language understanding that requires machines to read and understand a text passage and answer relevant questions about it. Many approaches based on CFN/FN have been proposed to improve MRC, which can be divided into two categories: feature-based methods and deep learning-based methods. Feature-based methods aim to map texts into a frame-based representation to facilitate question answering. For example, the latent-variable classifier[70] uses features of the target word, frame, and frame element to capture sentences that have semantic overlap with the question and correct answer. The CFN-based MRC method[71] retrieves answer-related sentences using frames, frame-frame relations, and discourse frame-frame relations, and also selects highly relevant frame elements using null instantiation and frame relations. Feature-based methods often rely on elaborate design, which can be time-consuming and difficult to apply in other tasks. Recently, some deep learning-based work has attempted to enhance sentence representation with CFN/FN and assist in question answering. The frame-based neural network for MRC[72] uses frame features, including lexical units, target, frame, and frame relations, to improve question answering. Additionally, the frame knowledge and syntax for MRC[73] propose to explore the schema of fusing syntax and frame semantics into an end-to-end neural network for MRC. Frame-based neural models can automatically model semantic knowledge and achieve relatively high performance.

Text summarization (TS) aims to condense a text into a short version while preserving its essential semantic information[74]. Some work has shown that using CFN/FN can benefit the performance of text summarization. The unsupervised extractive method[75] utilizes syntactic and semantic concepts from FrameNet to enhance sentence representation. The extractive summarization model[76] leverages frame features to model sentences from both the intra-sentence and inter-sentence levels, facilitating the text summarization task. This work also introduces frame semantics into the abstractive summarization task, allowing for better text semantic representation by selecting more relevant frame information from the text. Additionally, a dual graph network[77] is proposed for abstractive sentence summarization, which models structural and word-level information of articles by constructing frame scenario graphs and word relation graphs, respectively.

Information extraction (IE) aims to automatically extract valuable information from large-scale unstructured, semi-structured, or structured text and process it into human-machine readable structured data[78]. Some research has explored the use of CFN/FN in information extraction tasks. For event detection, Li et al.[79] extract frames expressing event information from FrameNet and leverage the frame-to-frame relations to build a hierarchy of more fine-grained event schemas that have wider coverage than ACE. The multilingual information extraction system LOME[80] first uses the FrameNet parser to identify all frames and their roles, as well as trigger spans in a sentence. It then performs coreference resolution, fine-grained entity typing, and temporal relation prediction between events. Some researchers have also explored event-causal relation extraction based on frame knowledge[81] and found that the introduction of external knowledge is effective in identifying the causal relation between verb-noun pairs. For relation extraction, the frame-based relation extraction model CFSRE[82] uses frame semantic construction rules to eliminate some noise instances in distant supervision, and the experimental results show that fusing frames can significantly improve the noise problem.

4.3 Case study and analysis

Frame semantic knowledge facilitates reasoning and question answering of MRC. Fig. 9(a) presents a typical example of machine reading comprehension question answering, where the frame and frame relations in FN provide crucial information for the reasoning and question answering of the model[83]. Specifically, the Food frame filters out key information related to snacks, and the frame relation between the Locating and Commerce_buy in FN establishes the association of target words Found and Buy in the given passage/question, which helps models find the final answer.

Incorporating frame semantic knowledge is an effective approach to improving the performance of summarization systems. DG-ABS[77] is a based dual-graph summarization system that generates summaries based on scenario graphs and word relation graphs, where scenario graphs and word relation graphs are from FN. As Table 2 and Fig. 9(b) shown, ablation experiments of DG-ABS and its case study example are presented. From ablation experiments results, we can observe that compared to the model without the dual graph, DG-ABS can provide better performance. In addition, from the case study example, we can observe that compared to the –w/o SSG and –w/o SWRG models, DG-ABS captures more complete key information and generates new words.

Semantic scenarios based on frame representation help filter out noise information in information extraction. Fig. 9(c) presents a case study example of frame-based semantic models and non-frame semantic models in information extraction. Among the three sentences, the second one is invalid since it does not contain related content with label semantics. We can find that for identifying invalid sentences, the frame-based model CFSRE[82] can exhibit lower weights, indicating that incorporating semantic scenario information can better distinguish invalid noisy sentences. Additionally, for valid sentences, we found that CFSRE exhibits higher weights compared to BiLSTM+ATT, suggesting that utilizing semantic scenario information about entities can effectively enhance the identification of valid sentences in information extraction, which is helpful for further extraction of entities and relations.

5 Conclusions and future work

As a crucial scenario CKB for natural language understanding, CFN is built from a cognitive perspective and contains a wealth of scenario commonsense knowledge, including frames, frame elements, lexical units, and frame

