# Exploiting Lexical Semantic Resource for Tree Kernel-Based Chinese Relation Extraction

Liu Dandan, Hu Yanan, and Qian Longhua*

Natural Language Processing Lab
School of Computer Science and Technology, Soochow University
Suzhou, China, 215006
qianlonghua@suda.edu.cn

**Abstract.** Lexical semantic resources play an important role in semantic relation extraction between named entities. This paper exploits lexical semantic information based on HowNet to convolution tree kernels via two methods: incorporating lexical semantic similarity and embedding lexical sememes, and systematically investigates its effects on Chinese relation extraction. The experimental results on the ACE 2005 Chinese corpus show that the incorporation of lexical semantic similarity can significantly improve the performance whether entity-related information is known or not, while embedding lexical sememes can also improve the performance, but only when entity types are unknown. This demonstrates the effectiveness of lexical resources for Chinese relation extraction. In addition, the experiments also suggest that lexical semantic similarity facilitates the relation extraction, particularly the fine-grained subtype extraction, more than that of relation detection.

**Keywords:** Relation Extraction, Convolution Tree Kernel, Lexical Semantic Similarity, Lexical Sememe, HowNet.

## 1 Introduction

Relation extraction (RE) is an important information extraction task in natural language processing (NLP), with many practical applications, including learning by reading, automatic question answering, text summarization and so on. The goal of relation extraction is to detect and characterize semantic relationships between pairs of named entities in text. For example, a typical relation extraction system needs to extract a Person-Social relationship between the person entities "他" and "妻子" in the Chinese phrase "他 的 妻子" (his wife).

Generally, machine learning-based methods are adopted in relation extraction due to their high accuracy. In terms of the expression of learning examples (i.e., relation instances) they can be divided into feature-based methods and kernel-based ones. The key issue of feature-based RE is how to extract various lexical, phrasal, syntactic, and semantic features [1-7], which are important for relation extraction, from the sentence

* Corresponding author.

M. Zhou et al. (Eds.): NLPCC 2012, CCIS 333, pp. 213–224, 2012.

involving two entities, while for kernel-based RE, the structured representation of relation instances, such as syntactic parse trees[8-10], dependency trees[11], and dependency paths [12-13] etc., becomes the central problem. In Chinese relation extraction, many studies focus on feature-based methods, such as [14-16] while kernel-based methods, such as edit distance kernel [17], string kernel[19], convolution tree kernels over parse trees [20-21], have gained wide popularity.

It is widely held that lexical semantic information plays an important role in relation extraction between named entities, since two words, different in surface but similar in semantic, may represent the same relationship. For example, the two phrases "他 的 妻子" (his wife) and "她 的 丈夫" (her husband) convey the same relationship "PER-SOC.Family" in the ACE terminology, though "他" (he) and "她" (she), "妻子" (wife) and "丈夫" (husband) are two distinctive, yet semantically similar words. Therefore, different approaches are proposed to exploit this lexical semantic similarity in relation extraction. Chan et al.[6] and Sun et al.[7]use the corpus-based clustering techniques [21] to obtain the semantic codes of entity headwords, and then embed them as semantic features into the framework of feature-based relation extraction. However, it is difficult to determine the level of generality for semantic codes as regards different levels of relation types to be extracted.

In Chinese relation extraction, Che et al.[17] and Liu et al.[18] embed lexical semantic similarity based on TongYiCi CiLin [22] or HowNet to an edit distance kernel or a sequence kernel respectively for relation extraction. However, as the convolution tree kernels[23] exhibits their potential for relation extraction and witness wide applications far beyond the RE domain, the unresolved issue is whether or not the widely adopted convolution tree kernel can benefit from such lexical semantic resources. Bloehdorn and Moschitti[24] propose a generalized framework for syntactic and semantic tree kernels for Question Classification, which incorporate semantic information when computing structural similarity between two parse trees. Following their work, we incorporate lexical semantic similarity based on HowNet (abbreviated as HN), into convolution tree kernels, and compare its effect on Chinese relation extraction with that of directly embedding lexical sememes in tree structures.

