Computers, Materials & Continua

Tech Science Press

DOI: 10.32604/cmc.2023.046475

ARTICLE

A Joint Entity Relation Extraction Model Based on Relation Semantic Template Automatically Constructed

Wei Liu, Meijuan Yin*, Jialong Zhang and Lunchong Cui

Key Laboratory of Cyberspace Situation Awareness of Henan Province, Zhengzhou, 450001, China

*Corresponding Author: Meijuan Yin. Email: raindot_ymj@163.com

Received: 03 October 2023 Accepted: 21 November 2023 Published: 30 January 2024

ABSTRACT

The joint entity relation extraction model which integrates the semantic information of relation is favored by relevant researchers because of its effectiveness in solving the overlapping of entities, and the method of defining the semantic template of relation manually is particularly prominent in the extraction effect because it can obtain the deep semantic information of relation. However, this method has some problems, such as relying on expert experience and poor portability. Inspired by the rule-based entity relation extraction method, this paper proposes a joint entity relation extraction model based on a relation semantic template automatically constructed, which is abbreviated as RSTAC. This model refines the extraction rules of relation semantic templates from relation corpus through dependency parsing and realizes the automatic construction of relation semantic templates. Based on the relation semantic template, the process of relation classification and triplet extraction is constrained, and finally, the entity relation triplet is obtained. The experimental results on the three major Chinese datasets of DuIE, SanWen, and FinRE show that the RSTAC model successfully obtains rich deep semantics of relation, improves the extraction effect of entity relation triples, and the F1 scores are increased by an average of 0.96% compared with classical joint extraction models such as CasRel, TPLinker, and RFBFN.

KEYWORDS

Natural language processing; deep learning; information extraction; relation extraction; relation semantic template

1 Introduction

The purpose of entity relation extraction technology is to extract the triple structure <head entity, relation, tail entity> from an unstructured or semi-structured natural language text, which is convenient for computer identification and utilization so that people can better extract key data from the vast amount of information on the Internet. As an important pre-task of many natural language tasks, entity relation extraction provides data support for downstream tasks such as knowledge graph, automatic question answer, intelligent search, and sentiment analysis, and is widely used in public opinion control, biomedicine, news finance, machine translation, and other fields, with important theoretical research value and broad application prospects [1,2].

The research on entity relation extraction technology has a long history and is generally divided into three development stages. The rule-based entity relation extraction technology relies on manual labeling rules, which is time-consuming and laborious, needs the assistance of domain knowledge, and has poor portability and insufficient applicability. The entity relation extraction technology based on traditional machine learning does not need to design rules, but it also needs to select features manually. The effectiveness of features directly affects the extraction results, and there is a problem of error propagation in feature extraction, which is not effective on large-scale corpus datasets [3]. The entity relation extraction technology based on deep learning can automatically learn sentence features, reduce the steps of manual feature selection, and improve the problem of error accumulation in feature extraction, so it has become a research hotspot in the field of entity relation extraction in recent years. According to the different completion sequences of the two subtasks of entity identification and relation classification, the entity relation extraction methods based on deep learning can be divided into pipeline method and joint extraction method. The pipeline method divides the entity relation extraction task into two subtasks: entity identification and relation classification, which is flexible, but there are problems of error accumulation and entity redundancy. The joint extraction method integrates the two subtasks by unified coding, which alleviates the error accumulation problem in the pipeline method, but there are still some problems such as nested entities, entity overlapping, and entity redundancy.

To solve the above problems, many scholars have made various attempts. Among them, the joint extraction strategy of integrating relation semantic information is widely adopted because of its outstanding performance in solving entity redundancy [4]. For example, a relation-oriented cascade extraction model proposed by Hu et al. [5] integrates the relation coding as a priori knowledge into the sentence coding for entity recognition, but this method only uses the surface semantic information of the relation and does not dig the semantic information of the relation in depth. Li et al. [6] proposed a joint extraction method based on a relation semantic template. The model transformed the relation into a template containing fine-grained relation semantic information by manually defining the relation semantic template, and modeled the entity identification task as a template filling task, which solved the problem of entity overlap, realized the deep expression of relation semantics and improved the accuracy of entity relation extraction. Although the method of manually defining the relation semantic template realizes the deep expression of relation semantics, there are still two shortcomings. First, the natural language is rich and varied in expression, and the artificially defined relation semantic template relies too much on expert experience and cannot completely contain all the semantic information of the relation. For example, not only the relation "wife" can be obtained from the semantic template of "[MASK] is the wife of [MASK]", but also can be obtained from the template of "[MASK] is the husband of [MASK]". Second, the semantic template of artificially defined relations is not portable and cannot be applied to large-scale data sets. The above two problems lead to the existing joint extraction model of fusion relation semantics being improved in effect.

Aiming at the problems of the existing joint extraction model integrating relation semantics, such as the artificial definition of relation semantic template depends on expert experience and poor portability, this paper proposes a joint entity relation extraction model based on a relation semantic template automatically constructed, which is called RSTAC model for short. This model first analyzes the relation corpus based on dependency parsingand automatically constructs the relation semantic template according to the rules. Then, based on the relation semantic template, the relation classification and triplet extraction are carried out, and finally, the entity relation triplet is obtained.

The main contributions of this paper are as follows:

1. An automatic construction method of relation semantic template based on dependency parsing is proposed. In this method, the relation corpus is statistically analyzed by dependency parsing and part-of-speech tagging, and the corresponding rules for extracting relation semantic templates are refined. According to the corresponding rules, templates containing fine-grained relation semantic information are automatically constructed, and a relation semantic template library is established for each relation, which overcomes the problems of dependence on expert knowledge and poor portability in manually defining relation semantic templates.
2. A relation classification method based on a relation semantic template is proposed. Starting from the deep semantic information contained in the relation semantic template, through the attention mechanism, according to the correlation between the relation semantic template and the sentence, only the relation associated with the sentence is identified, which improves the effect of relation classification.
3. A method of triple extraction based on a relation semantic template is proposed. This method starts from the correlation between triples and sentences, and the dependence between head and tail entities and relations within triples. Based on the relation semantic template, this method extracts entity relation triples through the attention mechanism, which improves the effect of triples extraction.

The organizational structure of this paper mainly includes: Section 1 mainly introduces the research background of this paper and the contributions made. Section 2 summarizes the work in the area of entity relation extraction. Section 3 describes our model in detail. Section 4 provides the dataset as well as the baseline model, and analyzes the experimental results. Conclusion of the paper and plans in Section 5.

2 Related Work

The related research on entity relation extraction has a long history. Since 1998, the development of entity relation extraction technology has gone through three stages: rule-based method, traditional machine learning method, and deep learning method.

The rule-based entity relation extraction method is also called the pattern-matching-based entity relation extraction method. This method needs to build a pattern rule set based on words, phrases, or semantics in related fields, and then match it with sentences to identify triples. Cunningham et al. [7] proposed and developed a rule-based GATE system, which can write rules to extract entity relations. Wen et al. [8] proposed an extended association rule method to extract Chinese unclassified relations. This kind of method needs to design a large number of rules manually and a lot of expert knowledge, and it performs well in the extraction effect of specific fields [9]. However, because there is no uniform format in the expression of natural language, it is of course impossible to encapsulate all the rules in the artificially constructed rule base, so there are problems such as narrow scope of use and poor portability.

An entity relation extraction method based on traditional machine learning classifies the entity relations in sentences through machine learning. Rink et al. [10] used SVM for classification, and the features used include 8 groups and 45 types of features such as vocabulary. Bunescu et al. [11] observed that the shortest dependency between two entities is related to the entity relation through dependency parsing, and improved it by using the shortest dependency tree core. The effect of this method depends heavily on the selection of features such as vocabulary and syntax, and a large number of manual selections make it impossible to apply this method to a large-scale corpus data set.

The method of entity relation extraction based on deep learning can automatically learn sentence features without complex feature engineering and has gradually become a research hotspot in the field of entity relation extraction in recent years [12–14]. According to the different ways of subtasks, the entity relation extraction methods based on deep learning can be mainly divided into pipeline method and joint extraction method.

The pipeline method [15,16] divides the entity relation extraction task into two independent entity recognition and relation classification subtasks, which are independent of each other and do not interfere with each other. Although the pipeline method has the advantages of freedom, flexibility, and convenient deployment, it has the problems of error accumulation and entity redundancy because it ignores the correlation between two subtasks.

