Dictionary Alignment for Context-sensitive Word Glossing

Willy Yap and **Timothy Baldwin**
Department of CSSE
University of Melbourne
{willy,tim}@csse.unimelb.edu.au

Abstract

This paper proposes a method for automatically sense-to-sense aligning dictionaries in different languages (focusing on Japanese and English), based on structural data in the respective dictionaries. The basis of the proposed method is sentence similarity of the sense definition sentences, using a bilingual Japanese-to-English dictionary as a pivot during the alignment process. We experiment with various embellishments to the basic method, including term weighting, stemming/lemmatisation, and ontology expansion.

1 Introduction

In a multi-lingual environment such as the Internet, users often stumble across webpages authored in an unfamiliar language which potentially contain information of interest. While users can consult dictionaries to help them understand the content of the webpages, the process of looking up words in unfamiliar languages is at best time-consuming, and at worst impossible due to a range of reasons. First, the writing system of the language may be unfamiliar to the user, e.g. the Cyrillic alphabet for a monolingual English speaker. Second, the user may not be familiar with the non-segmenting nature of languages such as Chinese and Japanese, and hence be incapable of delimiting the words to look up in the dictionary in the first place. Third, the user may be unable to lemmatise the word to determine the form in which it is listed in a dictionary.

There are several alternatives to help decipher webpages in unfamiliar languages. The first one is to use an online machine translation system such as Altavista's Babel Fish[1] or Google Translate.[2]
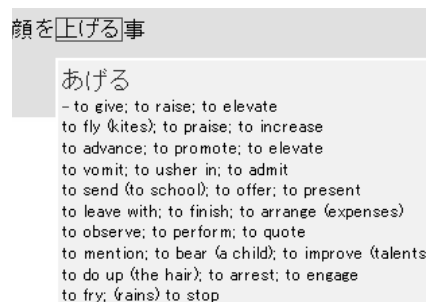

Figure 1: Multiple translations for the Japanese word 上げる [*ageru*] produced by rikai.com. The correct translation in this context is "to raise".

While web-based machine translation services occasionally produce good translations for linguistically-similar languages such as English and French, they do not perform very well in translating languages which are removed from one another (Koehn, 2005).

The second alternative is a pop-up glossing application. The application takes raw text or a URL, parses the words, and returns the pop-up translation of each word as the mouse hovers over it. Some example pop-up glossing applications for Japanese source text and English glosses are Rikai[3] and POPjisyo.[4] With the aid of these pop-up translations, the manual effort of segmenting words (if necessary) and looking up each can be avoided. This application is also useful as an educational aid for learners of that language.

The drawback with these applications is they display all possible translations of a given word irrespective of context. Faced with the task of determining the correct translation themselves, users frequently misinterpret words. An illustration of this situation is given in Figure 1.

---

[1] http://babelfish.altavista.com/
[2] http://www.google.com/translate_t

[3] http://www.rikai.com/perl/Home.pl
[4] http://www.popjisyo.com

We propose a context-sensitive dictionary glossing application to enhance the utility of on-line glossing applications by sensitising the presented glosses to the context of use. The proposed method works by combining a monolingual word sense disambiguation (WSD) system (Baldwin et al., to appear) with an automatically induced cross-lingual sense alignment table. Based on the prediction(s) of the WSD system, our application presents the corresponding set of context-sensitive glosses to the user dictionary glossing by analysing the output of the alignment process.

This paper focuses on the cross-lingual sense alignment aspect of the application. We take separate sense inventories for two distinct languages (Japanese and English in our case) and align the senses between the two. The basis of the alignment process is overlap in sense definitions. By adjusting a threshold for the required level of match, we are able to adjust the precision and recall of the alignment. In preliminary experimentation, we achieve promising results.

The remainder of this paper is structured as follows. We review previous research on dictionary alignment in Section 2, and outline the various resources we utilise during the alignment process in Section 3. We then describe the proposed basic sense-to-sense alignment method, along with various enhancements (Section 4), and present our experimental method and the results of our experiments (Sections 5 and 6, respectively). Finally we discuss our results and future research in Section 7.