The rest of the paper is organized as follows. In Section 2, we review the related studies while Section 3 introduces the tree structure used for RE in this paper. Section 4 elaborates two methods of exploiting a lexical resource for Chinese relation extraction. Section 5 reports experimental results and analysis. Finally, Section 6 concludes the paper and points out future directions.

## 2 Related Work

Due to the focus in this paper, this section only reviews the previous studies on the applications of semantic information to relation extraction.

In English relation extraction, there are three previous studies considering some kind of semantic information all in feature-based methods. Zhou et al. [1] first extract a country name list and a personal relative trigger word list from WordNet. They demonstrate that these two lists are helpful to distinguish the relations of

"ROLE.Residence" in ACE 2003 and of "PER-SOC.Family" in ACE 2004. Chan et al. [6] combine various relational predictions and background knowledge, including word clusters automatically gathered from unlabeled texts, through a global inference procedure called ILP (Integer Linear Programming) for relation extraction. They demonstrate that these background knowledge significantly improves the RE performance, particularly when the training data are scarce. Sun et al.[7] present a simple semi-supervised relation extraction system with large-scale word clustering. What they mean by semi-supervised learning is that the additional features are induced through word clustering from large-scale unlabeled texts, similar to Chan et al.[6]. Nevertheless, the subtle semantic commonality between words seems inherently difficult to be captured by feature-based methods.

On the other hand in Chinese RE, Che et al. [17] employ the Improved-Edit-Distance (IED) to calculate the similarity between two Chinese strings, and further considering lexical semantic similarity between words based on TongYiCi CiLin, their experiments show that the lexical semantic-embedded IED kernel method performs well for the person-affiliation relation extraction. Liu et al.[18] acquire lexical semantic similarity scores based on HowNet, a widely used Chinese lexical resource, and incorporate them into a sequence string kernel. Experiments on some ACE-defined fine-grained relationships show promising results. Up till now, no attempt has been made to incorporate such semantic information into tree kernel-based relation extraction, which seems more natural than feature-based methods and is exactly the focus of this paper.

## 3 Structured Representation for RE

The two key issues of tree kernel-based relation extraction are the representation of tree structure and the similarity calculation between trees. This section deals with the former while the next section discusses the latter.

Since the focus of this paper is the exploitation of lexical semantic resources to tree kernel-based RE, we directly adopt a state-of-the-art tree structure--the Unified Parse and Semantic Tree (UPST-FPT)[10]as the tree structure, which incorporates entity-related semantic information, such as entity types and subtypes in the Feature-Paired Tree (FPT) manner, into the Dynamic Syntactic Parse Tree (DSPT).

Figure 1 illustrates such a tree structure derived from the phrase "银行 总裁" (bank president) for a relation instance between the "银行" ORG (organization) entity and the

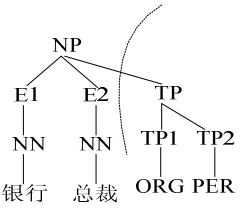


**Fig. 1.** Unified Parse and Semantic Tree with Feature-Paired Tree (UPST-FPT)

"总裁" PER (person) entity, where "TP1" denotes the entity type of the 1st entity and likewise "TP2" denotes that of the 2nd entity. The tree structure on the left of the dotted line is the DSPT while other entity-related semantic information, such as entity subtypes etc., is omitted for brevity.

## 4 Exploiting Semantic Resource in Tree kernels

This section first introduces convolution tree kernel for Chinese semantic relation extraction, then discusses two methods of using lexical semantic resource: incorporating into tree kernel computation and directly embedding into tree structures.

The first question for exploiting lexical semantic resource is, given too many words as leaf nodes in a parse tree, which of them are useful for relation extraction? Sun et al. [7] conduct a series of experiments using word clusters for different words in a relation instance, such as entity headwords, bag of headwords, the words before and after the entities etc., in a feature-based framework of RE and find that only entity headwords are important for RE. Therefore, we first consider the semantic information of lexical items corresponding to two entities involved in a relation instance. Moreover, we also consider the semantic information of the verb as Qian et al [10], if exists, between two entities, as some relation instances are clearly expressed in a verbal form.

