Exploring Local and Global Semantic Information for Event Pronoun Resolution

Fang Kong[1,2] *Guodong Zhou*[1*]
(1) School of Computer Science and Technology, Soochow University of China, 1 Shizi Street, Suzhou, China 215006
(2) Department of Computer Science, National University of Singapore, 13 Computing Drive, Singapore 117417
kongfang@suda.edu.cn, gdzhou@suda.edu.cn

ABSTRACT

Event anaphora resolution plays a critical role in discourse analysis. This paper focuses on improving event pronoun resolution using both local and global semantic information. In particular, a predicate-argument structure is proposed to represent the local semantic information about an event while the global semantic information is represented by the entity coreference chains related with various arguments in the predicate-argument structure to complement its locality. Evaluation on the OntoNotes English corpus shows the effectiveness of local and global semantic information for event pronoun resolution.

KEYWORDS: event pronoun, semantic information, corefernce resolution, predicate-argument structure

*Corresponding author

Proceedings of COLING 2012: Technical Papers, pages 1475–1488,
COLING 2012, Mumbai, December 2012.

1 Introduction

As one of the most important techniques in discourse analysis, anaphora resolution aims to resolve a given mention to its referred expression in a text and has been a focus of research in Natural Language Processing (NLP) for decades. According to the nature of the referred expression, anaphora resolution can be categorized into entity anaphora resolution and event anaphora resolution. While most studies focus on entity anaphora resolution and have achieved much success recently (e.g. Soon et al. 2001; Ng and Cardie 2002; Ng 2007, 2009; Yang et al.2004, 2006, 2008; Kong et al. 2009,2010), there are only a few studies on event anaphora resolution (Byron, 2002; Pradhan et al., 2007; Chen et al. 2010a, 2010b; Kong and Zhou, 2011).

In this paper, we address event pronoun resolution, the most difficult type of event anaphora resolution due to the least discriminative information that an event pronoun can provide. Here, an event pronoun is a pronoun whose antecedent refers to an event. In particular, we focus on improving event pronoun resolution using both local and global semantic information. For local semantic information, we employ a shallow semantic parser to extract the predicate-argument structure in a sentence to represent an involved event. In order to complement the locality of the predicate-argument structure, we consider the global semantic information via the entity coreference chains related with various arguments in the predicate-argument structure.

The rest of this paper is organized as follows. Section 2 describes the event anaphora resolution task. Section 3 briefly introduces the related work on event anaphora resolution. Section 4 presents our baseline framework which combines various kinds of flat features and a structured parse tree for event pronoun resolution. Section 5 explores both local and global semantic information for event pronoun resolution. Section 6 reports the experimental results. Finally, we conclude our work in Section 7.

2 Task Description

While entity anaphora focuses on mentions of an entity, event anaphora looks into mentions that refer to an event. In this paper, we consider event anaphora resolution. Consider the following examples:

a) In Yemen today, where the ship was ***[attacked]****$_1$*, the deliberate, well-organized familiar effort to find out who did *[it]$_2$*, and how ***[it]****$_3$* happened.

b) Two F Tomcats ***[struck]****$_1$* the targets. After ***[today's air strikes]****$_2$*, 13 Iraqi soldiers abandoned their posts and surrendered to Kurdish fighters.

c) Yes, it took a while last night to sort out precisely what the court had ***[decided]****$_1$* by such a narrow margin. ***[This]****$_2$* was a stabilizing decision that restored order to a very chaotic situation.

Example (a) shows the importance of event anaphora resolution in understanding the discourse. In example (a), the three mentions in italic and bold font form a chain of event "the ship was attacked". While entity anaphora resolution is capable of linking up mention 2 and mention 3, e.g. using the default proximity preference rule as widely adopted in the literature (Soon et al. 2001), this chain will only contain two pronouns without linking mention 2 to actual event mention 1. Event anaphora resolution provides an essential role to bridge the understanding gap in such a discourse by linking mention 2 to actual event mention 1. Here, we employ the predicate of a clause to represent an event mention.

