AN ASSESSMENT OF SEMANTIC INFORMATION AUTOMATICALLY EXTRACTED FROM MACHINE READABLE DICTIONARIES

Jean Véronis[1,2] and Nancy Ide[1]

[1]Department of Computer Science
VASSAR COLLEGE
Poughkeepsie, New York 12601 (U.S.A.)

[2]Groupe Représentation et Traitement des Connaissances
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
31, Ch. Joseph Aiguier
13402 Marseille Cedex 09 (France)

ABSTRACT

In this paper we provide a quantitative evaluation of information automatically extracted from machine readable dictionaries. Our results show that for any one dictionary, 55-70% of the extracted information is garbled in some way. However, we show that these results can be dramatically reduced to about 6% by combining the information extracted from five dictionaries. It therefore appears that even if individual dictionaries are an unreliable source of semantic information, multiple dictionaries can play an important role in building large lexical-semantic databases.

1. INTRODUCTION

In recent years, it has become increasingly clear that the limited size of existing computational lexicons and the poverty of the semantic information they contain represents one of the primary bottlenecks in the development of realistic natural language processing (NLP) systems. The need for extensive lexical and semantic databases is evident in the recent initiation of a number of projects to construct massive generic lexicons for NLP (project GENELEX in Europe or EDR in Japan).

The manual construction of large lexical-semantic databases demands enormous human resources, and there is a growing body of research into the possibility of automatically extracting at least a part of the required lexical and semantic information from everyday dictionaries. Everyday dictionaries are obviously not structured in a way that enables their immediate use in NLP systems, but several studies have shown that relatively simple procedures can be used to extract taxonomies and various other semantic relations (for example, Amsler, 1980; Calzolari, 1984; Chodorow, Byrd, and Heidorn, 1985; Markowitz, Ahlswede, and Evens, 1986; Byrd *et al.*, 1987; Nakamura and Nagao, 1988; Véronis and Ide, 1990; Klavans, Chodorow, and Wacholder, 1990; Wilks *et al.*, 1990).

However, it remains to be seen whether information automatically extracted from dictionaries is sufficiently complete and coherent to be actually usable in NLP systems. Although there is concern over the quality of automatically extracted lexical information, very few empirical studies have attempted to assess it systematically, and those that have done so have been restricted to consideration of the quality of grammatical information (e.g., Akkerman, Masereeuw, and Meijs, 1985). No evaluation of automatically extracted semantic information has been published.

The authors would like to thank Lisa Lassek and Anne Gilman for their contribution to this work.

In this paper, we report the results of a quantitative evaluation of automatically extracted semantic data. Our results show that for any one dictionary, 55-70% of the extracted information is garbled in some way. These results at first call into doubt the validity of automatic extraction from dictionaries. However, in section 4 we show that these results can be dramatically reduced to about 6% by several means--most significantly, by combining the information extracted from five dictionaries. It therefore appears that even if individual dictionaries are an unreliable source of semantic information, multiple dictionaries can play an important role in building large lexical-semantic databases.

2. METHODOLOGY

Our strategy involves automatically extracting hypernyms from five English dictionaries for a limited corpus. To determine where problems exist, the resulting hierarchies for each dictionary are compared to an "ideal" hierarchy constructed by hand. The five dictionaries compared were: the *Collins English Dictionary (CED)*, the *Oxford Advanced Learner's Dictionary (OALD)*, the *COBUILD Dictionary*, the *Longman's Dictionary of Contemporary English (LDOCE)* and the *Webster's 9th Dictionary (W9)*.

We begin with the most straightforward case in order to determine an upper bound for the results. We deal with words within a domain which poses few modelling problems, and we focus on hyperonymy, which is probably the least arguable semantic relation and has been shown to be the easiest to extract. If the results are poor under such favorable constraints, we can foresee that they will be poorer for more complex (abstract) domains and less clearly cut relations.

An ideal hierarchy probably does not exist for the entire dictionary; however, a fair degree of consensus seems possible for carefully chosen terms within a very restricted domain. We have therefore selected a corpus of one hundred kitchen utensil terms, each representing a concrete, individual object--for example, *cup, fork, saucepan, decanter*, etc. All of the terms are count nouns. Mass nouns, which can cause problems, have been excluded (for example, the mass noun *cutlery* is not a hypernym of *knife*). Other idiosyncratic cases, such as *chopsticks* (where it is not clear if the utensil is one object or a pair of objects) have also been eliminated from the corpus. This makes it easy to apply simple tests for hyperonymy, which, for instance, enable us to say that Y is a hypernym of X if "this is an X" entails but is not entailed by "this is a Y" (Lyons, 1963).