### 4.1 Convolution Tree Kernel

The convolution tree kernel [23] counts the number of common sub-trees between two parse trees T1 and T2 as their similarity measure without explicitly considering the whole tree space. It can be computed as follows:

$$K_{CTK}(T_1, T_2) = \sum_{n_1 \in N_1, n_2 \in N_2} \Delta(n_1, n_2) \tag{1}$$

where $N_1$ and $N_2$ are the sets of nodes for $T_1$ and $T_2$ respectively, and $\Delta(n_1, n_2)$ evaluates the number of two common sub-trees rooted at $n_1$ and $n_2$. It can be computed recursively as follows:

1. If the productions at $n_1$ and $n_2$ are different then $\Delta(n_1, n_2)=0$; otherwise go to Step 2;

2. If both $n_1$ and $n_2$ are part of speech (POS) tags, then $\Delta(n_1, n_2)=\lambda$; otherwise go to Step 3;

3. Calculate recursively as follows:

$$\Delta(n_1, n_2) = \lambda \prod_{k=1}^{\#ch(n_1)} (1 + \Delta(ch(n_1, k), ch(n_2, k)) \tag{2}$$

where $\#ch(n)$ is the number of children of the node n, $ch(n,k)$ is the k-th child of the node n, and $\lambda$ ($0<\lambda<1$) is a decay factor, which is used for preventing the similarity of sub-trees exceedingly depending on the size of sub-trees.

### 4.2 Semantic Convolution Tree Kernel: Incorporating Lexical Semantic Similarity

While convolution tree kernels exhibit promising results in the task of relation extraction [8-10, 25], they disregard lexical semantic similarity between words in parse trees, which is critical for relation extraction in some scenarios. Following the successful application of the syntactic and semantic convolution tree kernel [24] to the task of Question Classification (QC), we adopt a similar Semantic Convolution Tree Kernel (SCTK) to Chinese relation extraction with the lexical semantic similarity being calculated using Chinese lexical semantic resource.

The computation process of the SCTK is largely the same as that of the standard CTK except that in Step 1, one additional case should be considered as follows:

1. If the productions at $n_1$ and $n_2$ are the same, then go to Step 2; otherwise, if both $n_1$ and $n_2$ are the parents of entity headword nodes, then $\Delta(n_1, n_2)= \lambda *$ LexSim(*HW1*,*HW2*); otherwise $\Delta(n_1, n_2)=0$;

where *HW1* and *HW2* denote the headwords corresponding to two entities immediately under $n_1$ and $n_2$ respectively and LexSim(*HW1*, *HW2*) denotes the lexical semantic similarity between these two headwords which can be calculated using lexical resources such as HowNet.

HowNet[1], a commonly used Chinese lexical resource, is a lexical knowledge base with rich semantic information, where a word is described as a group of sememes in a complicated multi-dimensional knowledge description language, and the first sememe reflects the major feature of one concept. For example, the Chinese word "暗箱(camera obscura)" is described as: "part|部件, #TakePicture|拍摄, %tool|用具, body|身", "部件" is the first sememe of "暗箱". Due to its richness in lexical semantics, it has been widely exploited in various NLP researches [29, 30].

We adopt the software package by Liu and Li [31] to calculate lexical semantic similarity scores based on HowNet. The similarity score between content words (entity headwords or verbs) is a linear interpolation of four different similarity scores, i.e. similarity between primary sememes, that between other sememes, that between sets, and that between feature structures.

It is worth noting that in most cases, the entity headwords can be used directly to calculate their lexical similarity scores, e.g., in the Chinese relation instance "他 的 妻子" (his wife), both "他" (he) and "妻子" (wife) could be passed to the similarity calculation module since as common names they can be found in lexical resources. However, take the entity mention "大安森林公园" (DaAn Forest Park) as an example, since this headword is not a well-known proper noun and can not be found in HowNet, any similarity score involving this entity calculated using HN will be zero. Our solution to this problem is to first segment the entity headword into sequential words using the segmentation package and then to take the rightmost word as the new headword. For example, the entity mention "大安森林公园" is segmented into "大安 森林 公园" and then the word "公园" is passed to the lexical similarity calculation module.

