Aligning WordNet with Additional Lexical Resources

Oi Yee Kwong
Computer Laboratory, University of Cambridge
New Museums Site, Cambridge CB2 3QG, U.K.
oyk20@cl.cam.ac.uk

Abstract

This paper explores the relationship between WordNet and other conventional linguistically-based lexical resources. We introduce an algorithm for aligning word senses from different resources, and use it in our experiment to sketch the role played by WordNet, as far as sense discrimination is concerned, when put in the context of other lexical databases. The results show how and where the resources systematically differ from one another with respect to the degree of polysemy, and suggest how we can (i) overcome the inadequacy of individual resources to achieve an overall balanced degree of sense discrimination, and (ii) use a combination of semantic classification schemes to enrich lexical information for NLP.

1 Introduction

Lexical resources used in natural language processing (NLP) have evolved from handcrafted lexical entries to machine readable lexical databases and large corpora which allow statistical manipulation. The availability of electronic versions of linguistic resources was a big leap. Among these resources we find conventional dictionaries as well as thesauri.

However, it does not often suffice to depend on any single resource, either because it does not contain all required information or the information is not organised in a way suitable for the purpose. Merging different resources is therefore necessary. Calzolari's (1988) Italian lexical database, Knight and Luk's (1994) PANGLOSS ontology, and Klavans and Tzoukermann's (1995) bilingual lexicon are some responses to this need.

Many attempts have also been made to transform the implicit information in dictionary definitions to explicit knowledge bases for computational purposes (Amsler, 1981; Calzolari, 1984; Chodorow et al., 1985; Markowitz et al., 1986; Klavans et al., 1990; Vossen and Copestake, 1993). Nonetheless, dictionaries are also infamous for their non-standardised sense granularity, and the taxonomies obtained from definitions are inevitably ad hoc. It would therefore be a good idea if we could integrate such information from dictionaries with some existing, and widely exploited, classifications such as the system in Roget's Thesaurus (Roget, 1852), which has remained intact for years.

We can see at least the following ways in which an integration of lexical resources would be useful in NLP:

- Most NLP functions, notably word sense disambiguation (WSD), need to draw information from a variety of resources and cannot sufficiently rely on any single resource.
- When different systems of sense tagging are used in different studies, we need a common ground for comparison. Knowing where one sense in one resource stands in another would enable better evaluation.
- In attempting integration, we can discover how one resource differs from another and thus identify their individual limitations. This can guide improvement of the resources.

An approach to the integration problem is offered by WordNet. WordNet is designed to enable conceptual search (Miller et al., 1993), and therefore it should provide a way of linking word level senses as those in dictionaries with semantic classes as those in thesauri. However, one important question is whether WordNet, a psycholinguistically-based resource, will work the same way as conventional linguistic resources do.

We can divide this question into two parts. First, we are interested in how similar the sense discrimination is in WordNet and in a conventional dictionary. Second, WordNet has a classificatory structure, but the principle of classification is somehow different from that of a thesaurus. As a result, terms which are close in a thesaurus, thus allowing contextual sense disambiguation, may be found further apart in the WordNet taxonomy, which may therefore not be informative enough. For example, "car" and "driver" are located in two different branches in the WordNet hierarchy and the only way to relate them is through the top node "entity". This fails to uncover the conceptual closeness of the two words as Roget's Thesaurus does, for they are put in adjacent semantic classes ("Land travel" and "Traveller" respectively). Nevertheless, we believe that there must be some relation between the classes in WordNet and those in a thesaurus, which provides some means for us to make an association between them.

We have therefore proposed an algorithm to link up the three kinds of resources, namely a conventional dictionary, WordNet and a thesaurus. This is made possible with the WordNet taxonomic hierarchy as the backbone because traversing the hierarchy gives many kinds of linking possibility. The resulting integrated information structure should then serve the following functions:

- enhancing the lexical information in a dictionary with the taxonomic hierarchy in WordNet, and vice versa
- complementing the taxonomic hierarchy in WordNet with the semantic classification in a thesaurus, and vice versa

We have carried out an experiment, using the algorithm, to map senses in a dictionary to those in WordNet, and those in WordNet to the classes in a thesaurus. Our aim has been to (i) assess the plausibility of the algorithm, and to (ii) explore how the various resources differ from one another. The results suggest that mappings are in general successful (i.e. when links can be made, they are appropriate) while failures mostly arise from the inadequacy of individual resources. Based on these findings, we have also proposed some ways to overcome such inadequacies.

