Integrating Conceptual Density with WordNet Domains and CALD Glosses for Noun Sense Disambiguation

Davide Buscaldi[1,2], Paolo Rosso[2], and Francesco Masulli[3]

[1] Dipartimento di Informatica e Scienze dell'Informazione (DISI),
Università di Genova, Italy
buscaldi@disi.unige.it

[2] Dpto. de Sistemas Informáticos y Computación (DSIC),
Universidad Politecnica de Valencia, Spain
prosso@dsic.upv.es

[3] INFM - Genova and Dipartimento di Informatica,
Università di Pisa, Italy
masulli@disi.unige.it

Abstract. The lack of large, semantically annotated corpora is one of the main drawbacks of Word Sense Disambiguation systems. Unsupervised systems do not need such corpora and rely on the information of the WordNet ontology. In order to improve their performance, the use of other lexical resources need to be investigated. This paper describes the effort to integrate the Conceptual Density approach with sources of lexical information different from WordNet, particularly the WordNet Domains and the Cambridge Advanced Learner's Dictionary. Unfortunately, enriching WordNet glosses with samples of another lexical resource did not provide the expected results.

1 Introduction

The lack of large, semantically annotated corpora is one of the main drawbacks of supervised Word Sense Disambiguation(*WSD*) approaches. Our unsupervised approach does not need such corpora: it relies only on the WordNet (*WN*) lexical resource, and it is based on *Conceptual Density* and the frequency of WordNet senses[7]. Conceptual Density *(CD)* is a measure of the correlation among the sense of a given word and its context. The foundation of this measure is the *Conceptual Distance*, defined as the length of the shortest path which connects two concepts in a hierarchical semantic net.

Our approach gave good results, in terms of precision, for the disambiguation of nouns over SemCor (81.55% with a context window of only two nouns, compared with the MFU-baseline of 75.55%), and in the recent all-words task in the Senseval-3 (73.40%, compared with the MFU-baseline of 69.08%) [2]. Unfortunately, although the precision achieved by our system is above that of the baseline, we still need to improve the recall, since there are nouns, whose senses

J. L. Vicedo et al. (Eds.): EsTAL 2004, LNAI 3230, pp. 183–194, 2004.

are close in meaning, that are left undisambiguated by our system. We investigated the use of other lexical resources, the *WordNet Domains*[1] [5] and the Cambridge Advanced Learner's Dictionary[2] (*CALD*) to improve our approach.

2 Combining Conceptual Density and Frequency

In our approach the noun sense disambiguation is carried out by means of the formula presented in [7]. This formula has been derived from the original Conceptual Density formula described in [1]:

$$CD(c, m) = \frac{\sum_{i=0}^{m-1} nhyp^i}{\sum_{i=0}^{h-1} nhyp^i} \tag{1}$$

where c is the synset at the top of subhierarchy, m the number of word senses falling within a subhierarchy, h the height of the subhierarchy, and $nhyp$ the averaged number of hyponyms for each node (synset) in the subhierarchy. The numerator expresses the expected area for a subhierarchy containing m marks (word senses), whereas the divisor is the actual area.

Due to the fact that the averaged number of hyponyms for each node in WN2.0 (the version we used) is greater than in WN1.4 (the version which was used originally by Agirre and Rigau), we decided to consider only the *relevant* part of the subhierarchy determined by the synset paths (from c to an ending node) of the senses of both the word to be disambiguated and its context, and not the portion of subhierarchy constituted by the synsets that do not belong to the synset paths. The base formula takes into account the M number of relevant synsets, corresponding to the *marks* m in Formula 1 ($|M|=|m|$, even if we determine the subhierarchies before adding such marks instead of vice versa like in [1]), divided by the total number nh of synsets of the subhierarchy.