[1] http://www.keenage.com/

However, when the entity is a person, no segmentation is performed since it is meaningless to calculate the similarity between the Chinese characters in person names. This strategy of segmenting the entity headword also applies to the method of embedding lexical sememes in tree structures.

### 4.3 Incorporating Lexical Sememes into Tree Structures

The alternate method to exploit the lexical resource is to directly embed semantic information to tree structures for relation instances, thus avoiding the intensive computation cost brought about by semantic convolution tree kernel. For HowNet, since the first sememe of a lexical item reflects the major propery of one concept, we only extract its first sememe as the semantic information and embed it into the tree structures. For example, in the relation instance"*台北 大安森林公园*"(Taipei DaAn forest park), the first sememes of HowNet corresponding to "台北"(Taipei) and "公园"(park, the head word) are "地方"(place) and "设施"(facility) respectively. After the sememes are extracted from HowNet, namely, "地方"(place) and "设施"(facility), they are attached to the root of the parse tree as shown in Figure 2, where "SHN1" and "SHN2" denote semantic information (the first sememes) based on HowNet corresponding to the 1st entity and the 2nd entity.

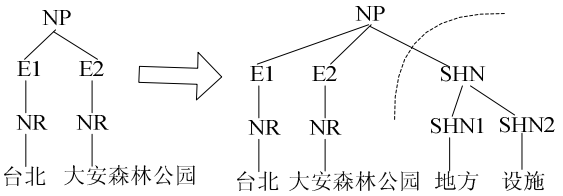


**Fig. 2.** Parse tree embedded with the first sememes of two entities

In addition, if there is a verb nearest to the 2nd entity along the path connecting two entities, a node "SHNV" followed by the verb's first sememe is also attached to the root node.

The first sememes of two entities or the verb are extracted from HowNet as follows:

1) find the lexical HW1、HW2 corresponding to the 1st and 2nd entity, and find the verb VLEX which near the 2nd entity;
2) search HowNet for the first sememes of HW1, HW2 and VLEX;
3) if the first sememe of a word does not exist, then the word will be further segmented, and again search the first sememe of the rightmost word after segmentation. Suppose that the first sememes are HCODE1、HCODE2 and HVCODE separately;
4) attach HCODE1、HCODE2 and HVCODE to the nodes SHN1、SHN2 and SHNV, which are further attached to the root of the parse tree.

# 5 Experimentation

This section experimentally investigates the effect of lexical semantic resources on Chinese relation extraction.

## 5.1 Experimental Setting

The ACE RDC 2005 Chinese corpus is used as the experimental datasets for Chinese semantic relation extraction. The corpus contains 633 documents, which were collected from newswires, broadcasts and weblogs. It defines 7 entity major types, 45 entity subtypes, 6 major relations types and 18 relation subtypes.

The corpus is first word-segmented using the ICTCLAS package, and then the corpus is parsed using the state-of-the-art Charniak's parser [32] with the boundaries of all the entity mentions kept. Finally, relation instances are generated by iterating over all pairs of entity mentions occurring in the same sentence, extracting corresponding tree structures and incorporating optional entity-related information (e.g., entity types, subtypes). In total, we obtain 9,147 positives and 97,540 negatives for Chinese relation instances.

In our experimentations, SVMLight-TK toolkit is adopted as our classifier since we usually treat RE as a classification problem. The package is modified to incorporate the lexical similarity calculation module. We apply the one vs. others strategy, which builds *K* classifiers so as to separate one class from the others. Particularly, the SubSet Tree (SST) kernel is used since it yields the best performance, while the decay factor $\lambda$ (tree kernel) is set to the default value (0.4).

