Employing the Centering Theory in Pronoun Resolution from the Semantic Perspective

KONG Fang   ZHOU GuoDong[*]   **ZHU Qiaoming**

JiangSu Provincial Key Lab for Computer Information Processing Technology

School of Computer Science and Technology

Soochow University. Suzhou, China 215006

Email: {kongfang, gdzhou, qmzhu}@suda.edu.cn

Abstract

In this paper, we employ the centering theory in pronoun resolution from the semantic perspective. First, diverse semantic role features with regard to different predicates in a sentence are explored. Moreover, given a pronominal anaphor, its relative ranking among all the pronouns in a sentence, according to relevant semantic role information and its surface position, is incorporated. In particular, the use of both the semantic role features and the relative pronominal ranking feature in pronoun resolution is guided by extending the centering theory from the grammatical level to the semantic level in tracking the local discourse focus. Finally, detailed pronominal subcategory features are incorporated to enhance the discriminative power of both the semantic role features and the relative pronominal ranking feature. Experimental results on the ACE 2003 corpus show that the centering-motivated features contribute much to pronoun resolution.

1 Introduction

Coreference accounts for cohesion in a text and is, in a sense, the hyperlink for a natural language. Especially, a coreference instance denotes an identity of reference and holds between two referring expressions, which can be named entities, definite noun phrases, pronouns and so on. Coreference resolution is the process of linking together multiple referring expressions of a given entity in the world. The key in coreference resolution is to determine the antecedent for each referring expression in a text. The ability of linking referring expressions both within a sentence and across the sentences in a text is critical to discourse and language understanding in general. For example, coreference resolution is a key task in information extraction, machine translation, text summarization, and question answering.

There is a long tradition of research on coreference resolution within computational linguistics. While earlier knowledge-lean approaches heavily depend on domain and linguistic knowledge (Carter 1987; Carbonell and Brown 1988) and have significantly influenced the research, the later approaches usually rely on diverse lexical, syntactic and semantic properties of referring expressions (Soon et al., 2001;Ng and Cardie, 2002; Zhou et al., 2004). Current research has been focusing on exploiting semantic information in coreference resolution. For example, Yang et al (2005) proposed a template-based statistical approach to compute the semantic compatibility between a pronominal anaphor and an antecedent candidate, and Yang and Su (2007) explored semantic relatedness information from automatically discovered patterns, while Ng (2007) automatically induced semantic class knowledge from a treebank and explored its application in coreference resolution.

Particularly, this paper focuses on the centering theory (Sidner,1981;Grosz et al.,1995; Tetreault,2001), which reveals the significant impact of the local focus on referring expressions in that the antecedent of a referring expression usually depends on the center of attention throughout the local discourse segment (Mitkov,1998). Although the centering theory has been considered as a critical theory and the driving force behind the coreferential phenomena since its proposal, its application in coreference resolution (in particular pronoun resolution) has been somewhat disappointing: it fails to improve or even harms the performance of the state-of-

---

[*] Corresponding author

the-art coreference resolution systems in previous research (e.g. Yang et al. 2004). This may be due to that centering was originally proposed as a model of discourse coherence instead of coreference.

The purpose of this paper is to employ the centering theory in pronoun resolution by extending it from the grammatical level to the semantic level. The intuition behind our approach is that, via determining the semantic roles of referring expressions in a sentence, such as agent and patient, we can derive various centering theory-motivated features in tracking the continuity or shift of the local discourse focus, thus allowing us to include document-level event descriptive information in resolving the coreferential relations between referring expressions.

To the best of our knowledge, this is the first research, which successfully applies the centering theory in pronoun resolution from the semantic perspective.

The rest of this paper is organized as follows. Section 2 briefly describes related work in employing the centering theory and semantic information in coreference resolution. Then, the centering theory is introduced in Section 3 while Section 4 details how to employ the centering theory from the semantic perspective. Section 5 reports and discusses the experimental results. Finally, we conclude our work in Section 6.

2 Related Work

This section briefly overviews the related work in coreference resolution from both the centering theory and semantic perspectives.

2.1 Centering Theory

In the literature, there has been much research in the centering theory and its application to coreference resolution.

In the centering theory itself, since the original work of Sidner (1979) on immediate focusing of pronouns and the subsequent work of Joshi and Weinstein (1981) on centering and inferences, much research has been done, including centering and linguistic realizations (Cote 1993; Prince and Walker 1995), empirical and psycholinguistic evaluation of centering predictions (Gordon et al 1993,1995; Brennan 1995; Walker et al 1998; Kibble 2001), and the cross-linguistic work on centering (Ziv and Crosz1994).