The algorithm is described in the next section. The test materials and the design are given in Section 3. The results are presented in Section 4. They are analysed and discussed in Section 5, where we also suggest some ways to apply them.

2 Integrating Different Resources

2.1 Relations between Resources

The three lexical resources we used are the 1987 revision of Roget's Thesaurus (ROGET) (Kirkpatrick, 1987), the Longman Dictionary of Contemporary English (LDOCE) (Procter, 1978) and the Prolog version of WordNet 1.5 (WN) (Miller et al., 1993). The linking of LDOCE and WN is in principle quite similar to Knight and Luk's (1994) approach in the PANGLOSS project. But our form of comparison between LDOCE and WN, motivated by the organisation of individual resources in relation to one another, was simpler but as effective. Figure 1 shows how word senses are organised in the three resources and the arrows indicate the direction of mapping.

In the middle of the figure is the structure of WN, a hierarchy with the nodes formed from the synsets. If we now look up ROGET for word $x2$ in synset X, since words expressing every aspect of an idea are grouped together in ROGET, we can expect to find not only words in synset X, but also those in the coordinate synsets (i.e. M and P, with words m_1, m_2, p_1, p_2, etc.) and the superordinate synsets (i.e. C and A, with words c_1, c_2, etc.) in the same ROGET paragraph. In other words, the thesaurus class to which $x2$ belongs should roughly include $X \cup M \cup P \cup C \cup A$. On the other hand, the LDOCE definition corresponding to the sense of synset X (denoted by D_x) is expected to be similar to the textual gloss of synset X (denoted by $Gl(X)$). Nevertheless, given that it is not unusual for dictionary definitions to be phrased with synonyms or superordinate terms, we would also expect to find words from X and C, or even A, in the LDOCE definition. That means we believe $D_x \approx Gl(X)$ and $D_x \cap (X \cup C \cup A) \neq \phi$. We did not include coordinate terms (called "siblings" in Knight and Luk (1994)) because we found that while nouns in WN usually have many coordinate terms, the chance of hitting them in LDOCE definitions is hardly high enough to worth the computation effort.

2.2 The Algorithm

Our algorithm defines a mapping chain from LDOCE to ROGET through WN. It is based on shallow processing within the resources themselves, exploiting their inter-relatedness, and does not rely on extensive statistical data (e.g. as suggested in Yarowsky (1992)). Given a word with part of speech, $W(p)$, the core steps are as follows:

Step 1: From LDOCE, get the sense definitions $D_1, ..., D_t$ under the entry $W(p)$.

Step 2: From WN, find all the synsets $S_n\{w_1, w_2, ...\}$ such that $W(p) \in S_n$. Also collect the corresponding gloss definitions, $Gl(S_n)$, if any, the hypernym synsets $Hyp(S_n)$, and the coordinate synsets $Co(S_n)$.

Step 3: Compute a similarity score matrix $\mathcal{A}$ for the LDOCE senses and the WN synsets. A similarity score $\mathcal{A}(i,j)$ is computed for the i^{th} LDOCE sense and the j^{th} WN synset using a weighted sum of the overlaps between the LDOCE sense and the WN synset, hypernyms, and gloss respectively, that is

$$\mathcal{A}(i,j) = a_1|D_i \cap S_j| + a_2|D_i \cap Hyp(S_j)| + a_3|D_i \cap Gl(S_j)|$$

For our tests, we just tried setting $a_1 = 3$, $a_2 = 5$ and $a_3 = 2$ to reveal the relative significance of finding a hypernym, a synonym, and any word in the textual gloss respectively in the dictionary definition.