We adopt the five-fold cross validation strategy for training and testing, and the averages of 5 runs are taken as the final performance scores. The commonly used evaluation metrics are Precise, Recall, F-measure, which can be abbreviated as P/R/F1 respectively. Finally, in order to determine whether an improvement of performance is statistically significant or not, we perform approximate randomization tests similar to [33] using a Perl script adapted from Randomized Parsing Evaluation Comparator[2]. Conventionally, the performance difference is considered significant or very significant if $p \leq 0.01$ or $0.01 < p \leq 0.05$ respectively.

## 5.2 Experimental Results and Analysis

We first investigate the impacts of two different methods of using HowNet on the task of relation extraction. Then we compare our system with other state-of-the-art Chinese relation extraction systems.

### Impact of Incorporating Lexical Semantic Similarity

Table 1 compares the performance of P/R/F1 for relation detection (2 types) and major type extraction (6 types) and subtype extraction (18 types) respectively on the ACE

[2] http://www.cis.uppen.edu/~dbikel/software/html#comparator

2005 Chinese corpus when lexical semantic similarity is incorporated in tree kernels for Chinese relation extraction. The DSPT (Dynamic Syntactic Parse Tree) [7] structure is used as the baseline (BL) without any semantic information. "ET" denotes that entity types (namely, major types and subtypes) are augmented into DSPT in the FPT (Feature-Paired Tree) [10] fashion while "HN" or "HNV" means either entity lexical similarity or verb lexical similarity based on HowNet is considered in the kernel computation. The 2nd column represents systems which incorporate various features or lexical similarity. For example, "(1)+HN" denotes considering the entity similarity on System 1, while "(3)+HN+HNV" considers both entity and verb similarity on System 3. The significance tests are conducted between a certain system (e.g., "(1)+HN") with its base system (i.e., System 1) and the performance increase, which is significant or very significant, is underlined or double-underlined respectively. Additionally, the best scores of P/R/F1 for each subtask are also highlighted respectively.

**Table 1.** Contributions of lexical semantic similarity for relation detection and extraction on the ACE 2005 Chinese corpus

| No | Systems | Detection | | | Major types | | | Subtypes | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | P | R | F1 | P | R | F1 | P | R | F1 |
| 1 | Baseline | **86.8** | 54.5 | 67.0 | 72.6 | 46.2 | 56.5 | 70.1 | 43.1 | 53.4 |
| 2 | (1)+HN | 81.0 | 57.3 | 67.1 | 70.0 | 50.0 | 58.3 | 68.4 | 47.8 | 56.3 |
| 3 | (1)+ET | 85.9 | 62.5 | 72.3 | 80.1 | 58.9 | 67.9 | 76.7 | 56.1 | 64.8 |
| 4 | (3)+HN | 86.5 | 62.7 | 72.7 | 81.1 | 59.5 | 68.7 | 78.7 | 57.4 | 66.4 |
| 5 | (3)+HN+HNV | 86.4 | **63.2** | **73.0** | **81.1** | **60.0** | **69.0** | **79.1** | **57.5** | **66.6** |

The table shows that, in general, with the incorporation of lexical similarity, the Chinese RE systems achieve better performance no matter whether the entity type information is considered, though in different degrees. Specifically, the table also shows that:

- when the entity similarity is incorporated into the baseline, the R/F1 scores for relation extraction on major types and subtypes obtain very significant improvements (3.8/1.8 for major types, 4.7/2.9 for subtypes), while their P scores decrease moderately(-2.6 for major types, -1.7 for subtypes). This means more positive instances are recalled when considering the entity similarity, but at the expense of precision.
- when the entity types are provided, the incorporation of entity similarity significantly improves all the P/R/F1 scores for relation extraction on both major types and subtypes(1.0/0.6/0.8 for major types, 2.0/1.3/1.6 for subtypes). This suggests that entity similarity assisted by entity types is very helpful for Chinese relation extraction.

- when both the entity and verb similarity are considered on System 3, System 5 achieves very significant improvements for all the P/R/F1 scores on the three relation extraction subtasks, furthermore, all the P/R/F1 scores except P in relation detection, reach their peaks. This implies the verbs do help for relation extraction in a certain degree.