The joint extraction method [17–20] combines the two subtasks of entity identification and relation classification into a model, trying to improve the effect by strengthening the dependence between the two subtasks. Miwa et al. [21] adopted the unified coding method of entity recognition and relation classification for the first time to carry out joint extraction. Wei et al. [22] proposed a cascade extraction method based on head entity orientation, which newly modeled the relation extraction from $f(s,o) \to r$ to $f_r(s) \to o$, that is, first extracted the head entity in the sentence, and then identified its tail entity for each relation type. Li et al. [23] introduced an external language knowledge base combined with word information, but the effect of this method also depends heavily on the quality of the external language knowledge base. Wang et al. [24] put forward a joint model of entity relations based on structured prediction. This model transforms the joint extraction task into the problem of etymological pair linking, which can solve the overlapping of entities and relations, but cannot take into account the redundancy of entities.

To solve the problem of entity redundancy, the relation-oriented joint extraction strategy is widely used. Takanobu et al. [25] used hierarchical reinforcement learning to extract relations first, and then entities. Zheng et al. [26] divided the joint extraction task into three subtasks and extracted triples in the order of relation classification, entity recognition, and head-tail entity alignment. Hu et al. [5] put forward a relation-oriented cascade extraction model, which integrated the relation information into entity recognition as prior knowledge, thus alleviating the problem of entity redundancy. Li et al. [6] proposed a joint extraction method based on a relation semantic template. In this model, the relation was transformed into a template containing deep relation semantic information by manually defining the relation semantic template according to the order of relation before the entity and the entity identification task was modeled as a template-filling task, which solved the problem of entity overlap and improved the accuracy of entity relation extraction. However, due to the problems of relying on expert knowledge and poor portability in manually defining relation semantic templates, the existing related models still need to be improved in extraction effect.

From the triples basic format <head entity, relation, tail entity>, we can derive a strong connection between the relation information and the head-tail entity and the triplet whole. Unfortunately, the current model does few pay attention to this point, and always focuses too much on sentence and entity features such as syntax, entity position, and part of speech in feature utilization and selection, and does not organically combine relation features with triplet extraction, ignoring the important role of relation features in entity relation extraction. Although the relation-oriented joint extraction model has done some research on this, it still cannot obtain the deep semantic information of the relations. The relation semantic template designed by Li et al. [6] expands the simple relational word into a text that uses entities to limit the scope of relational semantics, giving us a channel to obtain the deep semantics of relation. Previously, due to the limitations of the relational word itself, no matter what method was used to encode, it was impossible to obtain more fine-grained relational semantic information, but this descriptive template that combines relation and entities provides a possibility. However, due to the diversity of natural language expressions, artificially defined relation semantic templates cannot contain all language patterns of relations, and there is a problem that the deep semantics of the acquired relations are not comprehensive. In addition, the method of manual participation also relies on expert experience and poor portability. So this article will focus on how to automate the construction of relation semantic templates and how to strengthen the connection between relation and triplet extraction.

The model designed in this paper is mainly divided into five stages, namely, automatic construction of relation semantic templates, sentence encoding, entity recognition, relation classification based on relation semantic templates, and triplet extraction based on relation semantic templates. Among them, the automatic construction of relation semantic templates solves the problems of relying on expert knowledge, poor portability, and incomplete relation semantic information of manually defined relational semantic templates. Relation classification based on a relation semantic template improves the effect of relation classification. Triplet extraction based on a relation semantic template improves the extraction effect of entity relation triples by strengthening the relation between relation and triplet extraction.

3 Model

3.1 Formal Definition of Entity Relation Extraction

For a sentence $S = \{x_1,\ x_2, \ldots, x_n\}$ with a given word length of n, x_i represents the i-th word in the sentence S; There is a predefined relation set $R = \{r_1, r_2, \ldots, r_m\}$. Then the task of entity relation extraction can be described as extracting all the relation triples (e_1, r, e_2) in the sentence S, where (e_1, e_2) represents the entity pair in the sentence, e_1 represents the head entity, e_2 represents the tail entity, r represents the relation type between them, $r \in R$.

3.2 Framework of the RSTAC

From the perspective of the automatic construction of relation semantic templates, this paper proposes a joint entity relation extraction model based on the automatic construction of relation semantic templates. The framework of the model is shown in Fig. 1. Firstly, the model extracts the rules for extracting relation semantic templates from the relation corpus through dependency parsing and automatically constructs relation semantic templates according to the rules. On this basis, the sentence and relation semantic template are coded by the BERT model, respectively; Then, based onsentence coding, entity recognition is carried out; And based on sentence coding and relation semantic template coding, the relation classification is carried out synchronously; Finally, the whole matrix of entity pairs-relations is established, and the triplet extraction is realized by using the correlation between candidate triplets and relation semantic templates and sentences.

The main steps of the model are as follows:

Step 1: Automatic construction of relation semantic template. According to part-of-speech tagging and dependency parsing, the extraction rules of relation semantic templates in relation corpus are extracted, and a candidate relation semantic template library is constructed for each relation according to the rules. Through the analysis of similar relation semantic templates in relation corpus, the final relation semantic template library $R_i^{M'} = \left\{R_{i1}^m, R_{i2}^m, \ldots, R_{if}^m\right\}$ is determined, in which R_{ig}^m is the g-th relation semantic template in the relation semantic template library of relation r_i.

Step 2: Code sentences. Because the pre-trained BERT model [27] can get the rich semantic information contained in the sentence, this paper directly uses the pre-trained BERT model to encode the sentence $S = \{x_1,\ x_2, \ldots, x_n\}$, and obtains the coded representation of the sentence H_e.

Step 3: entity extraction. Based on obtaining the coding representation H_e of the sentence, the starting position and ending position of the entity are identified by two identical entity markers, and then all the entities contained in the sentence are obtained.

Step 4: Relation classification based on relation semantic template. Based on sentence coding, the relation classification task is modeled as a multi-label two-classification task, and the candidate relation set $R_s = \{r_1, r_2, \ldots, r_s\}$ $(s \leq m)$ in the sentence is identified by the sigmod function; BERT is used to encode the relation semantic template to obtain the deep semantic information of the relation, and on this basis, through the attention mechanism, the correlation between the candidate relation set and the sentence coding and the relation semantic template corresponding to the candidate relation set is calculated, and then the real relation set $R_r = \{r_1, r_2, \ldots, r_t\}$ $(t \leq s)$ contained in the sentence is obtained.

Step 5: Triple extraction based on relation semantic template. Based on extracting entities and relations, the whole matrix of entity pair-relation is constructed, and the correlation between candidate triplets and relation semantic templates and sentences is calculated through the attention mechanism, and finally, the entity relation triplet is obtained.

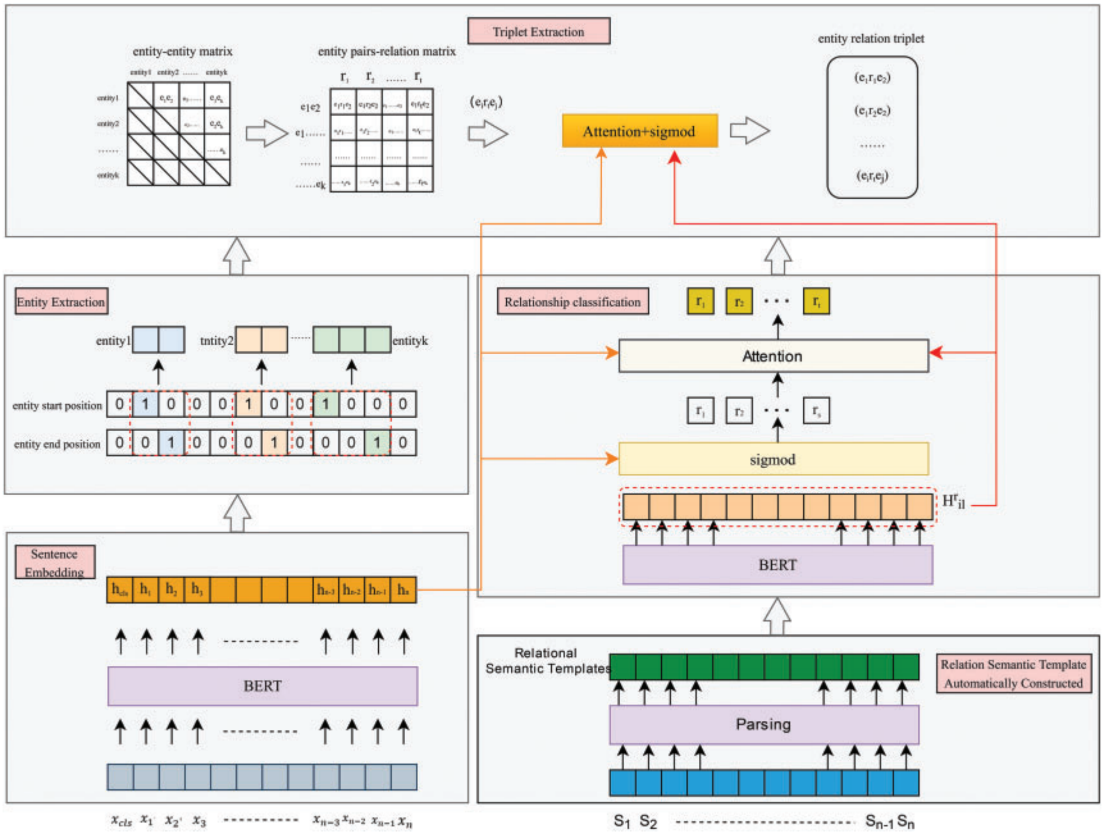

Figure 1: The flowchart of RSTAC model

The above steps are described in detail below.

