Intra-sentential Zero Anaphora Resolution using Subject Sharing Recognition

Ryu Iida   Kentaro Torisawa   Chikara Hashimoto
Jong-Hoon Oh   Julien Kloetzer
National Institute of Information and Communications Technology
Kyoto 619-0289, Japan
{ryu.iida,torisawa,ch,rovellia,julien}@nict.go.jp

Abstract

In this work, we improve the performance of intra-sentential zero anaphora resolution in Japanese using a novel method of recognizing subject sharing relations. In Japanese, a large portion of intra-sentential zero anaphora can be regarded as subject sharing relations between predicates, that is, the subject of some predicate is also the unrealized subject of other predicates. We develop an accurate recognizer of subject sharing relations for pairs of predicates in a single sentence, and then construct a *subject shared predicate network*, which is a set of predicates that are linked by the subject sharing relations recognized by our recognizer. We finally combine our zero anaphora resolution method exploiting the subject shared predicate network and a state-of-the-art ILP-based zero anaphora resolution method. Our combined method achieved a significant improvement over the the ILP-based method alone on intra-sentential zero anaphora resolution in Japanese. To the best of our knowledge, this is the first work to explicitly use an independent subject sharing recognizer in zero anaphora resolution.

1 Introduction

In 'pro-dropped' languages such as Japanese, Chinese and Italian, pronouns are often unrealized in text. For example, the subject of *nomu* (take) is omitted in example (1).

(1) *Tom$_i$-wa   infuruenza-ni   natta-node   ,*
      Tom-TOP   flu-IOBJ   had-since   punc

   *(ϕ_i-ga)   kusuri-o   **non-da**   .*
   he$_i$-SUBJ   medicine-OBJ   took   period
   Since Tom$_i$ had the flu, (he$_i$) took medicine.

Such unrealized pronouns are regarded as *zero anaphors*, which are indicated using ϕ in literature, like ϕ_i-ga in example (1). Zero anaphor refers to its antecedent somewhere. This phenomenon of the reference is called *zero anaphora*. In Japanese, about 60% of subjects appear as zero anaphors in newspaper articles (Iida et al., 2007b), and thus zero anaphora resolution is an essential task for developing highly accurate machine translation and information extraction systems.

In this paper, we propose a novel method of resolving intra-sentential zero anaphora, in which a subject zero anaphor refers to its antecedent inside *a single sentence*. This work does not address inter-sentential zero anaphora, in which a zero anaphor in a sentence refers to its antecedent in *another sentence*. The novelty of our method is in the use of *subject sharing relations*, which are relations between two predicates that share a subject by (zero) anaphora or coreference. For example, in example (2), there are two subject sharing relations for predicate pairs, *advance-plan* and *plan-dispatch*, as illustrated in Figure 1.

(2) *seifu$_i$-wa   (ϕ_i-ga)   hisaichi-ni*
      government$_i$-TOP   it$_i$-SUBJ   disaster site-IOBJ

   *50 nin-o   **hakensuru**   koto-o   (ϕ_i-ga)*
   50 people-OBJ   dispatch   COMP-OBJ   it$_i$-SUBJ

   ***keikakusi**   junbisagyo-o   **susumeru**   .*
   plan   preparation-OBJ   advance   period
   The government$_i$ plans that (it$_i$) will dispatch 50 people to the disaster site and (it$_i$) is advancing its preparations.

The most straightforward method to recognize subject sharing relations is to apply a (zero) anaphora resolution system to a sentence and detect such relations by recognizing (zero) anaphora, like the relations represented by *seifu$_i$* and two zero anaphors ϕ_i in Figure 1. However, to our surprise, we found that a simple supervised classifier that exploits the local contexts surrounding
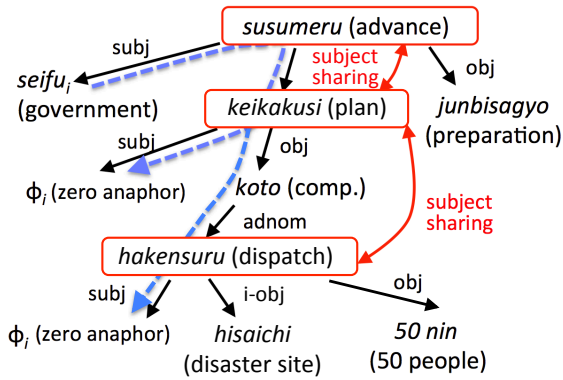
Figure 1: Example of subject shared predicate network

predicates achieved a higher accuracy than that of the straightforward method. This suggests that just *propagating* the realized subject of a predicate to the subject zero anaphor of other predicates through recognized subject sharing relations (e.g., propagating subject *government* of *advance* to the subject positions of *plan* and *dispatch* in Figure 1) might lead to a higher accuracy in zero anaphora resolution than the existing zero anaphora resolution methods. In addition, a large portion of zero anaphora can be regarded as subject sharing relations (e.g., 39% of the intra-sentential zero anaphora in the NAIST Text Corpus (Iida et al., 2007b) are such cases). Hence, just by combining our subject zero anaphora method with an existing general anaphora resolution method that covers other types of anaphora, significant improvement of accuracy over all types of anaphora might be achieved. This paper empirically shows that this is actually the case through a series of experiments in which we combine our method with an existing ILP-based zero anaphora resolution method (Iida and Poesio, 2011).

