VERB SENSE DISAMBIGUATION BASED ON THESAURUS OF PREDICATE-ARGUMENT STRUCTURE

An Evaluation of Thesaurus of Predicate-argument Structure for Japanese Verbs

Koichi Takeuchi, Suguru Tsuchiyama, Masato Moriya, Yuuki Moriyasu and Koichi Satoh

Graduate School of Natural Science and Technology, Okayama University, Tsushimanaka, Okayama-shi, Japan

Keywords: Thesaurus, Predicate-argument structure, Japanese verbs, Word sense disambiguation.

Abstract: This paper presents a system for word sense disambiguation based on a manually constructed thesaurus of predicate-argument structure, which is an ontology on the linguistic side providing essential information for mapping form texts to verb concepts. This system can be effective for word sense disambiguation even though the target word sense system is different from the thesaurus. We applied the proposed word sense disambiguation system to the test corpus of SemEval-2010 Japanese tasks. Experimental results showed that the thesaurus-based disambiguation system outperformed a CRFs-based system in recall rates of verb sense disambiguation. From the results of verb sense disambiguation, we clarified that the abstracted verb classes (709 types) in our proposed system were effective sets for verb sense disambiguation.

1 INTRODUCTION

This paper presents a system for word sense disambiguation based on a manually constructed thesaurus of predicate-argument structure. The system can be effective for word sense disambiguation even though the target word sense system is different from the thesaurus.

We are manually constructing a verb thesaurus of predicate-argument structure (Takeuchi et al., 2010) to deal with verbal paraphrases as well as verb sense disambiguation, such as "He employed/used a method" or "He employed/hired Mr. Smith as an accountant".Since each verb is highly polysemous, verbs can be categorized into several verb classes, and verbs in each verb class have shared concepts.Thus, our dictionary describes set verb classes with argument structures to make a correlation between their arguments[1]. From the view of ontology, our thesaurus can be regarded as an ontology on the linguistic side, which can provide essential information for mapping form texts to verb concepts (i.e., verb classes). Thus one of the issues in creating a verb thesaurus is making sure how effective the constructed thesaurus is for verb sense disambiguation.

On the other hand, the first Balanced Corpus of

---

[1]The thesaurus is freely downloadable at http://vsearch.cl.cs.okayama-u.ac.jp/.

Contemporary Written Japanese, BCCWJ, has been developed (Maekawa, 2008), and SemEval-2010 (Okumura et al., 2010) Japanese word sense disambiguation tasks are organized based on the sense annotated corpus of the BCCWJ. In regard to verb sense, identifying verb sense indicates the finding of a group of verb synonyms. Thus, the verb sense tagged corpus can be regarded as a test bed of a task to detect a verb synonym group. This task would be the first step in detecting the argument structures of verbs for dealing with verbal paraphrases.

Given the above background, in this paper we clarified the usefulness of the verb thesaurus by applying a thesaurus-based shallow semantic analyzer to WSD for verb senses, regarding the task of detecting a verb class, i.e., a verb synonym group, in our thesaurus. The definitions of verb classes in our thesaurus and those of verb sense tags in SemEval-2010 are not equivalent. Thus, our verb classes were converted to verb sense tags by a conversion table constructed independently for this task. Since the SemEval-2010 corpus provides a training corpus and test corpus, the performance of tasks of the thesaurus-based analyzer was compared with the performance of Conditional Random Fields (CRFs), a statistical learning approach-based model. It was also compared with the performance of the best WSD system for SemEval-2010 tasks.

Experimental results of verb sense disambigua-

tion showed that the thesaurus-based analyzer outperformed CRFs in recall rates for all genres as well as for white paper domains. For precision rates, the thesaurus-based analyzer provided lower accuracy than the best WSD system; however, most of the errors were caused not by the analyzer, but by the conversion table. Thus, we clarified that the performance of the thesaurus-based analyzer was almost equal to that of the best WSD system when we discounted conversion table errors.

2 BACKGROUND: LANGUAGE RESOURCES FOR PREDICATE ARGUMENT STRUCTURE

The position of our thesaurus of predicate-argument structure for Japanese verbs can be described from the following three viewpoints: language resources, Japanese dictionaries and ontologies.