3.2.1 Automatic Construction of Relation Semantic Template

To avoid the problems of relying on expert knowledge and poor portability in manually defining relation semantic templates, this paper designs automatic extraction rules of relation semantic templates based on syntactic analysis and finally establishes a relation semantic template library for each relation.

To extract the relation semantic template more accurately, considering that syntactic analysis has noise on complex and long sentences, and the relation semantics constructed from simple sentences are more concise and more applicable, this paper takes simple sentences in three datasets as relation corpus (this paper sets the sentence length $\leq$40 as simple sentences), as shown in Fig. 2, and automatically constructs the relation semantic template according to the following steps:

1. Part-of-Speech Tagging

Based on the part-of-speech tagging function of the NLTK toolkit, the part-of-speech tagging of simple sentence $S^{RM} = \{c_1, c_2, \ldots c_n\}$ (c_i is the i-th word in sentence S) in relation corpus is carried out to determine the part-of-speech composition of the sentence, and noun phrases and verb phrases are divided into two categories, which are replaced by np and vp symbols respectively to obtain $S_A^{RM} = \left\{c_1, \ldots np_1, np_2, np_3 \ldots vp_1, \ldots np_j, \ldots vp_k, \ldots c_n\right\}$.

2. Determining Noun Concept Set Based on Location

Due to the limitations of part-of-speech tagging tools and the diversity of nouns in real data, there is a great probability that nouns in adjacent positions are a whole set of noun concepts. For example, “American Apple Computer Inc.” is easily labeled as three different nouns: “American”, “Apple” and “Computer Inc.”. Therefore, based on the position information, the sentence S_A^{RM} combines similar noun phrases into a noun concept set NP to get the sentence $S_B^{RM} = \left\{c_1, \ldots NP_1 \ldots vp_1, \ldots np_j, \ldots vp_k, \ldots c_n\right\}$.

3. Analyzing the Blank Position of Semantic Template of Tag Relation according to Dependency Parsing

As shown in Fig. 3, the purpose of dependency parsing is to identify the dependency between words. Different from syntactic analysis, dependency syntactic analysis holds that there is a subordinate relation between words, that is, dependency relation. Through a large number of observation experiments, this paper proves that the relation trigger words in simple sentences are generally marked as HED (core words) by dependency parsing, and there are SBV (subject-predicate relation) and VOB (verb-object relation) dependencies between entity pairs and HED in most entity relation triplets, so dependency parsing is used to mark the position of entity pairs. The sentence S_A^{RM} is analyzed by using Baidu dependency parsing tool DDParser, and np or NP with the dependency between SBV and VOB in the sentence is determined, and compared with the marked entity pair in the sentence. If it is different, the sentence is abandoned and the next sentence is analyzed. If they are the same, replace np or NP with [MASK] to get the sentence $S_C^{RM} = \left\{c_1, \ldots [MASK] \ldots vp_f, \ldots [MASK], \ldots vp_k, \ldots c_n\right\}$, where vp_f refers to the core word of the sentence.

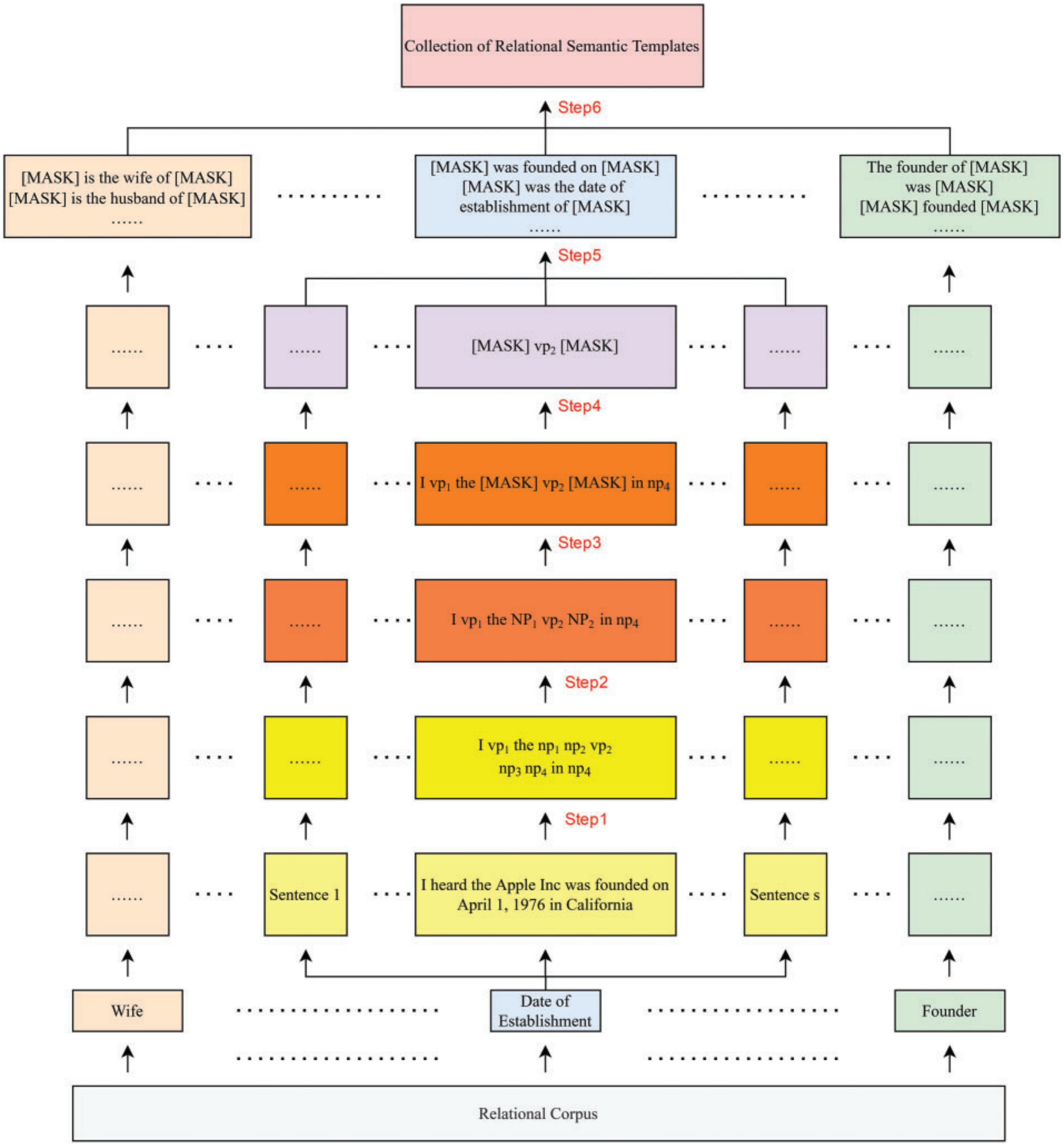
Collection of Relational Semantic Templates
Step6
[MASK] is the wife of [MASK]
[MASK] is the husband of [MASK]
......
[MASK] was founded on [MASK]
[MASK] was the date of establishment of [MASK]
......
The founder of [MASK] was [MASK]
[MASK] founded [MASK]
......
Step5
[MASK] vp_2 [MASK]
Step4
I vp_1 the [MASK] vp_2 [MASK] in np_4
Step3
I vp_1 the NP_1 vp_2 NP_2 in np_4
Step2
I vp_1 the np_1 np_2 vp_2 np_3 np_4 in np_4
Step1
Sentence 1
I heard the Apple Inc was founded on April 1, 1976 in California
Sentence s
Wife
Date of Establishment
Founder
Relational Corpus

Figure 2: The flowchart of the relation semantic template automatically constructed

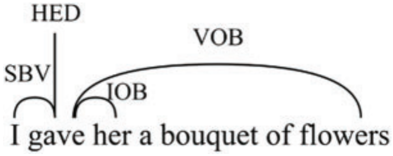

Figure 3: Dependency parsing

4. Pruning Based on Blank Marks

To avoid the noise influence of irrelevant words on the relation semantic information, and to refine the relation semantic information as much as possible, according to the position information, the words in the sentence S_C^{RM} that are not between [MSAK] and HED are pruned to get the relation semantic template $R^m = \{[MASK], \ldots vp_f, \ldots [MASK]\}$, and the R^m is classified into the relation semantic template library of this relation. By analogy, the relation semantic template library $R^M = \{R_1^m, R_2^m, \ldots R_f^m\}$ $(f \leq q)$ containing all relation semantic templates of the relation is finally obtained, in which R_g^m is the g-th relation semantic template in the relation semantic template library R^M.

