

Exploiting Syntactic Patterns as Clues in Zero-Anaphora Resolution

Ryu Iida, Kentaro Inui and Yuji Matsumoto

Graduate School of Information Science,

Nara Institute of Science and Technology

8916-5 Takayama, Ikoma, Nara, 630-0192, Japan

{ryu-i,inui,matsu}@is.naist.jp

Abstract

We approach the zero-anaphora resolution problem by decomposing it into intra-sentential and inter-sentential zero-anaphora resolution. For the former problem, syntactic patterns of the appearance of zero-pronouns and their antecedents are useful clues. Taking Japanese as a target language, we empirically demonstrate that incorporating rich syntactic pattern features in a state-of-the-art learning-based anaphora resolution model dramatically improves the accuracy of intra-sentential zero-anaphora, which consequently improves the overall performance of zero-anaphora resolution.

1 Introduction

Zero-anaphora is a gap in a sentence that has an anaphoric function similar to a pro-form (e.g. pronoun) and is often described as “referring back” to an expression that supplies the information necessary for interpreting the sentence. For example, in the sentence “*There are two roads to eternity, a straight and narrow, and a broad and crooked,*” the gaps in “*a straight and narrow (gap)*” and “*a broad and crooked (gap)*” have a zero-anaphoric relationship to “*two roads to eternity.*”

The task of identifying zero-anaphoric relations in a given discourse, *zero-anaphora resolution*, is essential in a wide range of NLP applications. This is the case particularly in such a language as Japanese, where even obligatory arguments of a predicate are often omitted when they are inferable from the context. In fact, in our Japanese newspaper corpus, for example, 45.5% of the nominative arguments of verbs are omitted. Since such gaps

can not be interpreted only by shallow syntactic parsing, a model specialized for zero-anaphora resolution needs to be devised on the top of shallow syntactic and semantic processing.

Recent work on zero-anaphora resolution can be located in two different research contexts. First, zero-anaphora resolution is studied in the context of anaphora resolution (AR), in which zero-anaphora is regarded as a subclass of anaphora. In AR, the research trend has been shifting from rule-based approaches (Baldwin, 1995; Lappin and Leass, 1994; Mitkov, 1997, etc.) to empirical, or corpus-based, approaches (McCarthy and Lehnert, 1995; Ng and Cardie, 2002a; Soon et al., 2001; Strube and Müller, 2003; Yang et al., 2003) because the latter are shown to be a cost-efficient solution achieving a performance that is comparable to best performing rule-based systems (see the Coreference task in MUC¹ and the Entity Detection and Tracking task in the ACE program²). The same trend is observed also in Japanese zero-anaphora resolution, where the findings made in rule-based or theory-oriented work (Kameyama, 1986; Nakaiwa and Shirai, 1996; Okumura and Tamura, 1996, etc.) have been successfully incorporated in machine learning-based frameworks (Seki et al., 2002; Iida et al., 2003).

Second, the task of zero-anaphora resolution has some overlap with Propbank³-style semantic role labeling (SRL), which has been intensively studied, for example, in the context of the CoNLL SRL task⁴. In this task, given a sentence “*To attract younger listeners, Radio Free Europe intersperses the latest in Western rock groups*”, an SRL

¹http://www-nlpir.nist.gov/related_projects/muc/

²<http://projects ldc.upenn.edu/ace/>

³http://www.cis.upenn.edu/~mpalmer/project_pages/ACE.htm

⁴<http://www.lsi.upc.edu/~srlconll/>

model is asked to identify the NP *Radio Free Europe* as the A0 (Agent) argument of the verb *attract*. This can be seen as the task of finding the zero-anaphoric relationship between a nominal gap (the A0 argument of *attract*) and its antecedent (*Radio Free Europe*) under the condition that the gap and its antecedent appear in the same sentence.

In spite of this overlap between AR and SRL, there are some important findings that are yet to be exchanged between them, partly because the two fields have been evolving somewhat independently. The AR community has recently made two important findings:

- A model that identifies the antecedent of an anaphor by a series of comparisons between candidate antecedents has a remarkable advantage over a model that estimates the absolute likelihood of each candidate independently of other candidates (Iida et al., 2003; Yang et al., 2003).
- An AR model that carries out antecedent identification *before* anaphoricity determination, the decision whether a given NP is anaphoric or not (i.e. discourse-new), significantly outperforms a model that executes those subtasks in the reverse order or simultaneously (Poesio et al., 2004; Iida et al., 2005).