In applications of the centering theory to coreference resolution, representative work includes Brennan et al. (1987), Strube (1998), Tetreault (1999) and Yang et al. (2004). Brennan et al. (1987) presented a centering theory-based formalism in modeling the local focus structure in discourse and used it to track the discourse context in binding occurring pronouns to corresponding entities. In particular, a BFP (Brennan, Friedman and Pollard) algorithm is proposed to extend the original centering model to include two additional transitions called smooth shift and rough shift. Strube (1998) proposed an S-list model, assuming that a referring expression prefers a hearer-old discourse entity to other hearer-new candidates. Tetreault (1999) further advanced the BFP algorithm by adopting a left-to-right breadth first walk of the syntactic parse trees to rank the antecedent candidates. However, the above methods have not been systematically evaluated on large annotated corpora, such as MUC and ACE. Thus their effects are still unclear in real coreference resolution tasks. Yang et al (2004) presented a learning-based approach by incorporating several S-list model-based features to improve the performance in pronoun resolution. It shows that, although including S-list model-based features can slightly boost the performance in the ideal case (i.e. given the correct antecedents of anaphor's candidates), it deteriorates the overall performance in F-measure with slightly higher precision but much lower recall, in real cases, where the antecedents of anaphor's candidates are determined automatically by a separate coreference resolution module.

2.2 Semantic Information

It is well known that semantic information plays a critical role in coreference resolution. Besides the common practice of employing a thesaurus (e.g. WordNet) in semantic consistency checking, much research has been done to explore various kinds of semantic information, such as semantic similarity (Harabagiu et al 2000), semantic compatibility (Yang et al 2005, 2007), and semantic class information (Soon et al 2001; Ng 2007). Although these methods have been proven useful in coreference resolution, their contributions are much limited. For example, Ng (2007) showed that semantic similarity information and semantic agreement information could only improve the performance of coreference resolution by 0.6 and 0.5 in F-measure respectively, on the ACE 2003 NWIRE corpus.

3 Centering Theory

The centering theory is a theory about the local discourse structure that models the interaction of referential continuity and the salience of discourse entities in the internal organization of a text. In natural languages, a given entity may be referred by different expressions and act as different grammatical roles throughout a text. For example, people often use pronouns to refer to the main subject of the discourse in focus, which can change over different portions of the discourse. One main goal of the centering theory is to track the focus entities throughout a text.

The main claims of the centering theory can be formalized in terms of C_b (the backward-looking center), C_f (a list of forward-looking centers for each utterance U_n) and C_p (the preferred center, i.e. the most salient candidate for subsequent utterances). Given following two sentences: 1) Susan[i] gave Betsy[j] a pet hamster[k]; 2) She[i] reminded her[j] that such hamsters[k] were quite shy. We can have U_b, U_f and U_p as follows: U_b= "Susan"; U_f={"Susan", "Betsy", "a pet hamster"}; U_p= "Susan".

	$C_b(U_n)=C_b(U_{n-1})$ or $C_b(U_{n-1})$ undefined	$C_b(U_n) \neq C_b(U_{n-1})$
$C_b(U_n)=C_p(U_n)$	Continue	Smooth Shift
$C_b(U_n) \neq C_p(U_n)$	Retain	Rough Shift

Table 1: Transitions in the centering theory

Constraints
C1. There is precisely one C_b.
C2. Every element of $C_f(U_n)$ must be realized in U_n.
C3. $C_b(U_n)$ is the highest-ranked element of $C_f(U_{n-1})$ that is realized in U_n.

Rules
R1. If some element of $C_f(U_{n-1})$ is realized as a pronoun in U_n, then so is $C_b(U_n)$.
R2. Transitions have the descending preference order of "Continue > Retain > Smooth Shift > Rough Shift".

Table 2: Constraints and rules in the centering theory

Furthermore, several kinds of focus transitions are defined in terms of two tests: whether C_b stays the same (i.e. $C_b(U_{n+1})=C_b(U_n)$), and whether C_b is realized as the most prominent referring expression (i.e. $C_b(U_n=C_p(U_n))$). We refer to the first test as cohesion, and the second test as salience. Therefore, there are four possible combinations, which are displayed in Table 1 and can result in four kinds of transitions, namely Continue, Retain, Smooth Shift, and Rough Shift. Obviously, salience, which chooses a proper verb form to make C_b prominent within a clause or sentence, is an important matter for sentence planning, while cohesion, which orders propositions in a text to maintain referential continuity, is an important matter for text planning.

