1. Introduction

The goal of this study is to automatically extract a large set of hyponymy relations, which play a critical role in many NLP applications such as Q&A systems (Fleischman et al., 2003) and specification retrieval (Yoshinaga and Torisawa, 2006). In this paper, a hyponymy relation is defined as a relation between a hypernym and a hyponym when “the hyponym is a (kind of) hypernym.”1 We acquired more than 1.34 million hyponymy relations in Japanese with a precision of 90.1%.

Many NLP researchers have attempted to automatically acquire hyponymy relations from texts (Hearst, 1992; Caraballo, 1999; Mann, 2002; Fleischman et al., 2003; Morin and Jacquemin, 2004; Shinzato and Torisawa, 2004; Etrzioni et al., 2005; Pantel and Pennacchiotti, 2006; Sumida et al., 2006; Sumida and Torisawa, 2008). Most of these methods, however, require tera-scale documents (e.g., a web repository) and powerful computational resources to acquire a wide range of hyponymy relations that include concept-instance relations. On the other hand, Sumida and Torisawa (2008) have shown that you could easily obtain numerous hyponymy relations from Wikipedia; in particular, they have acquired more than 0.63 million hyponymy relations only from hierarchical layouts in the 2.2GB Japanese version of Wikipedia (e.g., Figure 1 shows a hierarchical structure of a Wikipedia article shown in Figure 2). Although the reported precision (76.4%) is insufficient for practical applications, the hierarchical structures in Wikipedia are definitely a promising resource to mine hyponymy relations.

Figure 1: Hierarchical layout of an article shown in Figure 2

Asuka Sumida†, Naoki Yoshinaga†, Kentaro Torisawa†
†Japan Advanced Institute of Science and Technology,
†University of Tokyo,
†1-1, Asahidai, Nomi, Ishikawa, 923-1292, Japan,
†4-6-1 Komaba, Meguro-Ku, Tokyo, 153-8505, Japan,
{a-sumida,n-yoshi,torisawa}@jaist.ac.jp

Abstract

This paper proposes an extension of Sumida and Torisawa’s method of acquiring hyponymy relations from hierarchical layouts in Wikipedia (Sumida and Torisawa, 2008). We extract hyponymy relation candidates (HRCs) from the hierarchical layouts in Wikipedia by regarding all subordinate items of an item x in the hierarchical layouts as x’s hyponym candidates, while Sumida and Torisawa (2008) extracted only direct subordinate items of an item x as x’s hyponym candidates. We then select plausible hyponymy relations from the acquired HRCs by running a filter based on machine learning with novel features, which even improve the precision of the resulting hyponymy relations. Experimental results show that we acquired more than 1.34 million hyponymy relations with a precision of 90.1%.

1 Black tea is a variety of tea that is more oxidized than the green, oolong and white varieties.

In this paper, we extend Sumida and Torisawa’s method (2008) of acquiring hyponymy relations, and double the number of acquired hyponymy relations, while improving the precision by 13.7%. The key idea of our method is to enumerate a wider range of hyponymy relation candidates (HRCs) from the hierarchical layouts than Sumida and Torisawa’s method. While Sumida and Torisawa
extracted direct subordinate items of an item $x$ as the candidates of $x$’s hyponyms, we extract all subordinate items of an item $x$ as the candidates of $x$’s hyponyms. We then apply a filter based on machine learning with novel features to the acquired HRCs to select plausible hyponymy relations. The reminder of this paper is organized as follows. Section 2 briefly explains the structure of Wikipedia and describes previous studies on hyponymy relation acquisition from Wikipedia. Section 3 introduces our method of acquiring hyponymy relations from hierarchical structures in Wikipedia. Section 4 presents experimental results. Section 5 concludes this paper and mentions future research directions.

2. Research Background

In this section, we first explain the structure of Wikipedia, and then describe previous studies that attempted to acquire hyponymy relations from Wikipedia.

2.1. The Structure of Wikipedia

Wikipedia is a free, multilingual, open-content encyclopedia, and consists of numerous articles that convey comprehensive information in the headings (basically concepts or instances). The Wikipedia is built on the MediaWiki software package, which interprets source codes written in the MediaWiki syntax to produce human-readable web pages. Figure 2(a) shows an article on ‘Black tea’, which is the result of interpreting the source code in Figure 2(b). The elements of the hierarchical structures used in this study are as follows.