Similarly, in example (b), the anaphor of NP “today’s air strikes” refers to event “Two F Tomcats struck the targets” while in example (c), the anaphor of pronoun “this” refers to event “what the court had decided by such a narrow margin”. Obviously, compared with noun phrases, pronouns carry little information of their own. This indicates the difficulty of event pronoun resolution in event anaphora resolution. Besides, our statistics on the OntoNotes English corpus (Release 3.0) shows that event pronouns occupy about 40% of event anaphors. This indicates the importance of event pronoun resolution in event anaphora resolution.

For better understanding the paper, here we give some related terminologies:

- ***Entity***: an object or a set of objects in one of the semantic categories of interest, referred by one or more coreferential entity mentions in the document.
- ***Entity mention***: a reference to an entity (typically, a noun phrase).
- ***Event trigger***: the key word that most clearly expresses the occurrence of an event. In this paper, as mostly adopted in the literature, we take the main predicate (either verbal or nominal) of a clause as the event trigger to represent the corresponding event.
- ***Event arguments***: the entity mentions involved in an event.
- ***Event mention***: a clause within which an event is described, including event trigger and event arguments. Although some event pronouns can actually refer to a paragraph or larger chunks of texts, in this paper we only consider the cases taking clauses as antecedents.

3 Related Work

In comparison with entity anaphora resolution, there are few linguistic studies on event anaphora (e.g. Asher, 1993) and very initial explorations on event anaphora resolution. It was only until recently, with the increasing interest in discourse analysis, event anaphora resolution has begun to draw more and more attention for the natural language processing community. While some of them focus on hand-crafted constraints to resolve event anaphora of normally limited kinds of predicates (e.g. Byron, 2002), most of previous studies adopt a learning-based framework (e.g. Muller, 2007; Pradhan et al., 2007; Chen et al. 2010a, 2010b; Kong and Zhou, 2011).

As a representative to linguistic studies on event anaphora, Asher (1993) proposed a discourse representation theory to resolve the references to events. However, no computational system was proposed in his work.

As a representative of using hand-crafted knowledge to resolve specific kinds of predicates, Byron (2002) proposed a semantic filter as a complement to salience calculation in resolving event pronouns. However, since the semantic filter was constructed by using a set of hand-crafted constraints on some specific domains, this approach is not suitable for general event pronoun resolution.

Among learning-based methods to event anaphora resolution, Chen et al. (2010a) explored various kinds of positional, lexical and syntactic features for event pronoun resolution, which turned out quite different from entity pronoun resolution. Besides, they studied the importance of structured syntactic information by incorporating it into event pronoun resolution via a composite kernel. Finally, they explored the incorporation of negative instances from non-event anaphoric pronouns, the fine-tuning of the SVM model and the employment of the twin-candidate model (Yang et al. 2003) in event pronoun resolution. Chen et al. (2010b) extended their previous work from event pronoun resolution to general event anaphora resolution by considering other types of event anaphors. Kong and Zhou (2011) proposed a new tree expansion scheme to automatically determine a proper parse tree structure for event pronoun resolution by considering various kinds of competitive information related with the anaphor and the antecedent candidate and achieved a much better performance on the OntoNotes English corpus than Chen et al (2010a).

Besides, there are some studies which integrate event anaphora resolution with entity anaphora resolution. For example, Pradhan et al. (2007) proposed a unified event and entity anaphora resolution framework based on a set of widely-used features which have been proven to be effective for entity anaphora resolution. Evaluation on the OntoNotes English corpus shows that their unified framework achieved the performance of 51.2 in F1-measure on overall entity and event anaphora resolution. However, they did not report the performance of their unified framework on event anaphora resolution. Alternatively, Muller (2007) constructed a logistic regression model to resolve event and entity pronouns together. For event pronoun resolution, he achieved 11.94 in F1-measure and found that the types of information effective for event pronoun resolution were very different from those for entity pronoun resolution. From this aspect, it seems better to independently explore event anaphora resolution first and then explore its possible integration (e.g. joint learning) with entity anaphora resolution.