$$baseCD(M, nh) = M/nh \tag{2}$$

The original formula and the above one do not take into account sense frequency. It is possible that both formulas select subhierarchies with a low frequency related sense. In some cases this would be a wrong election. This pushed us to modify the CD formula by including also the information about frequency contained in WN:

$$CD(M, nh, f) = M^{\alpha}(M/nh)^{\log f} \tag{3}$$

where M is the number of relevant synsets, α is a constant (the best results were obtained over the SemCor corpus with α near to 0.10), and f is an integer representing the frequency of the subhierarchy-related sense in WN (1 means the most frequent, 2 the second most frequent, etc.). This means that the first sense of the word (i.e., the most frequent) gets at least a density of 1 and one of the

[1] Istituto per la Ricerca Scientifica e Tecnologica, Trento, Italy.
[2] http://dictionary.cambridge.org/

less frequent senses will be chosen only if it will exceed the density of the first sense. The M^{α} factor was introduced to give more weigth to those subhierarchies with a greater number of relevant synsets, when the same density is obtained among many subhierarchies.

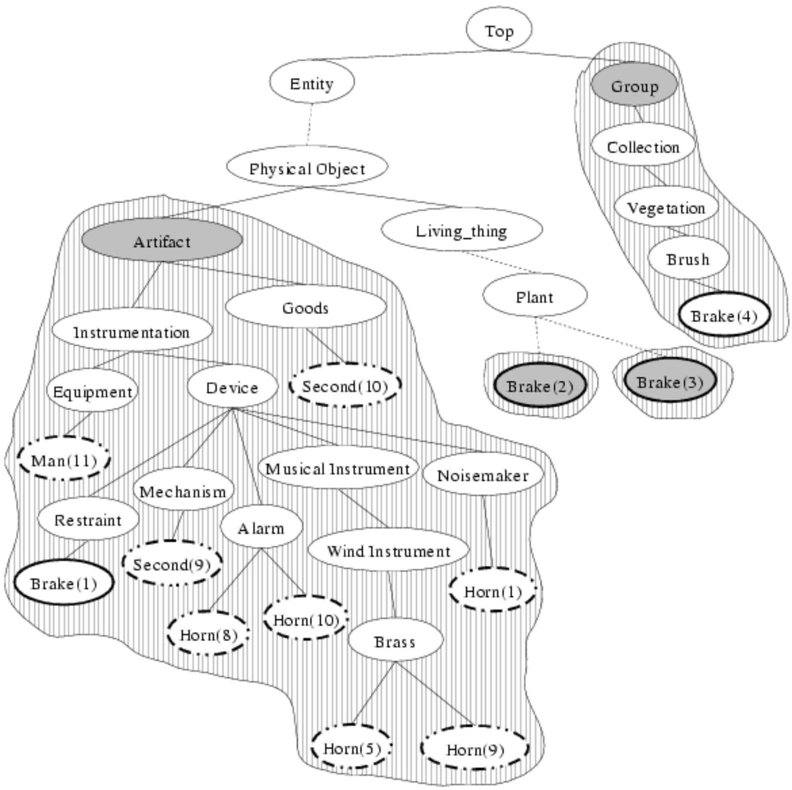

Fig. 1. Subhierarchies resulting from the disambiguation of *brake* with the context words $\{$*horn, man, second*$\}$. Example extracted from the Senseval-3 english-all-words test corpus ($M_1 = 9$, $nh_1 = 21$, $M_2 = M_3 = nh_2 = nh_3 = 1$ $M_4 = 1$, $nh_4 = 5$, where M_i and nh_i indicates, respectively, the M and nh values for the i-th sense)

In Figure 1 are shown the resulting WordNet subhierarchies from the disambiguation of *brake* with the context words $\{$*horn, man, second*$\}$ from the sentence: "Brakes howled and a horn blared furiously, but the man would have been hit if Phil hadn't called out to him a second before", extracted from the all-words test corpus of Senseval-3. The areas of subhierarchies are drawn with a dashed background, the root of subhierarchies are the darker nodes, while the nodes corresponding to the synsets of the word to disambiguate and those of the context words are drawn with a thicker border. Four subhierarchies have been identified, one for each sense of *brake*. The senses of the context words falling outside of these subhierarchies are not taken into account. The resulting

CDs are, for each subhierarchy, respectively: $9^{0.10} * (9/21)^{\log 1} = 1.27$, 1, 1 and $1^{0.10} * (1/5)^{\log 4} = 0.07$; therefore, the first one is selected and the first sense is assigned to *brake*.