2 Previous Research

There has been a significant amount of research on bilingual dictionary alignment using a third language as a pivot. For example, Shirai et al. (2001) built Japanese–French and Japanese–Korean dictionaries using English as the pivot language. In other research, Paik et al. (2001) used English and Chinese as pivots to generate a Korean–Japanese dictionary: English because of the accessibility of Korean–English and Japanese–English dictionaries, and Chinese because of the high overlap in orthography between Korean and Japanese, based on Chinese hanzi.

There have been numerous attempts to manually develop multilingual resources that include cross-lingual sense alignments (Vossen, 1998; Stamou et al., 2002), and the import of cross-lingual semantic alignment has been ably demonstrated by the high impact of these resources. Due to the high overhead in manually constructing such resources, there have been various attempts at automatic cross-lingual sense alignment. The methods are predominantly corpus-driven, based either on cross-lingual distributional similarity in a comparable corpus (e.g. Ngai et al. (2002)) or word alignment over a parallel corpus (e.g. Gliozzo et al. (2005)).

There is a lesser amount of research on cross-lingually aligning ontologies without using large-scale corpus data, which we discuss in greater detail as it is more closely related to that proposed in this research. Asanoma (2001) aligned the Japanese Goi-Taikei ontology with WordNet by first translating a significant subset of the WordNet synonym sets (synsets) into Japanese, automatically matching these based on (monolingual Japanese) lexical overlap, and "filling in the gaps" for the remaining classes based on their hierarchical positioning relative to the aligned classes. Knight and Luk (1994) aligned Spanish and English senses based on: (1) overlap in sets of translations corresponding to each sense of a given Spanish word, with synsets in WordNet; and (2) domain codes in the Spanish and English ontologies. They additionally aligned monolingual English dictionaries based on overlap in the definitions of each sense. The former cross-lingual case assumes a sense-discriminated bilingual dictionary, which we do not have access to. The latter case is similar to our research in that it compares definition sentences, but differs in that the definitions are in the same language. The most closely related work to our research is that of Nichols et al. (2005), who aligned Lexeed senses with WordNet synsets as a by-product of the Lexeed ontology induction task (see Section 3.1), although they do not provide an explicit evaluation of the Lexeed–WordNet alignment for direct comparison.

3 Resources

In this section, we review the key resources used in this research.

Word　竜　*ryuu*
POS　noun

Sense 1
- Lexical-type　`noun-lex`
- Definition　想像/上/の/動物/。　体/は/巨大/な/蛇/に/似/、/４/本/の/足/と/角/を/持つ 。
海中/や/湖/や/沼/に/棲み、/空/に/昇っ/て/雲/を/起こし/雨/を/降ら/せる/と/言う。
An imaginary animal. Dragons are like enormous snakes with 4 legs and horns.
Dragons live in the sea, lakes and ponds, and are said to form clouds and cause rain
when they fly up into the sky.
- Hypernym　ANIMAL

Sense 2,3,4　[...]

Sense 5
- Lexical-type　`noun-lex`
- Definition　将棋/で/、/飛車/の/成っ/た/もの。
In shogi, a promoted rook.
- Domain　SHOGI

Figure 2: A partial view of the Lexeed entry for 竜 [*ryuu*] (with English glosses)

3.1 The Lexeed semantic database of Japanese

The Lexeed Semantic Database of Japanese is a machine-readable dictionary consisting of the most commonly-used words in Japanese (Kasahara et al., 2004). In total, there are 28,000 words in Lexeed, and a total of 46,437 senses. Associated with each sense is a set of definition sentences, constructed entirely using the closed vocabulary of the 28,000 words found in Lexeed, such that 60% of the 28,000 words occur in the definition sentences (Tanaka et al., 2006). In addition to the definition sentences, Lexeed also contains part of speech (POS), lexical relations between the senses (if any) and an example sentence, also based on the closed vocabulary of 28,000 words. All content words in the definition and example sentences are sense annotated.

Automatic ontology acquisition methods have been applied to Lexeed to induce lexical relations between sense pairs, based on the sense-annotated definition sentences (Nichols et al., 2005) and comparison with both the Goi-Taikei thesaurus and WordNet 2.0.

An example Lexeed entry for the word *ryuu* is given in Figure 2.

3.2 EDICT

EDICT is a free machine-readable Japanese-to-English dictionary (Breen, 1995). The project is highly active and has been extended to other target languages such as German, French and Russian. EDICT contains more than 170,000 Japanese entries, each of which is associated with one or more English glosses. It also optionally contains information such as the pronunciation of the entry, POS, and domain of application.