In this paper, we focus on improving event pronoun resolution by exploring both local and global semantic information.

4 Baseline System

Our event pronoun resolution framework adopts the common learning-based one for entity anaphora resolution, as described by Soon et al. (2001). Specially, the way generating instances during training and testing procedures of event pronoun resolution is similar to Kong and Zhou (2011).

4.1 Flat Features

For entity pronoun resolution, Yang et al. (2004, 2005, 2006) explored various kinds of syntactic and semantic features to describe the information related with the antecedent candidate and the anaphor from their own and the relationship between them. However, few of these features can be adopted in event pronoun resolution. On one hand, since the antecedent candidate is an event trigger and the anaphor is a pronoun, both carry little obvious information about their own. On the other hand, the event anaphor and candidate pair in event pronoun resolution consists of a predicate and a pronoun. The difference in syntactic categories introduces extra difficulties. The features, such as number agreement, gender agreement, name alias, string matching and head matching, which have been proven to be effective for entity pronoun resolution, will no longer function here. Instead, we employ a list of flat features as shown in Table 1, inspired by Chen et al. (2010a).

In principle, the features in Table 1 can be grouped into three categories:

1) *Positional features*: the intuition is that the antecedent of an event pronoun should be close to each other. In particular, different kinds of distances between the anaphor and

the antecedent candidate are explored, i.e. over sentence, word, pronoun, predicate and main predicate.

2) *Grammatical features*: mainly used to describe the grammatical roles of the anaphor and the antecedent candidate.
3) *Similarity feature*: There is no doubt that semantic information is important for event pronoun resolution. However, since both event pronouns and event triggers carry little obvious semantics of their own, the context similarity is employed to measure the semantic compatibility between the anaphor and its antecedent candidate. In our baseline system, the similarity is calculated based on a list of nearby 10 contextual words (including previous 5 words and following 5 words) with proper stemming using the Porter stemmer and stop words (such as "in", "the" and etc.) filtered out.

	Features	Description
Positional Features	SenDist	Sentence distance between event anaphor and antecedent candidate
	WordDist	Word distance between event anaphor and antecedent candidate
	PredDist	Number of predicates between event anaphor and antecedent candidate
	PronDist	Number of pronouns between event anaphor and antecedent candidate
	MPredDist	Number of main predicate between event anaphor and antecedent candidate
Grammatical Features	isAnaInMC	Whether event anaphor occurs in main clause
	isAnaSub	Whether event anaphor is at subject position
	isPredInMC	Whether antecedent candidate occurs in main clause
	isMPred	Whether antecedent candidate is main predicate
Similarity Feature	ContextSim	Similarity between event anaphor's context and antecedent candidate's context

TABLE 1 – FLAT FEATURES FOR EVENT PRONOUN RESOLUTION

4.2 Structured Parse Tree

Besides various kinds of flat features described above, we also explore structured syntactic information via a parse tree structure. The commonly-used syntactic knowledge for anaphora resolution, such as the governing relations, can be directly described by the parse tree structure. Other syntactic knowledge that may be helpful for anaphora resolution could also be implicitly represented in the parse tree structure. Furthermore, tree kernel-based methods have been explored in entity anaphora resolution and achieved comparable performance with the dominated feature-based methods (Yang et al. 2006; Zhou et al. 2008). For structured syntactic information, we adopt the Dynamic Competitive Tree, as proposed in Kong and Zhou (2011), which takes the related competitive information, such as the event pronoun predicate (i.e. the predicate of the event pronoun), event antecedent competitors and event pronoun competitors between the anaphor and the considered antecedent candidate into consideration. For more details, please refer to Kong and Zhou (2011).

4.3 Polynomial Kernel, Tree Kernel and Composite Kernel

For two vectors of flat features, we compute their similarity using a polynomial kernel, while given any parse tree structure, e.g. the one as described above, the similarity between two parse trees is calculated using a convolution tree kernel. For more details about this kernel, please refer to Collins and Duffy (2002) and Moschitti (2004).