Finally, the centering theory imposes several constraints and rules over C_b/C_f and above transitions, as shown in Table 2.

Given the centering theory as described above, we can draw the following conclusions:
1) The centering theory is discourse-related and centers are discourse constructs.
2) The backward-looking center C_b of U_n depends only on the expressions that constitute the utterance. That is, it is independent of its surface position and grammatical roles. Moreover, it is not constrained by any previous utterance in the segment. While the elements of $C_f(U_n)$ are partially ordered to reflect relative prominence in U_n, grammatical role information is often a major determinant in ranking C_f, e.g. in the descending priority order of "Subject > Object > Others" in English (Grosz and Joshi, 2001).
3) Psychological research (Gordon et al. 1993) and cross-linguistic research (Kameyama 1986, 1988; Walker et al. 1990,1994) have validated that C_b is preferentially realized by a pronoun in English.
4) Frequent rough shifts would lead to a lack of local cohesion. To keep local cohesion, people tend to plan ahead and minimize the number of focus shifts.

In this paper, we extend the centering theory from the grammatical level to the semantic level in attempt to better model the continuity or shift in the local discourse focus and improve the performance of pronoun resolution via centering-motivated semantic role features.

4 Employing Centering Theory from Semantic Perspective

In this section, we discuss how to employ the centering theory in pronoun resolution from the semantic perspective. In Subsection 4.1, we introduce the semantic roles. In Subsection 4.2, we introduce how to employ the centering theory in pronoun resolution via semantic role features. Finally we compare our method with the previous work in Subsection 4.3.

4.1 Semantic Role

A semantic role is the underlying relationship that a participant has with a given predicate in a clause, i.e. the actual role a participant plays in

an event, apart from linguistic encoding of the situation. If, in some situation, someone named "John" purposely hits someone named "Bill", then "John" is the agent and "Bill" is the patient of the hitting event. Therefore, given the predicate "hit" in both of the following sentences, "John" has the same semantic role of agent and "Bill" has the same semantic role of patient: 1) John hit Bill. 2) Bill was hit by John.

In the literature, labeling of such semantic roles has been well defined by the SRL (Semantic Role Labeling) task, which first identifies the arguments of a given predicate and then assigns them appropriate semantic roles. During the last few years, there has been growing interest in SRL. For example, CoNLL 2004 and 2005 have made this problem a well-known shared task. However, there is still little consensus in the linguistic and NLP communities about what set of semantic role labels are most appropriate. Typical semantic roles include core roles, such as agent, patient, instrument, and adjunct roles (such as locative, temporal, manner, and cause). For core roles, only agent and patient are consistently defined across different predicates, e.g. in the popular PropBank (Palmer et al. 2005) and the derived version evaluated in the CoNLL 2004 and 2005 shared tasks, as ARG0 and ARG1.

In this paper, we extend the centering theory from the grammatical level to the semantic level for its better application in pronoun resolution via proper semantic role features due to three reasons:

Sentence	Grammatical Role	Semantic Role
Bob opened the door with a key.	Bob: SUBJECT	Bob: AGENT
The key opened the door.	The key: SUBJECT	The key : INSTRUMENT
The door opened.	The door: SUBJECT	The door: PATIENT

Table 3: Relationship between grammatical roles and semantic roles: an example

1) Semantic roles are conceptual notions, whereas grammatical roles are morph-syntactic. While the original centering theory mainly builds from the grammatical perspective and grammatical roles do not always correspond directly to semantic roles (Table 3 shows an example of various semantic roles which a subject can play), there is a close relationship between semantic roles and grammatical roles. The statistics in the CoNLL 2004 and 2005 shared tasks (Shen and Lapata, 2007) shows that the semantic roles of

ARG0/agent and ARG1/patient account for 85% of all arguments and most likely act as the centers of the local focus structure in discourse due to the close relationship between subject/object and agent/patient. Therefore, it is appropriate to model the centers of an utterance from the semantic perspective via semantic roles.

2) In a sense, semantic roles imply the information of grammatical roles, especially for subject/object. For example, the position of an argument and the voice of the predicate verb play a central role in SRL. In intuition, an argument, which occurs before an active verb and has the semantic role of Arg0/agent, tends to be a subject. That is to say, semantic roles (e.g. Arg0/agent and Arg1/patient) can be mapped into their corresponding grammatical roles (e.g. subject and object), using some heuristic rules. Therefore, it would be interesting to represent the centers of the utterances and employ the centering theory from the semantic perspective.

