SemEval-2012 Task 5: Chinese Semantic Dependency Parsing

Wanxiang Che[†], **Meishan Zhang**[†], **Yanqiu Shao**[‡], **Ting Liu**[†]
[†]Research Center for Social Computing and Information Retrieval
Harbin Institute of Technology, China
{car, mszhang, tliu}@ir.hit.edu.cn
[‡]Beijing City University, China
yqshao@bcu.edu.cn

Abstract

The paper presents the SemEval-2012 Shared Task 5: *Chinese Semantic Dependency Parsing*. The goal of this task is to identify the dependency structure of Chinese sentences from the semantic view. We firstly introduce the motivation of providing Chinese semantic dependency parsing task, and then describe the task in detail including data preparation, data format, task evaluation, and so on. Over ten thousand sentences were labeled for participants to train and evaluate their systems. At last, we briefly describe the submitted systems and analyze these results.

1 Introduction

Semantic analysis is a long-term goal of Natural Language Processing, and as such, has been researched for several decades. A number of tasks for encoding semantic information have been developed over the years, such as entity type recognition and word sense disambiguation. Recently, sentence-level semantics – in particular, semantic role labeling – has received increasing attention. However, some problems concerning the semantic representation method used in semantic role labeling continue to exist (Xue and Palmer, 2005).

1. Semantic role labeling only considers predicate-argument relations and ignores the semantic relations between a noun and its modifier.

2. The meaning of semantic roles is related to special predicates. Therefore, there are infinite semantic roles to be learned, as the number of predicates is not fixed. Although the PropBank (Xue and Palmer, 2003) normalizes these semantic roles into certain symbols, such as Arg0-Arg5, the same symbol can have different semantic meanings when paired with different predicates, and thus cannot be learned well.

Semantic dependency parsing is therefore proposed to solve the two problems above for Chinese. Firstly, the proposed method analyzes all the words' semantic roles in a sentence and specifies the concrete semantic relation of each word pair. Afterward, this work analyzes and summarizes all the possible semantic roles, obtaining over 100 of them, and then uses these semantic roles to specify the semantic relation for each word pair.

Dependency parsing (Kübler et al., 2009) is based on dependency grammar. It has several advantages, such as concise formalization, easy comprehension, high efficiency, and so on. Dependency parsing has been studied intensively in recent decades, with most related work focusing on syntactic structure. Many research papers on Chinese linguistics demonstrate the remarkable difference between semantics and syntax (Jin, 2001; Zhou and Zhang, 2003). Chinese is a meaning-combined language with very flexible syntax, and semantics are more stable than syntax. The word is the basic unit of semantics, and the structure and meaning of a sentence consists mainly of a series of semantic dependencies between individual words (Li et al., 2003). Thus, a reasonable endeavor is to exploit dependency parsing for semantic analysis of Chinese languages. Figure 1 shows an example of Chinese semantic dependency parsing.

*First Joint Conference on Lexical and Computational Semantics (*SEM)*, pages 378–384,
Montréal, Canada, June 7-8, 2012.

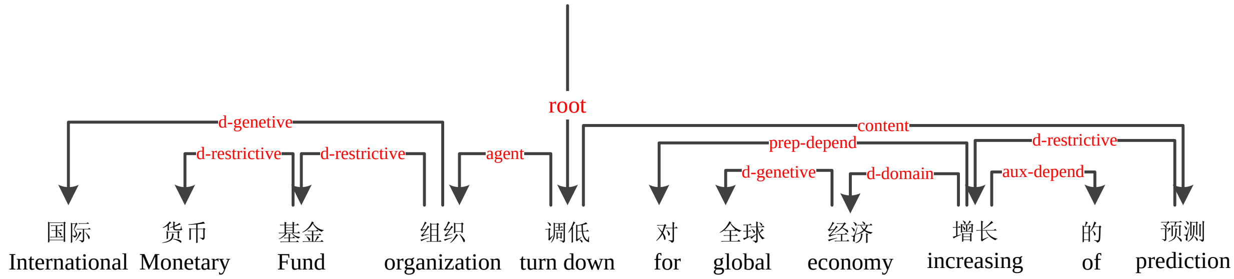

Figure 1: An example of Chinese Semantic Dependency Parsing.