In order to combine flat features and structured parse trees, a linear-interpolated composite kernel is adopted in this paper:

$$K_{comp}(x_1, x_2) = \frac{K_{tree}(x1, x2)}{|K_{tree}(x1, x2)|} + \frac{K_{flat}(x1, x2)}{|K_{flat}(x1, x2)|}$$

For simplicity, this paper equally weights the convolution tree kernel K_{tree} for the parse tree structure and the polynomial kernel K_{flat} (d=2) for flat features with proper normalization.

5 Incorporating Semantic Information

It is well proven that the semantic compatibility between the anaphor and the antecedent candidate is important for entity anaphora resolution. For event pronoun resolution, since the anaphor is a pronoun and the antecedent candidate is an event trigger, both carry little obvious semantic information about themselves. Therefore, it is more difficult to measure the semantic compatibility between the anaphor and the antecedent candidate in event pronoun resolution. A possible way to measure it is to explore the contexts where the event pronoun and the antecedent candidate occur.

In our baseline system, we use a bag-of-words method to represent the contexts of the anaphor and the candidate. However, such a bag-of-words method suffers from the dilemma between the noise and the necessary information covered in a context window. In order to resolve this problem, we propose a predicate-argument structure representation to capture the local semantic information related with an event. Besides, we explore the global semantic information via entity coreference chains related with various arguments of the predicate-argument structure to complement its locality.

5.1 Local Semantic Information

Obviously, various arguments closely related with an event trigger contain necessary semantic information for event pronoun resolution. Therefore, it is reasonable to represent an event mention using the event trigger and corresponding event arguments. Since the event trigger is normally the predicate of a clause and event arguments correspond to the arguments driven by the predicate of a clause, we employ a shallow semantic parser to extract the predicate-argument structure of a clause as the representation of the local semantic information of the event. Especially for an event pronoun, we retrieve its governing predicate and related arguments as its local representation.

Figure1 shows the algorithm for computing the semantic compatibility between an event pronoun and an antecedent candidate. In particular, only those core arguments as defined in the Propbank (Palmer et al. 2005) are included in the computation. Consider following example:

a) ***Both President Gore and President Bush*** $_{\boldsymbol{1}}$ have held a flurry of news conferences in recent weeks and with each one ***they*** $_{\boldsymbol{1}}$ have ***[increased]*** *the* number of Stars and Stripes ***they*** $_{\boldsymbol{1}}$ use as a backdrop. I think ***they*** $_{\boldsymbol{1}}$ think ***[it]*** makes ***them*** $_{\boldsymbol{1}}$ look patriotic and presidential.

which includes an inter-sentence event anaphora with pronoun "it" referring to event "they have increased the number of Stars and Stripes". For this pair of event pronoun and antecedent, the algorithm in Figure1 returns {makes, it, them} and {increased, they, the number of Stars and Stripes} as AnaSet and CandSet, respectively.

Algorithm:

computing the semantic compatibility between an anaphor and an antecedent candidate

Input:

an anaphor: current event anaphor, a pronoun

an antecedent candidate: an event trigger, a predicate

Steps:

1) Initialize Score, CandSet and AnaSet
2) Use a semantic role labeling (SRL) toolkit to get all the core arguments of the antecedent candidate (i.e., event arguments) and add the antecedent candidate (i.e., the event trigger) and all the core arguments to CandSet (except pronouns).
3) Get the governing predicate of the anaphor and use a SRL toolkit to get all the core arguments of the predicate. Add the governing predicate of the anaphor and all the core arguments to AnaSet (except pronouns).
4) For every element pair from CandSet (candi) and AnaSet (anaj), compute the similarity between them using WordNet as described in Satanjeev and Ted (2002). If the similarity is larger than a threshold (0.5 in this paper), increase Score by 1.
5) Return Score/sqrt(|CandSet|*|AnaSet|) as the semantic compatibility between the given anaphor and the given antecedent candidate

FIGURE1– ALGORITHM FOR COMPUTING THE SEMANTIC SIMILARITY BETWEEN AN ANAPHOR AND AN ANTECEDENT CANDIDATE

5.2 Global Semantic Information

Obviously, the algorithm in Figure1 can only retrieve local descriptions of an event mention. It may be difficult to correctly measure the global semantic compatibility between the anaphor and the antecedent candidate using the predicate-argument structure due to its locality, e.g. considering example (d) due to the occurrence of various pronouns.