Nouns are left undisambiguated when different senses are close in meaning, like the senses 2 and 3 of *brake* in the previous example, and no context senses fall into the subhierachies. In this case, the maximum density is the same (1) for more senses, therefore, we cannot assign a sense to the word. Initially, we investigated the opportunity of assigning the most frequent of those senses, but it gave no significant advantages with respect to selecting always the first sense. Consequently, we decided to integrate our approach with other resources, in order to retrieve more useful informations for the disambiguation of words.

3 Experiments with WordNet Domains

An useful information to be considered into the disambiguation process is the domain of words: e.g. it is more probable to find *bank(1)* ('*financial institution*') with the context word *money* than *bank(2)* ('*sloping land*'), because both *bank(1)* and *money* concern the domain of "economy". We observed that in the version 2.0 of WordNet three 'domain' relationships have been introduced: *category*, *usage* and *region*. Unfortunately, they are defined only for a small number of synsets: respectively 3643 (4.5% of the total number of noun synsets), 653 (0.8%) and 1146 (1.4%). Due to the fact that the WordNet Domains resource provides a wider coverage of synsets, we carried out some experiments to see if we could exploit both the new relationships and the WordNet Domains.

We performed some tests over the SemCor corpus, disambiguating all the words by assigning them the sense corresponding to the synset whose domain is matched by the majority of context words' domains (e.g. *bank* with context words *money, stock, savings, river* is assigned the first sense). We tried different size of the context window. The results are summarized in Figure 2.

The very low recall, even with large context windows, and the smaller precision, obtained by using the domains relationship in WN2.0 with respect to the WordNet Domains resource, suggested us to rely only on the latter for further experiments. Since WordNet Domains has been developed on the version 1.6 of WordNet, it has been necessary to map the synsets from the older version to the the last version. This has been done in a fully automated way, by using the WordNet mappings for nouns and verbs, and by checking the similarity of synset terms and glosses for adjectives and adverbs. Some domains have also been assigned by hand in some cases, when necessary.

Thereafter, additional weights (*Mutual Domain Weights*, *MDWs*) have been added to the densities of the subhierarchies corresponding to those senses having the same domain of context nouns' senses. Each weight is proportional to the frequency of such senses, and is calculated in the following way:

$$MDW(w_f, c_{ij}) = \begin{cases} 0 & \text{if } Dom(w_f) \neq Dom(c_{ij}) \\ 1/f * 1/j & \text{if } Dom(w_f) = Dom(c_{ij}) \end{cases} \tag{4}$$