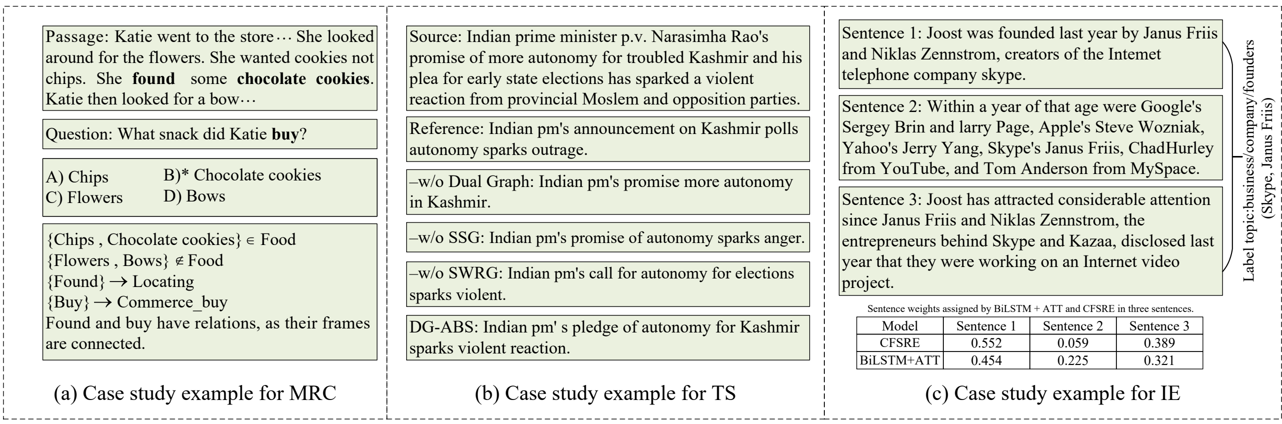


Fig. 9 Case study examples for MRC, TS and IE[77, 82, 83]. Note that * indicates correct answer in MRC example.

Table 2 Ablation study on Gigaword data[77]. Note that SSG, SWRG indicate semantic scenario graph and word relation graph, respectively.

Method	ROUGE-1	ROUGE-2	ROUGE-L
DG-ABS	41.94	23.58	38.97
–w/o SSG	41.02	22.67	38.01
–w/o SWRG	39.29	21.14	36.54
–w/o dual graph	36.74	19.83	35.17

relations. It also includes many annotated instances with frame semantic labels for Chinese sentences and discourses. In this paper, we provide a comprehensive overview of CFN from a commonsense perspective, covering the following topics: 1) Scenario commonsense. We introduce what is scenario commonsense, including its concept, examples, and representation methods, with a focus on the relationship between scenario commonsense and the frame concept in CFN/FN. 2) CFN resources. We present an overview of CFN resources, highlighting the newly proposed frame-based discourse representation and a human-machine collaboration framework for expanding the CFN corpus. 3) CFN/FN-based natural language understanding. We systematically review frame semantic parsing tasks and some FN/FN-based applications. In addition, we also introduce and discuss potential future research directions. After nearly two decades of development, scenario CKBs have a large research community and a wide range of methodologies and applications. Although many efforts have been made to build scenario CKBs and apply them in NLU, there are still many potential challenges and open problems that require attention and further research.

Frame representation in complex Chinese contexts. Metaphors are a complex problem in natural language processing and are considered a high-level reasoning process, i.e., creative thought[84]. They are also essential for representing complex contexts in Chinese. Some related research has emphasized that metaphors arise when one concept is viewed in terms of the properties of another, which aims to establish an association implying the presence of characteristics in common between two concepts[84, 85], e.g., How can I kill a process?[86] To model the way humans use language to frame metaphor reasoning processes, efforts have been made in both metaphor recognition and interpretation[87–91]. Although these existing studies have made significant progress in both benchmark and deep learning models, they lack linguistic knowledge and cognitive computing theory. Therefore, using CFN to describe Chinese complex contexts (e.g., metaphors) is a promising research direction for CFN in the future.

Spatio-temporal relation representation. Several efforts have been made to address spatio-temporal problems, such as building knowledge bases[92], improving models[93], and performing knowledge completion[94]. This highlights the importance of spatio-temporal representation in supporting reasoning tasks in NLU. Scenario CKBs contain information about spatio-temporal frames, such as "Spatial_co-location" and "Time_period_of_action", which are based on people′s everyday experiences and have the potential to benefit various spatio-temporal reasoning tasks, including MC-TACO[95], Temporal-NLI[96], TIMEDIAL[17], and SPARTQA[97]. However, the current spatio-temporal relations in scenario CKBs are insufficient, and there is a need to enhance the expression of space-time in the future to better support reasoning in NLU tasks.

Improving interpretability for MRC. Large-scale pre-trained language models are widely used in a variety of fields[98, 99]. In particular, on MRC tasks such as SQuAD[100] and CMRC[101], these models have achieved performance that is close to human performance[102]. However, these models can learn implicit knowledge but lack interpretability for the reasoning process. Interpretable conceptual attributes of a model typically include "transparency interpretability" and "post-hoc interpretability"[103], which focus on the internal operating mechanism of the model and extracting useful information after training, respectively. Some existing work with knowledge graphs has explored integrating pre-trained language models with knowledge graphs to provide interpretability while capturing reliable knowledge[104, 105]. Similar to a knowledge graph, CFN can link concepts through frame relations, providing potential explainability for downstream tasks such as MRC evidence retrieval. Due to the existence of frames and frame relations, CFN can provide a lot of knowledge inside scenarios and inference between concepts, which are very helpful for many reasoning problems to improve the interpretability of models.

Enhancing performance of text generation. Text generation is a crucial but challenging task in natural language processing. However, the input text alone often provides limited knowledge for generating the desired output, so the performance of text generation is still far from satisfactory in many real-world scenarios[106]. To address this issue, most existing work[107, 108] has explored incorporating knowledge bases into text generation models. However, the knowledge bases used in these models, such as ConceptNet[13] and ATOMIC[14], mainly describe knowledge of general concepts or inferential rules, lacking scenario-based characterization. That characterization in the scenario can supply more fine-grained and comprehensive knowledge based on a special scenario, which can facilitate richer knowledge representation in text generation. Therefore, combining scenario CKBs with text generation models is a promising direction.