3) Semantic role labeling has been well studied in the literature and there are good ready-to-use toolkits available. For example, Pradhan (2005) achieved 82.2 in F-measure on the CoNLL 2005 version of the Propbank. In contrast, the research on grammatical role labeling is much less with the much lower state-of-the-art performance of 71.2 in F-measure (Buchholz, 1999). Therefore, it may be better to explore the centering theory from the semantic perspective.

4.2 Designing Centering-motivated Features from Semantic Perspective

In this paper, the centering theory is employed in pronoun resolution via three kinds of centering-motivated features:

1) Semantic role features. They are achieved by checking possible semantic roles of referring expressions with regard to various predicates in a sentence. Due to the close relationship between subject/object and agent/patient, semantic role information should be also a major determinant in deciding the center of an utterance, which is likely to be the antecedent of a referring expression in the descending priority order of "Agent > Patient > Others" with regard to their semantic roles, corresponding to the descending priority order of "Subject > Object > Others" with regard to their grammatical roles.

2) Relative pronominal ranking feature. Due to the predominance of pronouns in tracking the local discourse structure[1], the relative ranking of a pronoun among all the pronouns in a sentence should be useful in pronoun resolution. This is realized in this paper according to its semantic roles (with regard to various predicates in a sentence) and surface position (in a left-to-right order) by mapping each pronoun into 5 levels: a) rank 1 for pronouns with semantic role ARG0/agent of the main predicate; b) rank 2 for pronouns with semantic role ARG1/patient of the main predicate; c) rank 3 for pronouns with semantic role ARG0/agent of other predicates; d) rank 4 for pronouns with semantic role ARG1/patient of other predicates; e) rank 5 for remaining pronouns. Furthermore, for those pronouns with the same ranking level, they are ordered according to their surface positions in a left-to-right order, motivated by previous research on the centering theory (Grosz et al. 1995).

3) Detailed pronominal subcategory features. Given a pronominal expression, its detailed pronominal subcategory features, such as whether it is a first person pronoun, second person pronoun, third person pronoun, neuter pronoun or others, are explored to enhance the discriminative power of both the semantic role features and the relative pronominal ranking feature, considering the predominant importance of pronouns in tracking the local focus structure in discourse.

4.3 Comparison with Previous Work

As a representative in explicitly employing semantic role labeling in coreference resolution, Ponzetto and Strube (2006) explored two semantic role features to capture the predicate-argument structure information to benefit coreference resolution: I_SEMROLE, the predicate-argument pairs of one referring expression, and J_SEMROLE, the predicate-argument pairs of another referring expression. Their experiments on the ACE 2003 corpus shows that, while the two semantic role features much improve the performance of common noun resolution by 3.8 and 2.7 in F-measure on the BNEWS and NWIRE domains respectively, they only

slightly improve the performance of pronoun resolution by 0.4 and 0.3 in F-measure on the BNEWS and NWIRE domains respectively.

In comparison, this paper proposes various kinds of centering-motivated semantic role features in attempt to better model the continuity or shift in the local discourse focus by extending the centering theory from the grammatical level to the semantic level. For example, the CAARG0MainVerb feature (as shown in Table 5) is designed to capture the semantic role of the antecedent candidate in the main predicate in modeling the discourse center, while, the AN-PronounRanking feature (as shown in Table 5) is designed to determinate the relative priority of the pronominal anaphor in retaining the discourse center.

Although both this paper and Ponzetto and Strube (2006) employs semantic role features, their ways of deriving such features are much different due to different driving forces/motivations behind. As a result, their contributions on coreference resolution are different: while the semantic role features in Ponzette and Strube (2006) captures the predicate-argument structure information and contributes much to common noun resolution and their contribution on pronoun resolution can be ignored, the centering-motivated semantic role features in this paper contribute much in pronoun resolution. This justifies our attempt to better model the continuity or shift of the discourse focus in pronoun resolution by extending the centering theory from the grammatical level to the semantic level and employing the centering-motivated features in pronoun resolution..

5 Experimentation and Discussion

We have evaluated our approach of employing the centering theory in pronoun resolution from the semantic perspective on the ACE 2003 corpus.

5.1 Experimental Setting

The ACE 2003 corpus contains three domains: newswire (NWIRE), newspaper (NPAPER), and broadcast news (BNEWS). For each domain, there exist two data sets, training and devtest, which are used for training and testing respectively. Table 4 lists the pronoun distributions with coreferential relationships in the training data and the test data over pronominal subcategories and sentence distances. Table 4(a) shows that third person pronouns occupy most and neu-

---

[1] According to the centering theory, the backward-looking center C_b is preferentially realized by a pronoun in the subject position in natural languages, such as English, and people tend to plan ahead and minimize the number of focus shifts.

tral pronouns occupy second while Table 4(b) shows that the antecedents of most pronouns occur within the current sentence and the previous sentence, with a little exception in the test data set of BNEWS.