Headings See lines 2-3, 6, 8, 11 of Figure 2(b). They are marked up as “=+title=+” in the MediaWiki syntax, where title is the subject of the paragraph. Note that “+” here means a finite number of repetitions of the preceding symbol, and we use this notation in the following explanation as well.

Bulleted lists See lines 4-5, 7 of Figure 2(b). They are marked as “**title**” in the MediaWiki syntax, where title is the subject of a listed item.

Ordered lists See lines 9-10 of Figure 2(b). They are marked up as “#title” in the MediaWiki syntax, where title is the subject of a numbered item.

Definition lists See lines 12-13 of Figure 2(b). They are marked as “*title” where title is a term. Although definition lists contain terms and their definitions, our method focuses only on the terms.

The basic hierarchical structures of Wikipedia articles are organized according to a pre-determined ordering among the above items. In general, items occupy a higher position in the hierarchy according to the order of headings, definition lists, bulleted lists, and ordered lists. In addition, note that headings, bullet lists and ordered lists allow the repetitions of the symbols “=”, “*” and “#”. The number of repetitions of the symbols indicates the position in the hierarchy, and the more repetitions of the symbol an item contains, the lower the position the item belongs to. For instance, “= Common tea brands =” occupies a higher position than “== England ==” as illustrated in Figure 2(b). Then, it is easy to extract a hierarchical structure from a Wikipedia article by parsing the source code of the article according to the above order among the mark-up items. Figure 1 illustrates the hierarchical structure obtained from the source code in Figure 2(b).

2.2. Hyponymy Acquisition from Wikipedia

Previous studies attempted to extract hyponymy relations from definition sentences (Kazama and Torisawa, 2007; Herbelot and Copestake, 2006; Ruiz-Casado et al., 2005) and category labels (Suchanek et al., 2007) included in Wikipedia articles.

Kazama and Torisawa (2007) considered the first sentence of a Wikipedia article as the definition sentence for the heading of the article, and extracted a hypernym of the heading from the definition sentence. Herbelot and Copestake (2006) parsed sentences in Wikipedia articles to find argument structures that represent the definition of concepts, and then obtain hypernym-hyponym pairs from the argument structures. They extracted 4,771 hyponymy relations from 12,200 animal-related articles with a precision of 88.5%. Ruiz-Casado et al. (2005) exploited WordNet (Fellbaum, 1998) to learn patterns for acquiring hyponymy relations. They acquired 1,204 hyponymy relations with a precision of 69%.

Suchanek et al. (2007) regarded the heading of a Wikipedia article as a hyponym and obtained category labels attached to the article as its hypernym candidates. A language-dependent heuristics then selected correct hypernyms from the hypernym candidates. They acquired more than 2.04 millions of hyponymy relations (relations SUBCLASSOF and TYPE in their paper) from 1.6 millions of Wikipedia articles with a precision of about 95%.

Although the above studies extracted hyponymy relations from the English version of Wikipedia, Sumida and Torisawa (2008) extracted hyponymy relations from definition sentences, category labels, and hierarchical structures in Wikipedia articles. They reported that the number of hyponymy relations acquired from the hierarchical structures was larger than the number of hyponymy relations acquired from the other resources. We thus focus on the hierarchical structures to acquire more hyponymy relations.

3. Proposed Method

Our method of acquiring hyponymy relations is an extension of the supervised method proposed by Sumida and Torisawa (2008), but differs in the way of enumerating hyponymy relation candidates (hereafter, HRCs) from the hierarchical layouts, and in the features of machine learning. Our method consists of the following two steps:

Step 1: We first extract HRCs from hierarchical layouts in Wikipedia articles.

Step 2: We then select proper hyponymy relations from the HRCs extracted in Step 1 by using Support Vector Machines (SVMs) as a classifier (Vapnik, 1998).
In what follows, we describe each step in detail.

3.1. Step 1: Extracting HRCs from the hierarchical structures in Wikipedia articles

We obtain HRCs by considering the title of each marked-up item as a hypernym candidate, and titles of its all subordinate marked-up items as its hyponym candidates; for example, we extract ‘England’, ‘France’, ‘Wedgwood’, ‘Lipton’, and ‘Fauchon’ as hyponym candidates of ‘Common tea brands’ from the hierarchical structure in Figure 1. Note that Sumida and Torisawa (2008) extracted HRCs by regarding the title of each marked-up item as a hypernym candidate and titles of its direct subordinate marked-up items as its hyponyms; for example, they extracted only ‘England’ and ‘France’ as hyponym candidates of ‘Common tea brands’ from the hierarchical structure in Figure 1. They also employed patterns shown in Figure 3 (e.g., “Xの一覧” (list of X)) to find plausible hypernyms denoted by X in the pattern. They regarded the HRCs whose hypernyms matched the patterns as correct hyponymy relations, and did not apply a filter based on machine learning to these HRCs.