View from Language Resources. In English thorough, well-organized language resources relating to predicate-argument structure are constructed and available, e.g., EVCA (Levin, 1993), Dorr's LCS (Dorr, 1997), PropBank (Palmer et al., 2005), VerbNet (Kipper-Schuler, 2005), WordNet (Fellbaum, 1998) and FrameNet (Baker et al., 1998). In addition, there is a research project (Pustejovsky and Meyers, 2005) to find a general description framework of predicate-argument structure by merging several lexical databases (i.e., PropBank, NomBank, TimeBank and PennDiscouse TreeBank).

On the other hand, our thesaurus provides several granularities of semantic verb classes with argument structure (see Section 3). Compared to the above language resources, our thesaurus corresponds partly to each lexical database, i.e., Frame and FrameElements in FrameNet correspond to our verb class and semantic role labels, and the way of organizing similar verb classes with a thesaurus corresponds to WordNet; however, these elements and the method of organization of similar verb classes with a thesaurus are not exactly the same as those of our approach and our proposed thesaurus.

View from Japanese Dictionaries. There are several Japanese lexicons: IPAL (IPAL, 1986) was constructed focusing on morpho-syntactic classes but IPAL is small[2]. EDR (EDR, 1995) is composed of a large-scale lexicon and corpus. EDR is a well thought out and wide ranging bilingual dictionary between Japanese and English, but EDR's semantic classes

---

[2]It contains 861 verbs and 136 adjectives.

were not designed with syntactical lexical relations between verbs, such as alternations and causative or transitive relations between verbs. In contrast, our thesaurus can deal with these relations.

Besides Japanese version of WordNet (Bond et al., 2008) and FrameNet (Ohara et al., 2006) have been proposed. Japanese WordNet are constructed by machine translation from English to Japanese and manual revision, and then we find that some of the basic verbs of Japanese (i.e., highly ambiguous verbs) are wrongly assigned to unrelated synsets. Japanese FrameNet currently has published fewer than 150 verbs, then it is much smaller than our thesaurus.

View from Ontologies. Previous upper ontology studies have discussed how to describe processes and events (Takeda, 2004) (Galton, 2010), and practical ontologies such as SUMO (Niles and Pease, 2001) and DOLCE (Masolo et al., 2002) have been published; however, because they are upper ontologies, they are too abstract to deal with practical events. In contrast, a more concrete event ontology that can give a formal framework to deal with relations between Japanese verb concepts with description logic has been proposed (Kaneiwa et al., 2007) (Kaneiwa and Iwazume, 2010). An event ontology gives clear definitions of different verb concepts; however, the target of the ontology is not to deal with practical texts, but to deal with logical semantic relations between concepts. Thus, there is no information for verb sense disambiguation that constitutes an essential technique for mapping from texts to concepts. Compared to this approach, our proposed verb thesaurus provides information on verb sense disambiguation as well as on verb classes (i.e., concepts), and thus, our thesaurus can bridge the gap between texts and event ontologies, which are well-organized according to description logic.

3 THESAURUS OF PREDICATE-ARGUMENT STRUCTURE

Since the details of our thesaurus are described in previous papers (Takeuchi et al., 2010) here we describe the basic design of our thesaurus.

The proposed thesaurus consists of a hierarchy of verb classes we defined. A verb class, a conceptual class, indicates a shared meaning of verbs in a verb class. A parent verb class includes concepts of a child verb class; thus a child verb class is a detailed concept of the parent verb class and multiple inheritance is prohibited in the hierarchy. Meaning

of a verb class is described as a semantic description that is a kind of semantic skeleton such as lexical conceptual structure (Jackendoff, 1990) (Kageyama, 1996) (Dorr, 1997). Thus, core semantic relations between arguments are represented in the semantic description.

We allow a verb has several senses, i.e., a verb can be a polysemous, and adopt verb senses defined in Lexeed (Fujita et al., 2006) as a verb senses repository of our thesaurus. Thus each verb sense is assigned to a verb class, and then a verb sense is linked to example sentences. Every example sentence is analyzed into its arguments and semantic role labels; the core arguments are linked to the semantic description via variables. This allows that if semantic role labels cannot capture the correct linking, the links of variables can designate corresponding arguments in example sentences (Figure 2). In addition, by linking one semantic description to several example sentences of a verb sense, our thesaurus can provide rich verb sense disambiguation information.