Our subject zero anaphora resolution method constructs a *subject shared predicate network (SSPN)*, which is a network of predicates in which subject sharing predicates are linked, from the results of our accurate *pairwise subject sharing recognizer*, which detects the predicate pairs that share a subject. Zero anaphora resolution is done by propagating the realized subject of a predicate to the subject zero anaphor of other predicates in the SSPN. An important point here is that SSPN was introduced to solve the issue related to our pairwise subject sharing recognizer. Our recognizer is applied only to the restricted

pairs of predicates in a sentence, such as predicates that have a *direct* dependency relation between them. This is because our current recognizer cannot achieve high accuracy for *any* pair of predicates. In Figure 1, for instance, our recognizer can detect a subject sharing relation between *advance* and *plan* and another between *plan* and *dispatch*, but it cannot detect one between *advance* and *dispatch*. However, in the SSPN, the undetected relations can be derived by connecting the two detected ones, and in the zero anaphora resolution subject *government* of *advance* can be successfully propagated to the subject position of *dispatch*.

The rest of our paper is organized as follows. In Section 2, we briefly overview previous work on zero anaphora resolution. In Section 3, we overview the procedure of our zero anaphora resolution method. We explain the three types of subject sharing relations on which we focus and propose a method of pairwise subject sharing recognition for the three types in Section 4. We evaluate how effectively our method recognizes subject sharing relations for these types in Section 5. After that, we investigate the impact of explicitly introducing SSPNs in Section 6 and compare our zero anaphora resolution method with a state-of-the-art ILP-based method on the task of intra-sentential subject zero anaphora resolution in Section 7. Finally, in Section 8 we summarize this work and discuss future directions.

2 Related work

Traditional approaches to zero anaphora resolution are based on manually created heuristic rules (Kameyama, 1986; Walker et al., 1994; Okumura and Tamura, 1996; Nakaiwa and Shirai, 1996), which are mainly motivated by the rules and preferences introduced in Centering Theory (Grosz et al., 1995). However, the research trend of zero anaphora resolution has shifted from such rule-based approaches to machine learning-based approaches because in machine learning we can easily integrate many different types of information, such as morpho-syntactic, semantic and discourse-related information. Researchers have developed methods of zero anaphora resolution for Chinese (Zhao and Ng, 2007; Chen and Ng, 2013), Japanese (Seki et al., 2002; Isozaki and Hirao, 2003; Iida et al., 2007a; Taira et al., 2008; Sasano et al., 2008; Sasano et al., 2009; Imamura

et al., 2009; Watanabe et al., 2010; Hayashibe et al., 2011; Iida and Poesio, 2011; Yoshikawa et al., 2011; Hangyo et al., 2013; Yoshino et al., 2013) and Italian (Iida and Poesio, 2011). One critical issue in zero anaphora resolution is optimizing the outputs of sub-problems (e.g., zero anaphor detection and antecedent identification). Recent works by Watanabe et al. (2010), Iida and Poesio (2011) and Yoshikawa et al. (2011) revealed that joint inference improves the overall performance of zero anaphora resolution. We employed one of these works as a baseline in Section 6.

Concerning subject sharing recognition, related methods have been explored for pronominal anaphora (Yang et al., 2005) or coreference resolution (Bean and Riloff, 2004; Bansal and Klein, 2012). In these methods, the semantic compatibility between the contexts surrounding an anaphor and its antecedent (e.g., the compatibility of verbs *kidnap* and *release* given some arguments) was automatically extracted from raw texts in an unsupervised manner and used as features in a machine learning-based approach. However, because the automatically acquired semantic compatibility is not always true or applicable in the context of any pair of an anaphor and its antecedent, the effectiveness of the compatibility features might be weakened. In contrast, we accurately recognize the explicit subject sharing relations and directly use them for propagating the subject of some predicate to the empty subject position of other predicates instead of indirectly using the relations as features.

3 Zero anaphora resolution using subject shared predicate network

In this section, we first give an overview of the procedure of our zero anaphora resolution method. Intra-sentential zero anaphora resolution in our method is performed in the following five steps, as depicted in Figure 2.

Step 1 The pairwise subject sharing relations between two predicates in a sentence are recognized by our subject sharing recognizer.

Step 2 A subject shared predicate network (SSPN) is constructed based on the results of pairwise subject sharing recognition.

