Optimizing Synonym Extraction Using Monolingual and Bilingual Resources

Hua WU, Ming ZHOU
Microsoft Research Asia
5F Sigma Center, No.49 Zhichun Road, Haidian District
Beijing, 100080, China
wu_hua_@msn.com, mingzhou@microsoft.com

Abstract

Automatically acquiring synonymous words (synonyms) from corpora is a challenging task. For this task, methods that use only one kind of resources are inadequate because of low precision or low recall. To improve the performance of synonym extraction, we propose a method to extract synonyms with multiple resources including a monolingual dictionary, a bilingual corpus, and a large monolingual corpus. This approach uses an ensemble to combine the synonyms extracted by individual extractors which use the three resources. Experimental results prove that the three resources are complementary to each other on synonym extraction, and that the ensemble method we used is very effective to improve both precisions and recalls of extracted synonyms.

1 Introduction

This paper addresses the problem of extracting synonymous English words (synonyms) from multiple resources: a monolingual dictionary, a parallel bilingual corpus, and a monolingual corpus. The extracted synonyms can be used in a number of NLP applications. In information retrieval and question answering, the synonymous words are employed to bridge the expressions gaps between the query space and the document space (Mandala et al., 1999; Radev et al., 2001; Kiyota et al., 2002). In automatic text summarization, synonymous words are employed to identify repetitive information in order to avoid redundant contents in a summary (Barzilay and Elhadad, 1997). In language generation, synonyms are employed to create more varied texts (Langkilde and Knight, 1998).

Up to our knowledge, there are few studies investigating the combination of different resources for synonym extraction. However, many studies investigate synonym extraction from only one resource. The most frequently used resource for synonym extraction is large monolingual corpora (Hindle, 1990; Crouch and Yang, 1992; Grefenstatte, 1994; Park and Choi, 1997; Gasperin et al., 2001 and Lin, 1998). The methods used the contexts around the investigated words to discover synonyms. The problem of the methods is that the precision of the extracted synonymous words is low because it extracts many word pairs such as "cat" and "dog", which are similar but not synonymous.

Other resources are also used for synonym extraction. Barzilay and Mckeown (2001), and Shimohata and Sumita (2002) used bilingual corpora to extract synonyms. However, these methods can only extract synonyms which occur in the bilingual corpus. Thus, the extracted synonyms are limited. Besides, Blondel and Sennelart (2002) used monolingual dictionaries to extract synonyms. Although the precision of this method is high, the coverage is low because the result of this method heavily depends on the definitions of words.

In order to improve the performance of synonym extraction, Curran (2002) used an ensemble method to combine the results of different methods using a monolingual corpus. Although Curran (2002) showed that the ensemble extractors outperformed the individual extractors, it still cannot overcome the deficiency of the methods using the monolingual corpus.

To overcome the deficiencies of the methods using only one resource, our approach combines both monolingual and bilingual resources to automatically extract synonymous words. By combining the synonyms extracted by the individual extractors using the three resources, our approach can combine the merits of the individual extractors to

improve the performance of synonym extraction. In fact, our approach can be considered as an ensemble of different resources for synonym extraction. Experimental results prove that the three resources are complementary to each other on synonym extraction, and that the ensemble method we used is very effective to improve both precisions and recalls of extracted synonyms.

The remainder of this paper is organized as follows. The next section presents our approach for synonym extraction. Section 3 describes an implementation of the three individual extractors. Section 4 presents the evaluation results. Section 5 discusses our method. In the last section, we draw the conclusions of this work.

2 Our Approach

Instead of using only one kind of resource to extract synonyms, we combine both monolingual and bilingual resources for synonym extraction. The resources include a monolingual dictionary, an English-Chinese bilingual corpus, and a large corpus of monolingual documents. Before combining them, we first propose three methods to extract synonyms from the three resources. Especially, a novel method is proposed to increase the coverage of the extracted synonyms using the bilingual corpus. Next, we develop an ensemble method to combine the individual extractors. The advantage of our approach is that it can combine the merits of the individual extractors to improve the precision and recalls of the extracted synonyms.