5. Relation Semantic Template Filtering

Due to the diversity of natural language expressions and the diversity of relation semantic information, R^M contains many similar relation semantic templates. According to the syntactic analysis function of the NLTK toolkit, only one similar relation semantic template with syntactic similarity greater than the threshold is kept in R^M, and the relation semantic template library $R^{M'}$ is obtained.

6. Constructing a Relation Semantic Template Based on TF-IDF

According to the results of dependency parsing, most core words are relation triggers, but because of the existence of synonyms, relation triggers are different. In this paper, based on TF-IDF [28], the frequency of the core words in the relation corpus is counted for each relation semantic template, and the core words whose word frequency is higher than the threshold are regarded as the standard core words of the relation semantic template, and finally the relation semantic template library $R_i^{M'} = \{R_{i1}^m, R_{i2}^m, \ldots R_{if}^m\}$ $(f \leq q)$ containing relation r_i $(1 \leq i \leq m)$ is obtained, where R_{ig}^m is the g-th relation semantic template in the relation semantic template library of relation r_i.

Examples of relation semantic templates are shown in Table 1.

Table 1: Examples of relation semantic templates

Graduation	Founder	Wife
[MASK] was admitted to [MASK]	The founder of [MASK] was [MASK]	[MASK] is the wife of [MASK]
[MASK] graduated from [MASK]	[MASK] founded [MASK]	[MASK] is the husband of [MASK]
[MASK] studied at [MASK]	[MASK] establish [MASK]	[MASK] married [MASK]

As can be seen from the example in Table 1, the template of each relation is different, there is no fixed template, and each template is related to how the relation is expressed in reality. Since the template is mainly used to reflect the fine-grained semantics of the relation, the structure of the template has nothing to do with the language used in the sentence, and since the method designed in this paper to automatically construct the relation semantic template is based on syntactic dependency analysis, the construction of the template is also independent of the tense, only related to the syntactic structure.

Table 2 shows the algorithm for automatically building a relational semantic template.

Table 2: Algorithms for automatic construction of relation semantic template

Algorithms for automatic construction of relation semantic template
Input: relation corpus
Output: relation semantic templates
1. Use the NLTK toolkit to mark word parts of speech on all sentences of the same relation in the relation corpus.
2. Depending on the position of the word, similar nouns are combined into a unified set of nouns.
3. The Baidu dependency syntactic analysis tool DDParser was used to analyze the sentences and mark the dependencies of nouns.
4. Determine which two nouns in the sentence have SBV and VOB dependencies with the core word HED.
5. Use [MASK] to replace the two nouns identified in the previous step.
6. Depending on the word position, all words that are not between the two [MASK] are removed.
7. Output the results to a specific collection of the relation.
8. Repeat this process until all sentences containing this relation have been entered.
9. The NLTK toolkit is used to perform syntactic analysis of all results in the relation set, and all results above the threshold are removed.
10. According to TF-IDF and syntactic dependency analysis, the HED with the most occurrences in sentences containing the same relationship is determined, and the word frequency above the threshold is determined as the relationship trigger word.

3.2.2 Sentence Coding

In the part of sentence coding, the feature information H_e is extracted from sentence S, which is the basis of subsequent entity recognition and relation classification. To extract the feature information in sentence S more fully, this paper chooses the pre-trained BERT model as the encoder.

The BERT model is a depth model trained in a large number of unlabeled datasets, which can significantly improve the accuracy of various natural language processing tasks. It uses the encoder structure of the Transformer as the feature extractor, which can realize the bidirectional encoding of the input sequence text. Compared with the one-way encoder which uses the pre-order text information to extract semantics, the BERT model has stronger semantic information extraction ability.

Given the above advantages, this paper chooses the pre-trained BERT model to encode sentences.

For a sentence with a given word length of n, add the [CLS] identifier at the beginning of the sentence to get the input S' of the model as follows:

$$S' = \{[CLS], c_1,\ c_2, \ldots, c_n\} \tag{1}$$

where [CLS] stands for classification identifier and c_i stands for the i-th word in the sentence S. The sentence S is encoded by the BERT model, and the expression vector H_e of the sentence is obtained as follows:

$$H_e = BERT\left(S'\right) \tag{2}$$

$$H_e = \{h_{cls}, h_1, h_2, h_3 \ldots h_n\} \tag{3}$$

where h_{cls} is the vector representation of the [CLS] classification identifier encoded by the BERT model, and h_i is the encoded vector representation of the word c_i in the sentence after the BERT model. Through Eq. (2), the representation vector of the sentence S encoded by BERT is obtained.

3.2.3 Entity Identification

This step is mainly based on the sentence coding vector H_e obtained in the previous step, and uses the cascade labeling strategy [22] to label the entity. Cascade's labeling strategy skillfully transforms the problem of entity recognition into the problem of labeling, which improves the accuracy of entity recognition. According to this strategy, two binary classifiers are used for the starting position and the ending position of the entity, and one classifier is named entity-start to mark the starting position of the entity. This classifier assigns a binary label to each word in the sentence, and if the word is considered as the starting position word of the entity, the corresponding label position of the word is marked as 1. Otherwise, it is marked as 0. Correspondingly, the binary classifier used to mark the end position of the entity is named entity-end, and the end position of the entity is marked by the same marking method.

$$p_i^{entity-start} = \sigma\left(W_{entity-start} h_i + b_{entity-start}\right) \tag{4}$$

$$p_i^{entity-end} = \sigma\left(W_{entity-end} h_i + b_{entity-end}\right) \tag{5}$$

where $p_i^{entity-start}$ and $p_i^{entity-end}$ represent the probability that the i-th word in the sentence is the start position and end position of the entity, respectively. If this probability is greater than a certain threshold, the label corresponding to its word is marked as 1, otherwise, it is marked as 0. h_i is the coded representation vector of the i-th word in the sentence, $W_{entity-end}$ is the trainable parameter, $b_{entity-end}$ is the offset vector, and they have the same dimension, and σ is the sigmod activation function. So, through Eq. (4) we can determine whether a word is the starting position of an entity, and through Eq. (5), we can determine whether a word is the end position of an entity.

3.2.4 Relation Classification Based on Relation Semantic Template

1. Relation Pre-Classification

Because a sentence may contain many different relations, and different sentences contain different numbers of relations, this paper models the relation classification task as a multi-label two-classification task. The relations contained in a sentence are classified by using the sentence coding vector H_e, and whether the relation is contained in a sentence is judged by assigning a label to each predefined relation.

$$H_e^{'} = avg\left\{h_{cls}, h_1, \ldots, h_n\right\} \tag{6}$$

$$p_r = \sigma(W_r \cdot H_e^{'} + b_r) \tag{7}$$

Among them, $H_e^{'}$ stands for the average expression vector of the sentence obtained by averaging all the word vectors in the sentence S, W_r is the training parameter, b_r is the bias vector, and p_r is the probability that the sentence contains a certain relation, which is similar to the entity identification step. When this probability is greater than a certain threshold, the relation label is marked as 1, otherwise, it is marked as 0, and then the candidate relation set $R_s = \{r_1, r_2, \ldots r_s\}$. The relation r_i in R_s indicates that the relation which label is assigned as 1 after this step. Through Eq. (6) we can get the average representation vector of the sentence, and Eq. (7) allows us to determine whether the sentence contains a relationship based on Eq. (6).

To screen the relation types contained in the sentence more accurately and reduce the subsequent consumption of the model, this paper aims at the candidate relation set R_s, combines the rich relation semantic information contained in the relation semantic template, and judges the correlation between each candidate relation R_s and the sentence through the attention mechanism, and finally gets the relation type set $R_r = \{r_1, r_2 \ldots r_t\}$ closely connected with the sentence.