In order to complement the locality of the predicate-argument structure in representing the anaphor and the antecedent candidate, we further explore the global semantic information via entity coreference chains related with various arguments for event pronoun resolution.

The basic idea is that, when measuring the semantic compatibility between the anaphor and the antecedent candidate, we not only consider the event trigger and involved arguments, but also include different entity mentions (except pronouns) related with those involved arguments. Consider CandSet={increased, they, the number of Stars and Stripes} in the above example. Although we can retrieve little semantic information from argument "they" itself, we can find that it actually refers to "Both President Gore and President Bush" via entity anaphora resolution.

Similarly, we can find the referred expressions of other pronouns in CandSet and AnaSet. In this way, CandSet and AnaSet can be better represented.

6 Evaluation and Discussion

This section systematically evaluates the performance of our event pronoun resolution framework on the OntoNotes English corpus (Release 3.0).

6.1 Experimental Setting

The OntoNotes English corpus (Release 3.0) contains 300K words of English newswire data (from the Wall Street Journal) and 200K words of English broadcast news data (from ABC, CNN, NBC, Public Radio International and Voice of America). Table 2 shows the statistics of the corpus. From Table 2 we can find that about 9% of coreference chains are event related. Among 3550 event pronoun candidates (i.e. all the occurrences of “this”, “that” and “it”, which function as pronoun), 504 are event pronouns, accounting for about 14%. This indicates the difficulty of identifying event pronouns.

Category	Num
Coreference chains	8154
Coreference chains related with events	737
Event pronoun candidates	3550
Event pronouns	504

TABLE 2 – STATISTICS ON ONTONOTES 3.0 ENGLISH CORPUS

System	P(%)	R(%)	F
Polynomial kernel(flat features)	34.84	53.08	42.07
Tree kernel(structured)	47.06	65.71	54.84
Composite kernel(flat+structured)	49.78	69.17	57.89

TABLE 3 – PERFORMANCE OF THE BASELINE SYSTEM

For preparation, all the documents in the corpus are pre-processed automatically using a pipeline of NLP components. In addition, the corpus is parsed using the Charniak parser, and a state-of-the-art semantic role labelling (SRL) toolkit as proposed by Li et al. (2009) is employed to extract the predicate-argument structure (i.e. various arguments of a predicate)[1]. Finally, we use the SVM-light toolkit (Joakim, 1998)[2] with the convolution tree kernel SVMlight–TK (Moschitti, 2004)[3] for computing the similarity between two parse trees, and the polynomial kernel (d=2) for computing the similarity between two vectors of flat features, with learning parameters same as Chen et al. (2010a). For performance evaluation, we report the performance of event pronoun resolution with 10-fold cross validation in terms of recall, precision, and F1-measure.

[1] In this paper, we only consider verbal predicates, since 97.3% of events in the OntoNotes English corpus (Release 3.0) are triggered by verbal predicates. Besides, only those core arguments as defined in the Propbank are explored in this paper. For reference, the toolkit developed by Li et al. (2009) achieved the performance of 81.8 in F1-measure on the CoNLL’2005 version of the Propbank.

[2] http://svmlight.joachims.org/

[3] http://ai-nlp.info.uniroma2.it/moschitti/

6.2 Experimental Results

6.2.1 Performance of baseline system

Table 3 shows the performance of our baseline system. It shows that the polynomial kernel with the flat features yields 42.07 in F1-measure while the convolution tree kernel with the parse tree structure achieves a much better performance of 54.84 in F1-measure due to direct modeling of commonly used syntactic knowledge and implicit including of other knowledge helpful for event pronoun resolution. It also shows that the flat features and the parse tree structure are quite complementary that their combination via a simple composite kernel achieves 57.89 in F1-measure. Although the employed dynamic competitive tree may lose some important contextual information, it prunes out potential noise as much as possible. At the same time, the flat features used in our baseline system mainly describe positional and grammatical information. This suggests the complementary nature of the flat features and the parse tree structure, which is justified by the effective use of the composite kernel. In all the following experiments, we only report the performance employing the composite kernel.