2.1 Synonym Extraction with a Monolingual Dictionary

This section proposes a method to extract synonyms from a monolingual dictionary. In a monolingual dictionary, each entry is defined by other words and may also be used in the definitions for other words. For a word in the dictionary, the words used to define it are called hubs and the words whose definitions include this word are called authorities as in (Blondel and Sennelart, 2002). We use the hubs and authorities of a word to represent its meaning. The assumption behind this method is that two words are similar if they have common hubs and authorities. In this paper, we only use content words as members of hubs and authorities.

We take these hubs and authorities as features of a word. The vector constructed with them is referred to as the feature vector of a word. The similarity between two words is calculated through their feature vectors with the cosine measure as shown in Equation (1).

$$sim_1(w_1, w_2) = \cos(F_1, F_2) = \frac{\sum_{w_{1i}=w_{2j}} (v_{1i} * v_{2j})}{\sqrt{\sum_i v_{1i}^2} * \sqrt{\sum_j v_{2j}^2}} \quad (1)$$

where

$$F_i = < (w_{i1}, v_{i1}), (w_{i2}, v_{i2}), ... (w_{im}, v_{im}) >$$

F_i is the feature vector of w_i;

$v_{ij} = 1$ if word w_{ij} is a hub or an authority of the word w_i; else, $v_{ij} = 0$;

2.2 Synonym Extraction with a Bilingual Corpus

This section proposes a novel method to extract synonyms from a bilingual corpus. It uses the translations of a word to express its meaning. The assumption of this method is that two words are synonymous if their translations are similar.

Given an English word, we get their translations with an English-Chinese bilingual dictionary. Each translation is assigned a translation probability, which is trained with a bilingual English-Chinese corpus based on the result of word alignment. The aligner use the model described in (Wang et al., 2001). In order to deal with the problem of data sparseness, we conduct a simple smoothing by adding 0.5 to the counts of each translation pair as in (2).

$$p(c \mid e) = \frac{count(c, e) + 0.5}{count(e) + 0.5 * |trans_c|} \quad (2)$$

where

$count(c, e)$ represents the co-occurring frequency of the Chinese word c and the English word e in the sentence pairs.

$count(e)$ represents the frequency of the English word e occurring in the bilingual corpus.

$|trans_c|$ represents the number of Chinese translations for a given English word *e*.

The translations and the translation probabilities of a word are used to construct its feature vector. The similarity of two words is estimated through their feature vectors with the cosine measure as shown in (3).

$$sim_2(w_1, w_2) = \cos(F_1, F_2) = \frac{\sum_{c_{1i}=c_{2j}} (p_{1i} * p_{2j})}{\sqrt{\sum_i p_{1i}^2} * \sqrt{\sum_j p_{2j}^2}} \quad (3)$$

where

$F_i =< (c_{i1}, p_{i1}), (c_{i2}, p_{i2}), ... (c_{im}, p_{im}) >$

F_i is the feature vector of w_i;

c_{ij} is the j[th] Chinese translation of the word w_i;

p_{ij} is the translation probability of the word w_i is translated into c_{ij}

For example, the feature vectors of two words "abandon" and "forsake" are:

forsake: < (放弃, 0.1333), (抛弃, 0.1333), (唾弃, 0.0667) (遗弃, 0.0667), (离开, 0.0667), …>

abandon: <(放弃, 0.3018), (抛弃, 0.1126), (遗弃, 0.0405), (舍弃, 0.0225), (抛开, 0.0135),…>

2.3 Synonym Extraction with a Monolingual Corpus

The context method described in Section 1 is also used for synonym extraction from large monolingual corpora of documents. This method relies on the assumption that synonymous words tend to have similar contexts. In this paper, we use the words which have dependency relationships with the investigated word as contexts. The contexts are obtained by parsing the monolingual documents. The parsing results are represented by dependency triples which are denoted as <w_1, Relation Type, w_2>. For example, the sentence "I declined the invitation" is transformed into three triples after parsing: <decline, SUBJ, I>, <decline, OBJ, invitation> and <invitation, DET, the>. If we name <Relation Type, w_2> as an attribute of the word w_1, the verb "decline" in the above sentence has two attributes <OBJ, invitation> and <SUBJ, I> . Thus, the contexts of a word can be expressed using its attributes. In this case, two words are synonymous if they have similar attributes.