In this study, we use these patterns only to justify the hypernym part of HRCs; namely, we just replace hypernyms that match the patterns shown in Figure 3 with the variable part, by discarding the non-variable part of the patterns. We then apply a filter based on machine learning to all the HRCs acquired in the manner described in the previous paragraph. This is because the hyponymy relations whose hypernyms matched these patterns were still too noisy to use in practical applications, and we would like to control the total quality of the acquired hyponymy relations by changing the threshold of the SVM value for each HRC.

3.2. Step 2: Selecting Proper Hyponymy Relations from the acquired HRCs

We select proper hyponymy relations from the HRCs obtained in Step 1 by using SVMs (Vapnik, 1998) as a classifier. In what follows, we briefly review the features proposed by Sumida and Torisawa (2008), and then explain the novel features introduced in this study. We expect that the readers will refer to the literature (Sumida and Torisawa, 2008) to see the effect of the features proposed by Sumida and Torisawa. In the following explanation, we refer to the hypernym candidate or the hyponym candidate of each HRC as hypernym or hyponym.

POS We assigned a unique dimension in the feature space to each part-of-speech (POS) tag. When the hypernym/hyponym consists of a morpheme with a particular POS tag, then the corresponding element of the feature vector is set to 1. When the hypernym/hyponym consists of multiple morphemes, the feature vectors for all the morphemes are simply summed (The resulting feature vector works as disjunction of each feature vector). The POS tag of the last morpheme is mapped to the dimension that is different from that of the POS tags of the other morphemes.

MORPH The morphemes are mapped to the dimensions of the feature vectors. The last morpheme is mapped to the dimension that is different from that of the other morphemes.

EXP The expression of a hypernym/hyponym itself is mapped to an element in a feature vector, and the corresponding element is set to 1.

ATTR Using the attribute set created by Sumida and Torisawa (2008), when a hypernym/hyponym is included as an element of the attribute set, we set a feature corresponding to the element to 1.

LAYER Each type of the marking items from which the hypernym/hyponym is extracted (namely, headings, bulleted lists, ordered lists, or definition lists) is mapped to an element of a feature vector, and the feature corresponding to the marking type for the hypernym/hyponym is set to 1.

In this study, we introduce the following three new features to improve the performance of the classifier.

DIST The distance $d$ between items from which the hypernym and the hyponym are acquired is mapped to two elements of the feature vector. When the distance $d = 1$, one element is set to 1, and otherwise (i.e., $d \geq 2$) the other element is set to 1. This feature reflects the tendency that HRCs acquired from items whose distance is $d = 1$ are more plausible than the other HRCs.

PAT This feature is set to 1 when the hypernym of the given HRC is obtained from a hyponym that matches the patterns in Figure 3. This reflects Sumida and Torisawa’s observation that HRCs whose hypernym matches the patterns are likely to be correct (Sumida and Torisawa, 2008).

LCHAR This feature is set to 1 when the hypernym and the hyponym share the last character. Such HRCs (e.g., “高等学校 (high school)”-“公立高校 (public high school)”) are likely to be correct, because the last characters are likely to convey major semantic contents of Japanese compound nouns.

Using the above features, we train an SVM classifier.

---

1 In Japanese, a morpheme takes a POS tag.
Table 1: Precision of the HRCs in the development set in terms of the distance

| DIST | PRE  |
|------|------|
| 1    | 35.0 (1.443/4.126) |
| 2    | 30.7 (915/2.981) |
| 3    | 23.2 (352/1.516) |
| 4    | 22.4 (68/304) |
| 5    | 18.2 (10/55) |
| 6    | 18.8 (3/16) |
| 7    | 0.0 (0/2) |
| TOTAL | 31.0 (2.79/19,000) |

4. Experiments

To evaluate our method, we used the Japanese version of the Wikipedia version of March 2007, which includes 276,323 articles (pages).\(^4\) In Step 2, we used TinySVM\(^5\) with a polynomial kernel of degree 2 as a classifier, and MeCab\(^6\) as a morphological analyzer.

