Research on Interpretation of Nominal Compound

Weiguang Qu[1,2] ,Rubing Dai[1],Taizhong Wu[1],Min Gu[1],Yanhui Gu[1],Junsheng Zhou[1]

1. School of Computer Science and Technology,
Nanjing Normal University
Nanjing, Jiang Su, P.R.China

2. Chinese Language and Literature,
Nanjing Normal University
Nanjing, Jiang Su, P.R.China

***Abstract*—Nominal compounds which constituted of two nouns together are very common in reading materials or web pages. The interpretation of these compounds can help us know the meaning of a text or sentences. Traditional approaches utilized the method based on verbs and rules to obtain the interpretation of compounds with low recall. So we investigate an interpretation method based on similarity which makes use of the interpretation templates and similar words to achieve the automatic interpretation. Experimental results show that our method can interpret these nominal compounds with a relatively high precision (84.91%), give an increase of 10.48% in recall than the general method, which contributes to the overall nominal compound recall improvement significantly.**

Keywords—Nominal Compound; Automatic Interpretation; Semantic Similarity

I. INTRODUCTION

With the rapid development of science and technology, a large number of new concepts, such as "卫星导航系统" (satellite navigation system), "网络警察" (network police), and "产业结构" (industrial structure), are emerging from the noun-noun combinations. These kinds of combinations belong to Nominal Compound (NC), which is the combination of noun and noun together, neither "的" (of), "和" (and) and other function words nor comma, spacers or other punctuations in them [1].

In Chinese, the nominal compound structure has been widely used in various kinds of texts. Zhu [2] has discovered 14249 nominal compound phrases in the corpus of ten thousand Chinese annotated sentences. Therefore research on nominal compound is of great significance in many fields of natural language processing, such as information retrieval, machine translation, automatic question and answer and so on.

There are mainly two ideas on English nominal compound automatic interpretation, namely, top-down strategy and bottom-up strategy. Top-down strategy first needs to define a clear set of relations and then assigns appropriate semantic relations for each nominal compound. But the sets of semantic relations proposed by different researchers are different [3-5]. In contrast, scholars who using bottom-up strategy believe that a fixed set of semantic relations are difficult to reflect this feature for the reason that they can't exhaust all nominal compounds which sometimes have more meanings. Hence scholars prefer to adopt an open approach without prior definition of the semantic relation set rather than find implicit semantic relations of nominal compound and interpret them through some certain modes based on large-scale corpus [6].

On Chinese nominal compound, the precision of current study has reached 94.2%, but the recall is relative low [7]. So we present an automatic interpretation approach which includes noun matching strategy based on *Cilin* and acquisition strategy of interpretation templates to improve the low recall.

II. RELATED WORK

For the interpretation of nominal compound, some external resources are frequently used in related research, such as Li et al. [8] obtains semantic knowledge from Google n-grams. In addition, the top-down strategy is applied to determine the predefined semantic relation category, for instance Su et al. [9] designs an automatic interpretation method on the basis of defined semantic relation classification.

While in the research of Chinese nominal compound, Wu [10] discusses the combination rules of N+N nominal compound in contemporary Chinese for the first time from the aspects of syntax structure, semantic structure as well as rhetoric and sums up some common combination patterns, which lay a foundation for the interpretation of the Chinese nominal compound.

It has been shown that most work on Chinese nominal compound is mainly based on the "predicate implicit" theory [11]. In the modifier-core structure which "的"(of) structure modifies the noun, some "certain predicate relations" exists between the noun of "的"(of) structure and the head noun, however these predicates that represent the transitive relation does not appear and hence we have to extend the interpretation by regaining the predicate between two nouns in the nominal compound to obtain the noun phrase. Zhou [12] makes a syntactic analysis on Chinese nominal compound, utilizing semantic grammar and downgraded predication structure theory with the purpose of extracting verb implicated between the N+N nominal compounds, finally to achieve the interpretation of it. Similarly, Wang et al. [13] adopts a bottom-up strategy with Chinese sketch engine [1] to search "implicit verb" for automatic interpretation, and then sorts the results in terms of hitting counts in order to filter the interpretation results of nominal compound. This study tested 391 nominal compounds and the precision has reached to 90% when selecting the top 3 interpretations of the search engine, yet the number and category of defined interpretation templates are limited.