We use a weighted version of the Dice measure to calculate the similarity of two words.

$$sim_3(w_1, w_2) = \frac{\sum_{att_k \in A(w_1) \cap A(w_2)} (W(w_1, att_k) + W(w_2, att_k))}{\sum_{att_i \in A(w_1)} W(w_1, att_i) + \sum_{att_j \in A(w_2)} W(w_2, att_j)} \quad (4)$$

where

att_i, att_j, att_k stands for attributes of words.

$W(w_i, att_j)$ indicates the association strength between the attribute att_j with the word w_i.

$A(w_i)$ denotes the attribute set of the word w_i.

The measure used to measure the association strength between a word and its attributes is weighted mutual information (WMI) (Fung and Mckeown, 1997) as described in (5).

$$W(w_i, att_j) = WMI(w_i, att_j) = p(w_i, att_j) * \log \frac{p(w_i, att_j)}{p(w_i) * p(att_j)} \quad (5)$$

where

$$p(w_i) = \frac{count(w_i, *, *)}{N}$$

$$p(att_j) = \frac{count(*, r, w)}{N}, \quad att_j = (r, w)$$

$count(*, r, w)$: frequency of the triples having dependency relation *r* with the word *w*.

$count(w_i, *, *)$: frequency of the triples including word w_i.

N: number of triples in the corpus.

We use it instead of point-wise mutual information in Lin (1998) because the latter tends to overestimate the association between two parts with low frequencies. Weighted mutual information meliorates this effect by adding $p(w_i, att_j)$.

2.4 Combining the Three Extractors

In terms of combining the outputs of the different methods, the ensemble method is a good candidate. Originally, the ensemble method is a machine learning technique of combining the outputs of several classifiers to improve the classification performance (Dietterich, 2000). It has been successfully used in many NLP tasks. For example, (Curran, 2002) proved that the ensembles of indi-

vidual extractors using different contexts in the monolingual corpus improve the performance of synonym extraction.

In fact, we can consider the extractors in the previous sections as binary classifiers. Thus, we use the ensemble method to combine the output of the individual extractors described in the previous sections for synonym extraction. The method is described in Equation (6).

$$sim(w_1, w_2) = \sum_{i=1}^{3} (a_i * sim_i(w_1, w_2)) \quad (6)$$

where

$sim_i(w_1, w_2)$ (i = 1, 2, 3) stands for the different similarity measure using different resources described in the previous sections.

a_i ($i = 1, 2, 3$, and $\sum_i a_i = 1$) is the weight for the individual extractors.

The reasons that we use the weighted ensemble method are as follows: (1) If the majority of three extractors select the same word as a synonym of a investigated word, it tend to be a real synonym. This method can ensure it has a high similarity score. Thus, it will improve the precision of the extracted synonyms. (2) With this method, it can improve the coverage of the extracted synonyms. This is because if the similarity score of a candidate with the investigated word is higher than a threshold, our method can select the candidate as a synonym even though it is only suggested by one extractor.

3 Implementation of Individual Extractors

For the extractor employing a monolingual dictionary, we use the same online dictionary as in (Blondel and Sennelart, 2002), which is named the Online Plain Text Dictionary. The dictionary consists of 27 HTML files, which is available from the web site http://www.gutenberg.net/. With the method described in Section 2.1, the result for the extracted synonyms is shown in Table 1 when the similarity threshold is set to 0.04. An example is shown as follows:

acclimatize:

(acclimate, 0.1481; habituate, 0.0976)

The numbers in the example are the similarity scores of two words.