Acknowledgements

We thank the reviewers for their helpful comments and suggestions. This work was supported by National Natural Science Foundation of China (Nos. 61936012, 62272285 and 61906111).

Declarations of conflict of interest

The authors declared that they have no conflicts of interest to this work.

References

[1] A. Zareian, Z. C. Wang, H. X. You, S. F. Chang. Learning visual commonsense for robust scene graph generation. In *Proceedings of the 16th European Conference on Computer Vision*, Springer, Glasgow, UK, pp. 642–657, 2020. DOI: 10.1007/978-3-030-58592-1_38.

[2] T. Wang, J. Q. Huang, H. W. Zhang, Q. R. Sun. Visual commonsense R-CNN. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Seattle, USA, pp. 10757–10767, 2020. DOI: 10.1109/CVPR42600.2020.01077.

[3] N. Tandon, A. S. Varde, G. de Melo. Commonsense knowledge in machine intelligence. *ACM SIGMOD Record*, vol. 46, no. 4, pp. 49–52, 2018. DOI: 10.1145/3186549.3186562.

[4] X. H. Zhou, Y. Zhang, L. Y. Cui, D. D. Huang. Evaluating commonsense in pre-trained language models. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, New York, USA, pp. 9733–9740, 2020. DOI: 10.1609/aaai.v34i05.6523.

[5] H. de Medeiros Caseli, B. A. Sugiyama, J. C. Anacleto. Using common sense to generate culturally contextualized machine translation. In *Proceedings of the NAACL HLT Young Investigators Workshop on Computational Approaches to Languages of the Americas*, Los Angeles, USA, pp. 24–31, 2010.

[6] T. Young, E. Cambria, I. Chaturvedi, H. Zhou, S. Biswas, M. L. Huang. Augmenting end-to-end dialogue systems with commonsense knowledge. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence*, Louisiana, USA, Article number 609, 2018.

[7] H. Zhou, T. Young, M. L. Huang, H. Z. Zhao, J. F. Xu, X. Y. Zhu. Commonsense knowledge aware conversation generation with graph attention. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, Stockholm, Sweden, pp. 4623–4629, 2018.

[8] S. X. Wu, Y. Li, D. W. Zhang, Y. Zhou, Z. H. Wu. TopicKA: Generating commonsense knowledge-aware dialogue responses towards the recommended topic fact. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence*, Yokohama, Japan, Article number 521, 2021.

[9] P. Persaud, A. S. Varde, S. Robila. Enhancing autonomous vehicles with commonsense: Smart mobility in smart cities. In *Proceedings of the 29th International Conference on Tools with Artificial Intelligence*, IEEE, Boston, USA, pp. 1008–1012, 2017. DOI: 10.1109/ICTAI.2017.00155.

[10] R. Vedantam, X. Lin, T. Batra, C. L. Zitnick, D. Parikh. Learning common sense through visual abstraction. In *Proceedings of IEEE International Conference on Computer Vision*, Santiago, Chile, pp. 2542–2550, 2015. DOI: 10.1109/ICCV.2015.292.

[11] M. Sap, V. Shwartz, A. Bosselut, Y. Choi, D. Roth. Commonsense reasoning for natural language processing. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts*, pp. 27–33, 2020. DOI: 10.18653/v1/2020.acl-tutorials.7.

[12] D. B. Lenat. CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, vol. 38, no. 11, pp. 33–38, 1995. DOI: 10.1145/219717.219745.

[13] R. Speer, J. Chin, C. Havasi. ConceptNet 5.5: An open multilingual graph of general knowledge. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, San Francisco, USA, pp. 4444–4451, 2017.

[14] M. Sap, R. Le Bras, E. Allaway, C. Bhagavatula, N. Lourie, H. Rashkin, B. Roof, N. A. Smith, Y. Choi. ATOMIC: An atlas of machine commonsense for if-then reasoning. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, Honolulu, USA, pp. 3027–3035, 2019. DOI: 10.1609/aaai.v33i01.33013027.

[15] Z. D. Dong, Q. Dong. HowNet – A hybrid language and knowledge resource. In *Proceedings of the International Conference on Natural Language Processing and Knowledge Engineering*, IEEE, Beijing, China, pp. 820–824, 2003. DOI: 10.1109/NLPKE.2003.1276017.

[16] Y. Wang, C. G. Cao, Z. W. Chen, S. Wang. CKGAC: A commonsense knowledge graph about attributes of concepts. In *Proceedings of the 15th International Conference on Knowledge Science, Engineering and Management*, Springer, Singapore, pp. 585–601, 2022. DOI: 10.1007/978-3-031-10983-6_45.

[17] L. H. Qin, A. Gupta, S. Upadhyay, L. H. He, Y. Choi, M. Faruqui. TIMEDIAL: Temporal commonsense reasoning in dialog. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, pp. 7066–7076, 2021. DOI: 10.18653/v1/2021.acl-long.549.

[18] X. Liu, D. Yin, Y. S. Feng, D. Y. Zhao. Things not written in text: Exploring spatial commonsense from visual signals. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, Dublin, Ireland, pp. 2365–2376, 2022. DOI: 10.18653/v1/2022.acl-long.168.

[19] H. Rashkin, A. Bosselut, M. Sap, K. Knight, Y. Choi. Modeling naive psychology of characters in simple commonsense stories. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, ACL, Melbourne, Australia, pp. 2289–2299, 2018. DOI: 10.18653/v1/P18-1213.

[20] M. Sap, H. Rashkin, D. Chen, R. Le Bras, Y. Choi. Social IQa: Commonsense reasoning about social interactions. In *Proceedings of Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, ACL,

Hong Kong, China, pp. 4463–4473, 2019. DOI: 10.18653/v1/D19-1454.