Chung studied Chinese noun-noun compounds and found

[1] http://wordsketch.ling.sinica.edu.tw/

that the two components(N1 and N2) linked either by semantic roles assigned by events (complex relations) or by static relations (simple relations). Based on it, several mappings to nine productive N2 categories are applied with moderate success[18]. Marius Pasca used interpretation patterns as query sessions, and then used a mapping between noun phrases to interpretations and a mapping via modifier variants to get candidate interpretations[19].

Wei [7] designs the approach of combination with top-down and bottom-up strategy to interpret nominal compound based on the interpretation rules. This study counts large numbers of nominal compounds, summarizes different semantic combination patterns and interpretation templates, points out that the predicate implied in nominal compound of modifier-core structure is generally the telic role or agentive role composed of noun. In addition, Noun Knowledge Base referred in this paper is established based on the *Semantic Knowledge-base of Modern Chinese* (SKCC) [14] and *HowNet* [15]. When given a nominal compound to be interpreted, we can rely on the Knowledge Base to find the corresponding semantic combinations and match the related interpretation templates, which are beneficial to search the implicit verb needed, ultimately generate interpretation of the phrase. The test generates 156 interpretation phrases in 245 nominal compounds with a relatively high precision (94.2%), while giving the recall about 60%. The main reasons for the low recall are the insufficient of the interpretation templates and the existence of Out Of Vocabulary (OOV).

We propose an automatic interpretation strategy of nominal compound which consist of three steps: First, we need to build a nominal compound interpretation database, including the basic information of nouns (such as semantic categories, collocated verbs, etc.) and information of interpretation templates (containing rules of nouns semantic combination and corresponding interpretation templates). Second, we get corresponding semantic category of nouns and the collocated verbs through the interpretation database, and then combine semantic category information of nouns to obtain the interpretation templates. Ultimately the implied verb (i.e., corresponding collocated verb) will be filled in the corresponding position of interpretation template in order to get final nominal compound results.

Due to the low recall of nouns acquired by direct matching method and the low recall of nominal compound interpretation, we utilize Cilin as an external semantic resource for those nouns that can't be obtained by directly matching to get the synonyms or similar semantic category information, which contributes to achieve the correct interpretation. Thus we develop the Automatic Interpretation Strategy of Nominal Compound based on Cilin (AISNCC).

We also find that the enhanced effect in recall is limited by examining the interpretation results of AISNCC (the test set will be discussed in section IV in detail). In this test, 11.04% of the nominal compound can obtain the information of the noun, but without the interpretation template of semantic category combination. So we design the algorithm of interpretation template acquisition based on calculating the similarity of nominal compound to be interpreted with annotated ones, and then select the interpretation template of annotated data whose similarity is higher than the threshold as the template of nominal compound to be interpreted. We incorporate the interpretation template acquisition algorithm based on similarity into AISNCC and the experiment proves that it has a significant effect on the recall of nominal compound interpretation. Therefore, with the combination of the noun information acquisition algorithm based on similarity and interpretation template acquisition algorithm, we put forward an Automatic Interpretation Strategy of Nominal Compound based on Similarity (AISNCS).

III. Interpretation Strategy of Nominal Compound Based on Similarity

Investigations show that the modifier-core relation in semantic relations occupies the largest proportion, so this paper chooses this structure of nominal compound as our research object.

Based on the predicate implicit theory and the strategy of similarity calculation, we present the AISNCS through AISNCC to get synonyms as noun expansion information and find the implied verb utilizing the most similar annotated nominal compound in order to generate final interpretation phrases. The overall flow chart is shown in Figure 1.