Figure 1 shows that Chinese semantic dependency parsing looks very similar to traditional syntax-dominated dependency parsing. Below is a comparison between the two tasks, dealing with three main points:

1. Semantic relations are more fine-grained than syntactic ones: the syntactic subject can either be the agent or experiencer, and the syntactic object can be the content, patient, possession, and so on. On the whole, the number of semantic relations is at least twice that of syntactic relations.

2. Semantic dependency parsing builds the dependency structure of a sentence in terms of semantics, and the word pairs of a dependency should have a direct semantic relation. This criterion determines many sizeable differences between semantics and syntax, especially in phrases formed by "XP+DEG", "XP+DEV" and prepositional phrases. For example, in "美丽 的 祖国" (beautiful country), the head of "美丽" (beautiful) is "祖国" (country) in semantic dependency parsing, whereas the head is "的" (de) in syntax dependency parsing.

3. Semantic relations are independent of position. For example, in "空气 被 污染" (the air is contaminated) and "污染 了 空气" (contaminate the air), the patient "空气" (the air) can be before or behind a predicate "污染" (contaminate).

The rest of the paper is organized as follows. Section 2 gives a short overview of data annotation. Section 3 focuses on the task description. Section 4 describes the participant systems. Section 5 compares and analyzes the results. Finally, Section 6 concludes the paper.

2 Data Annotation

2.1 Corpus Section

10,068 sentences were selected from the Penn Chinese Treebank 6.0[1] (Xue et al., 2005) (1-121, 1001-1078, 1100-1151) as the raw corpus from which to create the Chinese Semantic Dependency Parsing corpus. These sentences were chosen for the annotation for three reasons. First, gold syntactic dependency structures can be of great help in semantic dependency annotation, as syntactic dependency arcs are often consistent with semantic ones. Second, the semantic role labels in PropBank[2] can be very useful in the present annotation work. Third, the gold word segmentation and Part-Of-Speech can be used as the annotation input in this work.

2.2 Semantic Relations

The semantic relations in the prepared Chinese semantic dependency parsing corpus came mostly from HowNet[3] (Dong and Dong, 2006), a famous Chinese semantic thesaurus. We also referred to other sources. Aside from the relations from HowNet, we defined two kinds of new relations: reverse relations and indirect relations. When a verb modifies a noun, the relation between them is a reverse relation, and r-XXX is used to indicate this kind of relation. For instance, in "打 篮球 的 小 男孩" (the little boy who is playing basketball), the semantic relation between the head word "男孩" (boy)

[1] http://www.ldc.upenn.edu/Catalog/catalogEntry.jsp?catalog\\Id=LDC2007T36
[2] http://verbs.colorado.edu/chinese/cpb/
[3] http://www.keenage.com/

and “打” (playing) is the r-agent. When a verbal noun is the head word, the relation between it and the modifier is the indirect relation j-XXX. For instance, in “企业 管理” (business management), the head word is “管理” (management) and the modifier is “企业” (business), their relation is j-patient.

Finally, we defined 84 single-level semantic relations. The number of multi-level semantic relations that actually appear in the labeled corpus in this work is 39.

Table 1 summarizes all of the semantic relations used for annotation.

2.3 Annotation Flow

Our corpus annotation flow can be divided into the following steps.

1. Conversion of the sentences' constituent structures into dependency structures according to a set of rules similar with those used by the syntactic community to find the head of a phrase (Collins, 1999).

2. Labeling of the semantic relations for each dependency relation according to another set of rules using the functional tags in the Penn Chinese Treebank and the semantic roles in the Chinese PropBank.

3. Six human annotators are asked to check and adjust the structure and semantic relation errors introduced in Step 2.

The first two steps were performed automatically using rules. A high accuracy may be achieved with dependency structures when semantic labels are not considered. However, accuracy declines remarkably when the semantic label is considered. Unlabeled Attachment Score (UAS) and Labeled Attachment Score (LAS) can be used to evaluate the performance of the automatic conversion. Table 2 gives the detailed results.

	UAS	LAS
Conversion Result	90.53	57.38

Table 2: Accuracy after conversion from gold ProbBank.

3 Task Description

3.1 Corpus Statistics

We annotated 10,068 sentences from the Penn Chinese TreeBank for Semantic Dependency Parsing, and these sentences were divided into training, development, and test sections. Table 3 gives the detailed statistical information of the three sections.