To our best knowledge, however, existing SRL models do not exploit these advantages. In SRL, on the other hand, it is common to use syntactic features derived from the parse tree of a given input sentence for argument identification. A typical syntactic feature is the path on a parse tree from a target predicate to a noun phrase in question (Gildea and Jurafsky, 2002; Carreras and Marquez, 2005). However, existing AR models deal with intra- and inter-sentential anaphoric relations in a uniform manner; that is, they do not use as rich syntactic features as state-of-the-art SRL models do, even in finding intra-sentential anaphoric relations. We believe that the AR and SRL communities can learn more from each other.

Given this background, in this paper, we show that combining the aforementioned techniques derived from each research trend makes significant impact on zero-anaphora resolution, taking Japanese as a target language. More specifically, we demonstrate the following:

- Incorporating rich syntactic features in a state-of-the-art AR model dramatically im-

proves the accuracy of intra-sentential zero-anaphora resolution, which consequently improves the overall performance of zero-anaphora resolution. This is to be considered as a contribution to AR research.

- Analogously to inter-sentential anaphora, decomposing the antecedent identification task into a series of comparisons between candidate antecedents works remarkably well also in intra-sentential zero-anaphora resolution. We hope this finding to be adopted in SRL.

The rest of the paper is organized as follows. Section 2 describes the task definition of zero-anaphora resolution in Japanese. In Section 3, we review previous approaches to AR. Section 4 described how the proposed model incorporates effectively syntactic features into the machine learning-based approach. We then report the results of our experiments on Japanese zero-anaphora resolution in Section 5 and conclude in Section 6.

2 Zero-anaphora resolution

In this paper, we consider only zero-pronouns that function as an obligatory argument of a predicate for two reasons:

- Providing a clear definition of zero-pronouns appearing in adjunctive argument positions involves awkward problems, which we believe should be postponed until obligatory zero-anaphora is well studied.
- Resolving obligatory zero-anaphora tends to be more important than adjunctive zero-pronouns in actual applications.

A zero-pronoun may have its antecedent in the discourse; in this case, we say the zero-pronoun is *anaphoric*. On the other hand, a zero-pronoun whose referent does not explicitly appear in the discourse is called a *non-anaphoric* zero-pronoun. A zero-pronoun may be non-anaphoric typically when it refers to an extralinguistic entity (e.g. the first or second person) or its referent is unspecified in the context.

The following are Japanese examples. In sentence (1), zero-pronoun ϕ_i is anaphoric as its antecedent, ‘*shusho* (prime minister)’, appears in the same sentence. In sentence (2), on the other hand, ϕ_j is considered non-anaphoric if its referent (i.e. the first person) does not appear in the discourse.

- (1) *shusho_i-wa houbeyi-te ,*
 prime minister_i-TOP visit-U.S.-CONJ PUNC

ryoukoku-no *gaikou-o*
 both countries-BETWEEN diplomacy-OBJ

(ϕ_i -ga) *suishinsuru*
 (ϕ_i -NOM) promote-ADNOM

houshin-o *akirakanisi-ta* .
 plan-OBJ unveil-PAST PUNC

The prime minister visited the united states
 and unveiled the plan to push diplomacy
 between the two countries.

- (2) (ϕ_j -ga) *ie-ni* *kaeri-tai* .
 (ϕ_j -NOM) home-DAT want to go back PUNC
 (I) want to go home.

Given this distinction, we consider the task of zero-anaphora resolution as the combination of two sub-problems, antecedent identification and anaphoricity determination, which is analogous to NP-anaphora resolution:

For each zero-pronoun in a given discourse, find its antecedent if it is anaphoric; otherwise, conclude it to be non-anaphoric.

3 Previous work

3.1 Antecedent identification

Previous machine learning-based approaches to antecedent identification can be classified as either the *candidate-wise classification* approach or the *preference-based* approach. In the former approach (Soon et al., 2001; Ng and Cardie, 2002a, etc.), given a target anaphor, *TA*, the model estimates the absolute likelihood of each of the candidate antecedents (i.e. the NPs preceding *TA*), and selects the best-scored candidate. If all the candidates are classified negative, *TA* is judged non-anaphoric.

In contrast, the preference-based approach (Yang et al., 2003; Iida et al., 2003) decomposes the task into comparisons of the preference between candidates and selects the most preferred one as the antecedent. For example, Iida et al. (2003) proposes a method called the *tournament model*. This model conducts a tournament consisting of a series of matches in which candidate antecedents compete with each other for a given anaphor.