3.3 WordNet

WordNet is an electronic semantic lexical database of English (Fellbaum, 1998). It is made up of more than 100,000 synsets, with each synset representing a group of synonyms. Its entries are categorised into four POS categories: nouns, verbs, adjectives and adverbs. Each POS is described in a discrete lexical network.

Every synset in WordNet has a definition sentence, and sample sentence(s) are provided for most of the synsets; in combination, these are termed the WordNet gloss. Semantic relations connect one synset to another, and include relation types such as hypernym, hyponymy, antonymy and meronymy. The majority of these relations do not cross POS boundaries.

Since we only experiment with hypernyms (and, symmetrically, hyponyms), we provide a simple review of this relation. A synset A is a hypernym of a synset B iff B is a kind of A. For example, *vehicle* is a hypernym of *car*, while *perceive* is a hypernym of *hear, sight, touch, smell, taste*.[5]

---

[5] Strictly speaking, *hear*, etc. are troponyms of *perceive*, i.e. they denote specific ways of *perceiving*. Because WordNet

"promoted rook (shogi)"

⬇

"promoted rook shogi"

⬇

"promoted rook"

⬇

"rook"

Figure 3: Example of normalisation of the translation string; we stop at "rook" as WordNet has a matching entry for it

When building the baseline for our evaluation, we used the SemCor corpus—a subset of the Brown corpus annotated with WordNet senses—to derive the frequency counts of each WordNet synset (Landes et al., 1998). Section 5 discusses this process in more detail.

4 Proposed Methods

Our basic alignment method, along with various extensions, is outlined below.

4.1 Basic alignment method using cosine similarity

In this paper, we align a semantic database of Japanese (Lexeed) with a semantic network of English (WordNet) at the sense level. First, we use Lexeed to find all possible senses of a given word, and retrieve the definition sentences for each.

Since all the definition sentences are in Japanese, we use EDICT as a pivot to convert Lexeed definition sentences into English. In this process, all possible translations of all Japanese words found in the definition sentences are returned, along with their POS classes. For every translation returned, we find entries in WordNet that match the translation and POS category. If there is no match for the given POS, we relax this constraint and search for entries in WordNet that match the translation but not the POS.

Problems arise when WordNet does not have a matching entry for the translation. This situation

doesn't distinguish between hyponyms and troponyms, however, we treat the two identically.

usually happens when the translation returned by EDICT is comprised of more than one English word. For a Japanese verb, e.g., the English translation in EDICT almost always begins with the auxiliary *to* (e.g. *nomu* is translated as *to drink*). WordNet does not contain a verbal entry for *to drink*, but does contain an entry for *drink*. To handle this case of partial match, we locate the longest right word substring of the EDICT translation which is indexed in WordNet.

A related problem is when the translation contains domain or collocational information in parentheses. For example, *ryuu* is translated as both *dragon* and *promoted rook (shogi)*. The first translation has a matching entry in WordNet but the second translation does not. In this second case, there is no right word substring which matches in WordNet, as we end up with *rook (shogi)* and then *(shogi)*, neither of which is contained in WordNet. In order to deal with this situation, we first normalise the translation strings by removing all the brackets and query WordNet with the normalised string. Should there be a matching entry, we stop here. If not, we then remove all strings between brackets, and apply the longest right word substring heuristic as above. An illustration of this process is given in Figure 3.

In the worst case of WordNet not having a matching entry for any right word substring, we discard the translation.

At this point, we have aligned a given Japanese word with (hopefully) one or more English words, but are still no closer to inducing sense alignment pairs. In order to produce the sense alignments, we generate all pairings of Lexeed senses with WordNet synsets for each WordNet-matched word translation. For each such pair, we compile out the Lexeed definition sentence(s) word-translated into English, and the WordNet glosses, and convert each into a simple vector of term frequencies. We then measure the similarity of each vector pair using cosine similarity. An overview of this alignment process is presented in Figure 4.