2. Relation Semantic Template Coding

Firstly, the BERT model is used to encode the relation semantic template, in which the input of the l-th ($l \le q$) relation semantic template under the relation r_i can be expressed as Eq. (8).

$$T^r_{il} = \left(m^r_{i1}, t^r_{i1} \ldots m^r_{i2}, \ldots t^r_{in}\right) \tag{8}$$

where m^r_{i1} represents the blank of the head entity, m^r_{i2} represents the blank of the tail entity, and t^r_{in} is other words in the relation semantic template. After the input is encoded by the BERT model, the l-th relation semantic template encoding under the relation r_i is obtained to represent H^r_{il}, which contains some of the deep semantic information about the relation r_i.

$$H^r_{il} = BERT\left(T^r_{il}\right) \tag{9}$$

$$H^r_{il} = \{h_{m1}, h_{m2}, h_{m3}, h_{m4} \ldots h_{mn+2}\} \tag{10}$$

$$H^r_i = avg\left\{H^r_{i1}, H^r_{i2}, \ldots H^r_{il}, \ldots H^r_{io}\right\} (o \le q) \tag{11}$$

where H^r_{il} is the coded representation of the l-th relation semantic template under the relation r_i after being coded by the BERT model, h_{mi} is the feature representation of the i-th word in the relation semantic template, and H^r_i is the coded representation after performing a mean pooling operation on the feature vectors of all relation semantic templates to which the relation r_i belongs. After Eqs. (9)–(11), we can obtain the representation vector of all the relation semantic templates to which the relation r_i belongs, according to which a relation can be filtered in detail.

3. Relation Classification Filtering by Fusing Relation Semantic Templates

Then the weight value of the candidate relation in the sentence and relation semantic templates is obtained through the attention mechanism [29], and the Equation of the weight value is as follows:

$$e_{ik} = v^T tanh\left(W_r H^r_i + W_H H^{'}_e + W_h h_i\right) \tag{12}$$

$$\alpha_{ik} = \frac{exp\left(e_{ik}\right)}{\sum_{j=1}^{n} exp\left(e_{jk}\right)} \tag{13}$$

v, W_r, W_H and W_h are trainable parameters. In this way, the attention score can measure the relevance of each candidate's relation to the sentence and relation semantic templates. That is, with Eqs. (12)–(13), we get the attention score α_{ik}, and the score of the candidate relation can determine whether it is relevant to the sentence and relation semantic templates or not.

$$H^{r'}_i = \alpha_{ik}\left(W^r H^r_i\right) \tag{14}$$

$$p_r^{'} = \sigma\left(W_r^{'} \cdot H_i^{r'} + b_r^{'}\right) \tag{15}$$

where W^r and $W_r^{'}$ are trainable parameters, $b_r^{'}$ is the bias vector, $H_i^{r'}$ is the coded representation of the relation r_i obtained through the attention mechanism, and $p_r^{'}$ is the probability that the sentence contains a certain relation. If the probability is greater than the threshold, the relation is assigned a value of 1, otherwise, it is 0. After the relation semantic template Eq. (14), we obtain the relation embedding related to both the sentence and the relation semantic template, and on this basis, according to the Relation semantic template Eq. (15), we can further judge whether the sentence contains this relation.

Through the above steps, the real relation set $R_r = \{r_1, r_2, \ldots r_t\}$ contained in the sentence is obtained.

3.2.5 Triple Extraction Based on Relation Semantic Template

A correct entity relation triplet is bound to have a strong dependence on sentences. In addition, the relation between the head entity and the tail entity within the entity relation triad is also closely related. The relation semantic template constructed by dependency parsing on sentences cannot only be used to judge the degree of close connection within triples but also filter out entity relation triples unrelated to sentence semantics, thus screening out the correct entity relation triples contained in sentences.

Based on this cognition, this step is mainly to construct all the potential triplet sets in the sentence by enumerating and then to judge the close relation between triplets and sentences by using the relation semantic template through the attention mechanism, to get the correct entity relation triplet set contained in the sentence.

1. Establish the Whole Matrix of Entity Pair-Relation

First of all, based on extracting the entity pairs and relations, as shown in the upper left corner of Fig. 1, all possible entity pairs in the sentence are obtained by enumerating through establishing the entity-entity full matrix. Then the whole matrix of entity pairs and relations is constructed, and all the candidate triplet sets in this sentence are obtained by enumeration.

2. Triple Extraction of Semantic Template of Fusion Relation

Through the attention mechanism, according to the relation semantic template, the close relation between the candidate entity relation triplet and the sentence is judged.

$$h_i^{head} = avg\left\{h_1, \ldots h_p\right\} (p \le n) \tag{16}$$

$$h_i^{tail} = avg\{h_1, \ldots h_w\} (w \le n) \tag{17}$$

$$h_i^{entity} = \left[h_i^{head}; H_i^r; h_i^{tail}\right] \tag{18}$$

$$e_{ik}^{'} = v^T tanh\left(W_{entity} h_i^{entity} + W_H H_i^r + W_h H_e\right) \tag{19}$$

$$\alpha_{ik}^{'} = \frac{exp\left(e_{ik}^{'}\right)}{\sum_{j=1}^{n} exp\left(\alpha_{jk}^{'}\right)} \tag{20}$$

where h_i^{head} is the head entity representation vector obtained after the mean value pooling operation for all the word vectors in the head entity, h_i^{tail} is the head entity representation vector obtained after the mean value pooling operation for all the word vectors in the tail entity, and h_i^{entity} is the coded representation obtained by splicing the head entity, the relation, and the tail entity representation vectors as the feature representation of the i-th candidate triplet. W_{entity}, W_H and W_h are trainable parameters. Through the Eqs. (16)–(20), we obtain an attention score $\alpha_{ik}^{'}$, which can reflect the correlation between the candidate entity relation triplet and the relation semantic template and sentence, and judge whether the candidate entity relation triplet has a certain correlation with the relation semantic template and sentence according to the level of this score.

$$H_i^{entity} = \alpha'_{ik}\left(W^{entity} h_i^{entity}\right) \tag{21}$$

$$p_i^{entity} = \sigma(W'_{entity} \cdot H_i^{entity} + b_{entity}) \tag{22}$$

W^{entity} and W'_{entity} are trainable parameters, b_{entity} is the bias vector, H_i^{entity} is the coding representation corresponding to the i-th triplet after using the attention mechanism, and the probability that this triplet is represented by p_i^{entity} is the correct triplet. If the probability is greater than the set threshold, the corresponding position labels are marked as 1, and all of them are marked as 0. Through Eqs. (21), (22), we can finally determine whether this candidate entity relation triplet is the true entity relation triplet contained in the sentence.

Through the above steps, all the entity relations triplets contained in the sentence are finally obtained.

3.2.6 Loss Function

The joint extraction model proposed in this paper is mainly divided into three interrelated parts: entity extraction, relation classification, and triple extraction, so the loss function of the model should include three parts. In this model, the three parts need to be jointly trained and share the parameters of the encoder, so the loss function of the model should be the sum of the loss L_{ent} of entity extraction, the loss L_{rel} of relation classification and the loss L_{er} of triple extraction.

$$L = L_{ent} + L_{rel} + L_{er} \tag{23}$$

4 Experiment

To verify the feasibility and effectiveness of the proposed model, this paper carried out experiments on three Chinese public datasets [30] around the following four aspects and analyzed the experimental results in principle.

1. Correlation analysis between dependency parsing and triples. To fully reflect the rationality of this paper's automatic construction of relation semantic templates based on dependency parsing, this paper analyzes the dependency relation of triplet entity pairs in sentences in three large Chinese datasets: DuIE, SanWen, and FinRE.
2. A comparison between the model and the baseline method is proposed. The RSTAC model proposed in this paper is compared with the existing typical entity relation extraction methods.
3. The extraction effect of the model in a complex context. This paper simulates the phenomenon of entity overlap and long sentences encountered by the entity relation extraction model in reality, screens the relevant data in the data set, and carries out experiments on the performance effect of the model in complex situations.
4. Ablation experiment. To verify whether each part has positive benefits in this model, according to the idea of the ablation experiment, this paper verifies the role of different structures in this model by removing different parts of the model.

4.1 Experimental Setup

4.1.1 Dataset

In this paper, three Chinese entity relation datasets, DuIE, FinRE, and SanWen, are used to train and test the model. DuIE data set is a large-scale Chinese entity relation extraction data set. All the sentences in the data set are from the Baidu Encyclopedia and Baidu news corpus, and the overall stylistic expression is relatively standardized; The SanWen data set contains 837 Chinese essays, and the semantics of sentences are relatively more complicated, including 9 types of relations; FinRE data set is a manually marked data set from Sina Finance and Financial News, which contains 44 types of relations. Due to the particularity of the financial field, there are a large number of domain-specific entities, and the entity boundaries are difficult to divide. Table 3 shows the number of data in the dataset.

Table 3: Dataset information

Dataset	Train	Valid	Test	Relation
DuIE	173108	21639	17178	49
SanWen	17227	1793	2220	9
FinRE	13489	1489	3727	44

4.1.2 Evaluation Indicators

In this paper, the precision (P) commonly used in relation extraction, recall rate (R), and F1 scores as the evaluation index. Only when the head entity, relation type, and tail entity in the extracted triplet are completely consistent with the label, the triplet is judged to be correct. The calculation Relation semantic template is as follows:

$$\mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{24}$$

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{25}$$

$$\mathrm{F1} = \frac{2 \times \mathrm{P} \times \mathrm{R}}{\mathrm{P} + \mathrm{R}} \tag{26}$$

TP represents the number of samples whose positive samples are judged to be positive, *FP* represents the number of samples whose negative samples are judged to be positive, and *FN* represents the number of samples whose positive samples are judged to be negative, where the samples refer to triples.