We acquired 6,564,317 HRCs from the above articles in Step 1. The test set of 1,000 HRCs were randomly extracted from these HRCs, and the remaining HRCs were used to form the development set. We increased the size of the development set by adding the following four sets, while investigating the performance of the classifier on the development set. The first set was randomly chosen from the remaining HRCs, and consisted of 9,000 HRCs. The second set was chosen from the HRCs whose hypernyms did not match the patterns in Figure 3, and consisted of 10,000 HRCs. The third set was randomly chosen from the HRCs whose hypernym and hyponym are acquired from items with distance \(d = 1\) in the hierarchy, and consisted of 9,000 HRCs. The fourth set was chosen from the HRCs whose hypernyms matched the patterns in Figure 3, and consisted of 2,000 HRCs. The total number of HRCs in the development set was 29,900, when we eliminated duplicated entries. There is no overlap between the test set and the development set.

A human subject then manually judged whether HRCs in the test and development sets are correct or not using the same criteria as one in Hearst (1992); the subject checked whether the expression “a hyponym candidate is (a kind of) a hypernym candidate” is acceptable or not.

To investigate the quality of the input HRCs, we assessed the precision of the 9,000 development HRCs that were randomly extracted from all the HRCs excluding the test set. Table 1 shows the precision of the 9,000 HRCs according to the distance between items from which the hypernym and hyponym of each HRC are extracted. We can see that when the distance between the items from which the hypernym and the hyponym of HRCs are extracted increases, the precision of the HRCs decreases. However, the extraction of HRCs from distant items almost doubled the number of correct hyponymy relations in the HRCs.

Table 2 shows the performance of our method when we used the whole development set to train the SVM classifier. The columns titled ‘PRE’, ‘# RELS.’, and ‘# EST. CORR. RELS.’ show the precision of the hyponymy relations in the test set, the number of the acquired hyponymy relations, and the expected number of correct hyponymy relations estimated from the precision and the number of the acquired hyponymy relations, respectively. The row titled ‘S & T (2008)’ shows the performance of the method proposed by Sumida and Torisawa (2008). The following two rows show the precision of the HRCs acquired by the patterns in Figure 3 (PAT)\(^7\) and that of the results of machine learning (ML). We successfully obtained more than 1.73 million hyponymy relations with 85.2% precision, which greatly outperformed the results of Sumida and Torisawa (2008) in terms of both the precision and the number of acquired hyponymy relations.

Table 3 shows the performance of the classifier when we eliminated each type of features. The columns titled ‘ACC’, ‘PRE’, ‘REC’, and ‘F1’ show the accuracy, precision, recall, and F1-measure calculated on the test set. All the newly introduced features contributed to the accuracy, and improved the total accuracy by 1.1%. The features DIST and PAT improved the precision of the classifier, while the feature LCHAR improved the recall of the classifier.

To investigate the trade-off between precision and recall, we changed the threshold of the SVM values for the HRCs.

\(^4\)We excluded “user pages”, “special pages”, “template pages”, “redirection pages”, and “category pages”, since they are meant for internal purpose, and excluded “disambiguation pages”, since they only enumerate possible articles for the ambiguous headings.

\(^5\)http://chasen.org/~taku/software/TinySVM/

\(^6\)http://mecab.sourceforge.net/

\(^7\)The patterns marked by “*” in Figure 3 were not used in this acquisition.
Table 4: Precision and recall of the hyponymy relations in terms of the distance

| DIST | NUM | ACC   | PRE   | REC   | F1   |
|------|-----|-------|-------|-------|------|
| 1    | 446 | 88.3  | (394/446) | 84.2 | (139/165) | 84.2 |
| 2    | 345 | 90.7  | (313/345) | 87.3 | (62/71) | 79.5 |
| 3    | 161 | 90.1  | (145/161) | 87.5 | (14/16) | 63.6 |
| 4    | 39  | 97.4  | (38/39)   | 75.0 | (3/4)  | 85.7 |
| 5    | 9   | 77.8  | (7/9)    | 100.0| (1/1)  | 50.0 |
| TOTAL| 1,000 | 89.7 | (897/1,000) | 85.2 | (219/257) | 81.0 |

Figures 4 and 5 show the P-R curve of the hyponymy relation acquisition using the feature set in Table 3. We can observe that the newly introduced features improve the precision in the range of the recall greater than 60%. We can improve the precision of the acquired hyponymy relations by making the threshold of the SVM values larger. By setting the threshold to 0.36, we obtain 1,349,622 hyponymy relations with a precision of 90.1%, which cover 46,653 distinct hypernyms and 739,972 distinct hyponyms. We obtained on average 4.88 hyponymy relations from one Wikipedia article with a precision of 90.1%.