Step 3 For each predicate in the set of the subject shared predicates in the SSPN, a subject is detected by our subject detector, if one exists.
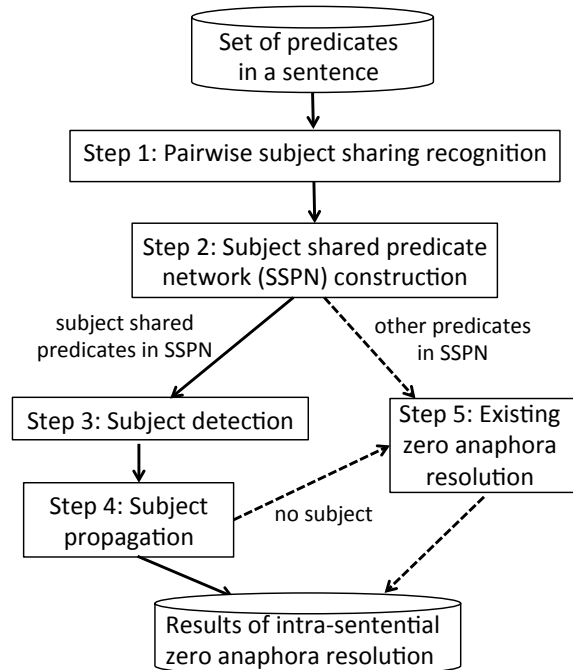

Figure 2: Procedure of our zero anaphora resolution method

Step 4 If a subject is detected, it is propagated to the empty subject position of each predicate in the subject shared predicates in the SSPN.

Step 5 For resolving the potential zero anaphora that were not resolved until Step 4, we apply the existing ILP-based method (Iida and Poesio, 2011).

We define subject sharing relations as follows. Two predicates have a subject sharing relation if and only if they share the same subject that is referred to by (zero) anaphora or coreference. Note that the shared subject does not need to be realized in the text; it can appear as inter-sentential zero anaphora or exophora. In Step 1, the pairwise subject sharing relations between two predicates are recognized, but recognizing the relations between any two predicates in a sentence remains difficult. We thus focus on some typical types of predicate pairs. The details of the predicate pair types will be explained in Section 4.1.

Given the results of pairwise subject sharing recognition, we construct an SSPN in Step 2. In an SSPN, every predicate in a sentence is a node and only the predicate pairs that were judged to be subject sharing are connected by a link. The major advantage of explicitly constructing an SSPN is that it enables us to resolve zero anaphora even

if a predicate with a subject zero anaphor does not have any direct subject sharing relation with a predicate with a subject, like predicates *susumeru* (advance) and *hakensuru* (dispatch) in Figure 1. By traversing the paths of the subject sharing relations in the SSPN, such predicates can be connected to successfully propagate the subject. The effect of introducing SSPNs is empirically evaluated in Section 6.

For use in Step 3, we create a subject detector, which judges whether an argument to a predicate is its subject using SVM^{light} [1], an implementation of Support Vector Machine (Vapnik, 1998), with a polynomial kernel of 2nd degree. The training instances of the subject detector are extracted from the predicate-argument relations [2] in the NAIST Text Corpus. The numbers of positive and negative instances are 35,304 and 104,250 respectively. As features, we used the morpho-syntactic information about the lemmas of the predicate and its argument and the functional words following the predicate and its argument. The results of subject detection with 5-fold cross-validation demonstrate that our subject detector accurately detects subjects with performances of 0.949 in recall, 0.855 in precision, and 0.899 in F-score.

Note that our subject detector checks whether each predicate in an SSPN has a syntactic subject among its arguments. An SSPN can include more than one predicate, and each predicate may have its own subject [3]. In this step, if two or more distinct subjects are detected for predicates in an SSPN, we use the most likely subject (i.e., the subject with the highest SVM score outputted by our subject detector) for subject propagation. Note that subject propagation is not performed if the subject position of a predicate is already filled.

Up to this point, the zero anaphora of the following three cases cannot be resolved: (i) no subject was detected for any predicate in a group linked by the subject sharing relations in the SSPN, (ii) no subject sharing relation was recognized for a predicate in the SSPN and (iii) non-

---

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

[2] Note that if a predicate appears in a relative clause and a noun modified by the clause is the *semantic* subject of the predicate, the noun is not regarded as subject by our subject detector.

[3] The subject sharing recognizer is likely to regard two predicates, each of which has its own subject, as non-subject sharing predicate pairs, but it is still logically possible that they are judged as subject sharing predicate pairs hence as a part of an SSPN.
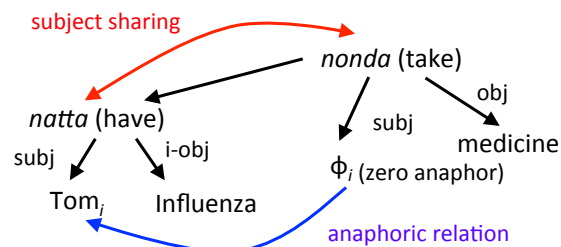

Figure 3: Example of DEP type

subject arguments were omitted as zero anaphors. To resolve zero anaphora in these cases, we apply a state-of-the-art ILP-based zero anaphora resolution method (Iida and Poesio, 2011) in Step 5. This method determines zero anaphor and its antecedent by joint inference using the results of subject detection, zero anaphor detection and intra- and inter-sentential antecedent identification. In the original method by Iida and Poesio (2011), the inter-sentential zero anaphora was resolved, but in this work we focus on intra-sentential zero anaphora. To adapt their method for our problem setting, we simply removed the inter-sentential antecedent identification model from their method.