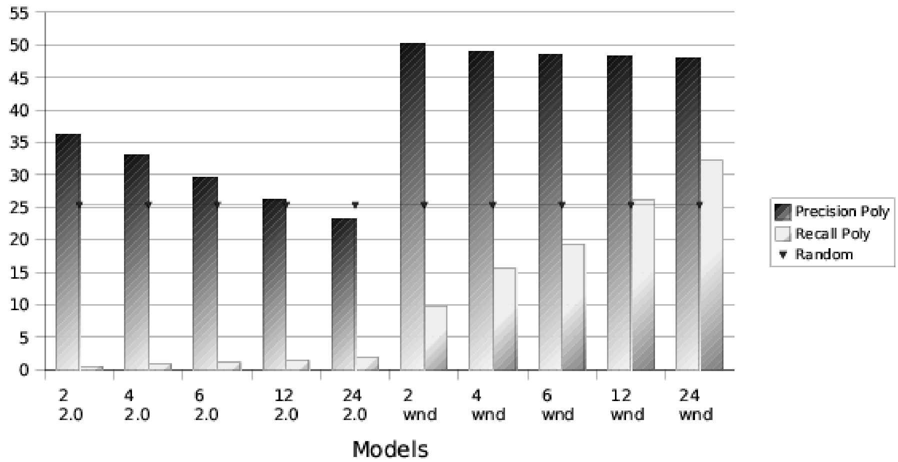

Fig. 2. Confrontation between WN2.0 domains relationships (2.0 columns) and WordNet domains (wnd columns) with different window sizes over polysemic nouns in the SemCor corpus. Precision and recall are given as percentages

where f is an integer representing the frequency of the sense of the word to be disambiguated, j gives the same information for the i-th context word, $Dom(x) : Synsets \longrightarrow Domains$ is the function returning the domain(s) corresponding to synset x, w_f and c_{ij} are, respectively, the synsets corresponding to the f-th sense of the word to be disambiguated and the j-th sense of the i-th context word.

E.g. if the word to be disambiguated (w) is *doctor*, we obtain $Dom(w_1) =$ "$Medicine$" and $Dom(w_4) =$ "*School*". Therefore, if the context word (c_i) is *university*, for which $Dom(c_{i3}) =$ "*School*", the resulting weight for *doctor(4)* and *university(3)* is $1/4 * 1/3$. Therefore, after the inclusion of MDWs, the formula (3) becomes as follows:

$$CD(M, nh, w, f, C) = M^{\alpha}(M/nh)^{\log f} + \sum_{i=0}^{|C|} \sum_{j=1}^{k} MDW(w_f, c_{ij}) \qquad (5)$$

where C is the vector of context words, k is the number of senses of the context word c_i, and c_{ij} is the synset corresponding to the j-th sense of the context word c_i.

With the introduction of MDWs, however, we did not obtain the desired improvements (70.79% in precision and 67.91% for recall, below the MFU baseline of 75.5% for both measures). The reason is that many of the correspondances in domains are found for the domain *Factotum*, that is too generic and, consequently, it does not provide any useful information about the correlation of two word senses. Our solution was to reduce by a 10 factor the relevance of the *Factotum* domain, with the formula(4) modified as follows:

$$MDW(w_f, c_{ij}) = \begin{cases} 0 & \text{if } Dom(w_f) \neq Dom(c_{ij}) \\ 1/f * 1/j & \text{if } Dom(w_f) = Dom(c_{ij}) \\ & \wedge \ Dom(w_f) \neq \text{“}Factotum\text{”} \\ 10^{-1} * (1/f * 1/j) & \text{if } Dom(w_f) = Dom(c_{ij}) \\ & \wedge \ Dom(w_f) = \text{“}Factotum\text{”} \end{cases} \quad (6)$$

In this way we could obtain a precision of 78.33% and a recall of 62.60% over the whole SemCor, with a context window of 4 nouns. We also tried not to take into account the *Factotum* domain. In this case we got an higher precision (80.70%), but the recall was only 59.08%. This means that whereas the *Factotum* domain does not provide useful information to the disambiguation task, it can help in disambiguating a certain number of nouns with the most frequent sense, thanks to the weights assigned proportionally to the frequency of senses.

We used nearly the same method to disambiguate words of POS categories other than nouns. In these cases we could not take into account the Conceptual Density. For the following reasons: first of all, in WordNet there is not a hierarchy for adjectives and adverbs. With regard to verbs, the hierarchy is too shallow to be used efficiently. Moreover, since the disambiguation is performed one sentence at a time, in most cases only one verb for each sentence can be found (with the consequence that no density can be computed).