Here we explain this structure using verbs such as *buy, purchase, hire, rent, recapture*[3]. Each verb sense is designated by example sentences, e.g., "Hiroshi buys a bicycle to his son", "Jiro purchases a car form her", "Taro hires a car", "Kazuko rents a book" and "Yoshio recaptures the top position". As Figure 1 shows, all of the above verb senses are involved in the verb class *Moving_One's_Possession_From* [4]. The semantic description, which expresses the core meaning of the verb class is

([Agent] CAUSE) BECOME [1] BE AT [2],

where the brackets [] denote variables that can be filled with arguments in example sentences; [Agent] is a semantic role label that can be annotated to all example sentences; the parentheses () denote an component. The semantic description consists of roughly 3 components describing causer, manner and (change of) state. A manner component expresses various kinds of complex meanings of a verb concept such as condition, purpose, attitude, and so on[5].

Figure 1 shows the children of the verb class *Moving_One's_Possession_From*, e.g.,

---

[3]The actual verbs are Japanese in our thesaurus, but to make the explanation easy to understand we give their English translation.

[4]The actual verb class names correspond to the hierarchical system in the thesaurus, e.g., the full length of *Moving_One's_Possession_From* is *Change_of_State/ Change_of_Position_(Physical)/ Moving_One's_Possession_From*, where '/' denotes a delimiter between hierarchies. Then the verb class names in Figures 1 and 2 show the abbreviated verb class names.

[5]Currently about 800 kinds of words are used to describe a manner component.

Moving_One's_Possession_From/Buying, *Moving_One's_Possession_From/Renting* and *Moving_One's_Possession_From/Repossessing*. The verbs *buy* and *purchase* are in the *Buying* class, while the verbs *rent* and *hire* are in the *Renting* class. The semantic descriptions in the children verb classes are more detailed than those in the parent's description.
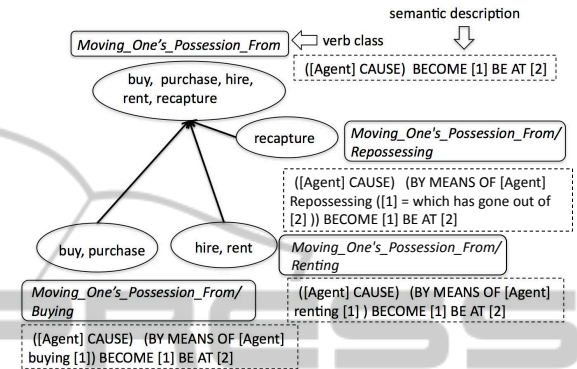

Figure 1: Example of verb classes, verbs and their semantic descriptions.

A semantic description in the *Renting* class, i.e.,

```
([Agent] CAUSE)
(BY MEANS OF [Agent] renting [1])
BECOME [1] BE AT [2]
```

describes semantic relations between [1], [2] and [Agent]. Since semantic role labels and variables are annotated to all of the example sentences, corresponding arguments can be linked via semantic role labels and variables in the semantic description. As we show
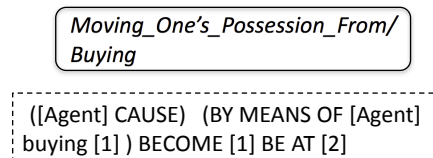
Figure 2: Linking between semantic description and example sentences.

in Figure 2, the semantic description contains only essential arguments. Thus, the key arguments such as [1] and [2] in example sentences have links, and another argument such as [Source] does not have a link. This indicates the two type of arguments we assumed,

i.e., arguments that are essential to a verb class and arguments that are not essential. This is because of the current perspective on arguments in linguistics, that is, there are several levels of arguments depending on closeness to a verb concept such as core, non-core, peripheral and adjunct arguments in FrameNet, and constructions (Goldberg, 1995). Since the variety of adjuncts and constructions is wide, the proposed description framework can deal with this variety, which is impossible to pre-compile, by adding analyzed example sentences. The current types of semantic roles are organized into 71 for the results of analyses of about 7,400 verb senses.