4 Pairwise subject sharing recognition

A key component in our zero anaphora resolution method is pairwise subject sharing recognition. In this work, we focus on three types of subject sharing relations (DEP, ADJ and PNP types) as a first step because the instances belonging to the three types occupy 62% of intra-sentential zero anaphora that can be regarded as subject sharing. We developed a method that recognizes each subject sharing type and evaluate it.

4.1 Three types of subject sharing relations

We first describe the three types of subject sharing relations we focus on.

DEP A typical type of subject sharing relation is one between two predicates that have a syntactical dependency relation. The relation between two predicates, *natta* (have) and *nonda* (take), in example (1) in Section 1 is classified as this type because the two predicates have the same subject Tom_i (ϕ_i), as illustrated in Figure 3. We call this type of subject sharing the DEP type.

ADJ This type is a subject sharing relation between two *adjacent* predicates, i.e., a predicate
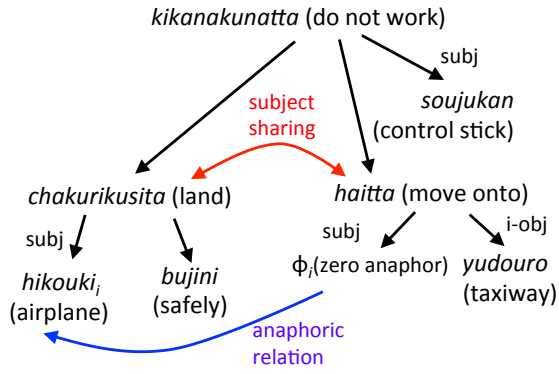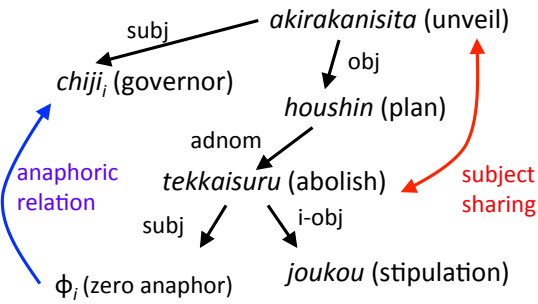
2182

Figure 4: Example of ADJ type

Figure 5: Example of PNP type

pairs that do not have any other predicate between them in the surface order of a sentence. Although two adjacent predicates in a sentence tend to share the same subject, they sometimes cannot be captured as the DEP type due to a long-distance dependency between predicates. For example, in example (3), two adjacent predicates, *land* and *move onto*, have the same subject but not a direct dependency relation, as illustrated in Figure 4.

(3) *hikouki-wa*   *bujini*   ***chakurikusi-ta-ga***
airplane-TOP   safely   land-PAST-but

(φ$_i$-ga)   *yudouro-ni*   ***hait-ta-atoni***
it$_i$-SUBJ   taxiway-IOBJ   move onto-PAST-after

soujukan-ga   *kikanakunat-ta*   .
control stick-SUBJ   do not work-PAST   period
The airplane safely landed, but its control stick did not work after (it$_i$) moved onto the taxiway.

To cover such cases, we also take into account the subject sharing relations of the ADJ type in which two predicates appear adjacently in the surface order.

PNP   In addition to the above two types of relations, in Japanese predicate pairs often have a subject sharing relation when one of the predicates syntactically depends on a noun (or noun phrase) that in turn syntactically depends on the other predicate. Example (4) is classified as such a type because noun *houshin* (plan) is placed between two predicates, *akirakanisita* (unveil) and *tekkaisuru* (abolish), in the dependency path and predicates share subject *chiji* (governor), as illustrated in Figure 5.

(4) *chiji$_i$-wa*   *(φ$_i$-ga)*   *joukou-o*
governor-TOP   he$_i$-SUBJ   stipulation-OBJ

tekkaisuru   *houshin-o*   ***akirakanisi-ta***   .
abolish   plan-TOP   unveil-PAST   period
The governor$_i$ unveiled his plan under which (he$_i$) will abolish the stipulation.

We call this type of subject sharing relation the PNP type.

In this work, we solve the problem of subject sharing recognition as a binary classification problem in which we classify whether two predicates share the same subject. We solve this problem using a supervised approach. We independently extract the training instances for each type from a corpus to which (zero) anaphora, coreference and subjects were annotated. The binary labels of the training instances are classified into the positive class if the subject of the two predicates in an instance is shared by coreference or (zero) anaphora, and negative otherwise. To create a classifier, we use SVM^{light} and experiment with both a linear kernel and a polynomial kernel of 2nd degree.

As features, we use the feature set shown in Table 1. Even though these features look simple, we expect them to work well to capture the characteristics of each subject sharing type. For example, as shown in example (5), the (subject) case marker of the argument (mother-SUBJ) between two predicates *natta* (have) and *katta* (buy) is a good indicator of non-subject sharing.