To investigate the contribution of the newly extracted HRCs to the acquired hyponymy relations, we classified the HRCs in the test set into subsets according to the distance between items from which the hypernym and hyponym are extracted. Table 4 shows the performance of the SVM classifier for the resulting subsets of the test set. The column DIST shows the distance between items from which the hypernym and hyponym are extracted, while NUM shows the number of the HRCs. The columns ACC, PRE, REC, and F1 show the accuracy, precision, recall, and F1-measure calculated on each subset of the test set. Although there is a larger number of noisy HRCs in the subsets of the test set which were acquired from distant items (DIST ≥ 2), we successfully maintained the precision of the acquired hyponymy relations above 75%. Boosting the recall of hyponymy relations acquired from the distant items will be the key to improve the performance of our method.

Figure 6 shows the performance of our method when varying the number of the training HRCs from 1,000 to 8,000.\(^8\) We can clearly observe that both precision and recall naturally improved with a larger size of the training data.

\(^8\)Here, all the training HRC samples are chosen from the 9,000 HRCs randomly extracted from all the HRCs excluding the test set. This is because the other training data used for constructing the final classifier were selected from a certain subset of the HRCs.
Table 5: Hyponym relations acquired from hierarchical structures in Wikipedia: incorrect hyponyms are marked as ‘*’. The hyponyms and hyponyms are followed by their English translations.

| HYPERNYM | HYPONYM |
|----------|---------|
| Lake (lake) | Lake Xie (Lake Xie) |
| City (city) | New York City (New York City) |
| Country (country) | United States (United States) |

Note: The hyponyms are manually selected and 10 hyponyms are randomly selected for each hypernym. For some classes such as ‘planet’ and ‘technique’, many fictional objects are marked as ‘*’ as extracted from Wikipedia.
We finally investigated details of the errors in the SVM classifier. We applied the SVM classifier to 1,000 HRCs that were randomly selected from all the HRCs excluding the test and development sets, and manually investigated the classification results. The classification accuracy of these HRCs was 89.1% (233 true positives, 658 true negatives, 22 false positives and 87 false negatives).

Table 6 summarizes the types of false positives. Meronymy (part-of relation; e.g., ‘car’-‘engine’) is the most frequent error, and the current classifier yields a high score for this type of error. To filter out meronymy correctly, we will need additional criteria to judge hyponymy relations, for example whether they have the same attributes in common (Dowty et al., 1980; Almuhareb and Poe-sio, 2004). The hierarchical structures also represented instance-attribute/value relations, and some instance-value pairs were wrongly regarded as hyponymy relations. We found that an attribute that specifies the relation between the instance and the value usually appeared between the nodes from which the instance and the value were extracted. For example, in the hierarchical structure that included ‘Studio Easter’ (a design studio) and ‘Uta Kata’ (TV animation series) as titles of nodes, there was a node titled ‘主な参加作品 (Major work)’ between them. We will be able to filter out these instance-value pairs by using information on the other nodes in the original hierarchical structures as features for machine learning. The other two cases, ‘concept-facet’ and ‘facet-instance’, are both related to a facet label, which is usually a value of a specific attribute to classify instances according to the attribute’s value (e.g., ‘England’ and ‘France’ in Figure 1 are values of the attribute ‘country’ of tea brands). For example, ‘Urawa Red Diamonds’ (a football club) is used to classify ‘supporter’s groups’ in terms of their origination. This paper presented an extended version of Sumida and Torisawa’s method (2008) of acquiring hyponymy relations from the hierarchical structures in Wikipedia. We extract more hyponymy relation candidates from the hierarchical structures than the original method to increase the number of hyponymy relations acquired by the method. We successfully acquired more than 1.34 million hyponymy relations, which doubled the number of hyponymy relations acquired by the method. We will exploit training samples for hyponymy candidates that are synonymous with or superclass of the infrequent hyponymy to solve the data sparseness problem.

Table 7 shows the classification of the 658 true negatives. We found that hierarchical structures in Wikipedia were mainly used to express instance-attribute-value relations, meronymy relations and concept-(facet-)instance relations (hyponymy relations). In Table 7, most of the HRCs classified as ‘other’ were extracted from items in distant positions in the hierarchical structures, and the hyponym and hyponym candidates were irrelevant. We will obtain instance-attribute-value triples from the hierarchical structures.