Pronoun Subcategory	NWIRE		NPAPER		BNEWS	
	Train	Test	Train	Test	Train	Test
First Person	263	103	283	120	455	258
Second Person	61	16	29	36	203	68
Third Person	618	179	919	263	736	158
Neuter	395	151	577	190	482	137
Reflexive	23	6	42	12	26	6
Other	0	0	2	0	2	3

(a) Distribution over pronominal subcategories

Distance	NWIRE		NPAPER		BNEWS	
	Train	Test	Train	Test	Train	Test
0	890	254	1281	347	1149	295
1	447	149	529	197	729	188
2	0	27	0	24	0	41
>2	0	19	0	41	0	**100**
Total	1337	449	1810	609	1878	624

(b) Distribution over sentence distances

Table 4: Pronoun statistics on the ACE 2003 corpus

For preparation, all the documents in the corpus are preprocessed automatically using a pipeline of NLP components, including tokenization and sentence segmentation, named entity recognition, part-of-speech tagging and noun phrase chunking. Among them, named entity recognition, part-of-speech tagging and noun phrase chunking apply the same Hidden Markov Model (HMM)-based engine with error-driven learning capability (Zhou and Su, 2000 & 2002). In particular for SRL, we use a state-of-the-art in-house toolkit, which achieved the precision of 87.07% for ARG0 identification and the precision of 78.97% for ARG1 identification, for easy integration. In addition, we use the SVM-light[2] toolkit with the radial basis kernel and default learning parameters. Finally, we report the performance in terms of recall, precision, and F-measure, where precision measures the percentage of correctly-resolved pronouns (i.e. correctly linked with any referring expression in the coreferential chain), recall measures the coverage of correctly-resolved pronouns, and F-measure gives an overall figure on equal harmony between precision and recall. To see whether an improvement is significant, we also conduct significance testing using paired t-test. In this paper, '>>>', '>>' and '>' denote p-values of an improvement smaller than 0.01, in-between (0.01, 0,05] and bigger than 0.05,

---

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

which mean significantly better, moderately better and slightly better, respectively.

5.2 Experimental Results

Table 5 details various centering-motivated features from the semantic perspective, which are incorporated in our final system. For example, the CAARG0MainVerb feature is designed to capture the semantic role of the antecedent candidate in the main predicate in modeling the discourse center, while the ANPronounRanking feature is designed to determinate the relative priority of the pronominal anaphor in retaining the discourse center. As the baseline, we duplicated the representative system with the same set of 12 basic features, as described in Soon et al (2001). Table 6 shows that our baseline system achieves the state-of-the-art performance of 62.3, 65.3 and 59.0 in F-measure on the NWIRE, NPAPER and BNEWS domains, respectively. It also shows that the centering-motivated features (from the semantic perspective) significantly improve the F-measure by 3.6(>>>), 4.5(>>>) and 7.7(>>>) on the NWIRE, NPAPER and BNEWS domains, respectively. This justifies our attempt to model the continuity or shift of the discourse focus in pronoun resolution via centering-motivated features from the semantic perspective. For comparison, we also evaluate the performance of our final system from the grammatical perspective. This is done by replacing semantic roles with grammatical roles in deriving centering-motivated features. Here, labeling of grammatical roles is achieved using a state-of-the-art toolkit, as described in Buchholz (1999). Table 6 shows that properly employing the centering theory in pronoun resolution from the grammatical perspective can also improve the performance. However, the performance improvement of employing the centering theory from the grammatical perspective is much lower, compared with that from the semantic perspective. This validates our attempt of employing the centering theory in pronoun resolution from the semantic perspective instead of from the grammatical perspective. This also suggests the great potential of applying the centering theory in pronoun resolution since the centering theory is a local coherence theory, which tells how subsequent utterances in a text link together.