The sense disambiguation of an adjective is performed only on the basis of the domain weights and the context, constituted by the *Closest Noun (CN)*, i.e., the noun the adjective is referring to (e.g. in *“family of musical instruments”* the CN of *musical* is *instruments*). Given one of its senses, we extract the synsets obtained by the *antonymy*, *similar_to*, *pertainymy* and *attribute* relationships. For each of them, we calculate the MDW with respect to the senses of the context noun. The weight assigned to the adjective sense is the average between these MDWs. The selected sense is the one having the maximum average weight.

In order to achieve the maximum coverage, the *Factotum* domain has been also taken into account to calculate the MDWs between adjective senses and context noun senses. However, due to the fact that in many cases this domain does not provide a useful information, the weights resulting from a *Factotum* domain are reduced by a 0.1 factor. E.g. suppose to disambiguate the adjective *academic* referring to the noun *credit*. Both *academic(1)* and *credit(6)* belong to the domain *School*. Furthermore, the *Factotum* domain contains the senses 1 4 and 7 of *credit*, and senses 2 and 3 of *academic*. The extra synsets obtained by means of the WN relationships are: *academia(1):Sociology*, pertainym of sense 1; *theoretical(3):Factotum* and *applied(2):Factotum*, similar and antonym of sense 2; *scholarly(1):Factotum* and *unscholarly(1):Factotum*, similar and antonym of sense 3. Since there are no senses of *credit* in the *Sociology* domain, *academia(1)* is not taken into account. Therefore, the resulting weights for *academic* are:
$1 * 1/6 = 0.16$ for sense 1;
$0.1 * (1/2 + 1/2 * 1/4 + 1/2 * 1/7 + [1/2 * 1/3 + 1/2 * 1/2])/5 \simeq 0.02$ for sense 2;
$0.1 * (1/3 + 1/3 * 1/4 + 1/3 * 1/7 + [1/3 * 1 + 1/3 * 1])/5 \simeq 0.02$ for sense 3.

The weights resulting from the extra synsets are represented within square brackets. Since the maximum weight is obtained for the first sense, this is the sense assigned to *academic*.

We tried to use the same idea to improve the results for the noun sense disambiguation task: if in the sentence *"family of musical instruments" musical* has a closer link to *instruments* than to *family*, it is also true that *instruments* has a close link to *musical*. This kind of link is also easy to be found, since in English adjectives always come before the noun they are referring to. Therefore, we investigated the utility of choosing only the *Closest Adjective* (CA) in the context for calculating MDWs for nouns, in a similar way to what we did for adjectives. Our experiments show that the precision and the recall values differ slightly from the base results obtained without domains.

Table 1. Comparison among approaches to noun sense disambiguation making use of WordNet Domains. The results are obtained over the whole SemCor with a context window size of 4 words (in all cases but the CA and the MFU ones)

	MFU	no WND	WND	WND (CA)
Precision	75.55%	80.70%	78.33%	80.45%
Recall	75.55%	59.07%	62.60%	59.42%
Coverage	100%	73.20%	79.91%	73.86%

The same experiments carried out over the Senseval-3 corpus showed a more significative difference between the CA technique and the other ones. Moreover, in this case the precision is even higher than the one obtained without taking into account the WordNet domains.

Table 2. Comparison among approaches to noun sense disambiguation making use of WordNet Domains. The results are obtained over the Senseval-3 All-Words Task corpus with a context window size of 4 words (in all cases but the CA and the MFU ones)

	MFU	no WND	WND	WND (CA)
Precision	69.08%	73.40%	65.39%	74.30%
Recall	69.08%	51.81%	58.28%	52.69%
Coverage	100%	70.58%	89.13%	70.91%

We tried to limit the search of the closest adjective for the noun only to the immediately preceding word or to the two preceding words, but results differ only of a $0.1 - 0.2\%$ (Tables 3 and 4) from those obtained without doing such distinction.

The results are very similar, since the approaches differ for a few hundreds nouns, as it can be observed from the coverage values (the corpus is made up of more than 70000 nouns).