One important trend for these three subtasks, exposed in the table, is that, although the absolute performance scores decrease with the increase of the number of relations types to be extracted (e.g., F1 scores range from 73.0, 69.0 to 66.6 on System 5), the improvements brought about by lexical similarity increase progressively (i.e., the F1 improvements are 0.7, 1.1 and 1.8 units respectively). This demonstrates that lexical similarity can do better in discerning fine-grained relation types than just binary relation types, and thus more helpful for subtype relation extraction. This can be intuitively explained by the example "他 的 妻子" (his wife), where a relationship "PER-SOC.Family" exists between the entities "他" and "妻子". The phrasal structure determines that certain relationship exists, while the lexical semantics of two entities determines the specific type of their relationship.

**Impact of Embedding Lexical Sememes**

Table 2 compares the P/R/F1 performance scores for relation detection and relation extraction on the ACE 2005 corpus when the first sememes are embedded in tree structures like in Figure 2. Different from Table 1, "+HN" denotes embedding the first sememes of two entities, rather than considering lexical similarity in kernel computation. Likewise "+HNV" denotes embedding the first sememe of the verb, all the other notations are the same as those in Table 1. Table 2 shows that:

**Table 2.** Contributions of the first sememes for relation detection and extraction on the ACE 2005 Chinese corpus

| No | Systems | Detection | | | Major types | | | Subtypes | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | P | R | F1 | P | R | F1 | P | R | F1 |
| 1 | Baseline | 86.8 | 54.5 | 67.0 | 72.6 | 46.2 | 56.5 | 70.1 | 43.1 | 53.4 |
| 2 | (1)+HN | **87.0** | 56.7 | 68.7 | 77.3 | 51.4 | 61.8 | 74.5 | 48.6 | 58.8 |
| 3 | (1)+ET | 85.9 | **62.5** | **72.3** | 80.1 | **58.9** | **67.9** | 76.7 | **56.1** | **64.8** |
| 4 | (3)+HN | 86.6 | 61.8 | 72.1 | 80.5 | 58.2 | 67.6 | 77.2 | 55.5 | 64.6 |
| 5 | (3)+HN+HNV | 86.9 | 61.2 | 71.8 | **80.7** | 57.6 | 67.2 | **77.3** | 55.1 | 64.3 |

- Similar to Table 1, when entity types are unknown, after the first sememes of two entities are embedded, the F1 scores achieve very significant improvements (1.7/5.3/5.4 units for detection, major types, and subtypes respectively), Moreover, the boost degree of performance is higher than that of lexical similarity incorporation, this implies that embedding the first sememes is likely

better than lexical similarity incorporation when the entity types are unknown. Particularly, the performance boost comes from both precision and recall, rather than lexical similarity incorporation boosts the recall performance at the cost of precision decrease.

- Different from Table 1, when the entity types are known, embedding the first sememes of entities or verb does not yield any F1 improvements, though the precision is increased in most cases. This shows that when the entity types are known, embedding the first sememes makes the structured information more accurate but at the great expense of recall decrease.

**Comparison with Other RE Systems**

Table 3 compares the performance of major type relation extraction of our SCTK method with other state-of-the-art systems for Chinese relation extraction on the ACE 2005 corpus. However, the comparison is only for reference as different parts of the corpus or evaluation strategies are adopted by different systems. For example, Li at el. [16] adopt a feature-based method using two-fold training/testing strategies while Yu et al. [20] experiment on a subset of the ACE 2005 corpus, though using the same 5-fold cross-validation evaluation strategy. Nevertheless, this table shows that our single-kernel method achieves promising results and has the potential to combine with other feature-based methods for better performance improvement.