Data Set	CTB files	# sent.	# words.
Training	1-10; 36-65;81-121; 1001-1078; 1100-1119; 1126-1140	8301	250311
Devel	66-80; 1120-1125	534	15329
Test	11-35; 1141-1151	1233	34311
Total	1-121; 1001-1078 1100-1151	10068	299951

Table 3: Statistics of training, development and test data.

3.2 Data Format

The data format is identical to that of a syntactic dependency parsing shared task. All the sentences are in one text file, with each sentence separated by a blank line. Each sentence consists of one or more tokens, and each token is represented on one line consisting of 10 fields. Buchholz and Marsi (2006) provide more detailed information on the format. Fields are separated from each other by a tab. Only five of the 10 fields are used: token id, form, pos tagger, head, and deprel. Head denotes the semantic dependency of each word, and deprel denotes the corresponding semantic relations of the dependency. In the data, the lemma column is filled with the form and the cpostag column with the postag. Figure 2 shows an example.

3.3 Evaluation Method

LAS, which is a method widely used in syntactic dependency parsing, is used to evaluate the performance of the semantic dependency parsing system. LAS is the proportion of “scoring” tokens assigned to both the correct head and correct semantic dependency relation. Punctuation is disregarded during the evaluation process. UAS is another important indicator, as it reflects the accuracy of the semantic dependency structure.

Main Semantic Roles	
Subject Roles	agent, experiencer, causer, possessor, existent, whole, relevant
Object Roles	isa, content, possession, patient, OfPart, beneficiary, contrast, partner, basis, cause, cost, scope, concerning
Auxiliary Semantic Roles	
Time Roles	duration, TimeFin, TimeIni, time, TimeAdv
Location and State Roles	LocationFin, LocationIni, LocationThru, StateFin, state, StateIni, direction, distance, location
Others Verb Modifiers	accompaniment, succeeding, frequency, instrument, material, means, angle, times, sequence, sequence-p, negation, degree, modal, emphasis, manner, aspect, comment
Attribute Roles	
Direct modifiers	d-genetive, d-category, d-member, d-domain, d-quantity-p, d-quantity, d-deno-p, d-deno, d-host, d-TimePhrase, d-LocPhrase, d-InstPhrase, d-attribute, d-restrictive, d-material, d-content, d-sequence, d-sequence-p, qp-mod
Verb Phrase	r-{Main Semantic Roles}, eg: r-agent, r-patient, r-possessor
Verb Ellipsis	c-{Main Semantic Roles}, eg: c-agent, c-content, c-patient
Noun as Predication	j-{Main Semantic Roles}, eg: j-agent, j-patient, j-target
Syntactic Roles and Others	
Syntactic Roles	s-cause, s-concession, s-condition, s-coordinate, s-or, s-progression, s-besides, s-succession, s-purpose, s-measure, s-abandonment, s-preference, s-summary, s-recount, s-concerning, s-result
Others	aux-depend, prep-depend, PU, ROOT

Table 1: Semantic Relations defined for Chinese Semantic Dependency Parsing.

ID	FORM	LEMMA	CPOS	PPOS	FEAT	HEAD	REL	PHEAD	PREL
1	钱其琛	钱其琛	NR	NR	_	2	agent	_	_
2	谈	谈	VV	VV	_	0	ROOT	_	_
3	香港	香港	NR	NR	_	4	d-genetive	_	_
4	前景	前景	NN	NN	_	7	s-coordinate	_	_
5	和	和	CC	CC	_	7	aux-depend	_	_
6	台湾	台湾	NR	NR	_	7	d-genetive	_	_
7	问题	问题	NN	NN	_	2	content	_	_

Figure 2: Data format of the Chinese Semantic Dependency Parsing corpus.

4 Participating Systems

Nine organizations were registered to participate in the Chinese Semantic Dependency Parsing task. Finally, nine systems were received from five different participating teams. These systems are as follows:

1. Zhou Qiaoli-1, Zhou Qiaoli-2, Zhou Qiaoli-3
   These three systems propose a divide-and-conquer strategy for semantic dependency parsing. The Semantic Role (SR) phrases are identified (Cai et al., 2011) and then replaced by their head or the SR of the head. The original sentence is thus divided into two types of parts that can be parsed separately. The first type is SR phrase parsing, and the second involves the replacement of SR phrases with either their head or the SR of the head. Finally, the paper takes a graph-based parser (Li et al., 2011) as the semantic dependency parser for all parts. These three systems differ in their phrase identification strategies.