While the candidate-wise classification model computes the score of each single candidate independently of others, the tournament model learns the relative preference between candidates, which is empirically proved to be a significant advantage over candidate-wise classification (Iida et al., 2003).

3.2 Anaphoricity determination

There are two alternative ways for anaphoricity determination: the *single-step model* and the *two-step model*. The single-step model (Soon et al., 2001; Ng and Cardie, 2002a) determines the anaphoricity of a given anaphor indirectly as a by-product of the search for its antecedent. If an appropriate candidate antecedent is found, the anaphor is classified as anaphoric; otherwise, it is classified as non-anaphoric. One disadvantage of this model is that it cannot employ the preference-based model because the preference-based model is not capable of identifying non-anaphoric cases.

The two-step model (Ng, 2004; Poesio et al., 2004; Iida et al., 2005), on the other hand, carries out anaphoricity determination in a separate step from antecedent identification. Poesio et al. (2004) and Iida et al. (2005) claim that the latter subtask should be done before the former. For example, given a target anaphor (*TA*), Iida et al.’s *selection-then-classification* model:

1. selects the most likely candidate antecedent (*CA*) of *TA* using the tournament model,
2. classifies *TA* paired with *CA* as either *anaphoric* or *non-anaphoric* using an anaphoricity determination model. If the *CA-TA* pair is classified as *anaphoric*, *CA* is identified as the antecedent of *TA*; otherwise, *TA* is conclude to be *non-anaphoric*.

The anaphoricity determination model learns the non-anaphoric class directly from non-anaphoric training instances whereas the single-step model cannot not use non-anaphoric cases in training.

4 Proposal

4.1 Task decomposition

We approach the zero-anaphora resolution problem by decomposing it into two subtasks: intra-sentential and inter-sentential zero-anaphora resolution. For the former problem, syntactic patterns in which zero-pronouns and their antecedents appear may well be useful clues, which, however, does not apply to the latter problem. We therefore build a separate component for each subtask, adopting Iida et al. (2005)’s selection-then-classification model for each component:

1. *Intra-sentential antecedent identification*: For a given zero-pronoun *ZP* in a given sentence *S*, select the most-likely candidate antecedent C_1^* from the candidates appearing in *S* by the intra-sentential tournament model

2. *Intra-sentential anaphoricity determination*: Estimate plausibility p_1 that C_1^* is the true antecedent, and return C_1^* if $p_1 \geq \theta_{intra}$ (θ_{intra} is a preselected threshold) or go to 3 otherwise
3. *Inter-sentential antecedent identification*: Select the most-likely candidate antecedent C_2^* from the candidates appearing outside of S by the inter-sentential tournament model.
4. *Inter-sentential anaphoricity determination*: Estimate plausibility p_2 that C_2^* is the true antecedent, and return C_2^* if $p_2 \geq \theta_{inter}$ (θ_{inter} is a preselected threshold) or return non-anaphoric otherwise.

4.2 Representation of syntactic patterns

In the first two of the above four steps, we use syntactic pattern features. Analogously to SRL, we extract the parse path between a zero-pronoun to its antecedent to capture the syntactic pattern of their occurrence. Among many alternative ways of representing a path, in the experiments reported in the next section, we adopted a method as we describe below, leaving the exploration of other alternatives as future work.

Given a sentence, we first use a standard dependency parser to obtain the dependency parse tree, in which words are structured according to the dependency relation between them. Figure 1(a), for example, shows the dependency tree of sentence (1) given in Section 2. We then extract the path between a zero-pronoun and its antecedent as in Figure 1(b). Finally, to encode the order of siblings and reduce data sparseness, we further transform the extracted path as in Figure 1(c):

- A path is represented by a subtree consisting of backbone nodes: ϕ (zero-pronoun), Ant (antecedent), Node (the lowest common ancestor), LeftNode (left-branch node) and RightNode.
- Each backbone node has daughter nodes, each corresponding to a function word associated with it.
- Content words are deleted.