4 ARGUMENT STRUCTURE ANALYZER

A predicate-argument structure analyzer, ASA[6], was constructed on the basis of our thesaurus of predicate-argument structure. ASA identifies the verb classes and semantic role labels of their arguments. In our thesaurus, for polysemous verbs, each verb sense categorized to a verb class has a few example sentences. Since each verb sense in a verb class only has a few example sentences, statistical learning methods do not work well in the preliminary tests. Thus, as the basic strategy for detecting the verb class (i.e., verb sense) of an input sentence, we take a nearest neighbor approach: find the most similar example sentence compared to the input sentence, and take the verb class of the example sentence as the word sense.

The similarity between an input sentence and an example sentence is evaluated on the similarity of the arguments between them; the similarity of the arguments is evaluated on three features: shallow syntactic position, noun categories, and surface words. Let *SimSnt* be this similarity function, the disambiguation of verb sense for an input sentence X is to detect the verb class $\hat{C}$ that gives the highest score of among the example sentences Y_C in a verb class C.

$$\hat{C} = \underset{C}{\text{argmax}} \, SimSnt(X, Y_C). \qquad (1)$$

The details of *SimSnt* are described in (Takeuchi et al., 2009). Since *SimSnt*() is calculated on the basis of the example sentences in the thesaurus, the performance of the ASA's verb class disambiguation depends on the quality and quantity of the thesaurus.

4.1 CRFs-based Word Sense Disambiguation System

We applied Conditional Random Fields (Lafferty et al., 2001) as a competitive alternative to the ASA. Conditional Random Fields is a probabilistic model for labeling sequence data, and we applied it to word sense disambiguation. The parameters of CRFs can be trained using the training corpus provided by SemEval-2010.

CRFs selects the best output sequence, i.e., a sequence of word senses $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, ..., \mathbf{Y}_n)$ given input word sequence $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n)$ by the following equations:

$$P(\mathbf{Y}|\mathbf{X}) = \frac{\exp(\Lambda \cdot \mathbf{F}(\mathbf{Y}, \mathbf{X}))}{\mathbf{Z_X}}, \qquad (2)$$

where $\mathbf{Z_X}$ denotes a normalized factor, $\mathbf{Y}_h$ denotes possible label (i.e., word sense) candidates and $\sum_{\mathbf{Y}_h}$ denotes the sum of all possible word sense sequences from an input word sequence $\mathbf{X}$. Λ is a weight for the feature vector $\mathbf{F}$. For the word sense disambiguation in Section 5 we apply as the features surface word, part-of-speech and combinations of previous and following words and part-of-speeches at from -3 to +1 positions[7] according to the results of CoNLL shared task[8].

5 VERB SENSE DISAMBIGUATION EXPERIMENTS AND DISCUSSIONS

SemEval-2010 annotated corpus consists of four genres (books, newspaper articles, white papers, and documents from a Q&A site on the WWW) and we used this corpus as the gold standard for a Japanese verb sense annotated corpus.

Table 1: Precision and recall of verb sense disambiguation in white papers.

	Precision	Recall
CRFs	0.971 (134/138)	0.244 (134/550)
ASA	0.660 (229/347)	0.416 (229/550)

Table 1 shows the results of verb sense disambiguation for the test data in white papers. The parameters in CRFs were trained on the training data in white papers, and the training data excluded documents that overlapped with test data. In Table 1 the

---

[7]See http://crfpp.sourceforge.net/.
[8]http://www.clips.ua.ac.be/conll2000/.

precision of CRFs is quite high; however, the number of correctly detected verb senses is much lower than that of the ASA—the ASA detected almost twice as many verb senses as the CRFs did. The point of the results is that the ASA does not use any domain dependent information, but the CRFs do. If we apply these systems to all genres, the differences become much more clearer.

Figure 2 shows the results of verb sense disambiguation in all genres. In Figure 2 RALI-2, a naive

Table 2: Precision and recall of verb sense disambiguation in books, newspaper articles, white papers and documents from a Q&A site on the WWW.