6.2.2 Contribution of local semantic information

Table 4 shows the contribution of local semantic information on the composite kernel. It shows that the local semantic information via the predicate-argument structure can significantly improve the performance on the composite kernel by 1.7 in F1-measure. This justifies the usefulness of the predicate-argument structure in representing the local semantic information of the anaphor and the antecedent candidate. It also shows that the inclusion of A0 (i.e. agent) improves the performance by 1.29 in F1-measure and further inclusion of A1 (i.e. recipient) contributes 0.36 in F1-measure while the effectiveness of other kinds of arguments is very limited due to their less number and their less possibility of being referred in a text.

System	P(%)	R(%)	F
Flat+Structured	49.78	69.17	57.89
+A0	53.04	66.92	59.18(+1.29)
++A1	53.59	66.98	59.54(+0.36)
+++Others	53.64	67.02	59.59(+0.05)

TABLE 4 – INCREMENTAL CONTRIBUTION OF DIFFERENT KINDS OF LOCAL SEMANTIC INFORMATION ON THE COMPOSITE KERNEL [4]

6.2.3 Contribution of global semantic information

While the global semantic information via entity coreference chains can complement the locality of the predicate-argument structure, entity anaphora resolution itself is a difficult task. Obviously, the effectiveness of the global semantic information will largely depend on the performance of entity anaphora resolution.

[4] Significance tests are conducted between each of them and the previous one. The *p*-values are all smaller than 0.01.

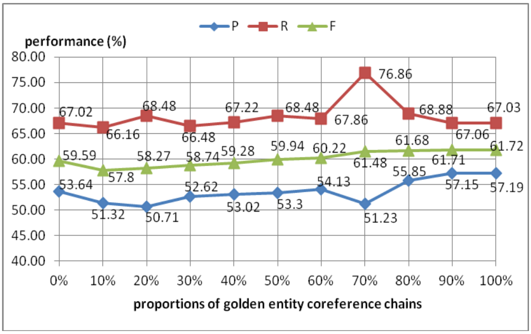

FIGURE 2– CONTRIBUTION OF GLOBAL SEMANTIC INFORMATION BY RANDOMLY CHOOSING DIFFERENT PROPORTIONS OF GOLDEN ENTITY COREFERENCE CHAINS

Figure 2 shows the contribution of global semantic information by randomly choosing different proportions of golden entity coreference chains. From Figure 2, we can find that:

1) Only considering a small proportion of even golden entity coreference chains (<50%) may harm the performance of event pronoun resolution. That is to say, only a small proportion of even golden entity coreference chains cannot complement the locality of the predicate-argument structure. In contrast, it may introduce too much uneven distribution across different events and thus harm the performance.
2) For golden entity coreference chains, including at least 50% begins to contribute.
3) When the used proportion of golden entity chains reaches a threshold (about 80%), the performance of event pronoun resolution stabilizes.

Systems	P(%)	R(%)	F
Soon et al. (2001) (duplicated)	61.5	45.9	52.57
Kong et al. (2009) (duplicated)	73.5	54.2	62.39

TABLE 5 – PERFORMANCE OF ENTITY ANAPHORA RESOLUTION

System	P(%)	R(%)	F
Flat+Structured+Local Semantic	53.64	67.02	59.59
+Global Semantic (Soon et al 2001)	53.6	67.36	59.70
+Global Semantic (Kong et al. 2009)	54.65	68.6	60.84