Table 1. Synonyms Extracted from the Monolingual Dictionary

Category	# Entries	# Average Synonyms
Noun	16963	4.7
Verb	5084	7.1

For synonym extraction from the bilingual corpus, we use an English-Chinese lexicon, which includes 219,404 English words with each source word having 3 translations on average. The word translation probabilities are estimated from a bilingual corpus that obtains 170,025 pairs of Chinese-English sentences, including about 2.1 million English words and about 2.5 million Chinese words. With the method described in Section 2.2, we extracted synonyms as shown in Table 2 when the similarity threshold is set to 0.04.

Table 2. Synonyms Extracted from the Bilingual corpus

Category	#Entries	#Average Synonyms
Noun	26253	10.2
Verb	7364	14.8

For synonym extraction from a monolingual corpus, we use the Wall Street Journal from 1987 to 1992, the size of which is about 500M bytes. In order to get contexts of words, we parse the corpus with an English parser –NLPWIN[1]. From the parsing results, we extracted the following four types of dependency triples.

(a) *<verb, OBJ, noun>*
(b) *<verb, SUBJ, noun>*
(c) *<noun, ATTRIB, adjective>*
(d) *<verb, MODS, adjunct>*

The statistics are shown in Table 3. *Token* means the total number of triples in the triple set and *type* means a unique instance of triple in the corpus. These triples are used as contexts of words to calculate the similarity between words as described in Section 2.3. The result is shown in Table 4 when the similarity threshold is set to 0.1.

[1] The NLPWIN parser is developed at Microsoft Research. Its output can be a phrase structure parse tree or a logical form which is represented with dependency triples.

Table 3. Statistics for Triples

	# Token	# Type
OBJ	7,041,382	1,487,543
SUBJ	7,180,572	2,116,761
ATTRIB	4,976,822	1,393,188
MODS	3,557,737	94,004
Total	22,756,512	5,937,496

Table 4. Synonyms Extracted from the Monolingual Corpus

Category	Entries	Average Synonyms
Noun	16963	4.6
Verb	5084	7.1

4 Evaluation

4.1 The Gold Standard

The simplest evaluation measure is direct comparison of the extracted synonyms with the manually created thesaurus. However, the thesaurus coverage is a problem. In this paper, we combined two thesauri as a gold stardard: WordNet 1.6 http://www.cogsci.princeton.edu/~wn/) and Roget (Roget's II: The New Thesaurus, 1995. http://www.bartleby.com/thesauri/).

In WordNet, one synset consists of several synonyms which represent a single sense. Therefore, a polysemous word occurs in more than one synsets. For example, the polysemous word "abandon" occur in five different synsets:

(abandon, forsake, desolate, desert, lurch)
(vacate, empty, abandon)
(abandon, give up, give)
(abandon, give up)
(abandon)

For a given word, we combine its synonyms from all synsets including the word. Thus, we get the synonyms of the word "abandon" as follows:

abandon forsake, desolate, desert, lurch, vacate, empty, give up, give

For synonyms in Roget, we also combine the synonyms in different synsets into one set as we do for WordNet. Thus, we get the synonyms of the word "abandon"as follows:

abandon break off, desist, discontinue, give up, leave off, quit, relinquish, remit, stop, desert, forsake, leave, throw over, abdicate, cede, demit, forswear, hand over, quitclaim, render, renounce, resign, surrender, waive, yield, give over, forgo, lay down

Combining the results of WordNet and Roget, we can get the synonyms of the word "abandon" as follows.

abandon desolate, lurch, vacate, empty, give, abdicate, break off, cede, demit, desert, desist, discontinue, forgo, forsake, forswear, give up, give over, hand over, lay down, lay off, leave off, leave, quit, quitclaim, relinquish, remit, stop, swear off, throw over, render, renounce, resign, surrender, waive, yield

4.2 Evaluation Measures

The evaluation metrics are precision, recall, and f-measure. If we use S to indicate the synonyms that our method extracts for a word and S_G to denote the synonyms of the word in WordNet and Roget, the methods to calculate the precision, recall, and f-measure of our methods are shown in Equation (7), (8), and (9). To investigate the results of more than one word, we calculate the average precision, recall and f-measure, which sum the individual values divided by the number of the investigated words.

$$precision = \frac{|S \cap S_G|}{|S|} \quad (7)$$

$$recall = \frac{|S \cap S_G|}{|S_G|} \quad (8)$$

$$\text{f - measure} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (9)$$

4.3 Test Set

In order to evaluate our methods, we build up a test set which includes three parts:

(a) high-frequency words: occurring more than 100 times;
(b) middle-frequency words: occurring more than 10 times and not greater than 100 times;
(c) low-frequency words: occurring no greater than 10 times.