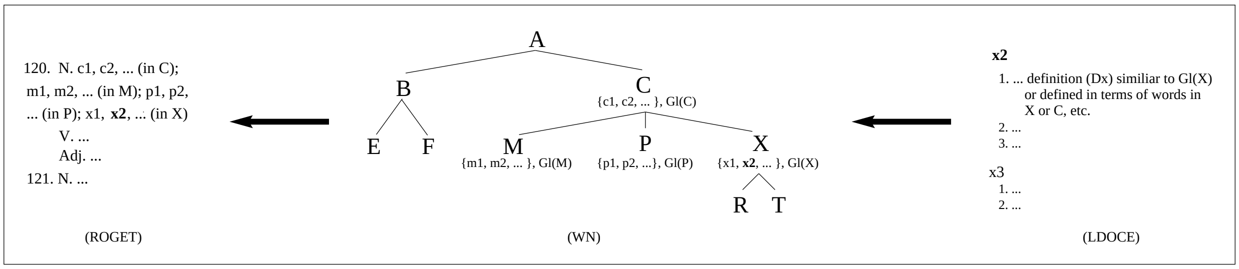

Figure 1: Organisation of word senses in different resources

Step 4: From ROGET, find all paragraphs $P_m\{w_1, w_2, ...\}$ such that $W(p) \in P_m$.

Step 5: Compute a similarity score matrix $\mathcal{B}$ for the WN synsets and the ROGET classes. A similarity score $\mathcal{B}(j,k)$ is computed for the j^{th} WN synset (taking the synset itself, the hypernyms, and the coordinate terms) and the k^{th} ROGET class, according to the following:

$$\mathcal{B}(j,k) = b_1|S_j \cap P_k| + b_2|Hyp(S_j) \cap P_k| + b_3|Co(S_j) \cap P_k|$$

We have set $b_1 = b_2 = b_3 = 1$. Since a ROGET class contains words expressing every aspect of the same idea, it should be equally likely to find synonyms, hypernyms and coordinate terms in common.

Step 6: For $i = 1$ to t (i.e. each LDOCE sense), find $max(\mathcal{A}(i,j))$ from matrix $\mathcal{A}$. Then trace from matrix $\mathcal{B}$ the j^{th} row and find $max(\mathcal{B}(j,k))$. The i^{th} LDOCE sense should finally be mapped to the ROGET class to which P_k belongs.

3 Testing

3.1 Materials

Three groups of test words (at present we only test on nouns), each containing 12 random samples, were prepared. Words have five or fewer WN senses in the "low polysemy group" (LO), between six to ten in the "medium polysemy group" (MED) , and 11 or more in the "high polysemy group" (HI). Table 1 shows the list of test words with the number of senses they have in the various lexical resources.

3.2 Design and Hypotheses

Our investigation was divided into three parts. While the application of the algorithm was in the third part, the first two parts were preliminaries to gather some information about the three resources so as to give a guide for expecting how well the mapping algorithm would work and such information would also help explain the results.

Part 1: First, we just consider whether the bulk of information, measured by a simple count of the number of senses, for each word captured by different resources is similar in size. Basically if it was not, it would mean that words are treated too differently in terms of sense distinction in different resources for the mapping to succeed. A crude way is to look for some linear relationship for the number of senses per word in different resources. If WN is as linguistically valid as other lexical databases, we would expect strong positive correlation with other resources for the "amount" of information.

Part 2: Second, we look for some quantitative characterisation of the relationship we find for the resources in Part 1. We achieve this by performing some paired-sample t-tests. If the resources do in fact capture similar "amount" of information, we would not expect to find any statistically significant difference for the mean number of senses per word among different resources.

Part 3: Third, we now apply the mapping algorithm and try to relate the results to the information regarding the resources found in the previous parts. We map LDOCE senses to WN synsets, and WN synsets to ROGET classes (i.e. we skip *step 6* in this study).