(5) *Tom$_i$-ga*   *infuruenza-ni*   ***natta-node***   ,
Tom-SUBJ   flu-IOBJ   had-since   punc

haha-ga   *kusuri-o*   ***katta***   .
mother-SUBJ   medicine-OBJ   buy-PAST   period
Since Tom had the flu, his mother bought medicine.

For recognizing the PNP type of subject sharing relations, whether certain nouns appear between two predicates is an important clue, e.g., *koto* (complementizer) in example (6) and *nouryoku* (ability) in example (7).

Name	Description
PoS_i (PoS_j)	PoS of p_i (p_j)
$lemma_i$ ($lemma_j$)	lemma of p_i (p_j)
func_w$_i$ (func_w$_j$)	function words following p_i (p_j)
case$_i$ (case$_j$)	case marker of arguments of p_i (p_j)
btw_case	case marker of arguments that appeared between p_i and p_j
NpPoS*	PoS of np
Np_lemma*	lemma of np
func_w$_{np}$*	function words following np
case$_{np}$*	case marker of dependents of np
n_class*	noun class of np based on Kazama and Torisawa (2008)

p_i and p_j stand for the left and right predicates in predicate pairs. np is the noun phrase between p_i and p_j. b_i (b_j) stands for the *bunsetsu*-unit[4] including p_i (p_j). The features marked with * are only used for PNP type.

Table 1: Features of subject sharing recognition

(6) *seifu$_i$-wa* (ϕ_i-ga) *sono isetsu-o*
government-TOP it$_i$-SUBJ the relocation-OBJ

mitomeru **koto-o** *kime-ta* .
admit COMP-OBJ decide-PAST period
The government$_i$ decided that (it$_i$) admits the relocation.

(7) *sono fune-wa* (ϕ_i-ga) *hayaku*
the ship-TOP it$_i$-SUBJ fast

hashiru **nouryoku-o** *motteiru* .
run ability-OBJ have period
The ship$_i$ has an ability that (it$_i$) runs fast.

To robustly capture this characteristic, we use as features the discrete classes created by the noun clustering algorithm proposed by Kazama and Torisawa (2008). It follows the distributional hypothesis, which states that semantically similar words tend to appear in similar contexts (Harris, 1954). By treating the syntactic dependency relations between words as 'contexts,' the clustering method defines a probabilistic model of noun-verb dependencies with hidden classes:

$$p(n, \langle v, r \rangle) = \sum_c p(n|c)p(\langle v, r \rangle|c)p(c)$$

where n is a noun, v is a verb or noun on which n depends by grammatical relation r (post-positions in Japanese), and c is a hidden class. The dependency relation frequencies were obtained from a 600-million page web corpus, and model parameters $p(n|c)$, $p(\langle v, r \rangle|c)$ and $p(c)$ were estimated using the EM algorithm (Hofmann, 1999). We clustered one million nouns into 500 discrete classes

---

[4]A bunsetsu-unit is a Japanese base phrase consisting of at least one content word optionally followed by functional words.

by assigning noun n to class c when the model parameter $p(c|n) > \theta$ ($\theta = 0.2$).

5 Experiment 1: pairwise subject sharing recognition

We first empirically evaluate the performance of our pairwise subject sharing recognition for the DEP, ADJ and PNP types.

5.1 Experimental setting

The training data for the subject sharing recognizer were generated from the NAIST Text Corpus 1.4 (Iida et al., 2007b), in which (zero) anaphora, coreference and subjects were manually annotated. We automatically extracted pairs of predicates from the corpus. Since the original NAIST Text Corpus has a wide variety of annotation noise, we cleaned it up by the following strategy. According to the annotation scheme in the NAIST Text Corpus, predicate-argument relations were annotated for the 'bare predicates' even if the predicates appear in passive or causative sentences. In such cases, the annotation was difficult and caused inconsistencies because the annotators needed to imagine the predicate-argument relations for predicates that are not explicitly written, considering case alternation caused by changes of voice and so on. As such, to achieve a higher level of consistency, we modified the annotation scheme for predicate-argument relations by considering 'surface predicates' and re-annotated predicate-argument relations in passive and causative cases, thus reducing the risk of inconsistent annotations caused by case alternation.

type	method	Recall	Precision	F-score
DEP	baseline	0.161	0.505	0.244
	proposed (linear)	0.545	0.719	0.620
	proposed (poly-2d)	0.578	0.732	**0.646**
ADJ	baseline	0.143	0.414	0.212
	proposed (linear)	0.011	0.604	0.021
	proposed (poly-2d)	0.285	0.713	**0.407**
PNP	baseline	0.154	0.329	0.210
	proposed (linear)	0.028	0.844	0.053
	proposed (poly-2d)	0.159	0.723	**0.260**

Table 2: Results of subject sharing recognition

Another important point is that in the NAIST Text Corpus, if the antecedent of a zero anaphor is not explicitly written in the corpus, it is simply annotated as 'exophoric', and the subject sharing relations between two predicates whose subject was annotated as exophoric cannot be captured. In contrast, in our cleaning procedure, the annotators additionally annotated such 'exophoric' subject sharing relations to take into account all subject sharing relations in the corpus.