	Precision	Recall
RALI-2	0.719	no published data
CRFs	0.940 (608/648)	0.0372 (608/16332)
ASA	0.593 (6173/10417)	0.378 (6173/16332)

Bayes model, denotes the best results of the word sense disambiguation system in SemEval-2010 tasks. Both CRFs and RALI-2 used training data documents that overlapped with test data documents. The training data consisted of three genres i.e., books, newspaper articles and white papers, and there was no training data for Q&A documents. From the results, we found that the ASA outperformed CRFs in recall rates; CRFs seemed to have difficulty detecting correct verb senses in different genres. This tendency i.e., simple statistical models such as naive Bayes overcome more sophisticated statistical models such as support vector machines (SVMs) and maximum entropy models in word sense disambiguation, is also reported in the results of SemEval-2010 task (Okumura et al., 2010).

Table 2 shows that the ASA did not detect any verb sense for 5915 (16332 - 10417) examples due to the lack of verb entries in our thesaurus. From this result we can estimate the coverage of the verb entries in our thesaurus for all genres in Japanese text was 63.8% (10417/16633).

The output verb senses of the ASA are verb classes in our thesaurus, but the results of the ASA in Tables 1 and 2 were evaluated based on the transformed verb senses in SemEval-2010. To do this transformation, we use the conversion table that was manually created while doing the construction work of the thesaurs; this construction work was independently done for the purpose of this evaluation task. The total number of detected verb classes by the ASA in all genres was 10,417 words, but 1,637 words of these (about 16%) were not converted to any word sense because of the lack of sense identifications in

the conversion table. This data indicates that there is room for further improvement of correctly detecting verb sense by adding instances of verb senses to our thesaurus.

If we exclude non-converted examples, the precision rate of the ASA would be 70% (6173/8780) for all genres. This precision rate is almost the same as that of RALI-2. From the view of the conversion table, about 70% of verb classes in total can be successfully converted to SemEval-2010 word senses. The number of types of all verb classes is currently 709; and the verb classes are manually defined by summarizing over 7,400 verb senses in Lexeed (Fujita et al., 2006). Thus the success of the conversion indicates that the abstracted verb class is not too coarse but still keeps a granularity that can discriminate verb senses in SemEval-2010 tasks. Furthermore, since verb class disambiguation is done on the basis of the analyzed example sentences in the thesaurus, we can conclude that the proposed thesaurus provides effective linguistic data for verb sense disambiguation.

The methods of verb sense disambiguation depend on the availability of linguistically annotated resources. Table 3 shows the results of CRFs in a white paper genre that used overlapping training data. Comparing the results in Table 1, CRFs outperformed

Table 3: Precision and recall of verb sense disambiguation by CRFs in a white paper genre that used overlapping training data.

	Precision	Recall
CRFs	0.993	0.434

the ASA in both precision and recall rates. The results indicate that if we have enough sense annotated training data as test data for the same genre, a statistical learning approach will work well. Thus, we need to keep developing our thesaurus.

6 CONCLUSIONS

To evaluate a manually constructed Japanese thesaurus that is an ontology on the linguistic side providing essential information for mapping form texts to verb concepts, we constructed a system for word sense disambiguation (ASA) based on the thesaurus and applied the system to SemEval-2010 word sense disambiguation tasks. The definitions of verb classes in our thesaurus and verb sense tags in SemEval-2010 are not equal. Thus, we converted our verb classes to verb sense tags by a conversion table that we constructed independently for this task. Since the SemEval-2010 corpus provides training corpus

and test corpus the performance of the ASA was compared with that of Conditional Random Fields (CRFs), a statistical learning approach-based model. Experimental results of verb sense disambiguation showed that the ASA outperformed CRFs in recall rates for all genres as well as for the white paper domain. Regarding precision rates, the ASA provided lower accuracy than the best system for SemEval-2010 tasks; however, since most of the errors were caused by the conversion table, we found that the precision rate of the ASA was almost equal to the level (70%) of the best WSD system when we excluded conversion table errors. From the recall rate of the ASA for all genres the current coverage of the verb entries in our thesaurus can be estimated at 63.8%.

The key to success for the ASA will be proper working of the conversion table. Thus, the success of the conversion indicates the abstracted verb class is not too coarse, but still keeps a granularity that can discriminate verb senses in SemEval-2010 tasks. In addition, since verb class disambiguation is done based on analyzed example sentences in the thesaurus, we can conclude that the proposed thesaurus provides effective linguistic information for verb sense disambiguation.

REFERENCES

Baker, C. F., Fillmore, C. J., and Lowe, J. B. (1998). The Berkeley FrameNet project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics*, pages 86–90.