**Table 3.** Comparison of different systems on the ACE 2005 Chinese corpus

| Systems | P(%) | R(%) | F1 |
|---|---|---|---|
| Qian et al[10]: Composite kernel (linear+tree) | **80.9** | **61.8** | **71.1** |
| Li et al[16]: Feature-based | 81.7 | 61.7 | 70.3 |
| Qian et al[10]: CTK with USST | 79.8 | 61.0 | 69.2 |
| Ours: SCTK with UPST | 81.1 | 60.0 | 69.0 |
| Yu et al[20]: CTK with UPST | 75.3 | 60.4 | 67.0 |
| Zhang et al.[34]: Composite kernel | 81.83 | 49.79 | 61.91 |

## 6 Conclusion and Future Work

In this paper, we empirically demonstrate the impact of lexical semantic resources of HowNet on Chinese relation extraction. We explore two methods of exploiting lexical semantic resources, i.e., incorporating HowNet-based lexical semantic similarity into tree kernels and directly embedding lexical sememes into tree structures. A series of experiments on the ACE 2005 benchmark corpus indicate that HowNet can significantly improve the performance of Chinese relation extraction via incorporating lexical similarity with or without entity-related information. On the other hand, embedding lexical sememes directly into tree structures, though as an intuitive method, improve the performance only when entity types are unknown. We also find that lexical similarity is better at extracting fine-grained relation types than just binary

relationships. This suggests that when extracting more specific semantic relationships, lexical semantic resources are preferable.

For future work, different ways of calculating similarity based on lexical resources could be investigated to find the best one, and we will explore the corpus-based word similarity for Chinese relation extraction when lexical resources are not available in some domains other than the general domain.

**Acknowledgement.** This work is funded by China Jiangsu NSF Grants BK2010219 and 11KJA520003.

## References

1. Zhou, G.D., Su, J., Zhang, J., Zhang, M.: Exploring Various Knowledge in Relation Extraction. In: ACL 2005, pp. 427–434 (2005)
2. Zhou, G.D., Su, J., et al.: Modeling Commonality among Related Classes in Relation Extraction. In: COLING-ACL 2006, pp. 121–128 (2006)
3. Zhou, G.D., Zhang, M.: Extracting Relation Information from Text Documents by Exploring Various Types of Knowledge. Information Processing and Management 43, 969–982 (2007)
4. Zhou, G.D., Qian, L.H., Fan, J.X.: Tree kernel-based Semantic Relation Extraction with Rich Syntactic and Semantic Information. Information Sciences 18(8), 1313–1325 (2010)
5. Jiang, J., Zhai, C.X.: A Systematic Exploration of the Feature Space for Relation Extraction. In: NAACL-HLT 2007, Rochester, NY, USA, pp. 113–120 (2007)
6. Chan, Y.S., Roth, D.: Exploiting Background Knowledge for Relation Extraction. In: COLING 2010, pp. 152–160 (2010)
7. Sun, A., Grishman, R., Sekine, S.: Semi-supervised Relation Extraction with Large-scale Word Clustering. In: ACL 2011, pp. 521–529 (2011)
8. Zhang, M., Zhang, J., Su, J., Zhou, G.D.: A Composite Kernel to Extract Relations between Entities with both Flat and Structured Features. In: COLING-ACL 2006, pp. 825–832 (2006)
9. Zhou, G.D., Zhang, M., Ji, D.H., Zhu, Q.M.: Tree Kernel-based Relation Extraction with Context-Sensitive Structured Parse Tree Information. In: EMNLP/CoNLL 2007, pp. 728–736 (2007)
10. Qian, L.H., Zhou, G.D., Kong, F., Zhu, Q.M., Qian, P.D.: Exploiting Constituent Dependencies for Tree Kernel-based Semantic Relation Extraction. In: COLING 2008, Manchester, pp. 697–704 (2008)
11. Culotta, A., Sorensen, J.: Dependency tree kernels for relation extraction. In: Proceedings of the 42nd Annual Meeting of the Association of Computational Linguistics, ACL 2004, pp. 423–439 (2004)
12. Bunescu, R.C., Mooney, R.J.: A Shortest Path Dependency Kernel for Relation Extraction. In: EMNLP 2005, pp. 724–731 (2005)
13. Nguyen, T., Moschitti, A., Riccardi, G.: Convolution Kernels on Constituent, Dependency and Sequential Structures for Relation Extraction. In: EMNLP 2009, pp. 1378–1387 (2009)
14. Che, W.X., Liu, T., Li, S.: Automatic Entity Relation Extraction 19(2), 1–6 (2005)
15. Dong, J., Sun, L., Feng, Y.Y., Huang, R.H.: Chinese Automatic Entity Relation Extraction. Journal of Chinese Information 21(4), 80–85, 91 (2007) (in Chinese)