4.2 Weighting terms using TF-IDF mechanism

The basic alignment method does not use any form of term weighting, and thus overemphasises common function words such as *the, which* and *and*, and downplays the impact of rare words. As we expect to have a large amount of noise in the word-
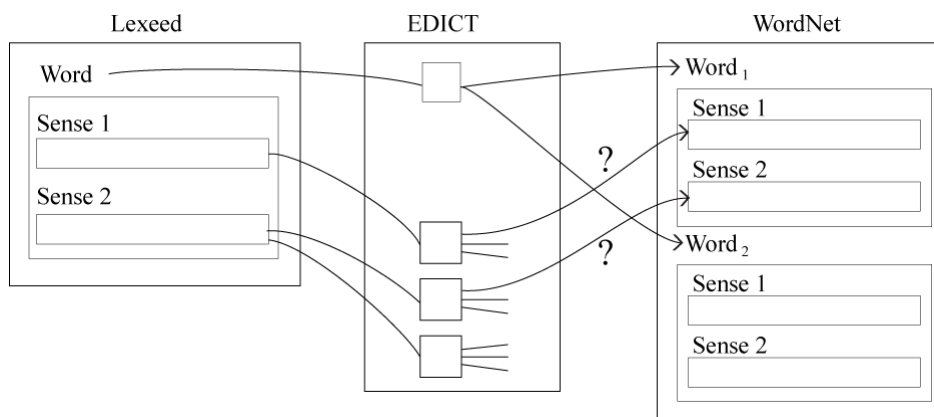
Figure 4: Overview of the Lexeed–WordNet sense alignment method

translated Lexeed definition sentences, including spurious translations for Japanese function words such as *ka, ga* and *no* that have no literal translation in English, we predict that an appropriate form of term weighting should improve the performance of our method.

As a first attempt at term weighting, we experimented with the classic SMART formulation of TF-IDF (Salton, 1971), treating the vector associated with each definition sentence as a single document.

4.3 Word stopping

As mentioned in the previous section, commonly-occurring semantically-bleached words are a source of noise in the naive cosine similarity scoring method. One conventional way of countering their impact is to filter them out of the vectors, based on a stop word list. For our experiments, we use the stop word list provided by the Snowball project.[6]

4.4 POS filtering

Another source of possible noise is the translations of Japanese function words. As all the Lexeed definition sentences are POS tagged, it is a relatively simple process to filter out all Japanese function words, focusing on prefixes, suffixes and particles.

4.5 Lemmatisation, stemming and normalisation

In its basic form, our vector space model treats distinct word as a unique term, including ignoring the

---

obvious similarity between inflectional variants of the same word, such as *dragon* and *dragons*. To remove such inflectional variation, we experiment with lemmatising all words found in both the Lexeed and WordNet vectors, using morph (Minnen et al., 2001). For similar reasons, we also experiment with the Porter stemmer, noting that stemming will further reduce the set of terms but potential introduce spurious matches.

As part of this process (with both lemmatisation and stemming), we remove all punctuation from the definition sentences.

4.6 Lexical relations

Both the Lexeed and WordNet sense inventories are described in the form of hierarchies, making it possible to complement the sense definitions with those from neighbouring senses. The intuition behind this is that the sense granularity in the two sense inventories can vary greatly, such that a single sense in Lexeed is split across multiple WordNet synsets, which we can readily uncover by considering each sense as not a single point in WordNet but a semantic neighbourhood. For example, the second sense of the word *kinou* in Figure 5, which literally means "near past", should be aligned with the second sense of *yesterday*, which is defined as "the recent past". This alignment is more self-evident, however, when we observe that the hypernym of each of the two senses is defined as "past".

In our current experiments, we only look at the utility of hypernymy. For a given sense Lexeed–

Lexeed		Word	WordNet
昨日 *(kinou)*			yesterday
Sense 1 [今日/*today, this day*][の/*indicates possessive, ...*][一日 /*one day, first of month*][前/*before, ago, ...*][の/*indicates possessive, ...*][日/*day, sun, sunshine*] Sense 2 [近い/ *near, close by, short*] [過去/*the past, bygone days, the previous*]		Definition sentences	Sense 1 *the day immediately before today* Sense 2 *the recent past*

Figure 5: The output of word-translating Japanese definition sentences to English

WordNet sense pairing, we extract the hypernyms of the respective senses and expand the definition sentences with the definition sentences from the hypernyms. The term vectors are then based on this expanded term set, similar to query expansion in information retrieval.