We first analyse the results by looking at mapping accuracy and failure, and secondly by characterising the relationship between the two pairs of source and target (i.e. LDOCE and WN, WN and ROGET) by means of the *Polysemy Factor* $\mathcal{P}$. This measure reflects the granularity of the senses in the source resource ($\mathcal{S}$) with respect to those in the target resource ($\mathcal{T}$), and is defined as follows:

$$\begin{array}{lcl} Let & & \\ f & : & \mathcal{S} \rightarrow \mathcal{T} \ (human\ judgement) \\ g & : & \mathcal{S} \rightarrow \mathcal{T} \ (algorithm) \\ \mathcal{T}' & = & \{t \in \mathcal{T} : t = g(s) = f(s) \ for\ some\ s \in \mathcal{S}\} \\ \mathcal{S}' & = & \{s \in \mathcal{S} : g(s) = f(s)\} \\ \mathcal{P} & = & \frac{|\mathcal{T}'|}{|\mathcal{S}'|} \end{array}$$

	Group 1: Low Polysemy Words											
	cheque	plan	party	bench	industry	society	chance	copy	music	current	letter	ceremony
W	1	3	5	5	3	3	4	4	4	3	4	3
L	2	5	9	5	4	6	6	4	5	4	5	3
R	1	5	10	6	5	6	2	9	2	3	5	4
	Group 2: Medium Polysemy Words											
	step	number	statement	strike	power	picture	sense	credit	note	title	cast	bridge
W	9	10	6	6	9	8	7	8	9	8	9	9
L	9	11	4	3	16	10	8	8	10	4	8	7
R	9	6	9	5	7	6	4	8	10	6	13	5
	Group 3: High Polysemy Words											
	life	service	field	card	hand	face	case	air	place	piece	call	block
W	13	14	15	11	13	13	16	13	15	11	12	11
L	20	16	12	8	12	8	17	13	21	16	15	8
R	8	10	11	10	13	7	14	7	5	10	5	11

Table 1: List of test words and number of senses

$\mathcal{P}$ is therefore the ratio between the minimum number of target senses covering all mapped source senses and all mapped source senses. In other words, the smaller this number, the more fine-grained are the senses in $\mathcal{S}$ with respect to $\mathcal{T}$. It tells us whether what is in common between the source and the target is similarly distinguished, and whether any excess information in the source (as found in Part 2) is attributable to extra fineness of sense distinction.

4 Results

Part 1: Table 2 shows the correlation matrix for the number of senses per word in different resources. The upper half of the matrix shows the overall correlations while the lower half shows correlations within individual groups of test words. (An asterisk denotes statistical significance at the .05 level.) Significant positive correlations are found in general, and also for the LO group between WN and the other two resources. Such abstract relations do not necessarily mean simple sameness, and the exact numerical difference is found in Part 2.

	LDOCE	WN	ROGET
LDOCE	–	0.8227*	0.4238*
WN	L: 0.7170* M: 0.6778* H: 0.4874	–	0.6133*
ROGET	L: 0.5658 M: 0.1026 H: -0.2424	L: 0.5831* M: 0.2605 H: 0.1451	–

Table 2: Correlation of number of senses per word among various resources

Part 2: The means for the number of senses per word in the three resources are shown in Table 3. LDOCE has an average of 1.33 senses more than WN for the LO Group, and this difference is statistically significant ($t = 3.75$, $p < .05$). Also, a significant difference is found in the HI group between WN and ROGET, with the latter having 3.83 senses fewer on average ($t = -4.24$, $p < .05$).

	LDOCE	WN	ROGET
LO	4.83	3.50	4.83
MED	8.17	8.17	7.33
HI	13.83	13.08	9.25

Table 3: Mean number of senses per word

Part 3: Apart from recording the accuracy, we also analysed the various types of possible mapping errors. They are summarised in Table 4. *Incorrectly Mapped* and *Unmapped-a* are both "misses", whereas *Forced Error* and *Unmapped-b* are both "false alarms". These errors are manifested either in a wrong match or no match at all. The performance of the algorithm on the three groups of nouns is shown in Table 5.

Target Exists	**Mapping Outcome**	
	Wrong Match	No Match
Yes	*Incorrectly Mapped*	*Unmapped-a*
No	*Forced Error*	*Unmapped-b*

Table 4: Different types of errors

Statistics from one-way ANOVA show that for the mapping between LDOCE and WN, there are significantly more *Forced Errors* in the HI group than both the LO and MED group ($F = 7.58$, $p < .05$).

For the mapping between WN and ROGET, the MED group has significantly more *Incorrectly Mapped* than the LO group ($F = 3.72$, $p < .05$). Also, there are significantly more *Forced Errors* in the HI group than the LO and MED group ($F = 8.15$, $p < .05$).