4.2 Experiment on the Effect of Entity Relation Extraction

4.2.1 Correlation Analysis between Dependency Syntactic Relations and Triples

To fully reflect the rationality of this paper's automatic construction of relation semantic templates based on dependency parsing, this paper analyzes the dependency relation of triplet entity pairs in sentences in three large Chinese datasets: DuIE, SanWen, and FinRE. Because the automatic construction of relation semantic template is based on simple sentences, this paper first preprocesses the three datasets according to the standard that the sentence length is less than or equal to 40, and then uses DDParser, a Baidu dependency parsing tool, to analyze the dependency of entity pairs in triples. The experimental results of the top four dependencies in each dataset are shown in Fig. 4.

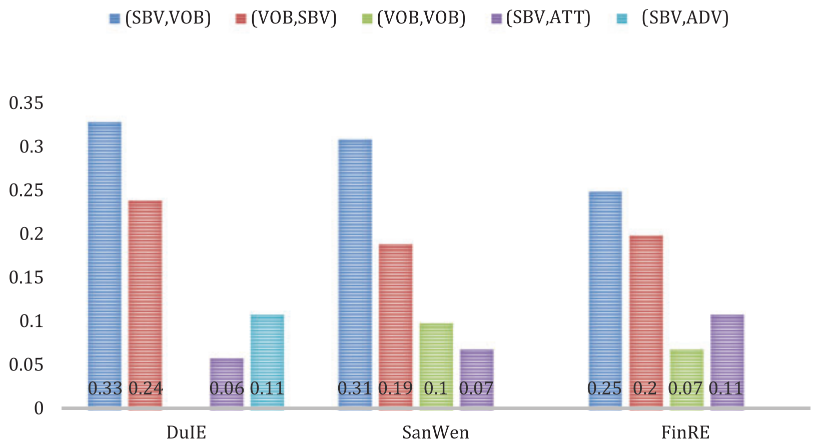

Figure 4: Top four dependencies in each dataset

Take the SanWen dataset, which contains a lot of prose information and has complex semantics, as an example. In the simple sentences of this data set, the number of dependencies between head and tail entities in the sentence is (SBV, VOB), which is 31%, followed by (VOB, SBV), which is 19%, and (VOB, VOB) which is 10%. That is to say, in the SanWen data set, the probability that entities are related to the two dependencies, SBV and VOB, exceeds 60%. This key information guarantees the rationality of using dependency parsing as a semantic template for automatically constructing relations in this paper.

4.2.2 Model Effect

To verify the validity of the model proposed in this paper, some models are selected here to carry out comparative experiments on three different datasets. Among them, DepCasRel [20] and MG Lattice [23] are both Chinese entity relation extraction models that integrate word features; CasRel [22] and TPLinker [23] are classical models for extracting entity relations. ROJER [5] is a relation-oriented extraction model; RFBFN [6] is an entity relation extraction model based on artificially defined relation templates. The basic principles of the six comparison models are as follows:

1. DepCasRel: A Chinese entity relation extraction model based on word features and combining word features, which solves the problem of entity overlap. However, this model ignores the feature that the smallest semantic unit of Chinese is a word, which leads to the Chinese semantic representation level is not optimal.
2. MG Lattice: A joint extraction model that uses an external language knowledge base to fuse word features to solve the ambiguity of Chinese polysemous words. However, the effect of this model depends heavily on the quality of the external language knowledge base, and it has poor performance and robustness for some Chinese words.
3. CasRel: An entity-oriented cascade extraction model. This model models the relation as a function from the head entity to the tail entity, which effectively solves the problem of entity overlap. However, due to the different language structures between Chinese and English, this model has a poor ability to capture Chinese semantics.

4. TPLinker: A joint model of entity relations based on structural prediction, which transforms the joint extraction task into the problem of etymological pair linking, and can solve the overlapping of entities and relations. However, this model pays more attention to the improvement of the model but ignores the difference in input granularity between Chinese and English. This model can not fully express Chinese semantics, and there are also problems with entity recognition errors caused by nested entities.
5. ROJER: A relation-oriented cascade extraction model, which adopts the method of classifying relations first and then extracting entity pairs, effectively taking into account the problems of entity overlap and entity redundancy. However, because the relation is simply integrated into entity identification as a priori knowledge, it is not necessary to use the relation type information to constrain entity identification in depth, so there is an entity identification error caused by the nested entity.
6. RFBFN: A joint extraction model based on artificially defined relation semantic templates, which transforms the entity extraction task into a template filling task and effectively improves the accuracy of entity relation extraction. However, because a single template is defined manually, the multi-level semantics of the relation cannot be completely summarized, so there are shortcomings in the semantic representation of the relation.

The experimental results of the RSTAC model and six comparative models in this paper are shown in Table 4.

Table 4: Experimental comparison of the model with the baseline model

Model	DuIE			SanWen			FinRE		
	P	R	F1	P	R	F1	P	R	F1
DepCasRel	81.2	80.0	80.6	72.1	63.0	67.2	52.3	47.7	49.9
MG Lattice	80.6	79.6	80.1	70.1	61.6	65.6	51.6	47.2	49.3
CasRel	80.1	77.2	78.6	65.4	66.6	66.0	48.9	46.2	47.5
TPLinker	78.4	80.6	79.5	64.2	**68.9**	66.5	48.4	49.0	48.7
ROJER	80.7	76.3	78.4	64.7	65.5	65.1	48.3	45.4	46.8
RFBFN	82.9	81.3	82.1	72.9	68.8	70.8	53.9	48.6	51.1
RSTAC (ours)	**84.9**	**81.8**	**83.3**	**73.8**	68.8	**71.2**	**55.8**	**49.3**	**52.4**

Analysis: It can beseen from Table 4, that the F1 scores of the model in this paper have reached the optimal values on three different datasets, which fully verifies the correctness of the idea of extracting entity relation triples based on the relation semantic template by automatically constructing the relation semantic template. The following is a detailed analysis of the experimental part.

On three completely different Chinese datasets, the DepCasRel model and MG Lattice model optimized for Chinese are better than the CasRel model, TPLinker model, and ROJER model which are not adjusted for Chinese, and the F1 scores are at least 0.6% higher. However, the effects of these two models are still inferior to those of the RFBFN model and RSTAC model, and the F1 scores have decreased by at least 1.2% on the three datasets, which fully proves the positive benefits of using relation semantic templates to extract entity relations. On the other hand, using the relation semantic template to obtain the deep semantic information of the relation also makes the model easily and quickly deployed on the data set of a specific language without special adjustment, which reflects the excellent universality of use. From the experimental results, the F1 scores of the RSTAC model are at least 0.4% higher than that of the RFBFN model on three datasets, which also proves that the automatically constructed relation semantic template can get more comprehensive relation semantic information than the manually constructed relation semantic template.

4.2.3 Experimental Results of the Model in Complex Context

To verify the effect of the model in a complex context, this paper simulates two experiments in a complex context. One is to classify the three datasets according to the fact that the entities are not overlapped, the single entity overlap(SEO), and the entity pair overlap(EPO), and experiments are carried out on this basis; One is to divide the SanWen data set according to the length of sentence words, and divide it into four different sentence length intervals with the length of 40 words, based on which experiments are carried out.

1. Influence of entity overlapping on model

The results F1 scores of the entity overlap experiment on three datasets are shown in Figs. 5–7.