This way of encoding syntactic patterns is used in intra-sentential anaphoricity determination. In antecedent identification, on the other hand, the tournament model allows us to incorporate three paths, a path for each pair of a zero-pronoun and left and right candidate antecedents, as shown in

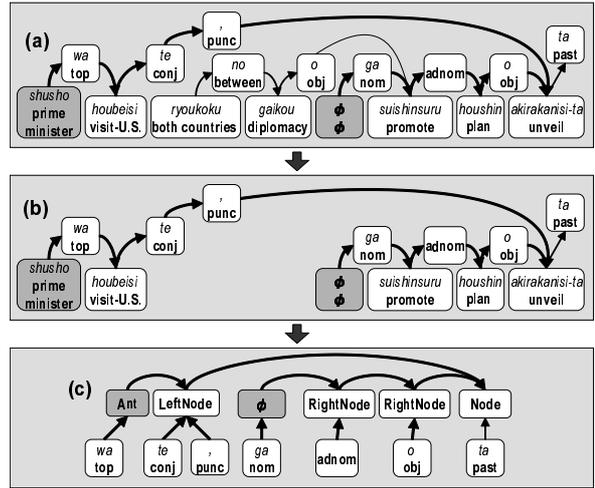


Figure 1: Representation of the path between a zero-pronoun to its antecedent

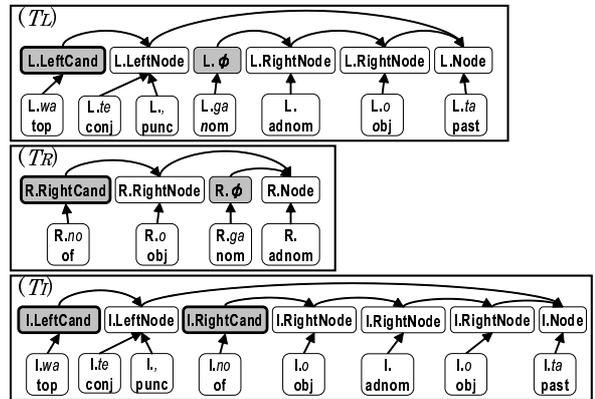


Figure 2: Paths used in the tournament model

Figure 2⁵.

4.3 Learning algorithm

As noted in Section 1, the use of zero-pronouns in Japanese is relatively less constrained by syntax compared, for example, with English. This forces the above way of encoding path information to produce an explosive number of different paths, which inevitably leads to serious data sparseness.

This issue can be addressed in several ways. The SRL community has devised a range of variants of the standard path representation to reduce the complexity (Carreras and Marquez, 2005). Applying Kernel methods such as Tree kernels (Collins and Duffy, 2001) and Hierarchical DAG kernels (Suzuki et al., 2003) is another strong option. The Boosting-based algorithm pro-

⁵To indicate which node belongs to which subtree, the label of each node is prefixed either with L, R or l.

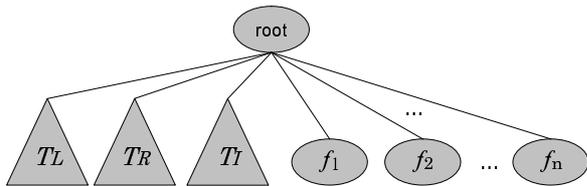


Figure 4: Tree representation of features for the tournament model.

posed by Kudo and Matsumoto (2004) is designed to learn subtrees useful for classification.

Leaving the question of selecting learning algorithms open, in our experiments, we have so far examined Kudo and Matsumoto (2004)’s algorithm, which is implemented as the BACT system⁶. Given a set of training instances, each of which is represented as a tree labeled either positive or negative, the BACT system learns a list of weighted decision stumps with a Boosting algorithm. Each decision stump is associated with tuple $\langle t, l, w \rangle$, where t is a subtree appearing in the training set, l a label, and w a weight, indicating that if a given input includes t , it gives w votes to l . The strength of this algorithm is that it deals with structured feature and allows us to analyze the utility of features.

In antecedent identification, we train the tournament model by providing a set of labeled trees as a training set, where a label is either left or right. Each labeled tree has (i) path trees T_L , T_R and T_I (as given in Figure 2) and (ii) a set nodes corresponding to the binary features summarized in Table 3, each of which is linked to the root node as illustrated in Figure 4. This way of organizing a labeled tree allows the model to learn, for example, the combination of a subtree of T_L and some of the binary features. Analogously, for anaphoricity determination, we use trees (T_C, f_1, \dots, f_n) , where T_C denotes a path subtree as in Figure 1(c).

5 Experiments

We conducted an evaluation of our method using Japanese newspaper articles. The following four models were compared:

1. **BM**: Ng and Cardie (2002a)’s model, which identify antecedents by the candidate-wise classification model, and determine anaphoricity using the one-step model.

2. **BM_STR**: BM with the syntactic features such as those in Figure 1(c).
3. **SCM**: The selection-then-classification model explained in Section 3.
4. **SCM_STR**: SCM with all types of syntactic features shown in Figure 2.