[21] R. Zellers, A. Holtzman, Y. Bisk, A. Farhadi, Y. Choi. HellaSwag: Can a machine really finish your sentence? In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, ACL, Florence, Italy, pp. 4791–4800, 2019. DOI: 10.18653/v1/P19-1472.

[22] A. Talmor, J. Herzig, N. Lourie, J. Berant. CommonsenseQA: A question answering challenge targeting commonsense knowledge. In *Proceedings of Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, ACL, Minneapolis, USA, pp. 4149–4158, 2019. DOI: 10.18653/v1/N19-1421.

[23] E. M. Ponti, G. Glavaš, O. Majewska, Q. C. Liu, I. Vulić, A. Korhonen. XCOPA: A multilingual dataset for causal commonsense reasoning. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, pp. 2362–2376, 2020. DOI: 10.18653/v1/2020.emnlp-main.185.

[24] C. J. Fillmore. Frame semantics. *Cognitive Linguistics: Basic Readings*, D. Geeraerts, Ed., New York, USA: De Gruyter Mouton, pp. 373–400, 2006. DOI: 10.1515/9783110199901.373.

[25] C. J. Fillmore. Frame semantics and the nature of language. *Annals of the New York Academy of Sciences*, vol. 280, no. 1, pp. 20–32, 1976. DOI: 10.1111/j.1749-6632.1976.tb25467.x.

[26] C. J. Fillmore, C. F. Baker. Frame semantics for text understanding. In *Proceedings of NAACL, WordNet and Other Lexical Resources Workshop*, Pittsburgh, USA, pp.59–64, 2001.

[27] L. J. Zang, C. Cao, Y. N. Cao, Y. M. Wu, C. G. Cao. A survey of commonsense knowledge acquisition. *Journal of Computer Science and Technology*, vol. 28, no. 4, pp. 689–719, 2013. DOI: 10.1007/s11390-013-1369-6.

[28] J. Ruppenhofer, M. Ellsworth, M. R. L. Petruck, C. R. Johnson, J. Scheffczyk. FrameNet II: Extended Theory and Practice, Technical Report, FrameNet Project, USA, 2006.

[29] G. Y. Lv, N. Su, R. Li, Z. Q. Wang. A study on Chinese discourse coherence based on CFN. *Journal of Chinese Information Processing*, vol. 31, no. 5, pp. 40–49, 2017. DOI: 10.3969/j.issn.1003-0077.2017.05.006. (in Chinese)

[30] S. H. Xu. Discourse: Combination of scenarios. *Journal of Foreign Languages*, no. 6, pp. 1–11, 1990. (in Chinese)

[31] B. R. Huang, X. D. Liao. *Modern Chinese*, 5th Ed., Beijing, China: Higher Education Press, 2011. (in Chinese)

[32] X. F. Su, R. Li, X. L. Li, J. Z. Pan, H. Zhang, Q. H. Chai, X. Q. Han. A knowledge-guided framework for frame identification. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, ACL, pp. 5230–5240, 2021. DOI: 10.18653/v1/2021.acl-long.407.

[33] H. F. Tu, R. Li, Z. Q. Wang, T. F. Zhou. Active learning for frame element labeling. *Journal of Chinese Information Processing*, vol. 30, no. 4, pp. 44–55, 2016. (in Chinese)

[34] D. Lenat, M. Prakash, M. Shepherd. CYC: Using common sense knowledge to overcome brittleness and knowledge acquistion bottlenecks. *AI Magazine*, vol. 6, no. 4, pp. 65–85, 1986.

[35] D. B. Lenat, R. V. Guha, K. Pittman, D. Pratt, M. Shepherd. Cyc: Toward programs with common sense. *Communications of the ACM*, vol. 33, no. 8, pp. 30–49, 1990. DOI: 10.1145/79173.79176.

[36] D. McDermott. Building large knowledge-based systems: Representation and inference in the CYC project: D. B. Lenat and R. V. Guha. *Artificial Intelligence*, vol. 61, no. 1, pp. 53–63, 1993. DOI: 10.1016/0004-3702(93)90093-Q.

[37] P. Singh, T. Lin, E. T. Mueller, G. Lim, T. Perkins, W. L. Zhu. Open mind common sense: Knowledge acquisition from the general public. *On the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE*, R. Meersman, Z. Tari, Eds., Berlin, Heidelberg, Germany: Springer, pp. 1223–1237, 2002. DOI: 10.1007/3-540-36124-3_77.

[38] L. P. You, K. Y. Liu. Building Chinese FrameNet database. In *Proceedings of International Conference on Natural Language Processing and Knowledge Engineering*, IEEE, Wuhan, China, pp. 301–306, 2005. DOI: 10.1109/NLPKE.2005.1598752.

[39] F. C. Qi, C. H. Yang, Z. Y. Liu, Q. Dong, M. S. Sun, Z. D. Dong. OpenHowNet: An open sememe-based lexical knowledge base, [Online], Available: https://arxiv.org/abs/1901.09957, 2019.

[40] T. Liebig, O. Noppens. OntoTrack: Combining browsing and editing with reasoning and explaining for OWL lite ontologies. In *Proceedings of the 3rd International Conference on Semantic Web Conference*, Springer, Hiroshima, Japan, pp. 244–258, 2004. DOI: 10.1007/978-3-540-30475-3_18.