5. Conclusion

This paper presented an extended version of Sumida and Torisawa’s method (2008) of acquiring hyponymy relations from the hierarchical structures in Wikipedia. We extract more hyponymy relation candidates from the hierarchical structures than the original method to increase the number of hyponymy relations acquired by the method. We successfully acquired more than 1.34 million hyponymy relations, which doubled the number of hyponymy relations acquired by the method, and we also increased the precision by 13.7% (from 76.4% to 90.1%). Since the number of Wikipedia articles increases day by day (cf. 276,323 articles in March 2007 to 449,233 articles in March 2008), we can obtain a larger number of hyponymy relations by simply applying our method to the latest version of Wikipedia.
In future research, we plan to apply the SVM classifier to HRCs acquired from the definition sentences and category labels in Wikipedia articles. We will apply our method to the Wikipedia in other languages, such as English. We will also evaluate the acquired hyponymy relations in practical application contexts.

6. References

Abdulrahman Almuhareb and Massimo Poesio. 2004. Attribute-based and value-based clustering: an evaluation. In Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 158–165.

Sharon A. Caraballo. 1999. Automatic construction of a hyponym-labeled noun hierarchy from text. In Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL), pages 120–126.

David A. Cruse. 1998. Lexical Semantics. Cambridge Textbooks in Linguistics.

David R. Dowty, Robert E. Wall, and Abdessamad Echihabi. 2003. Offline strategies for online question answering: Answering questions before they are asked. In Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL), pages 1–7.

Christiane Fellbaum, editor. 1998. WordNet: an electronic lexical database. MIT Press.

Michael Fleischman, Eduard Hovy, and Abdessamad Echihabi. 2003. Offline strategies for online question answering: Answering questions before they are asked. In Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL), pages 1–7.

Marti A. Hearst. 1992. Automatic acquisition of hyponyms from large text corpora. In Proceedings of the 14th conference on Computational linguistics (COLING), pages 539–545.

Aurelie Herbelot and Ann Copestake. 2006. Acquiring ontological relationships from Wikipedia using RMRS. In Proceedings of the ISWC 2006 Workshop on Web Content Mining with Human Language Technologies.

Jun'ichi Kazama and Kentaro Torisawa. 2007. Exploiting wikipedia as external knowledge for named entity recognition. In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP), pages 698–707.

Gideon S. Mann. 2002. Fine-grained proper noun ontologies for question answering. In Proceedings of the Workshop on SemaNet: building and using semantic networks at the 19th International Conference on Computational Linguistics (COLING), pages 1–7.

Emmanuel Morin and Christian Jacquemin. 2004. Automatic acquisition and expansion of hyponym links. Computers and the Humanities, 38(4):363–396.

Patrick Pantel and Marco Pennacchiotti. 2006. Espresso: Leveraging generic patterns for automatically harvesting semantic relations. In Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics (COLING-ACL), pages 113–120.

Maria Ruiz-Casado, Enrique Alfonseca, and Pablo Castells. 2005. Automatic extraction of semantic relationships for WordNet by means of pattern learning from Wikipedia. In Natural Language Processing and Information Systems: 10th International Conference on Applications of Natural Language to Information Systems, NLDB 2005, pages 67–79.

Keiji Shinzato and Kentaro Torisawa. 2004. Acquiring hyponymy relations from Web documents. In Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL), pages 73–80.

Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. 2007. YAGO: A core of semantic knowledge unifying WordNet and Wikipedia. In Proceedings of the 16th International World Wide Web Conference (WWW), pages 697–706.

Asuka Sumida and Kentaro Torisawa. 2008. Hacking Wikipedia for hyponym relation acquisition. In Proceedings of the Third International Joint Conference on Natural Language Processing (IJCNLP). 883–888.

Asuka Sumida, Kentaro Torisawa, and Keiji Shinzato. 2006. Concept-instance relation extraction from simple noun sequences using a search engine on a web repository. In Proceedings of the Web Content Mining with Human Language Technologies workshop on the fifth International Semantic Web.

Vladimir N. Vapnik. 1998. Statistical Learning Theory. Wiley-Interscience.

Naoki Yoshinaga and Kentaro Torisawa. 2006. Finding specification pages according to attributes. In Proceedings of the 15th International World Wide Web Conference (WWW), pages 1021–1022.