5.1 Setting

We created an anaphoric relation-tagged corpus consisting of 197 newspaper articles (1,803 sentences), 137 articles annotated by two annotators and 60 by one. The agreement ratio between two annotators on the 197 articles was 84.6%, which indicated that the annotation was sufficiently reliable.

In the experiments, we removed from the above data set the zero-pronouns to which the two annotators did not agree. Consequently, the data set contained 995 intra-sentential anaphoric zero-pronouns, 754 inter-sentential anaphoric zero-pronouns, and 603 non-anaphoric zero-pronouns (2,352 zero-pronouns in total), with each anaphoric zero-pronoun annotated to be linked to its antecedent. For each of the following experiments, we conducted five-fold cross-validation over 2,352 zero-pronouns so that the set of the zero-pronouns from a single text was not divided into the training and test sets.

In the experiments, all the features were automatically acquired with the help of the following NLP tools: the Japanese morphological analyzer *ChaSen*⁷ and the Japanese dependency structure analyzer *CaboCha*⁸, which also carried out named-entity chunking.

5.2 Results on intra-sentential zero-anaphora resolution

In both intra-anaphoricity determination and antecedent identification, we investigated the effect of introducing the syntactic features for improving the performance. First, the results of antecedent identification are shown in Table 1. The comparison between BM (SCM) with BM_STR (SCM_STR) indicates that introducing the structural information effectively contributes to this task. In addition, the large improvement from BM_STR to SCM_STR indicates that the use of the preference-based model has significant impact on intra-sentential antecedent identification. This

⁶<http://chasen.org/~taku/software/bact/>

⁷<http://chasen.naist.jp/hiki/ChaSen/>

⁸<http://chasen.org/~taku/software/cabocho/>

Figure 3: Feature set.

Feature Type	Feature	Description
Lexical	HEAD_BF	characters of right-most morpheme in <i>NP</i> (<i>PRED</i>).
Grammatical	PRED_IN_MATRIX	1 if <i>PRED</i> exists in the matrix clause; otherwise 0.
	PRED_IN_EMBEDDED	1 if <i>PRED</i> exists in the relative clause; otherwise 0.
	PRED_VOICE	1 if <i>PRED</i> contains auxiliaries such as ' <i>(ra)reru</i> '; otherwise 0.
	PRED_AUX	1 if <i>PRED</i> contains auxiliaries such as ' <i>(sa)seru</i> ', ' <i>hosii</i> ', ' <i>morau</i> ', ' <i>itadaku</i> ', ' <i>kudasaru</i> ', ' <i>yaru</i> ' and ' <i>ageru</i> '.
	PRED_ALT	1 if PRED_VOICE is 1 or PRED_AUX is 1; otherwise 0.
	POS	Part-of-speech of <i>NP</i> followed by IPADIC (Asahara and Matsumoto, 2003).
	DEFINITE	1 if <i>NP</i> contains the article corresponding to DEFINITE 'the', such as ' <i>sore</i> ' or ' <i>sono</i> '; otherwise 0.
Semantic	DEMONSTRATIVE	1 if <i>NP</i> contains the article corresponding to DEMONSTRATIVE 'that' or 'this', such as ' <i>kono</i> ', ' <i>ano</i> '; otherwise 0.
	PARTICLE	Particle followed by <i>NP</i> , such as ' <i>wa</i> (topic)', ' <i>ga</i> (subject)', ' <i>o</i> (object)'. Named entity of <i>NP</i> : PERSON, ORGANIZATION, LOCATION, ARTIFACT, DATE, TIME, MONEY, PERCENT or N/A.
	NE	Named entity of <i>NP</i> : PERSON, ORGANIZATION, LOCATION, ARTIFACT, DATE, TIME, MONEY, PERCENT or N/A.
	EDR_HUMAN	1 if <i>NP</i> is included among the concept 'a human being' or 'attribute of a human being' in EDR dictionary (Jap, 1995); otherwise 0.
	PRONOUN_TYPE	Pronoun type of <i>NP</i> . (e.g. ' <i>kare</i> (he)' → PERSON, ' <i>koko</i> (here)' → LOCATION, ' <i>sore</i> (this)' → OTHERS)
	SELECT_REST	1 if <i>NP</i> satisfies selectional restrictions in Nihongo Goi Taikai (Japanese Lexicon) (Ikehara et al., 1997); otherwise 0.
	COOC	the score of well-formedness model estimated from a large number of triplets $\langle \text{Noun}, \text{Case}, \text{Predicate} \rangle$ proposed by Fujita et al. (2004)
Positional	SENTNUM	Distance between <i>NP</i> and <i>PRED</i> .
	BEGINNING	1 if <i>NP</i> is located in the beginning of sentence; otherwise 0.
	END	1 if <i>NP</i> is located in the end of sentence; otherwise 0.
	PRED_NP	1 if <i>PRED</i> precedes <i>NP</i> ; otherwise 0.
	NP_PRED	1 if <i>NP</i> precedes <i>PRED</i> ; otherwise 0.
	DEP_PRED	1 if NP_i depends on <i>PRED</i> ; otherwise 0.
	DEP_NP	1 if <i>PRED</i> depends on NP_i ; otherwise 0.
IN_QUOTE	1 if <i>NP</i> exists in the quoted text; otherwise 0.	
Heuristic	CL_RANK	a rank of <i>NP</i> in forward looking-center list based on Centering Theory (Grosz et al., 1995)
	CL_ORDER	a order of <i>NP</i> in forward looking-center list based on Centering Theory (Grosz et al., 1995)