Table 5. Statistics for Nouns and Verbs

	High Frequency	Middle Frequency	Low Frequency
Noun	600	2000	1000
Verb	340	1300	800

The frequency counts are estimated from Wall Street Journal (1987-1992), from which we randomly extracted 3600 nouns and 2440 verbs. These

Table 6. Evaluation Results for Nouns

	High-Frequency Nouns			Middle-Frequency Nouns			Low-Frequency Nouns		
	Pre	Rec	F	Pre	Rec	F	Pre	Rec	F
1	0.174	0.140	0.155	0.212	0.137	0.167	0.198	0.119	0.149
2	0.225	**0.209**	0.217	0.242	0.212	0.226	0.207	0.212	0.209
3	0.118	0.109	0.114	0.117	0.104	0.109	0.099	0.096	0.098
1+2+3	**0.240**	0.201	**0.219**	**0.271**	**0.220**	**0.243**	**0.222**	**0.232**	**0.227**

Table 7. Evaluation Results for Verbs

	High-Frequency Verbs			Middle-Frequency Verbs			Low-Frequency Verbs		
	Pre	Rec	F	Pre	Rec	F	Pre	Rec	F
1	0.228	0.243	0.235	0.272	0.233	0.251	0.209	0.216	0.212
2	0.226	0.312	0.262	0.224	0.292	0.253	0.184	0.275	0.220
3	0.143	0.166	0.154	0.162	0.127	0.142	0.128	0.135	0.132
1+2+3	**0.295**	**0.323**	**0.308**	**0.311**	**0.304**	**0.307**	**0.238**	**0.302**	**0.266**

Note: 1, 2, and 3 represent the extractor using the monolingual dictionary, the bilingual corpus, and the monolingual corpus respectively. The symbols "Pre", "Rec", and "F" represent precision, recall, and f-measure scores.

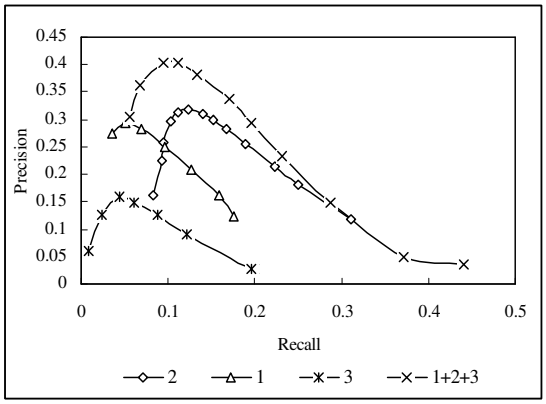

Figure 1. Recall-Precision curves for nouns

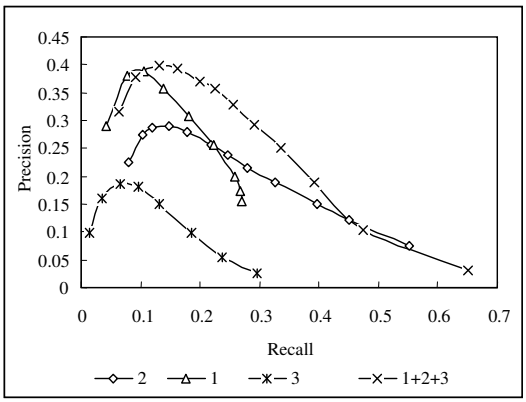

Figure 2. Recall-Precision curves for verbs

words have synonyms both in our results extracted from the three resources and in the thesauri WordNet and Roget. The statistics of the test set are shown in Table 5.