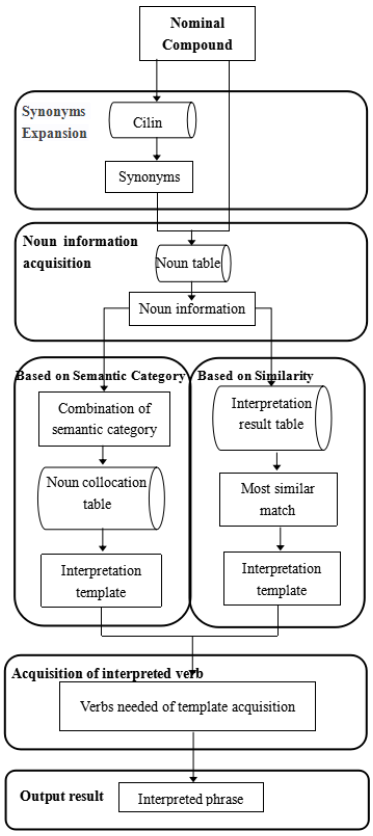

Figure.1: Flow chart of automatic interpretation strategy

In the strategy above, Cilin is treated as an external semantic resource used for the nouns expand matching and similarity computing. On the basis of Cilin, with the code for corresponding entries of words, we can calculate the similarity according to the semantic distance of two entries.

Cilin takes the hierarchical classification system with the five layers code, in terms of the tree structure to encode nouns. For example, "Al02B01= 专家 专门家 家 师 学者 大方 大家 土专家 学家" (the expert and its synonyms), specific coding table as shown in Table 1.

TABLE I: CODING FORMAT TABLE OF CILIN

Number	1	2	3	4	5	6	7	8
Symbol	A	l	0	2	B	0	2	=
Property of symbol	Class	Division	Section		Word chunk	Word chunk of atoms		
Level	First	Second	Third		Fourth	Fifth		

The 8 bits can be "=", "@" or "#", where, "=" represents words are equal, that is synonym; "@" means words have neither synonyms nor related words; "#" indicates words are unequal but belong to the same class, which are near-synonyms.

Based on the Generative Lexicon Theory [16], the noun information acquisition module uses the noun table included the Part Of Speech (POS), semantic category, all the telic roles and agentive roles of nouns. The verbs that are implied in the nominal compound are generally either the telic roles or agentive roles [7]. Among them, the telic role is connected to the function and usage of the object while the agentive role is used to describe how the object is formed or produced. The examples are shown in Table 2:

TABLE II: EXAMPLE OF NOUN TABLE

Noun	**POS**	**Semantic Category**	**Agentive role**	**Telic role**
教师 (teacher)	n	职业 (occupation)	(none)	教 (teach)
病人 (patient)	n	身份 (identity)	罹患 (suffer)	医治 (treat)

"Teach" is the teacher's telic role and "suffer" is the patient's agentive role. These nouns are common both in *SKCC* and *HowNet* and the basic semantic information such as semantic category is from *SKCC*, while telic role and agentive role come from *HowNet*.

Besides, the semantic category combination information and the corresponding interpretation template are stored in the noun collocation table where the template acquisition module based on the semantic category, with 433 semantic categories combination information available. For instance the semantic category combination "occupation + abstract things" corresponding to the interpretation template as shown in Table 3.

TABLE III:THE CORRESPONDING INTERPRETATION TEMPLATE OF THE "OCCUPATION+ ABSTRACT THINGS"

Semantic category 1	**Semantic category 2**	**Number of template**	**Template**	**Type of Verb**
职业 (occupation)	抽象事物 (abstract things)	1	v1+n1+of+n2	2

In the template of "v1+n1+of+n2", "v1" indicates the verb we need is the collocated verb of "n1" and the value of the verb type is 2, which means that "v1+n1+of+n2" requires the agentive role of n1.

By looking up combination of the semantic category, the corresponding interpretation templates can be obtained for the generation of the results. To prove the effectiveness of this method, we take out 1518 high frequency nominal compounds to carry on the automatic interpretation in the first half year of 1998 from the People's Daily Corpus and get the correct interpretation result after the manual proofreading. At present, the table of interpretation results contains 1084 records, including nominal compounds and their corresponding semantic categories as well as the positions of entries and interpretation templates in *Cilin*. In view of the result above, we can calculate the similarity between nominal compounds to be interpreted and the annotated ones then get the most similar nominal compound which contributes to selecting the correct interpretation. The next section will give a detailed description of AISNCC and AISNCS.

A. The Expansion of Nouns Based on Cilin

In the noun information acquisition module, the corresponding information can be obtained through direct matching, including the semantic category and the collocated verb. This paper adopts nouns expanding acquisition strategy based on *Cilin* to get synonyms of the noun *w*hich can't be matched directly for the purpose of obtaining the corresponding semantic category or collocated verb. The flow chart of the algorithm is shown in Figure 2.