NP and *PRED* stand for a bunsetsu-chunk of a candidate antecedent and a bunsetsu-chunk of a predicate which has a target zero-pronoun respectively.

finding may well contribute to semantic role labeling because these two tasks have a large overlap as discussed in Section 1.

Second, to evaluate the performance of intra-sentential zero-anaphora resolution, we plotted recall-precision curves altering threshold parameter and θ_{inter} for intra-anaphoricity determination as shown in Figure 5, where recall R and precision P were calculated by:

$$R = \frac{\# \text{ of detected antecedents correctly}}{\# \text{ of anaphoric zero-pronouns}},$$

$$P = \frac{\# \text{ of detected antecedents correctly}}{\# \text{ of zero-pronouns classified as anaphoric}}.$$

The curves indicate the upperbound of the performance of these models; in practical settings, the parameters have to be trained beforehand.

Figure 5 shows that BM_STR (SCM_STR) outperforms BM (SCM), which indicates that incorporating syntactic pattern features works remarkably well for intra-sentential zero-anaphora

Table 1: Accuracy of antecedent identification.

BM	BM_STR	SCM	SCM_STR
48.0%	63.5%	65.1%	70.5%
(478/995)	(632/995)	(648/995)	(701/995)

resolution. Furthermore, SCM_STR is significantly better than BM_STR. This result supports that the former has an advantage of learning non-anaphoric zero-pronouns (181 instances) as negative training instances in intra-sentential anaphoricity determination, which enables it to reject non-anaphoric zero-pronouns more accurately than the others.

5.3 Discussion

Our error analysis reveals that a majority of errors can be attributed to the current way of handling quoted phrases and sentences. Figure 6 shows the difference in resolution accuracy between zero-pronouns appearing in a quotation

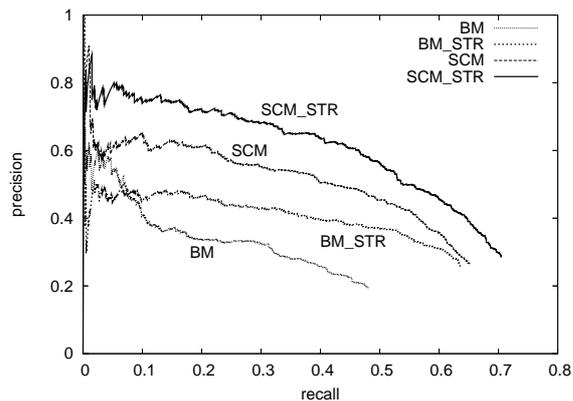


Figure 5: Recall-precision curves of intra-sentential zero-anaphora resolution.

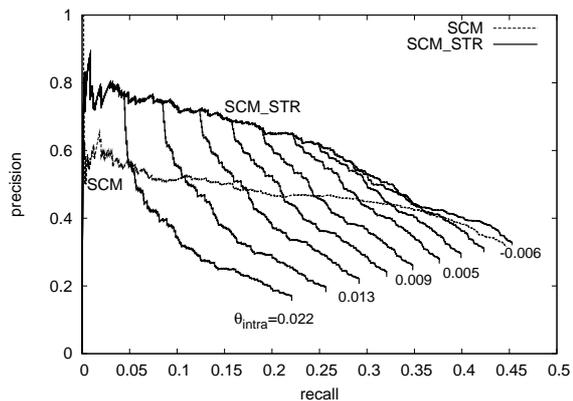


Figure 7: Recall-precision curves of overall zero-anaphora resolution.