5 Experimental Setup

5.1 Gold-standard data

To evaluate the performance of our system, we randomly selected 100 words from Lexeed, extracted out the Lexeed–WordNet sense pairings as described above, and manually selected the gold-standard alignments from amongst them. The 100 words were associated with a total of 268 Lexeed senses and 772 WordNet senses, creating a total of 206,896 possible alignment pairs. Of these, 259 alignments were selected as our gold-standard.

We encountered a number of partial matches that were caused by the Japanese word being more specific than its English counterparts (as identified by our WordNet matching method). For example, *kakkazan* is translated as "active volcano". Since WordNet does not have any entry for *active volcano*, the longest right word substring that matches in WordNet is simply *volcano*. The definition sentences returned by Lexeed describe *kakkazan* as "a volcano which still can erupt" and "a volcano that will soon erupt", while *volcano* is described as "a fissure in the earth's crust (or in the surface of some other planet) through which molten lava and gases erupt" and "a mountain formed by volcanic material". Although there is some similarity between these definitions (namely key words such as *erupt*

and *volcano*), we do not include this pairing in our gold-standard alignment data.

5.2 Baseline

As a baseline, we take the most-frequent sense of each of the 100 random words from Lexeed, and match it with the synset with the highest SemCor frequency count out of all the candidate synsets.

5.3 Thresholding

All our calculations are based on cosine similarity, which returns a similarity between 0 and 1, with 1 being an exact match. In its simplest form, we would identify the unique WordNet sense with highest similarity to each Lexeed sense, irrespective of the magnitude of the similarity. This has the dual disadvantage of allowing only one WordNet sense for each Lexeed sense, and potentially forcing alignments to be made on low similarity values. A more reasonable approach is to apply a threshold x, and treat all WordNet senses with similarity greater than x as being aligned with the Lexeed sense. Thresholding also gives us more flexibility in terms of tuning the performance of our method: at higher threshold values, we can hope to increase precision at the expense of recall, and at lower threshold values, we can hope to increase recall at the expense of precision.

5.4 Evaluation metrics

To evaluate the performance of our system, we use precision, recall and F-score. In an alignment context, precision is defined as the proportion of correct alignments to all alignments returned by the system, and recall is defined as the proportion of the correct alignments returned by our system to all the align-

	Basic	Stopping	POS filtering	Lemmatisation	Stemming	Normalisation
Basic model	0.228	0.334	0.256	0.240	0.243	0.221
Basic+TF-IDF	0.292	0.335	0.295	0.344	0.330	0.288

Table 1: Best system F-score of combination of all features using the basic model vs. the basic model with TF-IDF weighting

ments in our gold-standard. F-score is the harmonic mean of precision and recall, and provides a single figure-of-merit rating of the balance between these two factors. We evaluate our system using unbiased ($\beta = 1$) F-score.

6 Results

Throughout our experimentation, we evaluate relative to the 100 manually sense-aligned Japanese words.

Our baseline method predicts 100 alignments (as it is guaranteed to produce a unique alignment per source-language word), of which 60 are correct. Hence, the precision is $\frac{60}{100} = 0.600$, the recall is $\frac{60}{259} = 0.231$, and the F-score is 0.334.

With the basic alignment model, the highest F-score achieved with thresholding is 0.228 at a threshold value of 0.19, well below the baseline F-score. The recall and precision value at this threshold are 0.263 and 0.202, respectively.

The basic model with TF-IDF weighting performed considerably better, scoring the highest F-score of 0.292 (recall = 0.382 and precision = 0.236) at a threshold value of 0.04, but is still well below the baseline F-score. To confirm that TF-IDF term weighting is always beneficial to overall alignment performance, we took the unweighted model combined with each of the proposed extensions, and compared it with the same extension but with the inclusion of TF-IDF (without lexical relations at this point). The result of these experiments can be found in Table 1. As we can see, TF-IDF weighting constantly improves alignment performance. Also note that, with the exception of simple (punctuation) normalisation, all extensions improve over the basic model both with and without TF-IDF weighting.