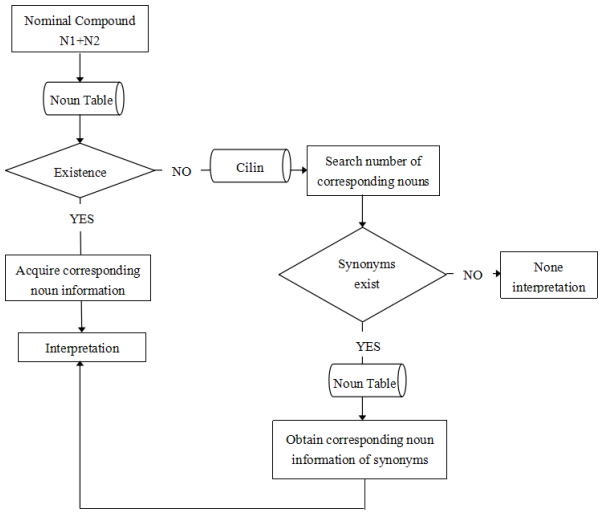

Figure.2: flow chart of noun information acquisition

Figure 2 shows the searching process for a given noun. If the noun exists, the semantic category and the verb can be obtained directly. If the noun does not exist, we will find the corresponding code for entry in *Cilin* and search the synonymy to get the semantic category information of synonyms as the expansion information of the noun.

For example, the word of "一等功" (the First-Class Merit) failed to retrieved in noun table and its code for entry in *Cilin* is "Da14B02=", which has the synonym of "头等功" and "头功", then we refer to these two synonyms as the expansion of "一等功" to obtain synonyms semantic category and collocated verb for subsequent interpretation.

B. Interpretation Templates Acquisition Based on Similarity

On the basis of noun expansion strategy described in last Section, we propose the AISNCC strategy, yet the experiment reveals its deficiency that the matched result of a single noun is far from to meet the needs of the interpretation of nominal compound, which requires us to analyze the composition pattern of them. For the N1+N2 nominal compound that has already been obtained, we still need to get the interpretation template to form the corresponding phrases. Template acquisition algorithm described in Algorithm1.

Algorithm1 Interpretation template acquisition algorithm based on similarity

Input: component nouns N1 and N2 of nominal compound NC and which corresponding Semantic category set S1 and S2

Output : Interpretation template set Template

```
Function GetTempalte(S1,S2)
    for Each S ∈S1 do
        for Each S' ∈S2 do
            if Exist template of S+S' then
                Template.add( )
            else
                GetTemplate(S1.UpperSemanticClass,S2)
                GetTemplate(S1.S2.UpperSemanticClass)
            end if
        end for
    end for
end function

Function SimilarTemplate(N1, N2)
    If Template is empty then
        similarity= Sim(N1, N2)
        Get the most similar N1+N2 template
        If similarity>0.8 then
            Template.add()
        end if
    else Template=NULL
    end if
end function
```

In the algorithm above, "the most similar templates" is defined as for nominal compound that can't obtain the template or all templates are invalid, we calculate the similarity of all the nominal compounds in its interpretation results table to get maximum similarity of nominal compounds, and the most similar template is the corresponding interpretation template.

This paper presents AISNCC using the code for entry and calculates the semantic distance of two words for their similarity [17]. We can abstract *Cilin* as a tree which belongs to the same kind of set with a total of 5 layers. The calculation method of the similarity between the words A and the word B is:

$$S(A,B)=\begin{cases}\lambda_0 & \text{A, B is not in the same tree}\\ \lambda_i \times \frac{n-k+1}{n} & \text{A, B is in the i-th branch}\\ 1 & \text{the number of A,B is same and the end is =}\\ 0.5 & \text{the number of A,B is same and the end is \#}\end{cases} \quad (1)$$

The initial value of the layer is λ_i and through the experiment, the value is: λ_1=0.65, λ_2=0.8, λ_3=0.9, λ_4=0.96, λ_0=0.1. $\frac{n-k+1}{n}$ is the control coefficient, where n is the total number of nodes in different positions of branching layer and the k is the distance between the two branches.