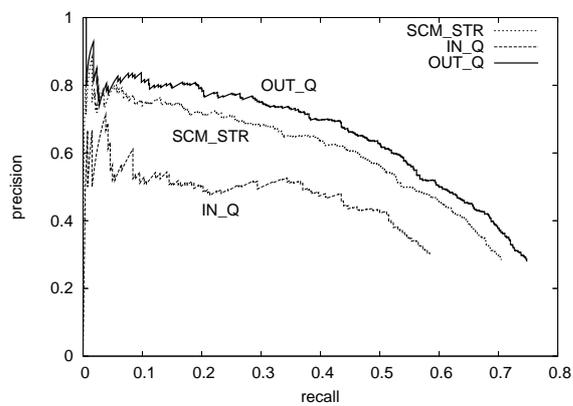


Figure 6: Recall-precision curves of resolving in-quote and out-quote zero-pronouns.

(262 zero-pronouns) and the rest (733 zero-pronouns), where “IN_Q” denotes the former (in-quote zero-pronouns) and “OUT_Q” the latter. The accuracy on the IN_Q problems is considerably lower than that on the OUT_Q cases, which indicates that we should deal with in-quote cases with a separate model so that it can take into account the nested structure of discourse segments introduced by quotations.

5.4 Impact on overall zero-anaphora resolution

We next evaluated the effects of introducing the proposed model on overall zero-anaphora resolution including inter-sentential cases.

As a baseline model, we implemented the original SCM, designed to resolve intra-sentential zero-anaphora and inter-sentential zero-anaphora simultaneously with no syntactic pattern features. Here, we adopted Support Vector Machines (Vapnik, 1998) to train the classifier on the baseline

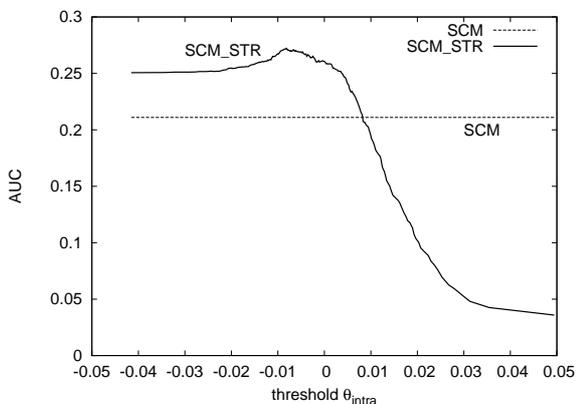


Figure 8: AUC curves plotted by altering θ_{intra} .

model and the inter-sentential zero-anaphora resolution in the SCM using structural information.

For the proposed model, we plotted several recall-precision curves by selecting different value for threshold parameters θ_{intra} and θ_{inter} . The results are shown in Figure 7, which indicates that the proposed model significantly outperforms the original SCM if θ_{intra} is appropriately chosen.

We then investigated the feasibility of parameter selection for θ_{intra} by plotting the AUC values for different θ_{intra} values. Here, each AUC value is the area under a recall-precision curve. The results are shown in Figure 8. Since the original SCM does not use θ_{intra} , the AUC value of it is constant, depicted by the SCM. As shown in the Figure 8, the AUC-value curve of the proposed model is not peaky, which indicates the selection of parameter θ_{intra} is not difficult.

6 Conclusion

In intra-sentential zero-anaphora resolution, syntactic patterns of the appearance of zero-pronouns and their antecedents are useful clues. Taking Japanese as a target language, we have empirically demonstrated that incorporating rich syntactic pattern features in a state-of-the-art learning-based anaphora resolution model dramatically improves the accuracy of intra-sentential zero-anaphora, which consequently improves the overall performance of zero-anaphora resolution.

In our next step, we are going to address the issue of how to find zero-pronouns, which requires us to design a broader framework that allows zero-anaphora resolution to interact with predicate-argument structure analysis. Another important issue is how to find a globally optimal solution to the set of zero-anaphora resolution problems in a given discourse, which leads us to explore methods as discussed by McCallum and Wellner (2003).