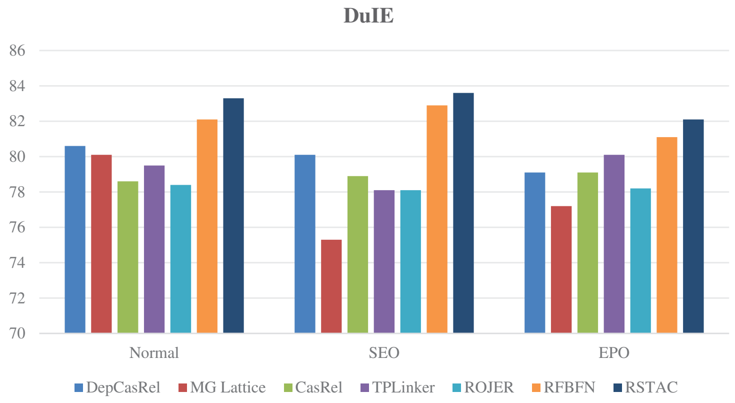

Figure 5: F1 scores on DuIE

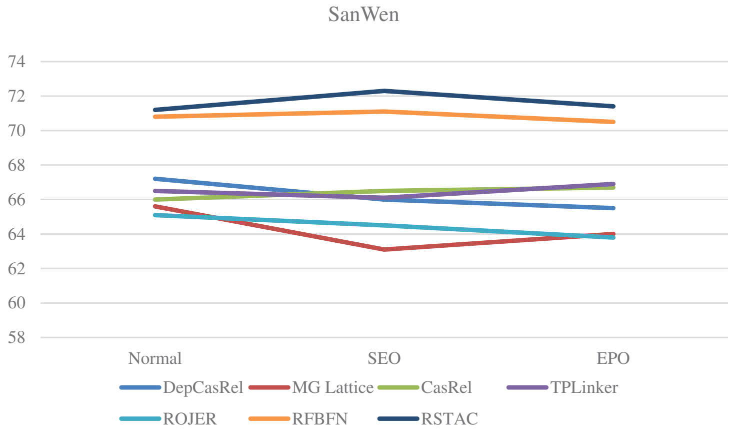

Figure 6: F1 scores on SanWen

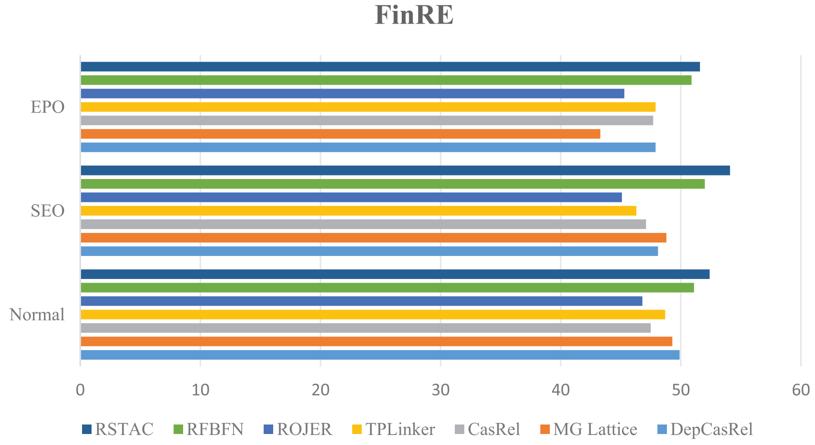

Figure 7: F1 scores on FinRE

2. The influence of sentence length on the model

To reflect the influence of sentence length on the model more intuitively, this paper chooses to experiment on the SanWen data set with more complicated semantic information. The experimental results of F1 scores are shown in Fig. 8.

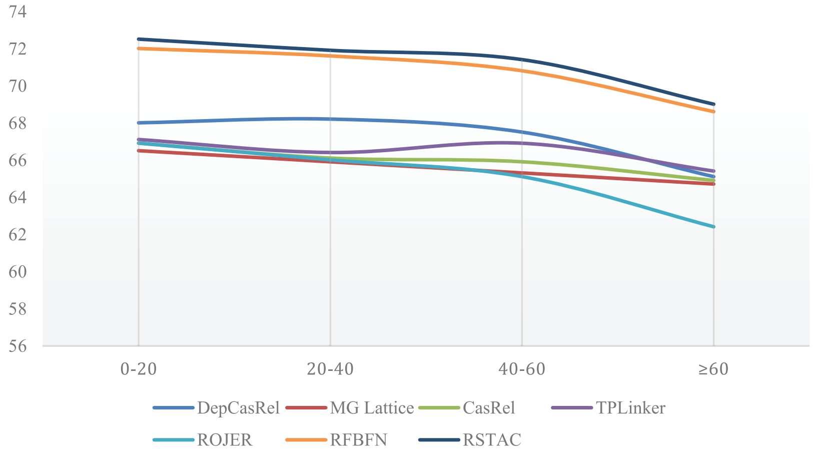

Figure 8: F1 scores of the model under complex sentences

3. Analysis of models in complex situations

From Figs. 6–8, through the above experiments, the model in this paper has achieved the best results in two different complex situations. Through comparative analysis, we can also see that the F1 scores of the RSTAC model and RFBFN model using the relation semantic template are higher than those of other models in two different complex situations, and the downward trend of F1 scores of the RSTAC model and RFBFN model is the slowest compared with the normal situation, which shows that the entity relation extraction model based on relation semantic template is not only more applicable but also verifies that the deep semantic information of relation can help the model improve the accuracy of identifying entity relation triples in complex situations. Compared with the normal situation, the F1 scores gap between the RSTAC model and RFBFN in this paper is more obvious in two different complicated situations and even reaches 2.1% in certain cases, which fully proves that the relation semantic template automatically constructed based on the dataset covers more abundant relation semantic information than the manually defined relation semantic template.

4.2.4 Ablation Experiment

The general idea of the ROJER model is to use the BERT model to encode the sentence, then classify the relation, splice the corresponding encoding of the relation to the sentence vector, extract the entity on this basis, and then obtain the triplet. Therefore, to verify the positive impact of each part of the model on the model, this paper takes the ROJER model as the baseline model and adds four parts to the ROJER model on three datasets according to the idea of the ablation experiment.

The first is to add the relation semantic template code, denoted ROJER+tem, to verify whether deeper relational semantic information can be obtained through the relation semantic template; The second is to add a relation classification part based on the relation semantic template, which is recorded as ROJER+rel, to verify whether this part can improve the effect of relation classification; The third is to add the triplet extraction part based on the relation semantic template, which is denoted as ROJER+ent-rel, to verify whether the internal connection of the triplet extraction can be enhanced through the relation semantic template; The fourth is to add the relation classification based on the relation semantic template and the triplet extraction part based on the relation semantic template, which are recorded as ROJER+rel and +ent-rel to verify whether the deep semantics of the relation have a positive impact on the model.

Ablation experiment results, as shown in the figure.

Analysis: We can get from Fig. 9 that the final improved ROJER+rel, +ent-rel models have a significantly improved effect compared with the original ROJER model, especially on the DuIE dataset, the F1 scores are increased by 4.8%, which fully verifies the positive impact of the three improvements proposed in this paper on the model.

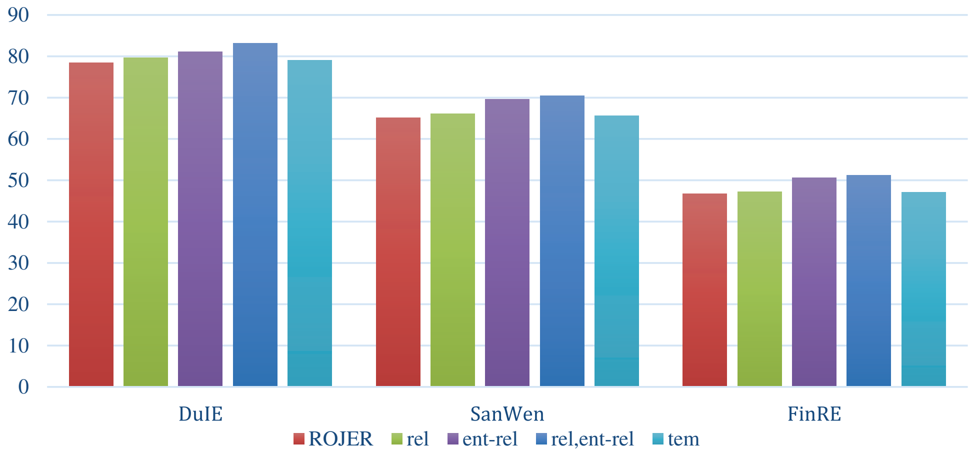

Figure 9: F1 scores for ablation experiments under different conditions

From the results of the ROJER+tem model and ROJER model that only added the relation semantic template encoding, the F1 scores on the three datasets increased by an average of 0.75%, indicating that encoding the relation semantic template can obtain deeper relational semantic information than the simple way of encoding the relation.

The results of the ROJER+rel model are analyzed and the F1 scores are improved by an average of 1.3% compared with the ROJER model, which proves that the relation semantic template can successfully constrain the relation classification process. Since this section includes steps to encode relational semantics, it is clear that this part is better than the ROJER+tem model.

By comparing the results of the four ablation experiments separately, it can be found that after adding the triplet extraction part based on the relation semantic template (ROJER+ent-rel model), the change of F1 scores is the largest, which proves that the relation semantic template can not only obtain the deep expression of relational semantic information but also reflect the dependence between the triples of entity relation from the side. At the same time, it also proves the correctness of the idea of automatic construction of entity relation extraction based on the relation semantic template designed in this paper.