The similarity between nominal compound N1 and N2 is defined as:

$$\mathrm{Sim}(N1,N2)=(\prod S(Ai,Bi))^{1/m} \quad (2)$$

The Ai is the i-th component nouns of N1 and Bi is the i-th component nouns of N2, m is the number of nouns included in nominal compound, in addition, the number of nouns in N1 and N2 must be equal.

Take the nominal compound "无产阶级革命家" (proletarian revolutionist) and "外国专家" (foreign expert) as an example, we search for the words "无产阶级" (proletarian) and "外国" (foreign) respectively in the Cilin, as shown in Table 4.

TABLE IV: EXAMPLE OF NOUN ITEM

Nouns	Semantic items in Cilin
无产阶级 (proletarian)	Di06A02=无产阶级 工人阶级
	Di06A05@无产阶级
	Af01A02=无产者 无产阶级
外国(foreign)	Di02A05=外国 外 外域 异国 异邦 异域 夷 别国

Select the shortest distance of two items, the "无产阶级" is the item "Di06A02=" and it is different from "外国" of which item is "Di02A05=" in the third layer, so n=569, k=4. The similarity between "无产阶级" and "外国" is:

$$S1 = S(A1, B1) = 0.8 * (569 - 4+1)/569 = 0.796$$

Similarly, the similarity between "革命家" (revolutionary) and "专家" (expert) is S2 =S(A2, B2) = 0.869. So the similarity of "无产阶级革命家" and "外国专家" is:

$$\mathrm{Sim}(\text{无产阶级革命家, 外国专家}) = \sqrt{(S1 * S2)} = 0.845$$

In order to ensure the validity of interpretation template, we set the threshold as 0. 8.

For the nominal compound "无产阶级革命家" (proletarian revolutionist), its semantic category combination is: "人群" (crowd) + "身份"(identity) which is unable to obtain its interpretation template directly through the semantic category combination, so the similar nominal compound matching algorithm is conducted. We need to traverse all nominal compounds in the interpretation results table to get "the most similar" nominal compound: "外国专家" (foreign expert) and similarity is 0.845, with the threshold condition satisfied, so the interpretation template of "外国专家" is "来自+n1+的+n2" (n2+come from+n1) treated as "the most similar"

template, thus to obtain the final interpretation “来自+无产阶级+的+革命家” (revolutionist+come from+ proletarian).

C. Interpreted Phrase Generation

For a majority of nominal compounds, we need to obtain the corresponding collocated verb to transform the interpretation template into the interpreted phrase.

With regard to each interpretation template, we obtain the corresponding telic role or agentive role from noun table as the implicit verb. Then the component nouns of the nominal compound and the corresponding implicit verb are inserted in the corresponding position of the template to obtain the interpretation of the phrase. Interpretation of “农民 收入” (farmer's income) as shown in table 5:

TABLE V: THE INTERPRETATION OF TEMPLATE AND PHRASE OF “农民收入”

N1	N2	Interpretation template	Interpretation phrase
农民(farmer)	收入(income)	n1 +v2+的+n2 (n2+n1+v2)	农民+赚+的+收入 (income + farmer+ earn)
		n1 +举行/组织+的+n2 (n2+n1+hold)	农民+举行/组织+的+收入 (income + farmer+hold)
		（性质）+是+n1+的+n2 ((property)+is+n1’s+n2)	（性质） +是+农民+的+收入 ((property)+is+farmer’s+income)

The interpretation of nominal compound “农民 收入” (farmer's income) corresponds to the template for n1+v2+的+n2 and verb types is 2, which means in this template we require the agentive role of N2 and the corresponding verb is “赚” (earn), filling the verb into V2 position of the template, then the “农民” (farmer), “收入” (income) are inserted into N1 and N2 position of the template thus finally get results of “农民+赚+的+收入” (income + farmer+ earn).

IV. EXPERIMENT AND RESULT ANALYSIS

In the research of Wei [7], the selected test data is 245 nominal compounds extracted from Baidu top searching words, which has a certain limitation due to the small scale of test set. This paper selects a N+N minimal sequence a total of 1500 with a higher word frequency in the first half year of 1998 from People's Daily Corpus and Xinhua Corpus from 2000 to 2002, then removes N+N sequence of which the left and the right collocated words are relatively fixed (frequency of left or right collocated words higher than 90%), finally obtains 1069 nominal compounds as the test set. With the same data set, both Wei's and our algorithms are conducted in this paper.