The predicates in the corpus and their dependency relations were detected based on the outputs of a Japanese dependency parser, J.DepP[5] (Yoshinaga and Kitsuregawa, 2009). We obtained 49,313 predicate pairs for the DEP type, 86,728 for the ADJ type, and 27,117 for the PNP types. The numbers of positive instances of DEP, ADJ and PNP types are 9,524, 13,104, and 2,363 respectively. To evaluate the subject sharing recognition, we conducted 5-fold cross-validation using these predicate pairs and measured the performance using recall, precision and F-score.

Note that we also evaluated a baseline method that recognizes subject sharing relations using the results of the state-of-the-art zero anaphora resolution method (Iida and Poesio, 2011) and the subject detector at Step 3 in Section 3.

5.2 Results: subject sharing recognition

We measured the performances of the baseline and our subject sharing recognition method using recall, precision and F-score for each of the three types of subject sharing relations, which are shown in Table 2. The results demonstrate that all of the proposed classifiers solved the problems with high precision. In particular, for each type, the classifier using a polynomial kernel achieved more than

70% precision. We thus used the classifiers with a polynomial kernel for evaluations in Section 6. The results also show that the classifier using a polynomial kernel for each type outperformed the baseline method based on the state-of-the-art zero anaphora resolution method. That is, the direct subject sharing recognition using our classifiers has the potential to lead to a significant improvement in zero anaphora resolution, which we confirm through the experiments in Section 7.

Table 2 also shows that the classifier for the DEP type outperformed those for all of the other types in F-score. The difference reflects the wider variations of the problems in both ADJ and PNP compared to the case of DEP. For example, to recognize the PNP type of subject sharing relation, our classifier needs to appropriately learn the complicated relationship between two predicates and the noun that intervenes between them, a problem we do not need to consider for the DEP type.

6 Experiment 2: intra-sentential zero anaphora resolution between subjects

We next investigate the effect of introducing SSPNs. In this experiment, we evaluated the performance of intra-sentential zero anaphora resolution only between subjects, i.e., the positive instances used in this experiment were limited to the cases where the antecedent of a zero anaphor is the realized subject of a predicate. We evaluated a method of zero anaphora resolution using only SSPNs, where intra-sentential zero anaphora is resolved by the first four steps (Steps 1 to 4) in Section 3. We compared it to a baseline that only used the results of pairwise subject sharing recognition without SSPNs: if the subject sharing relation between two predicates is recognized by our pairwise subject sharing recognizer and a

---

	Recall	Precision	F-score
DEP w/o SSPN	0.259	0.744	0.385
DEP with SSPN	0.284	0.744	0.411
ADJ w/o SSPN	0.182	0.554	0.274
ADJ with SSPN	0.193	0.561	0.288
PNP w/o SSPN	0.034	0.757	0.064
PNP with SSPN	0.033	0.780	0.064
DEP+ADJ w/o SSPN	0.315	0.602	0.413
DEP+ADJ with SSPN	0.354	0.604	0.447
DEP+PNP w/o SSPN	0.293	0.746	0.421
DEP+PNP with SSPN	0.324	0.749	0.453
ADJ+PNP w/o SSPN	0.191	0.558	0.285
ADJ+PNP with SSPN	0.203	0.566	0.299
DEP+ADJ+PNP w/o SSPN	0.324	0.604	0.422
DEP+ADJ+PNP with SSPN	0.365	0.607	**0.456**

Table 3: Results of intra-sentential zero anaphora resolution between subjects

single subject is detected by our subject detector for one of the two predicates, then the subject fills the empty subject position of the other predicate. Note that in this baseline method, *transitive* subject propagation through more than one subject sharing relation is not performed. Also, if multiple subjects are detected for a predicate, we used the most likely subject to fill the subject position of the predicate, as in our method.

We conducted 5-fold cross-validation using the modified version of the NAIST Text Corpus presented in Section 5.1. In this evaluation, we used the 8,473 subject zero anaphors that refer to the *subject* antecedents (46% of all the *intra-sentential subject zero anaphora*, in which a subject zero anaphor refers to the antecedent that are not limited to subject) in the corpus. We measured the performance using recall, precision and F-score for each of the three types of subject sharing relations and their combinations. When combining more than one subject sharing recognizer in our method, we construct the SSPN using the subject sharing relations recognized by at least one of those recognizers for transitive subject propagation. On the other hand, in the baseline method, the SSPN was not constructed and zero anaphoric relations were identified using only the outputs of our subject detector and one of those recognizers.

The experimental results shown in Table 3 clearly demonstrate that the method with SSPNs for each type or a combination of the three types consistently outperformed that without SSPNs except for the PNP type. This result suggests that

multi-step propagation of subjects through more than one subject sharing relation, as done in SSPNs, is an effective way to propagate a subject to a subject position that cannot be reached by a single subject sharing relation. Our results also show that the F-score is improved by combining different types of subject sharing relations, and the best F-score, 0.456, was achieved when we used all types of relations, i.e., in the case of DEP+ADJ+PNP with SSPNs.