4.4 Experimental Results

In this section, we compare the extracted synonyms of the nouns and verbs in the test set with those in WordNet and Roget. For each method, we select those as synonyms whose similarity scores with the investigated word are larger than a given threshold. A development set is used to determine the thresholds of each method. The thresholds for getting highest f-measure scores on the development set are selected. In our experiments, we get 0.04, 0.04, 0.1 and 0.04 for Method 1, Method 2, Method 3 and the combined approach, respectively. The evaluation results for the individual extractors and the ensemble extractor are shown in Table 6 and Table 7. We set a_1=0.4, a_2=0.4 and a_3=0.2 in Equation (6) for the ensemble to combine the results from the three resources. The weights are also obtained with the development set.

In order to examine the performance of each method in more details, we also get the precisions and recalls under different thresholds. Figure 1 and Figure 2 shows the precision values under different recall values (different thresholds) for all nouns and verbs, respectively.

Among all of the methods, the method combining all of the three resources gets the best results in terms of both precision and recall. The effect is similar to the ensemble methods for synonym

extraction in (Curran 2002). However, our method uses an ensemble of different resources instead of one single resource. During the experiments, we also find the ensemble combining all of the three extractors outperforms the ensembles only combining any two of the three extractors. This indicates that the extractors using the three different resources are complementary to each other. For example, the extractor using the monolingual dictionary gets a high precision and the extractor using the bilingual corpus gets a high recall. Although the extractor using the monolingual corpus achieved much lower precision and recall on synonym extraction, it is still useful to be included in the ensemble. This shows that the monolingual corpus is complementary to the other two resources on synonym extraction. The success of our method also indicates that our ensemble method by weighting all extractors is effective for synonym extraction.

Among the methods only using one kind of resource, Method 2, which uses the bilingual corpus, has the highest f-measure scores on both nouns and verbs. From the results in Figure 1 and Figure 2, we can see that the coverage of synonyms extracted by Method 2 is the highest. Although it has lower precisions than Method 1 under low recalls, its precisions are higher than those of Method 1 under higher recalls. This shows that Method 2 can get a good compromise between precision and recall. We also note that the maximum recall of Method 2 is much larger than that of Method 1. This is because (1) in Method 1, the words used in the definitions are highly limited. Thus, the coverage of the synonyms is limited; (2) the advantage of Method 2 is that the coverage of extracted synonyms is high because it can extract the synonyms not occurring in the corpus. It is different from the method in (Barzilay and Mckeown, 2001; Shimohata and Sumita, 2002), which can only extract the synonyms in the bilingual corpus.

The performance of Method 3 is the worst. It is caused by two factors: (1) the context model of Method 3 introduces much noise because of the errors of the parser; (2) this method is unable to distinguish synonyms, antonyms, and similar words because they tend to have similar contexts. From the contexts it uses, method 3 is suitable to extract related words which have the similar usages from the view of syntax.

5 Discussions

This paper uses three different methods and resources for synonym extraction. By using the corpus-based method, we can get some synonyms or near synonyms which can not be found in the hand-built thesauri. For Example: "handspring handstand", "audiology otology", "roisterer carouser" and "parmesan gouda". These kinds of synonyms are difficult for hand-built thesauri to cover because they occur too infrequent to be caught by humans. In addition, this corpus-based method can get synonyms in specific domains while the general thesauri don't provide such fine-grained knowledge.

Comparing the results with the human-built thesauri is not the best way to evaluate synonym extraction because the coverage of the human-built thesaurus is also limited. However, manually evaluating the results is time consuming. And it also cannot get the precise evaluation of the extracted synonyms. Although the human-built thesauri cannot help to precisely evaluate the results, they can still be used to detect the effectiveness of extraction methods.

Conclusion

This paper proposes a new method to extract synonyms from three resources: a monolingual dictionary, a bilingual corpus, and a large monolingual corpus. This method uses a weighted ensemble to combine all of the results of the individual extractors using one of the three resources respectively. Experimental results prove that the three resources are complementary to each other on synonym extraction, and that the ensemble method we used is very effective to improve both precisions and recalls when the results are compared with the manually-built thesauri WordNet and Roget.