There are three evaluation measures in our experiment, which are the Precision, Recall and F-measure.

A. Algorithm Validity Test

Compared with Wei’s algorithm [7] tested in this paper, we find that 32 nominal compounds failed to obtain the interpretations for the reason that they are unable to match the component nouns and 212 nominal compounds cannot obtain interpretation templates. With AISNCC getting synonyms of noun to obtain expand information, 21 nominal compounds get the interpretation in 32 which can’t be interpreted and give the 17 of the 21 nominal compounds are correct, which proves that the method is feasible and reasonable. Part of the semantic category information is obtained by the expansion of noun, as shown in Table 6, the interpretation results of nominal compounds are displayed in Table 7:

TABLE VI MATCHING RESULT OF WORD EXPANSION

Noun needed to expanded matching	synonym	Semantic category finally obtained
一等功(the First-Class Merit)	头等功|头功(the First-Class Merit)	事件(event)
财长(finance ministers)	厂长|校长|院长|馆长…… (Plant Manager |Principal| Dean| Curator)……	身份|职业 (identity| career)

TABLE VII MATCHING RESULT OF NOMINAL COMPOUND BY WORD EXPANSION

nominal compound	Interpretation phrase
财长会议（conference held by finance ministers）	财长+举行/组织+的+会议(conference+ held /organized by + finance ministers)
	（性质）+是+财长+的+会议((property) +is + financial+conference)
集体一等功 （collectivity’s the First-Class Merit）	产自+集体+的+一等功 （the First-Class Merit + produced by+ collectivity）

From the table above, we can observe that “财长” (finance ministers) by the direct searching can’t get the noun information in “财长会议” (conference held by finance ministers). But we can obtain its synonym “厂长” (Plant Manager) |“校长” (Principal) |“院长” (Dean) |“馆长” (Curator), which belong to the semantic category “身份 (identity) | 职业(career)” as the word semantic category of “财长” (finance ministers), through a combination of semantic categories to obtain interpretation template with generated phrase “the conference+held/organized by+finance ministers” and “(property)+is financial+conference”. However, there are some nominal compounds that can't get correct defined phrases, for example the interpretation of “集体 一等功” (collectivity’s the First-Class Merit) is “产自+集体+的+一等功” (the First-Class Merit+produced by+ collectivity) which is not reasonable while the correct interpretation should be “集体+获得+的+一等功” (the First-Class Merit+won by+ collectivity).

For the 212 nominal compounds that cannot get the interpretation template, the template acquisition algorithm based on similarity is tended to obtain the “most similar template” of nominal compounds and 128 nominal compounds based on the similarity of the template are gotten, of which 95 are correct, which shows that this interpretation template acquisition algorithm based on similarity can improve the coverage to some extent compared with previous method.

In contrast with the approach proposed by Wei [7], the three strategies are tested on the same test set and the interpretation results are shown in Table 8.

As we can see from the data in the table, the noun expansion strategy based on Cilin and the interpretation

template acquisition strategy based on similarity all have significantly improved in recall. In addition, although the precision of AISNCS strategy is 84.91%, which is lower than the that of the Wei's algorithm, our method gives a relatively high F-measure(about 80.96%), with recall from 66.88% to 77.36%, which indicates the effectiveness of the proposed algorithm.

TABLE VIII INTERPRETATION NUMBER OF NOMINAL COMPOUNDS

	NO. OF INTERPRET	NO OF CORRECT	P	R	F1
WEI(2014)	825	715	86. 67%	66. 88%	76. 25%
AISNCC	846	732	86. 52%	68. 48%	76. 45%
AISNCS	974	827	84. 91%	77. 36%	80. 96%

V. CONCLUSIONS AND FUTURE WORK

In this paper, we introduced the method of automatic interpretation of N+N nominal compound by combining top-down and bottom-up strategies. In view of the problem of lacking noun information and incomplete interpretation template, we incorporate interpretation strategy of nominal compound algorithm based on *Cilin* (AISNCC) into the automatic interpretation strategy of nominal compound based on similarity (AISNCS). Through the method of AISNCC, we can have access to the noun synonyms for the purpose of obtaining the corresponding interpretation template and then by calculating the similarity between nominal compounds, the final interpretation is got with the most similar template. Experimental results compared with Wei [7] demonstrate that our proposed approach has achieved substantial improvement on the recall of nominal compound. For future work, we plan to focus on the interpretation of the nominal compounds with the three or more component nouns.