	LDOCE → WN			WN → ROGET		
	LO	MED	HI	LO	MED	HI
Accurately Mapped	64.77%	65.63%	52.96%	78.89%	69.63%	69.79%
Incorrectly Mapped	15.37%	11.82%	15.17%	0.00%	6.94%	3.64%
Unmapped-a	2.08%	3.13%	2.57%	2.08%	1.39%	2.55%
Unmapped-b	11.25%	11.48%	8.46%	17.36%	14.03%	7.47%
Forced Error	6.53%	7.95%	20.83%	1.67%	8.01%	16.56%
Polysemy Factor	0.79	0.88	0.91	0.80	0.66	0.59

Table 5: Average mapping results

As far as the *Polysemy Factor* is concerned, it is found to be significantly lower in the HI group than the LO group ($F = 3.63$, $p < .05$) for the mapping between WN and ROGET.

5 Discussion

Though the actual figures in Table 5 may not be significant in themselves given the small sample size in the test, they nevertheless indicate some underlying relationships among the resources, and suggest it will be worth pursuing larger scale tests.

5.1 Resource Similarities and Differences

Results from the first two parts of the investigation give us a rough idea of the similarity and difference among the resources. Comparing the overall correlation between WN and LDOCE with respect to the number of senses per word with that between WN and ROGET, the much higher positive correlation found for the former suggests that WN, though organised like a thesaurus, its content is like that in a dictionary.

While the correlation results give us an idea of how strong the linear relationship is, the t-test results suggest to us that a conventional dictionary seems to capture relatively more meanings than WN when a word has fewer than five WN senses. On the other hand, a similar relation was found between WN and ROGET for words which have more than 10 WN senses. However, this could mean two things: either that WN does contain more information than a thesaurus, or that the WN senses are getting relatively more fine-grained.

In the experiment we divided the test nouns into three groups of different degree of polysemy. However, a rough count from WordNet 1.5 reveals that out of the 88200 nouns, only 0.07% belongs to the HI group, with an average of 13.18 senses; whereas 0.55% belongs to the MED group, with an average of 7.05 senses. Up to 99.37% of the nouns come from the LO group, averaging 1.18 senses. In other words, the idiosyncrasy found for the HI group may have been magnified in the test samples, and we can expect in general a rather consistent relationship between WN and the other two resources.

5.2 Aligning WN senses with others

The third part reveals more details of the interrelationship among the resources. Knight and Luk (1994) reported a trade-off between coverage and correctness. Our results, albeit for a smaller test and with different ambiguity grouping, are comparable with theirs. Thus our *Accurately Mapped* figures correspond effectively to their *pct correct* at their confidence level ≥ 0.0. A similar average of slightly more than 60% accuracy was achieved.

Overall, the *Accurately Mapped* figures support our hypothesised structural relationship between a conventional dictionary, a thesaurus and WN, showing that we can use this method to align senses in one resource with those in another. As we expected, no statistically significant difference was found for accuracy across the three groups of words. This would mean that the algorithm gives similar success rates regardless of how many meanings a word has.

In addition, we have also analysed the unsuccessful cases into four categories as shown earlier. It can be seen that "false alarms" were more prevalent than "misses", showing that errors mostly arise from the inadequacy of individual resources because there are no targets rather than from failures of the mapping process. Moreover, the number of "misses" can possibly be reduced, for example, by a better way to identify genus terms, or if more definition patterns are considered.

Forced Error refers to cases without any satisfactory target, but somehow one or more other targets score higher than the rest. We see that this figure is significantly higher in the HI group than in the other two groups for the mappings between WN and ROGET, showing that there are relatively more senses in WN which can find no counterpart in ROGET. So WN does have something not captured by ROGET.

The polysemy factor $\mathcal{P}$ can also tell us something regarding how fine-grained the senses are in one resource with respect to the other. The significantly lower $\mathcal{P}$ in the HI group implies that as more meanings are listed for a word, these meanings can nevertheless be grouped into just a few core meanings. Unless we require very detailed distinction, a cruder discrimination would otherwise suffice.