TABLE 6 – CONTRIBUTION OF AUTO ENTITY COREFERENCE CHAINS

In order to measure the effectiveness of global semantic information in practical environment, i.e. using automatic entity anaphora resolution, we duplicate two entity anaphora resolution systems with different levels of performance: the one proposed by Soon et al. (2001) and the other proposed by Kong et al. (2009). Table 5 shows the performance of these two duplicated systems on the OntoNote English corpus. Table 6 shows the contribution of global semantic information via entity coreference chains returned by the two automatic anaphora resolution systems. It shows that the effectiveness of global semantic information largely depends on the performance of an entity anaphora resolution system. Using the duplicated system proposed by Soon et al.(2001), we can only get a performance improvement of only 0.11 in F1-measure for event pronoun resolution while applying the duplicated system proposed by Kong et al. (2009) improve the performance by 1.25 in F1-measure. That is to say, the state-of-the-art entity anaphora resolution system can improve the performance of event pronoun resolution by filling the gap by about 50% (60.84-59.59 vs. 61.72-59.59). While most of the contribution of golden entity chains comes from gain in precision, the contribution of automatic entity chains comes from gain in both precision and recall.

6.2.4 Comparison with the State-of-the-Art

For comparison, Table 7 illustrates the performance of the state-of-the-art event pronoun resolution system developed by Chen et al. (2010a) using different schemes. From Table 7, we can find that Chen et al (2010a) achieved the performance of 40.6 in F1-measure via a feature-based method, and the best performance of 44.4 in F1-measure using the min-expansion tree via a tree kernel-based method. They further studied different ways of combining flat features and a parse tree structure to improve the performance and achieved the best performance of 47.2 in F1-measure when combining flat features with the simple-expansion tree structure. In our study, our feature-based system achieves 42.07 in F1-measure, and our tree kernel-based method with the dynamic competitive tree achieves the performance of 54.84 in F1-measure. By combining the flat features with the structured parse tree via a composite kernel, our system achieved the performance of 57.89 in F1-measure. The much better performance of our baseline system is mainly due to two reasons: 1) our better preprocessing in filtering out unnecessary negative instances by employing a set of constraints as described in Byron (2002). In Chen et al. (2010a), each event pronoun will generate 6.93 candidates while the number in our system is reduced to about 3; 2) employing the more effective structured tree span.

System	P(%)	R(%)	F
Flat	40.6	40.6	40.6
Min-Expansion	35.5	59.6	44.4
Simple-Expansion +Flat	42.3	53.4	47.2
+Negative Instances w/ Sampling	59.9	50.6	54.9
++SVM Fine-tuning	65.2	49.2	56.1
+++Twin-Candidate Modelling	62.6	54.0	57.9

TABLE 7 – PERFORMANCE OF CHEN ET AL. (2010A) ON EVENT PRONOUN RESOLUTION

Besides, Chen et al. (2010a) looked into the incorporation of negative instances from non-event anaphoric pronouns and achieved the best performance of 54.9 in F1-measure. They further improved the performance by keeping certain training data as the development data to help SVM select a more accurate hyper plane and achieved the performance of 56.1 in F1-measure. Finally, they proceeded to apply the twin-candidate model as proposed in Yang et al. (2003) to event pronoun resolution and achieved the performance of 57.9 in F1-measure. After employing so many strategies, Chen et al. (2010a) achieved the comparable performance with our baseline system. This justifies the strength of our baseline system.

We further improve the performance of event pronoun resolution by combining local and global semantic information via the predicate-argument structure and entity coreference chains and achieve the outstanding performance of 60.84 in F1-measure using automatic semantic role labelling and entity anaphora resolution.

Conclusion and Further Work

This paper studies the impact of both local and global semantic information for event pronoun resolution. In particular, a predicate-argument structure is proposed to represent the local semantic information related with an event while the global semantic information via entity coreference chains is further incorporated to complement the locality of the predicate-argument structure. Experimental results show that both the local and global semantic information are very effective for event pronoun resolution. We also study the influence of the performance of entity anaphora resolution on event pronoun resolution.

For further work, we will explore more structured syntactic information and semantic information in event anaphora resolution. In addition, we will study joint learning of entity anaphora resolution and event anaphora resolution.