Table 3. Comparison among approaches to noun sense disambiguation searching backwards in context for the closest adjective without restrictions, only within 2 words before the noun, and just the word before. The results are obtained over the whole SemCor

	unrestricted	2 words	1 word
Precision	80.45%	80.55%	80.57%
Recall	59.42%	59.32%	59.27%
Coverage	73.86%	73.64%	73.56%

Table 4. Comparison among approaches to noun sense disambiguation searching backwards in context for the closest adjective without restrictions, only within 2 words before the noun, and only the word before. The results are obtained over the whole Senseval-3 All-Words Task corpus

	unrestricted	2 words	1 word
Precision	74.30%	74.41%	74.49%
Recall	52.69%	52.57%	52.68%
Coverage	70.91%	70.58%	70.80%

The results obtained over the Senseval-3 All-Words Task corpus confirm that the precision of the CA approach can be slightly improved by considering adjectives with a stronger tie (i.e., closer) to the noun to disambiguate.

Another approach we evaluated was to add weights to subhierarchies only when nouns were left undisambiguated by the use of the "clean" CD formula (3). In other words, to include the sum in formula (5) only when CD and frequency are not able to disambiguate the noun. Unfortunately, the results are approximately the same of those in Tables 3 and 4, and, therefore, they are not worth mentioning.

The sense disambiguation of a verb is done nearly in the same way than for adjectives, but taking into consideration only the MDWs with the verb's senses and the context words (i.e., in the example above, if we had to disambiguate a verb instead of an adjective, the weights within the square brackets would not have been considered). In the Senseval-3 all-words and gloss disambiguation tasks the two context words were the noun before and after the verb, whereas in the lexical sample task the context words were four (two before and two after the verb), without regard to their POS category. This has been done in order to improve the recall in the latter task, whose test corpus is made up mostly by verbs, since our experiments carried out over the SemCor corpus showed that considering only the noun preceding and following the verb allows for achieving a better precision, while the recall is higher when the 4-word context is used. However, our results over verbs are still far from the most-frequent baseline. The sense disambiguation of adverbs (in every task) is carried out in the same way of the disambiguation of verbs for the lexical sample task.

4 Experiments with Glosses

Glosses have been used in the past as a resource for Word Sense Disambiguation by Lesk[4] and many other researchers. Usually WordNet glosses are composed of two parts:

- *definition* part;
- *sample* part.

E.g. the gloss of *light(7)* is: *used of vowels or syllables; pronounced with little or no stress; "a syllable that ends in a short vowel is a light syllable"; "a weak stress on the second syllable"*; the definition part in this case is *used of vowels or syllables; pronounced with littel or no stress*, while the sample part is *"a syllable that ends in a short vowel is a light syllable"; "a weak stress on the second syllable.*

We carried out some experiments over the WordNet glosses in order to understand which of these portions is more important to the task of Word Sense Disambiguation. Initially, we defined *Gloss Weights* (*GWs*) similarly to MDWs. Each GW is calculated as follows:

$$GW(w_f, c_i) = \begin{cases} 0 & \text{if } c_i \notin Gl(w_f) \\ 0.3 & \text{if } c_i \in Gl(w_f) \end{cases} \quad (7)$$

where c_i is the i-th word of the context, w_f is the f-th sense of the word to be disambiguated, and $Gl(x)$ is the function returning the set of words being in the gloss of the synset x without stopwords. E.g. $Gl(light_7)$=\{*used, vowels, syllables, pronounced, little, stress, syllable, ends, short, vowel, light, weak, stress, second*\}.

The GWs are added to the formula(3) as an alternative to MDWs:

$$CD(M, nh, w, f, C) = M^{\alpha}(M/nh)^{\log f} + \sum_{i=0}^{|C|} GW(w_f, c_i) \quad (8)$$

where C is the vector of context words, and c_i is the i-th word of the context.