ACKNOWLEDGMENT

This work is partially supported by Chinese National Fund of Natural Science under Grant 61272221, and Jiangsu Province Fund of Social Science under Grant 12YYA002.

REFERENCE

[1] H. Ma, "Semantic Investigation of 'Noun+Noun' Compounds," *Journal of Xinyang Teachers College (Philosophy and Social Science Edition)*, vol. 19, no. 1, pp. 117-120, 1999.

[2] H. Zhu, "Research on the Methods of Chinese None Compounds Identification and Classification," *Journal of Harbin Institute of Technology*, 2007.

[3] J. N. Levin, "The Syntax and Semantics of Complex Nominals," *Academic Press*, 1978.

[4] B. C. Warren, "Semantic Patterns of Noun-Noun Compounds," *Gothenburg Studies*, 1978.

[5] M. Lauer, "Designing Statistical Language Learners: Experiments on Noun Compounds," *Ph.D. thesis, Macquarie University.* Australia. 1995.

[6] N. Preslav and H. Marti, "Using Verbs to Characterize Noun-Noun Relations," *Lecture Notes in Computer Science*, vol. 4183, pp. 233-244, 2006.

[7] X. Wei and Y. Yuan, "Towards a Rule-based Approach to Automatic Interpretation of Chinese Noun Compounds," *Journal of Chinese Information Processing*, vol. 28, no. 3, pp. 1-10, 2014.

[8] G. Li, A. Lopez-Fernandez and T. Veale, "UCD-Goggle: A hybrid system for noun compound paraphrasing," in *Proceedings of the International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2010, pp. 230-233.

[9] N. K. Su and T. Baldwin, "Interpreting Semantic Relations in Noun Compounds via Verb Semantics," in *Proceedings of the International Conference on Computational Linguistics and, Meeting of the Association for Computational Linguistics*, 2006, pp. 491-498.

[10] J. Wu, "Schema Theory and the Interpretation of Noun-Noun Compounds in contemporary Chinese: A Database-Driven Study," *Shanghai International Studies University*, Shanghai, 2006.

[11] Y. Yuan, "Predicate Implicit and its Syntactic Consequences." *Studies of the Chinese Language*, vol. 4, pp. 241-255, 1995.

[12] R. Zhou, "A Syntactic and Semantic Study on Noun-Noun Constructions," *Jinan University*, Guangzhou, 2007.

[13] M. Wang, J. Huang, S. Yu and B. Li, "Chinese Noun Compound Interpretation Based on Paraphrasing Verbs," *Journal of Chinese Information Processing*, vol. 24, no. 6, pp. 3-9, 2010.

[14] H. Wang, W. Zhan and S. Yu, "Structure and Application of the Semantic Knowledge-base of Modern Chinese," *Applied Linguistics*, vol. 1, pp. 134-141, 2006.

[15] Z. Dong and Q. Dong, "Construction of A Knowledge System and its Impact on Chinese Research," Contemporary Linguistics, vol. 3, no. 1, pp. 33-44, 2001.

[16] J. Pustejovsky, "The generative lexicon," *Computational Linguistics*, vol. 17, no. 4, pp. 409-441, 1991.

[17] J. Tian and W. Zhao, "Words Similarity Algorithm Based on Cilin in Semantic Web Adaptive Learning System," *Journal of Jilin University (Information Science Edition)*, vol. 28, no. 6, pp. 602-608, 2010.

[18] Chung YS and Chen KJ, A Semantic-Based Approach to Noun-Noun Compound Interpretation. Chinese Journal of Computational Linguistics, 2013, 18(4): 45-62.

[19] Pasca M, Interpreting Compound Noun Phrases Using Web Search Queries. Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2015: 335–344.