Bond, F., Isahara, H., Kanzaki, K., and Uchimoto, K. (2008). Construction of Japanese WordNet from Multi-lingual WordNet. In *Proceedings of the 14th Annual Meeting of Japanese Natural Language Processing*, pages 853–856.

Dorr, B. J. (1997). Large-Scale Dictionary Construction for Foreign Language Tutoring and Interlingual Machine Translation. *Machine Translation*, 12(4):271–325.

EDR (1995). *EDR: Electric Dictionary the Second Edition*. Japan Electronic Dictionary Research Institute, Ltd.

Fellbaum, C. (1998). *WordNet an Electronic Lexical Database*. MIT Press.

Fujita, S., Tanaka, T., Bond, F., and Nakaiwa, H. (2006). An implemented description of japanese: The lexeed dictionary and the hinoki treebank. In *COLING/ACL06 Interactive Presentation Sessions*, pages 65–68.

Galton, A. (2010). On what goes on: The ontology of processes and events. In *Formal Ontology in Information Systems: Proceedings of the Fourth International Conference (FOIS2006)*, pages 4–11.

Goldberg, A. E. (1995). *Constructions*. The University of Chicago Press.

IPAL (1986). *IPA Lexicon of the Japanese Language for Computers*. IPA: Information-Technology Promotion Agency, Japan.

Jackendoff, R. (1990). *Semantic Structures*. MIT Press.

Kageyama, T. (1996). *Verb Semantics*. Kurosio Publishers. (In Japanese).

Kaneiwa, K. and Iwazume, M. (2010). An Event Ontology for the Semantic Web. *Computer Software*, 27(5):1–13. (in Japanese).

Kaneiwa, K., Iwazume, M., and Fukuda, K. (2007). An upper ontology for event classifications and relations. In *Proceedings of the Twentieth Australian Joint Conference on Artificial Intel ligence (AI2007), LNCS 4830, Springer-Verlag*, pages 394–403.

Kipper-Schuler, K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon*. PhD thesis, PhD Thesis, University of Pennsylvania.

Lafferty, J., McCallum, A., and Pereira, F. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proc. 18th International Conf. on Machine Learning*, pages 282–289.

Levin, B. (1993). *English Verb Classes and Alternations*. University of Chicago Press.

Maekawa, K. (2008). Balanced corpus of contemporary written japanese. In *Proceedings of the 6th Workshop on Asian Language Resources (ALR)*, pages 101–102.

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A., and Schneider, L. (2002). WonderWeb deliverable D17. the WonderWeb library of foundational ontologies and the DOLCE ontology.

Niles, I. and Pease, A. (2001). Towards a standard upper ontology. In *Proceedings of the 2nd Interna-tional Conference on Formal Ontology in Information Systems (FOIS-2001)*, pages 2–9.

Ohara, K. H., Fujii, S., Ohori, T., Suzuki, R., Saito, H., and Ishizaki, S. (2006). Frame-based contrastive lexical semantics and japanese framenet: The case of risk and kakeru. In *Proceeding of the Fourth International Conference on Construction Grammar*.

Okumura, M., Shirai, K., Komiya, K., and Yokono, H. (2010). Semeval-2010 task: Japanese wsd. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 69–74.

Palmer, M., Gildea, D., and Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–105.

Pustejovsky, J. and Meyers, M. P. A. (2005). Merging propbank, nombank, timebank, penn discourse treebank and coreference. In *Proceedings of the Workshop on Frontiers in Corpus Annotation II: Pie in the Sky*, pages 5–12.

Takeda, H. (2004). Upper Ontology. *Japanese Society of Artificial Intelligence*, 19(2):172–178. (in Japanese).

Takeuchi, K., Inui, K., Takeuchi, N., and Fujita, A. (2010). A Thesaurus of Predicate-Argument Structure for Japanese Verbs to Deal with Granularity of Verb Meanings. In *The 8th Workshop on Asian Language Resources*, pages 1–8.

Takeuchi, K., Tsuchiyama, S., Moriya, M., and Moriyasu, Y. (2009). Construction of Argument Structure Analyzer Towards Searching Same Situation and Actions. In *Technical Report of IEICE, NLC-2009-33*, pages 1–6. (in Japanese).