16. Li, W.J., Zhang, P., Wei, F.R., Hou, Y.X., Lu, Q.: A Novel Feature-based Approach to Chinese Entity Relation Extraction. In: ACL 2008, pp. 89–92 (2008)
17. Che, W.X., Jiang, J., Su, Z., Pan, Y., Liu, T.: Improved-Edit-Distance Kernel for Chinese Relation Extraction. In: IJCNLP 2005, pp. 132–137 (2005)
18. Liu, K.B., Li, F., Liu, L., Han, Y.: Implementation of a Kernel-Based Chinese Relation Extraction System. Computer Research and Development 44(8), 1406–1411 (2007) (in Chinese)
19. Huang, R.H., Sun, L., Feng, Y.Y., Huang, Y.P.: A Study on Kernel-based Chinese Relation Extraction. Journal of Chinese Information 22(5), 102–108 (2008) (in Chinese)
20. Yu, H.H., Qian, L.H., Zhou, G.D., Zhu, Q.M.: Chinese Semantic Relation Extraction Based on Unified Syntactic and Entity Semantic Tree. Journal of Chinese Information 24(5), 17–23 (2010) (in Chinese)
21. Brown, P.F., Vincent, J.: Class-based n-gram models of natural language. Computational Linguistics 18(4), 467–479 (1992)
22. Mei, J.J., Zhu, Y.M., Gao, Y.Q., Yin, H.X.: TongYiCi CiLin, 2nd edn. Shanghai Lexicographic Publishing House, Shanghai (1996) (in Chinese)
23. Collins, M., Duffy, N.: Covolution Tree Kernels for Natural Language. In: NIPS 2001, pp. 625–632 (2001)
24. Bloehdorn, S., Moschitti, A.: Exploiting Structure and Semantics for Expressive Text Kernels. In: Proceedings of the Sixteenth ACM Conference on Information and Knowledge Management, Lisbon, Portugal (2007)
25. Qian, L.H., Zhou, G.D., Zhu, Q.M.: Employing Constituent Dependency Information for Tree Kernel-based Semantic Relation Extraction between Named Entities. ACM Transaction on Asian Language Information Processing 10(3), Article 15, 24 pages (2011)
26. Jiang, J., Conrath, D.: Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy. In: Proceedings of International Conference on Research in Computational Linguistics, Taiwan (1997)
27. Resnik, P.: Semantic Similarity in a Taxonomy: An Information- Based Measure and its Applications to Problems of Ambiguity in Natural Language. Journal of Artificial Intelligence Research 11, 95–130 (1999)
28. Lin, D.: An Information-theoretic Definition of Similarity. In: Proceedings of the 15th International Conference on Machine Learning, Madison, WI (1998)
29. Suo, H.G., Liu, Y.S., Cao, S.Y.: A Keyword Selection Method Based on Lexical Chains. Journal of Chinese Information 20(6) (2006) (in Chinese)
30. Li, D., Ma, Y.T., Guo, J.L.: Words Semantic Orientation Classification Based on HowNet. The Journal of China Universities of Posts and Telecommunications 16, 106–110 (2009)
31. Liu, Q., Li, S.J.: Word Similarity Computing Based on How-net. Computational Linguistics, Chinese Information Processing, 59–76 (2002)
32. Charniak, E.: Immediate-head Parsing for Language Models. In: ACL 2001 (2001)
33. Chinchor, N.: The statistical significance of the MUC-4 results. In: MUC-4, pp. 30–50 (1992)
34. Zhang, J., Ouyang, Y., Li, W., Hou, Y.: A Novel Composite Kernel Approach to Chinese Entity Relation Extraction. In: Li, W., Mollá-Aliod, D. (eds.) ICCPOL 2009. LNCS, vol. 5459, pp. 236–247. Springer, Heidelberg (2009)