7 Experiment 3: intra-sentential subject zero anaphora resolution

Finally, we evaluate the performance of intra-sentential subject zero anaphora resolution. In the previous section, we evaluated just a part of our method, i.e., from Step 1 to Step 4 presented in Section 3. In this section, we evaluate the whole method, i.e., from Step 1 to Step 5, against 18,324 subject zero anaphors, which are all subject zero anaphors annotated in our modified version of the NAIST Text Corpus. As a baseline, we employed Iida and Poesio (2011)'s method that was tuned for intra-sentential zero anaphora resolution. The baseline method solves the problems by applying only Step 5 in Section 3 to all the predicates.

Our results in Table 4 show that all the methods using either each type or a combination of the three types significantly outperformed the baseline[6]. The best performing method was DEP+PNP, which achieved 0.380 in F-score, which is 3.6%

---

[6]The significance was tested using McNemar's testing ($p < 0.01$).

	Recall	Precision	F-score
Baseline (Step 5)	0.345	0.344	0.344
+DEP with SSPN	0.388	0.363	0.375
+ADJ with SSPN	0.374	0.347	0.360
+PNP with SSPN	0.351	0.347	0.349
+DEP+ADJ with SSPN	0.399	0.355	0.376
+DEP+PNP with SSPN	0.394	0.366	**0.380**
+ADJ+PNP with SSPN	0.376	0.347	0.361
+DEP+ADJ+PNP with SSPN	0.401	0.356	0.377

Table 4: Results of intra-sentential subject zero anaphora resolution (Steps 1 to 5 vs. Step 5)

	Recall	Precision	F-score
Baseline (Step 5)	0.345	0.344	**0.344**
DEP with SSPN	0.131	0.744	0.223
ADJ with SSPN	0.089	0.561	0.154
PNP with SSPN	0.015	0.780	0.030
DEP+ADJ with SSPN	0.164	0.604	0.258
DEP+PNP with SSPN	0.150	0.749	0.250
ADJ+PNP with SSPN	0.094	0.566	0.161
DEP+ADJ+PNP with SSPN	0.169	0.607	0.264

Table 5: Results of intra-sentential subject zero anaphora resolution (Steps 1 to 4 vs. Step 5)

higher than the baseline. This suggests that our method exploiting subject sharing relations and SSPNs has a positive impact on accuracy of *general* intra-sentential zero anaphora resolution methods because about 84% of zero anaphors of general intra-sentential zero anaphora appear as subject zero anaphor in our corpus.

We also estimate how accurately the method using only the SSPNs evaluated in Section 6 resolves intra-sentential subject zero anaphora in comparison to the baseline method. The results are shown in Table 5 and demonstrate that the performance of all the methods without Step 5 does not reach that of the baseline method in F-score. However, they retain high precision that ranges from 60% to 75%, preserving more than 10% of the recall on the DEP, DEP+ADJ, DEP+PNP and DEP+ADJ+PNP methods. Actually, in some of the potential applications of zero anaphora resolution, such as information extraction, methods with high precision and low recall are preferable to ones with low precision and high recall. Our methods with SSPNs alone might be usable in such applications because of their high precision.

8 Conclusion

In this paper, we introduced a subject shared predicate network (SSPN), which is a network of predicates that are linked by subject sharing relations for resolving typical intra-sentential zero anaphora. In our zero anaphora resolution method, zero anaphoric relations are identified by propagating a subject through subject sharing *paths* in the SSPN. To construct SSPNs, we developed a novel method of pairwise subject sharing recognition using the local contexts that surround two predicates and demonstrated that it can accurately recognize subject sharing relations. We combined our method of intra-sentential zero anaphora resolution with Iida and Poesio (2011)'s method and achieved significantly better F-score than Iida and Poesio (2011)'s method alone.

As future work, we are planning to use commonsense knowledge, such as causality (Hashimoto et al., 2014) and script-like knowledge (Sano et al., 2014), that has been automatically acquired from big data for accurate subject sharing recognition to improve inter-sentential zero anaphora resolution for cases not focused on in this work.

References

Mohit Bansal and Dan Klein. 2012. Coreference semantics from web features. In *Proceedings of the*

50th Annual Meeting of the Association for Computational Linguistics, pages 389–398.

David Bean and Ellen Riloff. 2004. Unsupervised learning of contextual role knowledge for coreference resolution. In *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pages 297–304.

Chen Chen and Vincent Ng. 2013. Chinese zero pronoun resolution: Some recent advances. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1360–1365.

Barbara J. Grosz, Scott Weinstein, and Aravind K. Joshi. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):203–225.

Masatsugu Hangyo, Daisuke Kawahara, and Sadao Kurohashi. 2013. Japanese zero reference resolution considering exophora and author/reader mentions. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 924–934.

Zellig Harris. 1954. Distributional structure. *Word*, 10(23):146–162.