[41] C. Matuszek, J. Cabral, M. Witbrock, J. DeOliveira. An introduction to the syntax and content of CYC. In *Proceedings of the Formalizing and Compiling Background Knowledge and its Applications to Knowledge Representation and Question Answering*, Stanford, USA, pp. 44–49, 2006.

[42] K. Coursey. Living in CYN: Mating AIML and CYC Together with Program N. Technical Report, Daxtron Laboratories, Inc., USA, 2004, [Online], Available: http://www.daxtron.com.

[43] W. X. Che, T. Liu, Y. Q. Li. Improving semantic role labeling with word sense. In *Proceedings of the Human Language Technologies: Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Los Angeles, USA, pp. 246–249, 2010.

[44] D. Das, D. S. Chen, A. F. T. Martins, N. Schneider, N. A. Smith. Frame-semantic parsing. *Computational Linguistics*, vol. 40, no. 1, pp. 9–56, 2014. DOI: 10.1162/COLI_a_00163.

[45] R. Johansson, P. Nugues. LTH: Semantic structure extraction using nonprojective dependency trees. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, ACL, Prague, Czech Republic, pp. 227–230, 2007.

[46] E. Bastianelli, A. Vanzo, O. Lemon. Encoding syntactic constituency paths for frame-semantic parsing with graph convolutional networks, [Online], Available: https://arxiv.org/abs/2011.13210, 2020.

[47] Z. C. Lin, Y. H. Sun, M. S. Zhang. A graph-based neural model for end-to-end frame semantic parsing. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, ACL, Punta Cana, Dominican Republic, pp. 3864–3874, 2021. DOI: 10.18653/v1/2021.emnlp-main.314.

[48] K. M. Hermann, D. Das, J. Weston, K. Ganchev. Semantic frame identification with distributed word representations. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, ACL, Baltimore, Maryland, pp. 1448–1458, 2014. DOI: 10.3115/v1/P14-1136.

[49] S. Hartmann, I. Kuznetsov, T. Martin, I. Gurevych. Out-of-domain FrameNet semantic role labeling. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, ACL, Valencia, Spain, pp. 471–482, 2017.

[50] A. Kabbach, C. Ribeyre, A. Herbelot. Butterfly effects in frame semantic parsing: Impact of data processing on model ranking. In *Proceedings of the 27th International Conference on Computational Linguistics*, ACL, Santa Fe, USA, pp. 3158–3169, 2018.

[51] S. S. Tan, J. C. Na. Positional attention-based frame identification with BERT: A deep learning approach to target disambiguation and semantic frame selection, [Online], Available: https://arxiv.org/abs/1910.14549, 2019.

[52] T. Y. Jiang, E. Riloff. Exploiting definitions for frame identification. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics*, ACL, pp. 2429–2434, 2021. DOI: 10.18653/v1/2021.eacl-main.206.

[53] M. Kshirsagar, S. Thomson, N. Schneider, J. Carbonell, N. A. Smith, C. Dyer. Frame-semantic role labeling with heterogeneous annotations. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, ACL, Beijing, China, pp. 218–224, 2015. DOI: 10.3115/v1/P15-2036.

[54] O. Täckström, K. Ganchev, D. Das. Efficient inference and structured learning for semantic role labeling. *Transactions of the Association for Computational Linguistics*, vol. 3, pp. 29–41, 2015. DOI: 10.1162/tacl_a_00120.

[55] N. FitzGerald, O. Täckström, K. Ganchev, D. Das. Semantic role labeling with neural network factors. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, ACL, Lisbon, Portugal, pp. 960–970, 2015. DOI: 10.18653/v1/D15-1112.

[56] S. Swayamdipta, S. Thomson, C. Dyer, N. A. Smith. Frame-semantic parsing with softmax-margin segmental RNNs and a syntactic scaffold, [Online], Available: https://arxiv.org/abs/1706.09528, 2017.

[57] S. Swayamdipta, S. Thomson, K. Lee, L. Zettlemoyer, C. Dyer, N. A. Smith. Syntactic scaffolds for semantic structures. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, ACL, Brussels, Belgium, pp. 3772–3782, 2018. DOI: 10.18653/v1/D18-1412.

[58] H. Peng, S. Thomson, S. Swayamdipta, N. A. Smith. Learning joint semantic parsers from disjoint data. In *Proceedings of Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, ACL, New Orleans, USA, pp. 1492–1502, 2018. DOI: 10.18653/v1/N18-1135.

[59] B. S. Yang, T. Mitchell. A joint sequential and relational model for frame-semantic parsing. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, Copenhagen, Denmark, pp. 1247–1256, 2017. DOI: 10.18653/v1/D17-1128.

[60] D. Marcheggiani, I. Titov. Graph convolutions over constituent trees for syntax-aware semantic role labeling. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, pp. 3915–3928, 2020. DOI: 10.18653/v1/2020.emnlp-main.322.

[61] X. D. Chen, C. Zheng, B. B. Chang. Joint multi-decoder framework with hierarchical pointer network for frame semantic parsing. In *Proceedings of the Findings of the Association for Computational Linguistics*, ACL, pp. 2570–2578, 2021. DOI: 10.18653/v1/2021.findings-acl.227.

[62] C. Zheng, X. D. Chen, R. X. Xu, B. B. Chang. A double-graph based framework for frame semantic parsing. In *Proceedings of Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, ACL, Seattle, USA, pp. 4998–5011, 2022. DOI: 10.18653/v1/2022.naacl-main.368.

[63] J. Ruppenhofer, P. Gorinski, C. Sporleder. In search of missing arguments: A linguistic approach. In *Proceedings of the International Conference Recent Advances in Natural Language Processing*, ACL, Hissar, Bulgaria, pp. 331–338, 2011.