Acknowledgment

This research was supported by Projects 61003153, 61273320 and 61272257 under the National Natural Science Foundation of China, Project 2012AA011102 under the National 863 Program of China, Project 11KJA520003 under the Natural Science Major Fundamental Research Program of the Jiangsu Higher Education Institutions, Project SYG201112 under the Applied Basic Research of Suzhou.

References

Asher N. 1993. Reference to Abstract Objects in Discourse. *Kluwer Academic Publisher*.

Byron D. 2002. Resolving Pronominal Reference to Abstract Entities. *ACL'2002*

Banerjee S. and Pedersen T. 2002. An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet .*CICLing'2002*.

Chen B., Su J. and Tan C.L. 2010(a). A Twin-Candidate Based Approach for Event Pronoun Resolution using Composite Kernel. *COLING'2010*

Chen B., Su J. and Tan C.L. 2010(b). Resolving Event Noun Phrases to Their Verbal Mentions. *EMNLP'2010*

Collins M. and Duffy N. 2001. Covolution kernels for natural language. *NIPS' 2001*

Kong F., Zhou G.D. and Zhu Q.M. 2009 Employing the Centering Theory in Pronoun Resolution from the Semantic Perspective. *EMNLP'2009*

Kong F., Li Y.C., Zhou G.D. and Zhu Q.M. 2009. Exploring Syntactic Features for Pronoun Resolution Using Context-Sensitive Convolution Tree Kernel. *IALP'2009*

Kong F., Zhou G.D., Qian L.H. and Zhu Q.M. 2010. Dependency-driven Anaphoricity

Determination for Coreference Resolution. *COLING'2010*

Kong F. and Zhou G.D. 2010. A Tree Kernel-based Unified Framework for Chinese Zero Anaphora Resolution. *EMNLP'2010*

Kong F. and Zhou G.D. 2011. Improve Tree Kernel-Based Event Pronoun Resolution with Competitive Information. *IJCAI'2011*

Laura H. and Constantin O. 2009. Do coreferential arguments make event mentions coreferential?. *DAARC'2009*

Li J.H, Zhou G.D., Zhao H., Zhu Q.M. and Qian P.D. 2009. Improving nominal SRL in Chinese language with verbal SRL information and automatic predicate recognition. *EMNLP '2009*

Moschitti A. 2004. A Study on Convolution Kernels for Shallow Semantic Parsing, *ACL'2004*

Muller C. 2007. Resoving it, this, and that in unrestricted multi-party dialog. *ACL'2007*

Ng V. and Cardie C. 2002. Improving machine learning approaches to coreference resolution. *ACL'2002*

Ng V. 2009. Graph-cut based anaphoricity determination for coreference resolution. *NAACL'2009*

Palmer M., Gildea D. and Kingsbury P. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics 31(1):71-105*

Pradhan S., Ramshaw L., Weischedel R., Macbride J. and Micciulla L. 2007. Unrestricted Coreference: Identifying Entities and Events in OntoNotes. *ICSC'2007*

Satanjeev B. and Ted P. 2002. An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet .*CICLing'2002*

Soon W.M., Ng H.T. and Lim D. 2001. A machine learning approach to coreference resolution of noun phrase. *Computational Linguistics, 27(4):521-544.*

Yang X.F., Su J., Zhou G.D. and Tan C.L. 2004. Improving pronoun resolution by incorporating coreferential information of candidates. *ACL'2004*

Yang X.F., Su J. and Chew C.L., 2006. Kernel-based pronoun resolution with structured syntactic knowledge. *COLING-ACL'2006*

Yang X.F., Su J. and Tan C.L. 2008. A Twin-Candidate Model for Learning-Based Anaphora Resolution. *Computational Linguistics 34(3):327-356*

Zhou G.D., Kong F. and Zhu Q.M. 2008. Context-sensitive convolution tree kernel for pronoun resolution. *IJCNLP'2008*

Zhou G.D. and Kong F. 2009. Global Learning of Noun Phrase Anaphoricity in Coreference Resolution via Label Propagetion. *EMNLP'2009*