Chikara Hashimoto, Kentaro Torisawa, Julien Kloetzer, Motoki Sano, István Varga, Jong-Hoon Oh, and Yutaka Kidawara. 2014. Toward future scenario generation: Extracting event causality exploiting semantic relation, context, and association features. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 987–997.

Yuta Hayashibe, Mamoru Komachi, and Yuji Matsumoto. 2011. Japanese predicate argument structure analysis exploiting argument position and type. In *Proceedings of 5th International Joint Conference on Natural Language Processing*, pages 201–209.

Thomas Hofmann. 1999. Probabilistic latent semantic indexing. In *Processing of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 50–57.

Ryu Iida and Massimo Poesio. 2011. A cross-lingual ILP solution to zero anaphora resolution. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 804–813.

Ryu Iida, Kentaro Inui, and Yuji Matsumoto. 2007a. Zero-anaphora resolution by learning rich syntactic pattern features. *ACM Transactions on Asian Language Information Processing*, Volume 6. Issue 4, Article 12.

Ryu Iida, Mamoru Komachi, Kentaro Inui, and Yuji Matsumoto. 2007b. Annotating a Japanese text corpus with predicate-argument and coreference relations. In *Proceedings of the ACL Workshop: 'Linguistic Annotation Workshop'*, pages 132–139.

Kenji Imamura, Kuniko Saito, and Tomoko Izumi. 2009. Discriminative approach to predicate-argument structure analysis with zero-anaphora resolution. In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, pages 85–88.

Hideki Isozaki and Tsutomu Hirao. 2003. Japanese zero pronoun resolution based on ranking rules and machine learning. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, pages 184–191.

Megumi Kameyama. 1986. A property-sharing constraint in centering. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pages 200–206.

Jun'ichi Kazama and Kentaro Torisawa. 2008. Inducing gazetteers for named entity recognition by large-scale clustering of dependency relations. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 407–415.

Hiromi Nakaiwa and Satoshi Shirai. 1996. Anaphora resolution of Japanese zero pronouns with deictic reference. In *Proceedings of the 16th International Conference on Computational Linguistics*, pages 812–817.

Manabu Okumura and Kouji Tamura. 1996. Zero pronoun resolution in Japanese discourse based on Centering Theory. In *Proceedings of the 16th International Conference on Computational Linguistics*, pages 871–876.

Motoki Sano, Kentaro Torisawa, Julien Kloetzer, Chikara Hashimoto, István Varga, and Jong-Hoon Oh. 2014. Million-scale derivation of semantic relations from a manually constructed predicate taxonomy. In *Proceedings of the 25th International Conference on Computational Linguistics*, pages 1423–1434.

Ryohei Sasano, Daisuke Kawahara, and Sadao Kurohashi. 2008. A fully-lexicalized probabilistic model for Japanese zero anaphora resolution. In *Proceedings of the 22nd International Conference on Computational Linguistics*, pages 769–776.

Ryohei Sasano, Daisuke Kawahara, and Sadao Kurohashi. 2009. The effect of corpus size on case frame acquisition for discourse analysis. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 521–529.

Kazuhiro Seki, Atsushi Fujii, and Tetsuya Ishikawa. 2002. A probabilistic method for analyzing Japanese anaphora integrating zero pronoun detection and resolution. In *Proceedings of the 19th International Conference on Computational Linguistics*, pages 911–917.

Hirotoshi Taira, Sanae Fujita, and Masaaki Nagata. 2008. A Japanese predicate argument structure analysis using decision lists. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pages 523–532.

Vladimir N. Vapnik. 1998. *Statistical Learning Theory*. Adaptive and Learning Systems for Signal Processing, Communications, and Control. John Wiley & Sons.

Marilyn Walker, Sharon Cote, and Masayo Iida. 1994. Japanese discourse and the process of centering. *Computational Linguistics*, 20(2):193–233.

Yotaro Watanabe, Masayuki Asahara, and Yuji Matsumoto. 2010. A structured model for joint learning of argument roles and predicate senses. In *Proceedings of the ACL 2010 Conference Short Papers*, pages 98–102.

Xiaofeng Yang, Jian Su, and Chew Lim Tan. 2005. Improving pronoun resolution using statistics-based semantic compatibility information. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics*, pages 165–172.

Katsumasa Yoshikawa, Masayuki Asahara, and Yuji Matsumoto. 2011. Jointly extracting Japanese predicate-argument relation with Markov logic. In *Proceedings of 5th International Joint Conference on Natural Language Processing*, pages 1125–1133.

Naoki Yoshinaga and Masaru Kitsuregawa. 2009. Polynomial to linear: Efficient classification with conjunctive features. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 1542–1551.

Koichiro Yoshino, Shinsuke Mori, and Tatsuya Kawahara. 2013. Predicate argument structure analysis using partially annotated corpora. In *Proceedings of the 6th International Joint Conference on Natural Language Processing*, pages 957–961.

Shanheng Zhao and Hwee Tou Ng. 2007. Identification and resolution of Chinese zero pronouns: A machine learning approach. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 541–550.