Table 7 shows the contribution of the semantic role features and the relative pronominal ranking feature in pronoun resolution when the detailed pronominal subcategory features are included:

Feature category	Feature	Remarks
Semantic Role-based Features	CAARG0	1 if the semantic role of the antecedent candidate is ARG0/agent; else 0
	CAARG0MainVerb	1 if the antecedent candidate has the semantic role of ARG0/agent for the main predicate of the sentence; else 0
	ANCASameTarget	1 if the anaphor and the antecedent candidate share the same predicate with regard to their semantic roles; else 0
Relative Pronominal Ranking Feature	ANPronounRanking	Whether the pronominal anaphor is ranked highest among all the pronouns in the sentence
Detailed Pronominal Subcategory Features	ANPronounType	Whether the anaphor is a first person, second person, third person, neuter pronoun or others
	CAPronounType	Whether the antecedent candidate is a first person, second person, third person, neuter pronoun or others

Table 5: Centering-motivated features incorporated in our final system
(with AN indicating the anaphor and CA indicating the antecedent candidate)

System Variation	NWIRE			NPAPER			BNEWS		
	R%	P%	F	R%	P%	F	R%	P%	F
Baseline System	*57.0*	***68.6***	*62.3*	*61.1*	*70.1*	*65.3*	*49.0*	*73.9*	*59.0*
Final System (from the semantic perspective)	***64.1***	*67.8*	***65.9***	***67.5***	***72.4***	***69.8***	***59.9***	***75.3***	***66.7***
Final System (from the grammatical perspective, for comparison)	*63.3*	*64*	*63.6*	*64.7*	*68.8*	*66.7*	*57.1*	*70.1*	*63.1*

Table 6: Contributions of centering-motivated features in pronoun resolution

System Variation	NWIRE			NPAPER			BNEWS		
	R%	P%	F	R%	P%	F	R%	P%	F
Baseline System	57.0	**68.6**	62.3	61.1	70.1	65.3	49.0	73.9	59.0
+SR and DC	**64.8**	67.8	**66.3**	67.2	**72.9**	**69.9**	59.1	**75.3**	66.3
+PR and DC	61.5	65.4	63.4	64.9	72.1	68.3	57.4	73.5	64.5
+SR, PR and DC (Final System)	64.1	67.8	65.9	**67.5**	72.4	69.8	**59.9**	**75.3**	**66.7**

Table 7: Contribution of the semantic role features (SR) and the relative pronominal ranking feature (PR) in pronoun resolution when the detailed pronominal subcategory features are included

1) The inclusion of the semantic role features improve the performance by 4.0(>>>), 4.6(>>>) and 7.3(>>>) in F-measure on the NWIRE, NPAPER and BNEWS domains, respectively. This suggests the impact of semantic role information in determining the local discourse focus. Since pronouns preferentially occur in the subject position and tend to refer to the main subject (Ehrlich 1980; Brennan 1995; Walker et al. 1998; Cahn 1995; Gordon and Searce 1995; Kibble et al. 2001), this paper only applies semantic features related with the semantic role of ARG0/agent, which is closely related with the grammatical role of subject, with regard to various predicates in a sentence. We have also explored features related with other semantic roles. However, our preliminary experimentation shows that they do not improve the performance, even for ARG1/patient, and thus are not included in the final system. This may be due to that other semantic roles are not discriminative enough to make a difference in deciding the local discourse structure.

2) It is surprising to notice that further inclusion of the relative pronominal ranking feature has only slight impact (slight positive impact on the BNEWS domain and slight negative impact on the NWIRE and NPAPER domains) on the ACE 2003 corpus. This suggests that most of information in the relative pronominal ranking feature has been covered by the semantic role features. This is not surprising since the semantic role of ARG0/agent, which is explored to derive the semantic role features, is also applied to decide the relative pronominal ranking feature.

The inclusion of the relative pronominal ranking feature improve the performance by 1.1(>>>), 3.0(>>>) and 5.5(>>>) in F-measure. Our further evaluation reveals that the performance improvement difference among different domains of the ACE 2003 corpus is due to the distribution of pronouns' antecedents occurring over different sentence distances, as shown in

Table 4. This suggests the usefulness of the relative pronominal ranking feature in resolving pronominal anaphors over longer distance. This is consistent with our observation that, as the percentage of pronominal anaphors referring to more distant antecedents increase, its impact turns gradually from negative to positive, when further including the relative pronominal ranking feature after the semantic role features. The reason that we include the detailed pronominal subcategory information is due to predominant importance of pronouns in tracking the local focus structure in discourse and that such detailed pronominal subcategory information is discriminative in tracking different subcategories of pronouns. This suggests the usefulness of considering the distribution of the local discourse focus over detailed pronominal subcategories. One interesting finding in our preliminary experimentation is that the inclusion of the detailed pronominal subcategory features alone even harms the performance. This may be due to the reason that the detailed pronominal subcategory features do not have the discriminative power themselves and that the semantic role features and the relative pronominal ranking feature provide an effective mechanism to explore the role of such detailed pronominal subcategory features in helping determine the local discourse focus.