Thus, the *Forced Error* and *Polysemy Factor* data show that both "more information" (in the sense of more coverage of the range of uses of a word) and "more granularity" contribute to the extra senses in WN in the HI group. However, no precise conclusion can be drawn because this is rather variable even within one resource.

Another observation was made regarding the mapping between WN and ROGET. Unlike the mapping between LDOCE and WN which is easy to check by comparing the definitions, synonyms and so on, judging whether a mapping from WN to ROGET is correct is not always straightforward. This is because either the expected target and the mapped target are not identical but are nevertheless close neighbours in the Roget class hierarchy, or because different targets would be expected depending on which part of the definition one's focus is on. For instance, "cast" has the sense of "the actors in a play". Strictly speaking it should be put under "assemblage", but we may be unwilling to say a mapping to "drama" is wrong. As we have said, WN and ROGET have different classificatory structures. Nevertheless, we may be able to take advantage of this difference as discussed in the next section.

5.3 Making use of the findings

Clearly successful mappings are influenced by the fineness of the sense discrimination in the resources. How finely they are distinguished can be inferred from the similarity score matrices generated from the algorithm for the two pairs of mappings. Reading the matrices row-wise shows how vaguely a certain sense is defined, whereas reading them column-wise reveals how polysemous a word is. The presence of unattached senses also implies that using only one single resource in any NLP application is likely to be insufficient.

This is illustrated for one of the test words (*note*) in Figure 2. ($\bullet$ = correctly mapped, $\circ$ = expected target, $\times$ = incorrectly mapped, and $*$ = forced error) Tables 6 and 7 show the corresponding WN synsets and LDOCE senses. It can be seen that D_1 and D_2 are both about *musical notes*, whereas D_7 and D_8 can both be construed as *letters*.

Consequently, using this kind of mapping data, we may be able to overcome the inadequacy of WN in at least two ways: (i) supplementing the missing senses to achieve an overall balanced sense discrimination, and (ii) superimposing the WN taxonomy with another semantic classification scheme such as that found in ROGET.

For the first proposal, we can, for example, conflate the mapped senses and complement them with the detached ones, thus resulting in a more complete but not redundant sense discrimination. In the above case, we can obtain the following new set of senses for *note*:

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
D_1					$\bullet$				
D_2					$\bullet$				
D_3				$\circ$					$\times$
D_4									$*$
D_5								$\bullet$	
D_6						$\bullet$			
D_7							$\bullet$		
D_8							$\bullet$		
D_9			$\bullet$						
D_{10}		$\bullet$							

Figure 2: LDOCE to WN mapping for *note*

Synset	WN Synsets
S_1	eminence, distinction, preeminence, note
S_2	note, promissory note, note of hand
S_3	bill, note, government note, bank bill, banker's bill, bank note, banknote, Federal Reserve note, greenback
S_4	note (tone of voice)
S_5	note, musical note, tone
S_6	note, annotation, notation
S_7	note, short letter, line
S_8	note (written record)
S_9	note (emotional quality)

Table 6: WordNet synsets for *note*

```
 1. promissory note     6. letter
 2. bank note           7. written record
 3. tone of voice       8. eminence
 4. musical note        9. emotional quality
 5. annotation         10. element
```

Note that 1 to 7 are the senses mapped and conflated, 8 and 9 are the unattached synsets in WN, and 10 is the unattached sense in LDOCE.

The second proposal is based on the observation that the classificatory structures in WN and ROGET may be used to complement each other because each of them may provide a better way to capture semantic information in a text at different times. As in our "cast" example, the WN taxonomy allows property inheritance and other logical inference from the information that "cast" is an assemblage, and thus is a social group; while the ROGET classification also captures the "drama" setting, so that we know it is not just any group of people, but only those involved in drama. Imagine we have another situation as follows:

> He sang a song last night. The notes were too high for a bass.