5 Summarize

From the basic form of the triplet <head entity, relation, tail entity>, we can infer that the entities and relations within the triplet should be closely related, and the relation plays the role of connecting the top and the bottom. However, the existing model focuses too much on sentence and entity features such as syntax and entity position in feature utilization, and ignores relation features. A small number of relation-oriented extraction models for the study of relation characteristics mostly stay on the shallow semantic information of the relation, and cannot obtain the deep semantic information of the relation. Although the artificial design of the relation semantic template can obtain the deep semantic information of the relation, due to the diversity of natural language expressions and the factors of artificial participation, this method still has problems such as relying on expert experience, poor portability, and incomplete acquisition of the deep semantic information of the relation. To solve this problem, this paper determines the focus of research on the automatic construction of relation semantic templates and the strengthening of inter-triplet connections through relation. On this basis, this paper proposes a joint entity relation extraction model based on a relation semantic template automatically constructed, RSTAC. The model refines the extraction rules of the relation semantic template in the relational corpus according to the dependency syntax analysis and automatically constructs the relation semantic template. Encode sentences for entity recognition; Based on the relation semantic template, relation classification and triplet recognition are carried out based on the attention mechanism. Experimental results show that the model effectively mines the deep semantic information of the relation, and the F1 scores are increased by an average of 0.96% compared with the baseline models such as CasRel and TPLinker on the three major Chinese datasets of DuIE, SanWen, and FinRE, which effectively improves the extraction effect of entity relation triples.

Since the relation semantic template comes from the summary of the language pattern in relation expectations, the more relations are expected for this model, the better, and the external language knowledge base is rich in relation expectations, and the next step we will carry out further research on how to introduce an external knowledge base to enrich the types of relation semantic templates.

Acknowledgement: None.

Funding Statement: This work was supported by the National Natural Science Foundation of China (Nos. U1804263, U1736214, 62172435) and the Zhongyuan Science and Technology Innovation Leading Talent Project (No. 214200510019).

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Wei Liu, Meijuan Yin; data collection: Jialong Zhang; analysis and interpretation of results: Wei Liu; draft manuscript preparation: Wei Liu, Jialong Zhang, Lunchong Cui. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: This statement should make clear how readers can access the data used in the study and explain why any unavailable data cannot be released.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] D. P. Xie and Q. Chang, "Review of relation extraction," *Application Research of Computer*, vol. 37, no. 7, pp. 1921–1924, 2020.

[2] E. H. H., W. J. Zhang, S. Q. Xiao, R. Cheng, Y. X. Hu *et al.,* "Survey of entity relationship extraction based on deep learning," *Journal of Software*, vol. 30, no. 6, pp. 1793–1818, 2019.

[3] C. Z. Zhuang, X. L. Jin, W. J. Zhu, J. W. Liu, L. Bai *et al.,* "Deep learning based relation extraction: A survey," *Journal of Chinese Information Processing*, vol. 33, no. 12, pp. 1–18, 2019.

[4] J. H. Wang, L. Lu, Z. L. Duan, L. Zhang and Y. K. Wang, "Relationship-oriented entity relationship extraction method combining dependent information," *Application Research of Computer*, vol. 40, no. 5, pp. 1410–1415+1440, 2023 (In Chinese).

[5] H. W. Hu, M. J. Yin and X. N. Liu, "A relation-oriented method for joint entity and relation extraction," in *Proc. of EITCE*, Xiamen, China, pp. 951–955, 2021.

[6] Z. Li, L. Y. Fu, X. B. Wang, H. S. Zhang and C. H. Zhou, "RFBFN: A relation-first blank filling network for joint relational triple extraction," in *Proc. of ACL*, Dublin, Ireland, pp. 10–20, 2022.

[7] H. Cunningham, D. Maynard, K. Bontcheva and V. Tablan, "GATE: An architecture for development of robust HLT applications," in *Proc. of ACL*, Philadelphia, Pennsylvania, USA, pp. 168–175, 2002.

[8] C. Wen, Z. X. Shi and Y. Xin, "Chinese non-taxonomic relation extraction based on extended association rule," *Computer Engineering*, vol. 35, no. 24, pp. 63–65, 2009.

[9] H. W. Hu, "Research on entity and relation extraction technology based on deep learning," M.S. dissertation, Information Engineering University, China, 2021.

[10] B. Rink and S. Harabagiu, "UTD: Classifying semantic relations by combining lexical and semantic resources," in *Proc. of SemEval*, Uppsala, Sweden, pp. 256–259, 2010.

[11] R. C. Bunescu and R. J. Mooney, "A shortest path dependency kernel for relation extraction," in *Proc. of EMNLP*, Vancouver, Canada, pp. 724–731, 2005.

[12] X. Han, T. Gao, Y. Lin, H. Peng, Y. L. Yang *et al.,* "More data, more relations, more context and more openness: A review and outlook for relation extraction," in *Proc. of AACL-IJCNLP*, Suzhou, China, pp. 745–758, 2020.

[13] M. E. Z. N. Kambar, A. Esmaeilzadeh and M. Heidari, "A survey on deep learning techniques for joint named entities and relation extraction," in *Proc. of AIIoT*, Seattle, WA, USA, pp. 218–224, 2022.

[14] Z. Nasar, S. W. Jaffry and M. K. Malik, "Named entity recognition and relation extraction: State-of-the-art," *ACM Computing Surveys*, vol. 54, no. 1, pp. 1–39, 2021.

[15] C. Y. Liu, W. B. Sun and W. H. Chao, "Convolution neural network for relation extraction," in *Proc. of ADMA*, Heidelberg, Berlin, Germany, pp. 231–242, 2013.

[16] R. Socher, B. Huval, C. D. Manning and A. Y. Ng, "Semantic compositionality through recursive matrix-vector spaces," in *Proc. of EMNLP-CoNLL*, Jeju Island, Korea, pp. 1201–1211, 2012.

[17] D. Sui, X. Zeng, Y. Chen, K. Liu and J. Zhao, "Joint entity and relation extraction with set prediction networks," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–12, 2023. https://doi.org/10.1109/TNNLS.2023.3264735

[18] K. Ding, S. Liu, Y. Zhang, H. Zhang, X. Zhang *et al.,* "A knowledge-enriched and span-based network for joint entity and relation extraction," *Computers, Materials & Continua*, vol. 68, no. 1, pp. 377–389, 2021.

[19] B. Ji, H. Xu, J. Yu, S. Li, J. Ma *et al.,* "A two-phase paradigm for joint entity-relation extraction," *Computers, Materials & Continua*, vol. 74, no. 1, pp. 1303–1318, 2023.

[20] M. Tuo, W. Yang, F. Wei and Q. Dai, "A novel chinese overlapping entity relation extraction model using word-label based on cascade binary tagging," *Electronics*, vol. 12, no. 4, pp. 1013, 2023.

[21] M. Miwa and M. Bansal, "End-to-end relation extraction using LSTMs on sequences and tree structures," in *Proc. of ACL*, Berlin, Germany, pp. 1105–1116, 2016.

[22] Z. Wei, J. Su and Y. Wang, "A novel cascade binary tagging framework for relational triple extraction," in *Proc. of ACL*, pp. 1476–1488, 2020.

[23] Z. Li, N. Ding, Z. Liu, H. Zheng and Y. Shen, "Chinese relation extraction with multi-grained information and external linguistic knowledge," in *Proc. of ACL*, Florence, Italy, pp. 4377–4386, 2019.

[24] Y. Wang, B. Yu, Y. Zhang, T. Liu, H. Zhu *et al.,* "TPLinker: Single-stage joint extraction of entities and relations through token pair linking," in *Proc. of COLING*, pp. 1572–1582, 2020.

[25] R. Takanobu, T. Zhang, J. Liu and M. Huang, "A hierarchical frame-work for relation extraction with reinforcement learning," in *Proc. of AAAI*, Hawaii, USA, pp. 7072–7079, 2019.

[26] H. Zheng, R. Wen, X. Chen, Y. Yang, Y. Zhang *et al.,* "PRGC: Potential relation and global correspondence based joint relational triple extraction," in *Proc. of ACL*, Bangkok, Thailand, pp. 6225–6235, 2021.

[27] J. Devlin, M. W. Chang, K. Lee and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. of NAACL*, Minneapolis, Minnesota, USA, pp. 4171–4186, 2019.

[28] A. Jalilifard, V. F. Caridá, A. F. Mansano, R. S. Cristo and F. P. C. Fonseca, "Semantic sensitive TF-IDF to determine word relevance in documents," in *Proc. of CoCoNet*, Singapore, Springer, pp. 327–337, 2021.

[29] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones *et al.,* "Attention is all you need," in *Proc. of NIPS'17*, New York, USA, pp. 6000–6010, 2017.

[30] D. Rao and R. Li, "Chinese entity relation extraction method based on schema enhancement," *Software Guide*, vol. 22, no. 2, pp. 47–52, 2023.