	Pronoun Subcategory	NWIRE	NPAPER	BNEWS
Baseline System	First Person	55.7	55.9	56.6
	Second Person	54.6	60.4	44.0
	Third Person	72.6	80.9	75.7
	Neuter	41.5	50.4	50.2
	Reflexive	**85.7**	**70.0**	60.0
	Total	*62.3*	*65.3*	*59.0*
Final System	First Person	**64.7**	**67.0**	**65.6**
	Second Person	**78.6**	**70.0**	**51.9**
	Third Person	**80.9**	**81.8**	**80.4**
	Neuter	**48.3**	**53.0**	**58.3**
	Reflexive	71.4	66.7	**80.0**
	Total	*65.9*	*69.8*	*66.7*

Table 8: Performance comparison of pronoun resolution in F-measure over pronoun subcategories

Table 8 shows the contribution of the centering-motivated features over different pronoun subcategories. It shows that the centering-motivated features contribute much to the resolution of the four major pronoun subcategories (i.e. first person, second person, third person and neuter) while its negative impact on the minor pronoun subcategories (e.g. reflexive) can be ignored due to their much less frequent occur-

rence in the corpus. In particular, the centering-motivated features improve the performance on the major three pronoun subcategories of third person / neuter / first person, by 8.3(>>>)/6.8(>>>)/9.0(>>>), 0.9(>>)/ 2.6 (>>>)/11.1(>>>) and 4.7(>>>)/8.1(>>>)/9.0 (>>>), on the NWIRE, NPAPER and BNEWS domains of the ACE 2003 corpus, respectively.

	Distance	NWIRE	NPAPER	BNEWS
Baseline System	<=0	61.6	64.5	68.7
	<=1	60.4	67.5	60.0
	<=2	62.9	67.4	63.7
	Total	*62.3*	*65.3*	*59.0*
Final System	<=0	**64.3**	**70.3**	**78.7**
	<=1	**66.8**	**72.3**	**72.5**
	<=2	**66.6**	**71.8**	**71.8**
	Total	*65.9*	*69.8*	*66.7*

Table 9: Performance comparison of pronoun resolution in F-measure over sentence distances

Table 9 shows the contribution of the centering-motivated features over different sentence distances. It shows that the centering-motivated features improve the performance of pronoun resolution on different sentence distances of 0/1/2, by 2.7(>>>) / 5.8(>>>) / 10.0 (>>>), 6.4(>>>) / 4.8(>>>) / 12.5(>>>) and 3.7 (>>>)/4.4(>>>)/8.1(>>>), on the NWIRE, NPAPER and BNEWS domains of the ACE 2003 corpus, respectively. This suggests that the centering-motivated features are helpful for both intra-sentential and inter-sentential pronoun resolution.

6 Conclusion and Further Work

This paper extends the centering theory from the grammatical level to the semantic level and much improves the performance of pronoun resolution via centering-motivated features from the semantic perspective. This is mainly realized by employing various semantic role features with regard to various predicates in a sentence, in attempt to model the continuity or shift of the local discourse focus. Moreover, the relative ranking feature of a pronoun among all the pronouns is explored to help determine the relative priority of the pronominal anaphor in retaining the local discourse focus. Evaluation on the ACE 2003 corpus shows that both the centering-motivated semantic role features and pronominal ranking feature much improve the performance of pronoun resolution, especially when the detailed pronominal subcategory features of both the anaphor and the antecedent candidate are included. It is not surprising due to the predomi-

nance of pronouns in tracking the local discourse structure in a text.

To our best knowledge, this paper is the first research which successfully applies the centering-motivated features in pronoun resolution from the semantic perspective.

For future work, we will explore more kinds of semantic information and structured syntactic information in pronoun resolution. In particular, we will further employ the centering theory in pronoun resolution from both grammatical and semantic perspectives on more corpora.

Acknowledgement

References

C. Aone and W.W. Bennett. 1995. Evaluating automated and manual acquisition of anaphora resolution strategies. *ACL'*1995:122-129.

S.E. Brennan, M.W. Friedman, and C.J. Pollard. 1987. A centering approach to pronoun. *ACL'1987:* 290-297.

S.E. Brennan 1995. Centering attention in discourse. *Language and Cognitive Process*, 10/2: 137-67.

S. Buchholz, J. Veenstra and W. Daelemans. 1999. Cascaded Grammatical Relation Assignment. *EMNLP-VLC'1999*: 239-246

S. Cote. 1983. Ranking and forward-looking centers. *In Proceedings of the Workshop on the Centering Theory in Naturally-Occurring Discourse*. 1993.