We initially used a weight of 0.5, considering as a good matching the fact that two context words were found in a gloss, but subsequently we obtained better results with a weight of 0.3 (i.e., at least three context words are needed to be found in a gloss to obtain a density close to 1). Two other Gloss Weights were defined, each making use of a different Gl function: GW_d, which, by using $Gl_d(x)$, returns the set of words in the definition part of the gloss of the synset x; and GW_s, which uses $Gl_s(x)$ to return the set of words in the sample part.

The obtained results show that disambiguation carried out by considering only the sample portion of WordNet glosses is more precise than working on the whole gloss and/or the definitions. This has been observed also in the experiments conducted over the Senseval-3 All-Words Task corpus (Table 5). Therefore, we looked for another machine-readable resource in order to expand the sample portions of WordNet glosses. We decided to use the Cambridge Advanced Learner's Dictionary (CALD), since it is one of the few available on-line, and its

Table 5. Results obtained over the whole SemCor and Senseval-3 All-Words Task corpora, with a window size of 4 nouns, by using whole glosses or their separate parts, and Gloss Weights of 0.3

SemCor	precision	recall
Whole gloss	79.31	60.22
Definition only (GW_d)	79.85	59.52
Samples only (GW_s)	80.12	59.96
Senseval-3 AWT		
Whole gloss	73.75	52.14
Definition only (GW_d)	73.98	51.81
Samples only (GW_s)	74.06	52.03

HTML pages are organized in a format that allows to easily retrieve information about the POS of a word, its definition, and the sample parts of glosses.

The heuristics we use to retrieve a CALD gloss corresponding to a WordNet's one compares its synset terms and the definition parts of the gloss. For each synset in WordNet, we search in CALD the synset's words, and then select the resulting entries as candidate glosses. If more than the 40% of the definition part of the WordNet gloss is found in one of the CALD definition parts of candidate glosses, then the corresponding sample part is added to the WordNet gloss.

E.g. for the WordNet synset *coherence, coherency, cohesion, cohesiveness, (the state of cohering or sticking together)*, we search in the CALD web page for: *coherence*, *coherency*, *cohesion*, *cohesiveness*, obtaining respectively 1, 0 and 1 entries (one is missing since the CALD returns the same entry for *cohesion* and *cohesiveness*). A matching greater than 40% is found only within the CALD definition of *coherence*: "the quality of cohering or being coherent". Since this definition shares 4 words (over 7) with the WN gloss (the, of, cohering, or), the resulting matching is $4/7 = 54\%$, with the result that the following sample sentence is added to the WordNet gloss: "There was no coherence between the first and the second half of the film". A drawback of this heuristics is that stopwords are taken into account in the calculation of similarity between glosses. This may be useful when they keep the same order and position in both definitions, like in the previous example, but in most cases they are actually adding noise into the disambiguation process. Therefore, we will need to determine a better way to check the matching of gloss definitions.

In WordNet 2.0 there are 8195 noun glosses with samples. With our heuristics, we found 7416 totally new sample sentences, raising the total number of sample sentences from 8195 to 15611. Moreover, new sample sentences were added to 2483 already existing samples. We used these "expanded" glosses to carry out some Word Sense Disambiguation tests over the SemCor and the Senseval-3 All-Words Task corpora. The results are shown in Table 6.

The experiments have been carried out initially looking for all the context words in the expanded gloss of the word to be disambiguated. Thereafter, due to the poor results obtained, we used only those samples which contained its context nouns. Finally, we tried to select the gloss in a more precise way by

Table 6. Results obtained using the glosses expanded with CALD samples over the whole SemCor and Senseval-3 AWT corpora. The window size was of 4 words

SemCor	precision	recall
All context words	73.84	67.07
Only context nouns	79.78	59.76
Context nouns and word to be disambiguated	78.97	59.58
Senseval-3 AWT		
Only context nouns	64.72	48.73

using only those containing the word to be disambiguated itself together with the context nouns. With respect to the Senseval-3 All-Words Task, we performed the test only with the context nouns, which gave the best results over the SemCor. Whereas for the SemCor corpus we obtained results comparable with those in Table5, for the Senseval-3 AWT we observed a precision decrease of about 10%.