We extended our experiments by considering all possible combinations of 2 or more proposed extensions (excluding lexical relations for the time being) with TF-IDF weighting. The purpose of this ex-

Method	Precision	Recall	F-score
Baseline	0.600	0.231	0.334
WS+PF+L+N	0.298	0.494	0.372
WS+PF+L	0.326	0.428	0.370
WS+L	0.305	0.455	0.365
WS+PF+L+S+N	0.275	0.540	0.364
WS+L+S	0.301	0.455	0.363

Table 2: Top-5 combinations of extensions, excluding lexical relations (WS = Word stopping, PF = POS filtering, L = Lemmatisation, S = Stemming, N = Normalisation)

periment is to investigate whether the proposed extensions are complementary in improving alignment performance. The 5 top-performing combinations are presented in Table 2.

The best result is achieved by combining all the proposed extensions, at an F-score of 0.364, which is significantly above baseline. It is also interesting to see that not all methods are fully complementary. By excluding stemming, e.g., the system actually performs better, producing a higher F-score of 0.372.

We then experimented with the addition of lexical relations to the different combinations of extensions explored above. The 5 top-performing combinations are presented in Table 3. The best F-score of 0.408 is achieved with the combination of all the extensions proposed. When lexical relations are used exclusively or combined with less than three of the proposed extensions, the performance tends to decline.

In our best performing combination, we outperformed the baseline F-score by 22%. 349 alignments were returned for this F-score, of which 124 matched the gold-standard. The precision and recall scores are 0.355 and 0.478, respectively.

We carried out more detailed analysis of the precision–recall trade-off. While we expect the pre-

Method	Precision	Recall	F-score
Baseline	0.600	0.231	0.334
WS+PF+L+S+N+H	0.355	0.478	0.408
WS+PF+L+S+H	0.344	0.490	0.404
WS+PF+N+S+H	0.342	0.478	0.399
WS+L+S+H	0.361	0.440	0.396
WS+PF+S+H	0.317	0.525	0.396

Table 3: Top-5 performing combinations of extensions, including lexical relations (WS = stopping, PF = POS filtering, L = Lemmatisation, S = Stemming, N = Normalisation, H = Hypernym)

cision to go up to 1 as we increase our threshold, we found out that it is in fact not the case. The precision peaks at 0.625 at a threshold level of 0.265. At this level, there are 10 correct alignments out of 16 alignments returned. Upon investigating the six non-matching entries, we found that they all contain similar words but that the literal meaning of the senses are very different. Below, we present two of the six non-matching entries.

The first one relates to a sense of the Japanese word *shanpuu* "shampoo". The definition sentences for this sense found in Lexeed are directly translated as "shampoo medicine, drug, or dose; detergent or washing material that is used to wash hair or fur". The corresponding match in WordNet is "the act of washing your hair with shampoo". We can see that there are similar terms in the two vectors, such as *shampoo, washing* and *hair*, but that the literal meaning of the two senses is quite different.

The second example is very similar to the *kakkazan* example presented in Section 5. One sense of *sengetsu* ("last month") is defined as "the previous month", and is aligned to the WordNet synset of *month* (WordNet does not have an entry for *last month*). It does not help that the hypernym of *sengetsu* is *tsuki* which translates to "month", boosting the similarity of this alignment.

7 Discussion

In terms of F-score, the best-performing combination of extensions performed better than the baseline. However, the recall seems to be the dominant factor in the F-score calculations for the proposed method. This is in sharp contrast to what we have in

our baseline, where precision dominates the F-score calculation. There are several reasons for the baseline scores. First, there are 259 alignments in our gold-standard for 100 random words, corresponding to approximately 2.6 alignment per word. Given how we created our baseline, with one alignment per word, the maximum recall that the baseline can achieve is $\frac{100}{259} = 0.386$.

On the other hand, the first-sense basis of the baseline method leads to high precision, largely due to the design process for ontologies and dictionaries. Namely, there is usually good coverage of frequent word senses in ontologies and dictionaries, and additionally, the translations for a given word are generally selected to be highly biased towards common senses (i.e. even if a polysemous word is chosen as a translation, its predominant sense is almost always that which corresponds to the source language word, for obvious accessibility/usability reasons). For this reason, there is a very high probability that these frequent senses for each of the two languages align with each other.

In this paper, we proposed a cross-lingual sense-to-sense alignment method, based on similarity of definition sentences as calculated via a bilingual dictionary. We explored various extensions to a simple lexical overlap method, and achieved promising results in preliminary experiments.

In future work, we plan to exploit more lexical relations, such as synonymy and hyponymy. We also plan to experiment with weighting up alignments where both the sense pairing and the hypernym pairing match well.