[64] S. Tonelli, R. Delmonte. VENSES++: Adapting a deep semantic processing system to the identification of null instantiations. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, ACL, Uppsala, Sweden, pp. 296–299, 2010.

[65] R. Li, J. Wu, Z. Q. Wang, Q. H. Chai. Implicit role linking on Chinese discourse: Exploiting explicit roles and frame-to-frame relations. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, ACL, Beijing, China, pp. 1263–1271, 2015. DOI: 10.3115/v1/P15-1122.

[66] E. Laparra, G. Rigau. Sources of evidence for implicit argument resolution. In *Proceedings of the 10th International Conference on Computational Semantics*, ACL, Potsdam, Germany, pp. 155–166, 2013.

[67] N. Schenk, C. Chiarcos. Unsupervised learning of prototypical fillers for implicit semantic role labeling. In *Proceedings of Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, ACL, San Diego, USA, pp. 1473–1479, 2016. DOI: 10.18653/v1/N16-1173.

[68] P. Gorinski, J. Ruppenhofer, C. Sporleder. Towards weakly supervised resolution of null instantiations. In *Proceedings of the 10th International Conference on Computational Semantics*, Potsdam, Germany, pp. 119–130, 2013.

[69] M. Le, A. Fokkens. Neural models of selectional preferences for implicit semantic role labeling. In *Proceedings of the 11th International Conference on Language Resources and Evaluation*, European Language Resources Association, Miyazaki, Japan, pp. 3062–3067, 2018.

[70] H. Wang, M. Bansal, K. Gimpel, D. McAllester. Machine comprehension with syntax, frames, and semantics. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, ACL, Beijing, China, pp. 700–706, 2015. DOI: 10.3115/v1/P15-2115.

[71] Z. Q. Wang, R. Li, J. Y. Liang, X. H. Zhang, J. Wu, N. Su. Research on question answering for reading comprehension based on Chinese discourse frame semantic parsing. *Chinese Journal of Computers*, vol. 39, no. 4, pp. 795–807, 2016. DOI: 10.11897/SP.J.1016.2016.00795. (in Chinese)

[72] S. R. Guo, Y. Guan, H. Y. Tan, R. Li, X. L. Li. Frame-based neural network for machine reading comprehension. *Knowledge-based Systems*, vol. 219, Article number 106889, 2021. DOI: 10.1016/j.knosys.2021.106889.

[73] S. R. Guo, Y. Guan, R. Li, X. L. Li, H. Y. Tan. Incorporating syntax and frame semantics in neural network for machine reading comprehension. In *Proceedings of the 28th International Conference on Computational Linguistics*, International Committee on Computational Linguistics, Barcelona, Spain, pp. 2635–2641, 2020. DOI: 10.18653/v1/2020.coling-main.237.

[74] M. F. Mridha, A. A. Lima, K. Nur, S. C. Das, M. Hasan, M. M. Kabir. A survey of automatic text summarization: Progress, process and challenges. *IEEE Access*, vol. 9, pp. 156043–156070, 2021. DOI: 10.1109/ACCESS.2021.3129786.

[75] N. Schluter, A. Søgaard. Unsupervised extractive summarization via coverage maximization with syntactic and semantic concepts. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, ACL, Beijing, China, pp. 840–844, 2015. DOI: 10.3115/v1/P15-2138.

[76] Y. Guan, S. R. Guo, R. Li, X. L. Li, H. Y. Tan. Frame semantic-enhanced sentence modeling for sentence-level extractive text summarization. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, ACL, Punta Cana, Dominican Republic, pp. 4045–4052, 2021. DOI: 10.18653/v1/2021.emnlp-main.331.

[77] Y. Guan, S. R. Guo, R. Li, X. L. Li, H. Zhang. Integrating semantic scenario and word relations for abstractive sentence summarization. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, ACL, Punta Cana, Dominican Republic, pp. 2522–2529, 2021. DOI: 10.18653/v1/2021.emnlp-main.196.

[78] M. F. Moens. *Information Extraction: Algorithms and Prospects in a Retrieval Context*, Dordrecht, The Netherlands: Springer, 2006.

[79] W. Li, D. Z. Cheng, L. He, Y. Z. Wang, X. L. Jin. Joint event extraction based on hierarchical event schemas from FrameNet. *IEEE Access*, vol. 7, pp. 25001–25015, 2019. DOI: 10.1109/ACCESS.2019.2900124.

[80] P. Xia, G. H. Qin, S. Vashishtha, Y. M. Chen, T. F. Chen, C. May, C. Harman, K. Rawlins, A. S. White, B. Van Durme. LOME: Large ontology multilingual extraction. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, ACL, pp. 149–159, 2021. DOI: 10.18653/v1/2021.eacl-demos.19.

[81] M. Riaz, R. Girju. Recognizing causality in verb-noun pairs via noun and verb semantics. In *Proceedings of EACL Workshop on Computational Approaches to Causality in Language*, ACL, Gothenburg, Sweden, pp. 48–57, 2014. DOI: 10.3115/v1/W14-0707.

[82] H. Y. Zhao, R. Li, X. L. Li, H. Y. Tan. CFSRE: Context-aware based on frame-semantics for distantly supervised relation extraction. *Knowledge-based Systems*, vol. 210, Article number 106480, 2020. DOI: 10.1016/j.knosys.2020.106480.

[83] S. R. Guo, R. Li, H. Y. Tan, X. L. Li, Y. Guan, H. Y. Zhao, Y. P. Zhang. A frame-based sentence representation for machine reading comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 891–896, 2020. DOI: 10.18653/v1/2020.acl-main.83.