The reason of these poor results is that many of the added glosses were "off-topic", like: *"The Chinese anthem was played after the Union Jack was lowered in Hong Kong for the last time."*, that has been added for the synset *jack,(an electrical device consisting of a connector socket designed for the insertion of a plug)*. Therefore, we tried to improve the quality of the added samples by setting an higher threshold (70%) for the matching of the WordNet and CALD definitions, and by selecting only those glosses sharing at least a (non stopword) lemma with the defintion part of the gloss to be expanded. In this case, only 264 sample sentences were added, a number not relevant with respect to the total number of gloss samples in WordNet (8195 samples over 79688 glosses). Even if a more precise study over the threshold parameter could be done, we suspect that it could be really difficult to select gloss samples from different lexical resources, since in each resource can be different the way definitions are given. Therefore, the heuristics we used, inspired by the one used for the mapping of different versions of WordNet, cannot be applied between different lexical resources.

5 Conclusions and Further Work

The experiments with the WordNet Domains show that using this resource allows for improving recall without losing too much in precision, although the conditions when this can be done are very few. This is mostly due to the small number of correspondances that can be found for domains different than "Factotum". We observed that a better precision can be obtained for the gloss approach if we consider only the sample part of WordNet glosses. Therefore, we tried to add further gloss samples from the Cambridge Advanced Learner's online Dictionary. However, due to the poor results obtained, we decided not to integrate in our approach also the CALD glosses, until we will not be able to add gloss samples in an appropriate way. Maybe it could be worthwhile to investigate the possibility of selecting the gloss samples by disambiguating glosses and find a matching between concepts. The use of other resources such as the Roget's Thesaurus or the Oxford Advanced Learner's Dictionary will be also investigated. At the moment, we are also investigating the possibility to use the web as a knowledge source for WSD [3], by using an approach inspired by [6].

Acknowledgments

This work was supported by the CIAO SENSO MCYT Spain-Italy (HI 2002-0140), the R2D2 CICYT (TIC2003-07158-C04-03) and the ICT for EU-India Cross-Cultural Dissemination (ALA/95/23/2003/077-054) projects.

References

1. Agirre, E. and Rigau, G.: A proposal for Word Sense Disambiguation using Conceptual Distance. In: Proceedings of the International Conference on Recent Advances in NLP. RANLP-2000, Tzhigov Chark, Bulgaria, (2000).
2. Buscaldi, D., Rosso,P., Masulli,F.: The upv-unige-CIAOSENSO WSD System. Senseval-3 Workshop, ACL - Association of Computational Linguistics. Barcelona, Spain (2004) (in press).
3. Gonzalo, J., Verdejo, F., Chugar, I.: The Web as a Resource for WSD. In: 1st MEANING workshop, Spain (2003).
4. Lesk, M.: Automatic Sense Disambiguation using Machine Readable Dictionaries: How to tell a Pine Cone from an Ice Cream Cone. In: Proceedings of SIGDOC'86, U.S.A. (1986)
5. Magnini, B. and Cavaglià, G.: Integrating Subject Field Codes into WordNet. In: Proceedings of LREC-2000, Second International Conference on Language Resources and Evaluation. (2000) 1413–1418.
6. Mihalcea, R., Moldovan, D.I.: A Method for Word Sense Disambiguation of Unrestricted Text. In: Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL-99). Maryland, NY, U.S.A. (1999)
7. Rosso, P., Masulli, F., Buscaldi, D., Pla, F., Molina, A.: Automatic Noun Disambiguation. Lecture Notes in Computer Science, Vol. 2588. Springer-Verlag (2003) 273–276