Nichols et al. (2005) linked Lexeed senses to WordNet in their evaluation on ontology induction. Comparison with their method would be very interesting and is an area for future research.

References

Naoki Asanoma. 2001. Alignment of ontologies: Word-Net and Goi-Taikei. In *Proc. of the NAACL 2001 Workshop on WordNet and Other Lexical Resources: Applications, Extensions and Customizations*, pages 89–94, Pittsburgh, USA.

Timothy Baldwin, Su Nam Kim, Francis Bond, Sanae Fujita, David Martinez, and Takaaki Tanaka. to appear. MRD-based word sense disambiguation: Further$_{\#2}$ extending$_{\#1}$ Lesk. In *Proc. of the Third International Joint Conference on Natural Language Prcoessing (IJCNLP 2008)*, Hyderabad, India.

Jim Breen. 1995. Building an electronic Japanese-English dictionary. In *Japanese Studies Association of Australia Conference*.

Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, USA.

Alfio Massimiliano Gliozzo, Marcello Ranieri, and Carlo Strapparava. 2005. Crossing parallel corpora and multilingual lexical databases for WSD. In *Proc. of the 6th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2005)*, pages 242–5, Mexico City, Mexico.

Kaname Kasahara, Hiroshi Sato, Francis Bond, Takaaki Tanaka, Sanae Fujita, Tomoko Kasunagi, and Shigeaki Amano. 2004. Construction of a Japanese semantic lexicon: Lexeed. In *Proc. of SIG NLC-159, IPSJ*, Tokyo, Japan.

Kevin Knight and Steve K. Luk. 1994. Building a large-scale knowledge base for machine translation. In *Proc. of the 12th Annual Conference on Artificial Intelligence (AAAI-94)*, pages 773–8, Seattle, USA.

Philipp Koehn. 2005. Europarl: A parallel corpus for statistical machine translation. In *Proc. of the Tenth Machine Translation Summit (MT Summit X)*, Phuket, Thailand.

Shari Landes, Claudia Leacock, and Randes Tengi. 1998. Building semantic concordances. In Christiane Fellbaum, editor, *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, USA.

Guido Minnen, John Carroll, and Darren Pearce. 2001. Applied morphological processing of English. *Natural Language Engineering*, 7(3):207–23.

Grace Ngai, Marine Carpuat, and Pascale Fung. 2002. Identifying concepts across languages: A first step towards a corpus-based approach to automatic ontology alignment. In *Proc. of the 19th International Conference on Computational Linguistics (COLING 2002)*, Taipei, Taiwan.

Eric Nichols, Francis Bond, and Daniel Flickinger. 2005. Robust ontology acquisition from machine-readable dictionaries. In *Proc. of the 19th International Join Conference on Artificial Intelligence (IJCAI-2005)*, pages 1111–6, Edinburgh, UK.

Kyonghee Paik, Francis Bond, and Shirai Satoshi. 2001. Using multiple pivots to align Korean and Japanese lexical resources. In *Proc. of the NLPRS-2001 Workshop on Language Resources in Asia*, pages 63–70, Tokyo.

Gerald Salton. 1971. *The SMART Retrieval System: Experiments in Automatic Document Processing*. Prentice-Hall.

Satoshi Shirai, Kazuhide Yamamoto, and Kyonghee Paik. 2001. Overlapping constraints of two step selection to generate a transfer dictionary. In *ICSP-2001*, pages 731–736.

Sofia Stamou, Kemal Oflazer, Karel Pala, Dimitris Christoudoulakis, Dan Cristea, Dan Tufiş, Svetla Koeva, George Totkov, Dominique Dutoit, and Maria Grigoriadou. 2002. BALKANET: A multilingual semantic network for the Balkan languages. In *Proc. of the International Wordnet Conference*, pages 12–4, Mysore, India.

Takaaki Tanaka, Francis Bond, and Sanae Fujita. 2006. The hinoki sensebank — a large-scale word sense tagged corpus of Japanese —. In *Proceedings of the Workshop on Frontiers in Linguistically Annotated Corpora 2006*, pages 62–9, Sydney, Australia.

Piek Vossen, editor. 1998. *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Kluwer Academic Publishers, Dordrecht, Netherlands.