The hypernym chains for the underlined nouns in WN are as follows (assuming that we have spotted the intended senses):

Sense	LDOCE Definition
D_1	musical sound, usu. of a particular length and PITCH
D_2	a written sign for any of these sounds
D_3	a quality of voice
D_4	any quality ; ELEMENT
D_5	a record or reminder in writing
D_6	a remark added to a piece of writing and placed outside the main part of the writing (as at the side or bottom of a page, or at the end)
D_7	a short usu. informal letter
D_8	a formal letter between governments
D_9	a piece of paper money
D_{10}	PROMISSORY NOTE

Table 7: LDOCE definitions for *note*

```
song -- (a short musical composition ...)
  => musical composition, opus, ...
    => music
      => art, fine art
        => creation
          => artifact, artefact
            => object, inanimate object, ...
              => entity

note, musical note, tone
  => musical notation
    => notation, notational system
      => writing, symbolic representation
        => written communication, ...
          => communication
            => social relation
              => relation
                => abstraction

bass, basso -- (an adult male singer ...)
  => singer, vocalist
    => musician, instrumentalist, player
      => performer, performing artist
        => entertainer
          => person, individual, ...
            => life form, organism, ...
              => entity
            => causal agent, cause, ...
              => entity
```

Again, it is important that *bass* should be able to inherit the properties from *person* or *note* from *written communication*, and so on, as WN now allows us to do. But at the same time, it can be seen that the nouns can hardly be related to one another in the WN hierarchical structure except at the top node *entity*, and it is then difficult to tell what the discourse is about. However, if we align the senses with the ROGET classes, we can possibly solve the problem. Consequently, the details of how we can flexibly use the two classifications together can be a future direction of this research.

6 Conclusion

In general we cannot expect that a single resource will be sufficient for any NLP application. WordNet is no exception, but we can nevertheless enhance its utility. The study reported here began to explore the nature of WordNet in relation to other lexical resources, to find out where and how it differs from them, and to identify possible ways to absorb additional information. Apart from linking WordNet and LDOCE, as Knight and Luk (1994) did, we also experimented with ROGET to broaden the amount and type of information. The results suggest that, with an algorithm like the one described, WordNet can be fine-tuned and combined with other available resources to better meet the various information requirement of many applications.

References

R. Amsler. 1981. A taxonomy for English nouns and verbs. In *Proceedings of ACL '81*, pages 133–138.

N. Calzolari. 1984. Detecting patterns in a lexical data base. In *Proceedings of COLING-84*, pages 170–173.

N. Calzolari. 1988. The dictionary and the thesaurus can be combined. In M.W. Evens, editor, *Relational Models of the Lexicon: Representing Knowledge in Semantic Networks*. Cambridge University Press.

M.S. Chodorow, R.J. Byrd, and G.E. Heidorn. 1985. Extracting semantic hierarchies from a large on-line dictionary. In *Proceedings of ACL '85*, pages 299–304.

B. Kirkpatrick. 1987. *Roget's Thesaurus of English Words and Phrases*. Penguin Books.

J. Klavans and E. Tzoukermann. 1995. Combining corpus and machine-readable dictionary data for building bilingual lexicons. *Machine Translation*, 10:185–218.

J. Klavans, M. Chodorow, and N. Wacholder. 1990. From dictionary to knowledge base via taxonomy. In *Proceedings of the Sixth Conference of the University of Waterloo*, Canada. Centre for the New Oxford English dictionary and Text Research: Electronic Text Research.

K. Knight and S.K. Luk. 1994. Building a large-scale knowledge base for machine translation. In *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94)*, Seattle, Washington.

J. Markowitz, T. Ahlswede, and M. Evens. 1986. Semantically significant patterns in dictionary definitions. In *Proceedings of ACL '86*, pages 112–119.

G.A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller. 1993. Introduction to WordNet: An online lexical database. *Five Papers on WordNet*.

P. Procter. 1978. *Longman Dictionary of Contemporary English*. Longman Group Ltd.

P.M. Roget. 1852. *Roget's Thesaurus of English Words and Phrases*. Penguin Books.

P. Vossen and A. Copestake. 1993. Untangling definition structure into knowledge representation. In T. Briscoe, A. Copestake, and V. de Paiva, editors, *Inheritance, Defaults and the Lexicon*. Cambridge University Press.

D. Yarowsky. 1992. Word-sense disambiguation using statistical models of Roget's categories trained on large corpora. In *Proceedings of COLING-92*, pages 454–460, Nantes, France.