2. NJU-Parser-1, NJU-Parser-2
   The NJU-Parser is based on the state-of-the-art MSTParser (McDonald, 2006). NJU-Parser applies three methods to enhance semantic dependency parsing. First, sentences are split into sub-sentences using commas and semicolons: (a) sentences are split using only commas and semicolons, as in the primary system, and (b) classifiers are used to determine whether a comma or semicolon should be used to split the sentence. Second, the last character in a Chinese word is extracted as the lemma, since it usually contains the main sense or semantic class. Third, the multilevel-label is introduced into the semantic relation, for example, the r-{Main Semantic Roles}, with NJU-Parser exploiting special strategies to handle it. However, this third method does not show positive performance.

3. Zhijun Wu-1
   This system extends the second-order of the MSTParser by adding third-order features, and then applying this model to Chinese semantic dependency parsing. In contrast to Koo and Collins (2010) this system does not implement the third-order model using dynamic programming, as it requires $O(n^4)$ time. It first first obtained the K-best results of second-order models and then added the third-order features into the results.

4. ICT-1
   The ICT semantic dependency parser employs a system-combining strategy to obtain the dependency structure and then uses the classifier from Le Zhang's Maximum Entropy Modeling Toolkit[4] to predict the semantic relation for each dependency. The system-combining strategy involves three steps:

   - Parsing each sentence using Nivre's arc standard, Nivre's arc eager (Nivre and Nilsson, 2005; Nivre, 2008), and Liang's dynamic algorithm (Huang and Sagae, 2010);
   - Combining parses given by the three parsers into a weighted directed graph;
   - Using the Chu-Liu-Edmonds algorithm to search for the final parse for each sentence.

5. Giuseppe Attardi-SVM-1-R, Giuseppe Attardi-SVM-1-rev
   We didn't receive the system description of these two systems.

5 Results & Analysis

LAS is the main evaluation metric in Chinese Semantic Dependency Parsing, whereas UAS is the secondary metric. Table 4 shows the results for these two indicators in all participating systems.

As shown in Table 4, the Zhou Qiaoli-3 system achieved the best results with LAS of 61.84. The LAS values of top systems are very closely. We performed significance tests[5] for top six results. Table 5 shows the results , from which we can see that the performances of top five results are comparative ($p > 0.1$) and the rank sixth system is significantly ($p < 10^{-5}$) worse than top five results.

[4] http://homepages.inf.ed.ac.uk/s0450736/maxenttoolkit.html

[5] http://www.cis.upenn.edu/~dbikel/download/compare.pl

	NJU-Parser-2	NJU-Parser-1	Zhijun Wu-1	Zhou Qiaoli-1	Zhou Qiaoli-2
Zhou Qiaoli-3	$\sim$	$\sim$	$\sim$	$\sim$	$>$
NJU-Parser-2	–	$\sim$	$\sim$	$\sim$	$>$
NJU-Parser-1	–	–	$\sim$	$\sim$	$>$
Zhijun Wu-1	–	–	–	$\sim$	$>$
Zhou Qiaoli-1	–	–	–	–	$>$

Table 5: Significance tests of the top five systems. $\sim$ denotes that the two systems are comparable ($p > 0.1$), and $>$ means the system of this row is significantly ($p < 10^{-5}$) better than the system of this column.

System	LAS	UAS
Zhou Qiaoli-3	61.84	80.60
NJU-Parser-2	61.64	80.29
NJU-Parser-1	61.63	80.35
Zhijun Wu-1	61.58	80.64
Zhou Qiaoli-1	61.15	80.41
Zhou Qiaoli-2	57.55	78.55
ICT-1	56.31	73.20
Giuseppe Attardi-SVM-1-R	44.46	60.83
Giuseppe Attardi-SVM-1-rev	21.86	40.47
Average	54.22	72.82

Table 4: Results of the submitted systems.

The average LAS for all systems was 54.22. Chinese Semantic Dependency Parsing performed much more poorly than Chinese Syntactic Dependency Parsing due to the increased complexity brought about by the greater number of semantic relations compared with syntactic relations, as well as greater difficulty in classifying semantic relations.