References

- M. Asahara and Y. Matsumoto. 2003. *IPADIC User Manual*. Nara Institute of Science and Technology, Japan.
- B. Baldwin. 1995. *CogNIAC: A Discourse Processing Engine*. Ph.D. thesis, Department of Computer and Information Sciences, University of Pennsylvania.
- X. Carreras and L. Marquez. 2005. Introduction to the conll-2005 shared task: Semantic role labeling. In *Proceedings of the Ninth CoNLL*, pages 152–164.
- M. Collins and N.I. Duffy. 2001. Convolution kernels for natural language. In *Proceedings of the NIPS*, pages 625–632.
- A. Fujita, K. Inui, and Y. Matsumoto. 2004. Detection of incorrect case assignments in automatically generated paraphrases of japanese sentences. In *Proceeding of the first IJCNLP*, pages 14–21.
- D. Gildea and D. Jurafsky. 2002. Automatic labeling of semantic roles. In *Computational Linguistics*, pages 245–288.
- B. J. Grosz, A. K. Joshi, and S. Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):203–226.
- R. Iida, K. Inui, H. Takamura, and Y. Matsumoto. 2003. Incorporating contextual cues in trainable models for coreference resolution. In *Proceedings of the 10th EACL Workshop on The Computational Treatment of Anaphora*, pages 23–30.
- R. Iida, K. Inui, and Y. Matsumoto. 2005. Anaphora resolution by antecedent identification followed by anaphoricity determination. *ACM Transactions on Asian Language Information Processing (TALIP)*, 4:417–434.
- S. Ikehara, M. Miyazaki, S. Shirai A. Yokoo, H. Nakaiwa, K. Ogura, Y. Ooyama, and Y. Hayashi. 1997. *Nihongo Goi Taikai (in Japanese)*. Iwanami Shoten.
- Japan Electronic Dictionary Research Institute, Ltd. Japan, 1995. *EDR Electronic Dictionary Technical Guide*.
- M. Kameyama. 1986. A property-sharing constraint in centering. In *Proceedings of the 24th ACL*, pages 200–206.
- T. Kudo and Y. Matsumoto. 2004. A boosting algorithm for classification of semi-structured text. In *Proceedings of the 2004 EMNLP*, pages 301–308.
- S. Lappin and H. J. Leass. 1994. An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 20(4):535–561.
- A. McCallum and B. Wellner. 2003. Object consolidation by graph partitioning with a conditionally trained distance metric. In *Proceedings of the KDD-2003 Workshop on Data Cleaning, Record Linkage, and Object Consolidation*, pages 19–24.
- J. F. McCarthy and W. G. Lehnert. 1995. Using decision trees for coreference resolution. In *Proceedings of the 14th IJCAI*, pages 1050–1055.
- R. Mitkov. 1997. Factors in anaphora resolution: they are not the only things that matter. a case study based on two different approaches. In *Proceedings of the ACL’97/EACL’97 Workshop on Operational Factors in Practical, Robust Anaphora Resolution*.
- H. Nakaiwa and S. Shirai. 1996. Anaphora resolution of japanese zero pronouns with deictic reference. In *Proceedings of the 16th COLING*, pages 812–817.
- V. Ng. 2004. Learning noun phrase anaphoricity to improve coreference resolution: Issues in representation and optimization. In *Proceedings of the 42nd ACL*, pages 152–159.
- V. Ng and C. Cardie. 2002a. Improving machine learning approaches to coreference resolution. In *Proceedings of the 40th ACL*, pages 104–111.
- M. Okumura and K. Tamura. 1996. Zero pronoun resolution in japanese discourse based on centering theory. In *Proceedings of the 16th COLING*, pages 871–876.
- M. Poesio, O. Uryupina, R. Vieira, M. Alexandrov-Kabadjov, and R. Goulart. 2004. Discourse-new detectors for definite description resolution: A survey and a preliminary proposal. In *Proceedings of the 42nd ACL Workshop on Reference Resolution and its Applications*, pages 47–54.
- K. Seki, A. Fujii, and T. Ishikawa. 2002. A probabilistic method for analyzing japanese anaphora integrating zero pronoun detection and resolution. In *Proceedings of the 19th COLING*, pages 911–917.
- W. M. Soon, H. T. Ng, and D. C. Y. Lim. 2001. A machine learning approach to coreference resolution of noun phrases. *Computational Linguistics*, 27(4):521–544.
- M. Strube and C. Müller. 2003. A machine learning approach to pronoun resolution in spoken dialogue. In *Proceedings of the 41st ACL*, pages 168–175.
- J. Suzuki, T. Hirao, Y. Sasaki, and E. Maeda. 2003. Hierarchical directed acyclic graph kernel: Methods for structured natural language data. In *Proceeding of the 41st ACL*, pages 32–39.
- V. N. Vapnik. 1998. *Statistical Learning Theory*. Adaptive and Learning Systems for Signal Processing Communications, and control. John Wiley & Sons.
- X. Yang, G. Zhou, J. Su, and C. L. Tan. 2003. Coreference resolution using competition learning approach. In *Proceedings of the 41st ACL*, pages 176–183.