[84] E. Shutova. Models of metaphor in NLP. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Uppsala, Sweden, pp. 688–697, 2010.

[85] S. Rai, S. Chakraverty. A survey on computational metaphor processing. *ACM Computing Surveys*, vol. 53, no. 2, Article number 24, 2021. DOI: 10.1145/3373265.

[86] J. H. Martin. Representing regularities in the metaphoric lexicon. In *Proceedings of the 12th Conference on Computational Linguistics*, pp. 396–401, 1988.

[87] A. Neidlein, P. Wiesenbach, K. Markert. An analysis of language models for metaphor recognition. In *Proceedings of the 28th International Conference on Computational Linguistics*, Barcelona, Spain, pp. 3722–3736, 2020. DOI: 10.18653/v1/2020.coling-main.332.

[88] O. Rohanian, M. Rei, S. Taslimipoor, L. A. Ha. Verbal multiword expressions for identification of metaphor. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 2890–2895, 2020. DOI: 10.18653/v1/2020.acl-main.259.

[89] X. Chen, Z. Hai, S. G. Wang, D. Y. Li, C. Wang, H. B. Luan. Metaphor identification: A contextual inconsistency based neural sequence labeling approach. *Neurocomputing*, vol. 428, pp. 268–279, 2021. DOI: 10.1016/j.neucom.2020.12.010.

[90] B. Xu, T. T. Li, J. Z. Zheng, M. Naseriparsa, Z. H. Zhao, H. F. Lin, F. Xia. MET-meme: A multimodal meme dataset rich in metaphors. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Madrid, Spain, pp. 2887–2899, 2022. DOI: 10.1145/3477495.3532019.

[91] O. Zayed, J. P. McCrae, P. Buitelaar. Figure me out: A gold standard dataset for metaphor interpretation. In *Proceedings of the 12th Language Resources and Evaluation Conference*, Marseille, France, pp. 5810–5819, 2020.

[92] J. Hoffart, F. M. Suchanek, K. Berberich, G. Weikum. YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. *Artificial Intelligence*, vol. 194, pp. 28–61, 2013. DOI: 10.1016/j.artint.2012.06.001.

[93] W. Chen, H. Y. Wan, S. N. Guo, H. Y. Huang, S. J. Zheng, J. M. Li, S. H. Lin, Y. F. Lin. Building and exploiting spatial-temporal knowledge graph for next POI recommendation. *Knowledge-based Systems*, vol. 258, Article number 109951, 2022. DOI: 10.1016/j.knosys.2022.109951.

[94] P. Jain, S. Rathi, Mausam, S. Chakrabarti. Temporal knowledge base completion: New algorithms and evaluation protocols. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, ACL, pp. 3733–3747, 2020. DOI: 10.18653/v1/2020.emnlp-main.305.

[95] B. Zhou, Q. Ning, D. Khashabi, D. Roth. Temporal common sense acquisition with minimal supervision. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 7579–7589, 2020. DOI: 10.18653/v1/2020.acl-main.678.

[96] S. Vashishtha, A. Poliak, Y. K. Lal, B. Van Durme, A. S. White. Temporal reasoning in natural language inference. In *Proceedings of the Findings of the Association for Computational Linguistics: EMNLP*, pp. 4070–4078, 2020. DOI: 10.18653/v1/2020.findings-emnlp.363.

[97] R. Mirzaee, H. R. Faghihi, Q. Ning, P. Kordjamshidi. SPARTQA: A textual question answering benchmark for spatial reasoning. In *Proceedings of Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 4582–4598, 2021. DOI: 10.18653/v1/2021.naacl-main.364.

[98] Z. Y. Zhang, X. Han, Z. Y. Liu, X. Jiang, M. S. Sun, Q. Liu. ERNIE: Enhanced language representation with informative entities. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, ACL, Florence, Italy, pp. 1441–1451, 2019. DOI: 10.18653/v1/P19-1139.

[99] H. W. Zha, Z. Y. Chen, X. F. Yan. Inductive relation prediction by BERT. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence*, Palo Alto, USA, pp. 5923–5931, 2022. DOI: 10.1609/aaai.v36i5.20537.

[100] P. Rajpurkar, R. Jia, P. Liang. Know what you don't know: Unanswerable questions for SQuAD. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, ACL, Melbourne, Australia, pp. 784–789, 2018. DOI: 10.18653/v1/P18-2124.

[101] Y. M. Cui, T. Liu, W. X. Che, L. Xiao, Z. P. Chen, W. T. Ma, S. J. Wang, G. P. Hu. A span-extraction dataset for Chinese machine reading comprehension. In *Proceedings of Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, ACL, Hong Kong, China, pp. 5883–5889, 2019. DOI: 10.18653/v1/D19-1600.

[102] H. Y. Tan, X. Y. Wang, Y. Ji, R. Li, X. L. Li, Z. W. Hu, Y. X. Zhao, X. Q. Han. GCRC: A new challenging MRC dataset from Gaokao Chinese for explainable evaluation. In *Proceedings of the Findings of the Association for Computational Linguistics*, ACL, pp. 1319–1330, 2021. DOI: 10.18653/v1/2021.findings-acl.113.

[103] Z. C. Lipton. The mythos of model interpretability. *Communications of the ACM*, vol. 61, no. 10, pp. 36–43, 2018. DOI: 10.1145/3233231.