In general, all the systems employed the traditional syntax-dominated dependency parsing frameworks. Some new methods were proposed for this task. Zhou Qiaoli's systems first identified the semantic role phrase in a sentence, and then employed graph-based dependency parsing to analyze the semantic structure of the sentence. NJU-Parser first split the sentence into sub-sentences, then trained and parsed the sentence based on these sub-sentences; this was shown to perform well. In addition, ensemble models were also proposed to solve the task using ICT systems.

6 Conclusion

We described the Chinese Semantic Dependency Parsing task for SemEval-2012, which is designed to parse the semantic structures of Chinese sentences.

Nine results were submitted by five organizations, with the best result garnering an LAS score of 61.84, which is far below the performance of Chinese Syntax. This demonstrates that further research on the structure of Chinese Semantics is needed.

In the future, we will check and improve the annotation standards while building a large, high-quality corpus for further Chinese semantic research.

Acknowledgments

We thank the anonymous reviewers for their helpful comments. This work was supported by National Natural Science Foundation of China (NSFC) via grant 61133012 and 61170144, and the National "863" Leading Technology Research Project via grant 2012AA011102.

References

Sabine Buchholz and Erwin Marsi. 2006. Conll-x shared task on multilingual dependency parsing. In *Proceedings of the Tenth Conference on Computational Natural Language Learning (CoNLL-X)*, pages 149–164, New York City, June. Association for Computational Linguistics.

Dongfeng Cai, Ling Zhang, Qiaoli Zhou, and Yue Zhao. 2011. A collocation based approach for prepositional phrase identification. *IEEE NLPKE*.

Michael Collins. 1999. *Head-Driven Statistical Models for Natural Language Parsing*. Ph.D. thesis, Pennsylvania University.

Zhendong Dong and Qiang Dong. 2006. *Hownet And the Computation of Meaning*. World Scientific Publishing Co., Inc., River Edge, NJ, USA.

Liang Huang and Kenji Sagae. 2010. Dynamic programming for linear-time incremental parsing. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1077–1086, Uppsala, Sweden, July. Association for Computational Linguistics.

Guangjin Jin. 2001. *Theory of modern Chinese verb semantic computation*. Beijing University Press.

Terry Koo and Michael Collins. 2010. Efficient third-order dependency parsers. In *Proceedings of the 48th Annual Meeting of the ACL*, number July, pages 1–11.

Sandra Kübler, Ryan McDonald, and Joakim Nivre. 2009. Dependency Parsing. In *Synthesis Lectures on Human Language Technologies*.

Mingqin Li, Juanzi Li, Zhendong Dong, Zuoying Wang, and Dajin Lu. 2003. Building a large chinese corpus annotated with semantic dependency. In *Proceedings of the second SIGHAN workshop on Chinese language processing - Volume 17*, SIGHAN '03, pages 84–91, Stroudsburg, PA, USA. Association for Computational Linguistics.

Zhenghua Li, Min Zhang, Wanxiang Che, Ting Liu, Wenliang Chen, and Haizhou Li. 2011. Joint models for chinese pos tagging and dependency parsing. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 1180–1191, Edinburgh, Scotland, UK., July. Association for Computational Linguistics.

Ryan McDonald. 2006. *Discriminative learning and spanning tree algorithms for dependency parsing*. Ph.D. thesis, University of Pennsylvania.

Joakim Nivre and Jens Nilsson. 2005. Pseudo-projective dependency parsing. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL)*.

Joakim Nivre. 2008. Algorithms for deterministic incremental dependency parsing. *Computational Linguistics*, 34(4):513–553.

Nianwen Xue and Martha Palmer. 2003. Annotating the propositions in the penn chinese treebank. In *Proceedings of the Second SIGHAN Workshop on Chinese Language Processing*.

Nianwen Xue and Martha Palmer. 2005. Automatic semantic role labeling for chinese verbs. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*.

Nianwen Xue, Fei Xia, Fu-Dong Chiou, and Martha Palmer. 2005. The penn chinese treebank: Phrase structure annotation of a large corpus. *Natural Language Engineering*, 11(2):207–238.

Guoguang Zhou and Linlin Zhang. 2003. *The theory and method of modern Chinese grammar*. Guangdong Higher Education Press.