D.M. Carter. 1987. Interpreting Anaphors in Natural Language Texts. *Ellis Horwood*, Chichester, UK.

J. Carbonell and R. Brown. 1988. Anaphora resolution: a multi-strategy approach. *COLING'1988*: 96-101.

B.J. Grosz, A.K. JoShi and S. Weinstein. 1995. Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics,* 21(2):203-225.

P.C. Gordon, B.J. Grosz and L.A. Gilliom. 1993. Pronouns, names and the centering of attention in discourse. *Cognitive Science*.1993.17(3):311-348

P.C. Gordon and K. A. Searce. 1995. Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory and Cognition*. 1995.

S. Harabagiu and S. Maiorano. 2000. Multiligual coreference resolution. *ANLP-NAACL* 2000:142-149.

A.K. Joshi and S. Weinstein. 1981. Control of inference: Role of some aspects of discourse structure-centering. *IJCAI'1981*:385-387

R. Kibble. 2001. A Reformulation of Rule 2 of Centering. *Computational Linguistics,* 2001,27(4): 579-587

M. Kameyama. 1986. Aproperty-sharing constraint in centering. *ACL* 1986:200-206

M. Kameyama. 1988. Japanese zero pronominal binding, where syntax and discourse meet. *In Proceeding of the Second International Workshop on Japanese Syntax.* 1988.

R. Mitkov. 1998. Robust pronoun resolution with limited knowledge. *COLING-ACL'*1998:869-875.

A. Moschitti and S. Quarteroni. 2008. Kernels on linguistics structures for answer entraction. *ACL'*08:113-116

V. Ng and C. Cardie. 2002. Improving machine learning approaches to coreference resolution. *ACL'2002*: 104-111

V. Ng. 2007. Semantic Class Induction and Coreference Resolution. *ACL'*2007 536-543.

M. Palmer, D. Gildea and P. Kingsbury. 2005. The proposition bank: A corpus annotated with semantic roles. *Computational Linguistics,* 31(1):71-106.

S. Pradhan, K. Hacioglu, V. Krugler, W. Ward, J.H. Martin, and D. Jurafsky. 2005. Support vector learning for semantic argument classification. *Machine Learning,* 2005.60:11-39.

S. P. Ponzetto and M. Strube. 2006. Semantic Role Labeling for Coreference Resolution. *EMNLP'*2006 143-146.

E. F. Prince and M. A. Walker. 1995. A bilateral apporach to givenness: a hearer-status algorithm and a centering algorithm. *In Proceedings of 4th International Pragmatics Conference.*

E. Rich and S. LuperFoy. 1988. An architecture for anaphora resolution. *In Proceedings of the 2nd Conference on Applied Natural Language Processing. ANLP'*1988: 18-24.

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

D. Shen and M. Lapata. 2007. Using semantic roles to improve question answering. *EMNLP-CoNIL* 2007:12-21

C. Sidner. 1979. Toward a computation of intrasentential coreference. *Technical Report TR-537,MIT. Artificial Intelligence Laboratory.*

C. Sidner. 1981. Focusing for interpretation of pronouns. *Computational Linguistics*,1981.7:217-231

J. Tetreault. 1999. Analysis of syntax-based pronoun resolution methods. *ACL* 1999:602-605

J. Tetreault. 2001. A corpus-based evaluation of centering and pronoun resolution. *Computational Linguistics*. 2001. 27(4):507-520.

995

M. Walker, A. K. Joshi and E. Prince. 1998. Centering in naturally occurring discourse: *An overview. Clarendon Press*:1-28

X.F. Yang, J. Su, G.D. Zhou and C.L. Tan. 2004. Improving pronoun resolution by incorporating coreferential information of candidates. *ACL'*2004:127-134.

X.F. Yang, J. Su and C.L. Tan. 2005. Improving Pronoun Resolution Using Statistics - Based Semantic Compatibility Information. *ACL'*2005:165 -172.

X.F. Yang and J. Su. 2007. Coreference Resolution Using Semantic Relatedness Information from Automatically Discovered Patterns. *ACL'2007*: 528-535.

G.D. Zhou and J. Su. 2004. A high- performance coreference resolution system using a multi- agent strategy. *COLING'* 2004:522- 528.

Y. Ziv and B.J. Grosz. 1994. Right dislocation and attentional state. *Israel Association of Theoretical Linguistics Meetings'*1994. 184-199.