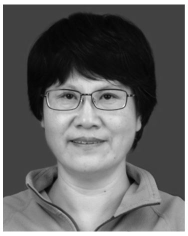
[104] X. K. Zhang, A. Bosselut, M. Yasunaga, H. Y. Ren, P. Liang, C. D. Manning, J. Leskovec. GreaseLM: Graph reasoning enhanced language models for question answering, [Online], Available: https://arxiv.org/abs/2201.08860, 2022.

[105] Y. Q. Sun, Q. Shi, L. Qi, Y. Zhang. JointLK: Joint reasoning with language models and knowledge graphs for commonsense question answering. In *Proceedings of Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, ACL, Seattle, USA, pp. 5049–5060, 2022. DOI: 10.18653/v1/2022.naacl-main.372.

[106] W. H. Yu, C. G. Zhu, Z. T. Li, Z. T. Hu, Q. Y. Wang, H. Ji, M. Jiang. A survey of knowledge-enhanced text generation. *ACM Computing Surveys*, vol. 54, no. 11s, Article number 227, 2022. DOI: 10.1145/3512467.

[107] J. Guan, F. Huang, Z. H. Zhao, X. Y. Zhu, M. L. Huang. A knowledge-enhanced pretraining model for commonsense story generation. *In Proceedings of the Transactions of the Association for Computational Linguistics*, Cambridge, USA, pp. 93–108, 2020. DOI: 10.1162/tacl_a_00302.

[108] Y. Liu, Y. Wan, L. F. He, H. Peng, P. S. Yu. KG-BART: Knowledge graph-augmented BART for generative commonsense reasoning. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence*, Palo Alto, USA, pp. 6418–6425, 2021. DOI: 10.1609/aaai.v35i7.16796.



Ru Li received the Ph. D. degree in computer application technology from Shanxi University, China in 2012. She is currently a professor with Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, School of Computer and Information Technology, Shanxi University, China. She has published numerous papers in important academic journals and conferences, such as IEEE TKDE, ACL, IJCAI, EMNLP, COLING, NLPCC, etc.

Her research interests include natural language processing, text semantics analysis and machine learning.

E-mail: liru@sxu.edu.cn
ORCID iD: 0000-0003-1545-5553

Yunxiao Zhao received the B. Sc. degree in computer and information technology from Shanxi University, China in 2020. He is currently a Ph. D. degree candidate in computer science and technology at School of Computer and Information Technology, Shanxi University, China.

His research interests include natural language processing and text data mining.

E-mail: yunxiaomr@163.com
ORCID iD: 0000-0002-9133-7324

Zhiqiang Wang received the Ph.D. degree in computer application technology from Shanxi University, China in 2018. He is currently an associate professor with School of Computer and Information Technology, Shanxi University, China. He received the Best Doctoral Dissertation of Chinese Information Processing Society of China (CIPS) in 2019. He has been serving as a reviewer for many academic journals and conferences such as *IEEE Transactions on Neural Networks and Learning Systems*, *Neural Networks*, *Neurocomputing*, ACL, KDD, COLING, and EMNLP.

His research interests include natural language processing, network big data mining and machine learning.

E-mail: wangzq@sxu.edu.cn (Corresponding author)
ORCID iD: 0000-0002-9269-3988

Xuefeng Su received the M.Sc. degree in computer application technology from Taiyuan University of Technology, China in 2007. He is an associated professor at the Shanxi Vocational University of Engineering Science and Technology, China and also a Ph.D. degree candidate in computer science and technology at the Shanxi University, China. Recently, he has published more than 10 papers in journals and conferences of NLP area, such as ACL, NLPCC, TALLIP, etc.

His research interests include data mining and semantic analysis on text.

E-mail: suexf@163.com
ORCID iD: 0000-0002-8954-4412

Shaoru Guo received the Ph.D. degree in computer science and technology from Shanxi University, China in 2021. She is currently a postdoctoral researcher in Institute of Automation, Chinese Academy of Sciences, China.

Her research interests include machine reading comprehension and information extraction.

E-mail: guoshaoru0928@163.com
ORCID iD: 0000-0003-4130-3924

Yong Guan received the Ph.D. degree in computer science and technology from Shanxi University, China in 2022. He is currently a postdoctoral researcher in Department of Computer Science and Technology, Tsinghua University, China.

His research interests include text generation and information extraction.

E-mail: guanyong0130@gmail.com
ORCID iD: 0000-0002-9044-2595

Xiaoqi Han received the B.Sc. degree in computer science and technology from Hefei University of Technology, China 2019. He is currently a Ph.D. degree candidate in computer science and technology at Shanxi University, China. During his Ph.D. study, he took part in Research Chinese FrameNet Computing Based on Language Cognitive Mechanism Project, mainly responsible for the semantic analysis.

His research interests include natural language processing and semantic parsing.

E-mail: xiaoqisev@163.com
ORCID iD: 0000-0002-8827-8474

Hongyan Zhao received the Ph.D. degree in computer application technology from Shanxi University, China in 2021. She is currently an associate professor with School of Computer Science and Technology, Taiyuan University of Science and Technology, China.

Her research interests include natural language processing, intelligent data mining and information extraction.

E-mail: hongyanzhao@tyust.edu.cn
ORCID iD: 0000-0002-1968-6239

Machine Intelligence Research

Citation: R. Li, Y. Zhao, Z. Wang, X. Su, S. Guo, Y. Guan, X. Han, H. Zhao. A comprehensive overview of cfn from a commonsense perspective. *Machine Inteligence Research*, vol.21, no.2, pp.239–256, 2024. https://doi.org/10.1007/s11633–023–1450–8