Further, we also propose a new method to extract synonyms from a bilingual corpus. This method uses the translations of a word to represent its meaning. The translation probabilities are trained with the bilingual corpus. The advantage of this method is that it can improve the coverage of the extracted synonyms. Experiments indicate that this method outperforms the other methods using a monolingual corpus or a monolingual dictionary.

The contribution of this work lies in three aspects: (1) develop a method to combine the results of individual extractors using the three resources on synonym extraction; (2) investigate the performance of the three extraction methods using different resources, exposing the merits and demerits of each method; (3) propose a new method to extract synonyms from a bilingual corpus, which greatly improves the coverage of the extracted synonyms.

References

Barzilay R. and Elhadad M. 1997. *Using lexical chains for text summarization.* In proceedings of the ACL Workshop on Intelligent Scalable Text Summarization, pp10-17.

Barzilay R. and McKeown K. 2001. *Extracting Paraphrases from a Parallel Corpus.* In Proc. of ACL/EACL.

Blondel V. D. and Sennelart P. 2002. *Automatic extraction of synonyms in a dictionary*. In Proc. of the SIAM Workshop on Text Mining.

Crouch C. J. and Yang B. 1992. *Experiments in automatic statistical thesaurus construction.* In Proc. of the 15th Annual International ACM SIGIR conference on Research and Development in Information Retrieval, pp77-88.

Curran J. 2002 *Ensemble Methods for Automatic Thesaurus Extraction.* In Proc. of the Conference on Empirical Methods in Natural Language Processing. pp. 222-229.

Dietterich T. 2000. *Ensemble Methods in Machine Learning*. In Proc. of the First International Workshop on Multiple Classier Systems. pp 1-15.

Fung P., Mckeown K. 1997. *A Technical Word- and Term- Translation Aid Using Noisy Parallel Corpora across Language Groups.* In: Machine Translation, Vol.1-2 (special issue), pp53-87.

Gasperin C., Gamallo P., Agustini A., Lopes G., Lima V. 2001 *Using Syntactic Contexts for Measuring Word Similarity.* Workshop on Knowledge Acquisition & Categorization, ESSLLI.

Grefenstette G. 1994 *Explorations in Automatic Thesaurus Discovery.* Kluwer Academic Press.

Hindle D. 1990. *Noun Classification from Predicate-Argument Structure.* In Proc. of the 28th Annual Meeting of the Association for Computational Linguistics.

Kiyota Y., Kurohashi S., Kido F. 2002. *"Dialog Navigator": A Question Answering System Based on Large Text Knowledge Base*. In Proc. of the 19th International Conference on Computational Linguistics.

Langkilde I. and Knight K. 1998. *Generation that Exploits Corpus-based Statistical Knowledge*. In Proc. of the COLING-ACL.

Lin D. 1998 *Automatic Retrieval and Clustering of Similar Words.* In Proc. of the 36th Annual Meeting of the Association for Computational Linguistics.

Mandala R., Tokunaga T. Tanaka H. 1999. *Combining Multiple Evidence from Different Type of Thesaurus for Query Expansion*. In Proc. of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval.

Park Y.C. and Choi K. S. 1997. *Automatic Thesaurus Construction Using Baysian Networks.* Information Processing & Management. Vol. 32.

Radev D., Qi H., Zheng Z., Goldensohn S., Zhang Z., Fan W., Prager J. 2001. *Mining the Web for Answers to Natural Language Questions*. In the Tenth International ACM Conference on Information and Knowledge Management.

Shimohata M. and Sumita E. 2002. *Automatic Paraphrasing Based on Parallel Corpus for Normalization.* In Proc. of the Third International Conference on Language Resources and Evaluation.

Wang W., Huang J., Zhou M., Huang C.N. 2001. *Finding Target Language Correspondence for Lexicalized EBMT System.* In Proc. of the 6th Natural Language Processing Pacific Rim Symposium.