Chodorow, Byrd, and Heidorn (1985) proposed a heuristic for extracting hypernyms which exploits the fact that definitions for nouns typically give a hypernym term as the head of the defining noun phrase. Consider the following examples:

dipper	a *ladle* used for dipping... [CED]
ladle	a long-handled *spoon*... [CED]
spoon	a metal, wooden, or plastic *utensil*... [CED]

In very general terms, the heuristic consists of extracting the word which precedes the first preposition, relative pronoun, or participle encountered in the definition text. When this word is "empty" (e.g. *one, any, kind, class*) the true hypernym is the head of the noun phrase following the preposition *of*:

slice	any of various *utensils*... [CED]

Automatically extracted hierarchies are necessarily *tangled* (Amsler, 1980) because many words are polysemous. For example, in the *CED*, the word ***pan*** has the following senses (among others):

pan[1]	1.a a wide metal *vessel*... [CED]
pan[2]	1 the *leaf* of the betel tree... [CED]

The *CED* also gives *pan* as the hypernym for *saucepan*, which taken together yields the hierarchy in figure 1.a. The tangled hierarchy is problematic because, following the path upwards from *saucepan*, we find that *saucepan* can be a kind of *leaf*. This is clearly erroneous. A hierarchy utilizing senses rather than words would not be tangled, as shown in figure 1.b.

In our study, the hierarchy was disambiguated by hand. Sense disambiguation in dictionary definitions is a difficult problem, and we will not address it here; this problem is the focus of much current research and is considered in depth elsewhere (e.g., Byrd *et al.*, 1987; Byrd, 1989; Véronis and Ide, 1990; Klavans, Chodorow, and Wacholder, 1990; Wilks *et al.*, 1990).

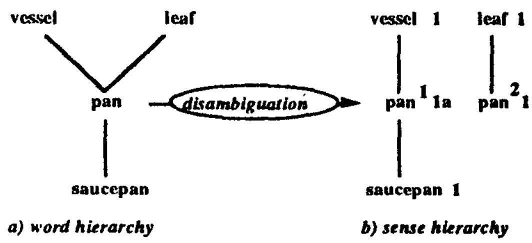

Figure 1 : Sense-tangled hierarchy

3. EVALUATION

Hierarchies constructed with methods such as those outlined in section 2 show, upon close inspection, several serious problems. In this section, we describe the most pervasive problems and give their frequency in our five dictionaries. The problems fall into two general types: those which arise because information in the dictionary is incomplete, and those which are the result of a lack of distinction among terms and the lack of a one-to-one mapping between terms and concepts, especially at the highest levels of the hierarchy.

3.1. Incomplete information

The information in dictionaries is incomplete for two main reasons. First, since a dictionary is typically the product of several lexicographers' efforts and is constructed, revised, and updated over many years, there exist inconsistencies in the criteria by which the hypernyms given in definition texts are chosen. In addition, space and readability restrictions, on the one hand, and syntactic restrictions on phrasing, on the other, may dictate that certain information is unspecified in definition texts or left to be implied by other parts of the definition.

3.1.1. Attachment too high : 21-34%

The most pervasive problem in automatically extracted hierarchies is the attachment of terms too high in the hierarchy. It occurs in 21-34% of the definitions in our sample from the five dictionaries (figure 8). For example, while *pan* and *bottle* are *vessels* in the *CED*, *cup* and *bowl* are simply *containers*, the hypernym of *vessel*. Obviously, "this is a cup" and "this is a bowl" both entail (and are not entailed by) "this is a vessel". Further, other dictionaries give *vessel* as the hypernym for *cup* and *bowl*. Therefore, the attachment of *cup* and *bowl* to the higher-level term *container* seems to be an inconsistency within the *CED*.

The problem of attachment too high in the hierarchy occurs relatively randomly within a given dictionary. In dictionaries with a controlled definition vocabulary (such as the *LDOCE*), the problem of attachment at high levels of the hierarchy results also from a lack of terms from which to choose. For example, *ladle* and *dipper* are both attached to *spoon* in the *LDOCE*, although "this is a dipper" entails and is not entailed by "this is a ladle". There is no way that *dipper* could be defined as a *ladle* (as, for instance, in the *CED*), since *ladle* is not in the defining vocabulary. As a result, hierarchies extracted from the *LDOCE* are consistently flat (figure 7).

3.1.2. Absent hypernyms : 0-3%

In some cases, strategies like that of Chodorow, Byrd and Heidorn yield *incorrect* hypernyms, as in the following definitions:

grill	A grill is a *part* of a cooker... [COBUILD]
corkscrew	a pointed spiral *piece* of metal... [W9]
dinner service	a complete *set* of plates and dishes... [LDOCE, not included in our corpus]

The words *part, piece, set*, are clearly not hypernyms of the defined concepts: it is virtually meaningless to say that *grill* is a kind of *part*, or that *corkscrew* is a kind of *piece*. In these cases, the head of the noun phrase serves to mark another relation: part-whole, member-class, etc. It is easy to reject these and similar words (*member, series*, etc.) as hypernyms, since they form a closed list (Klavans, Chodorow, and Wacholder, 1990). However, excluding these words leaves us with no hypernym. We call these "absent hypernyms"; they occur in 0-3% of the definitions in our sample corpus (figure 8).

The absence of a hypernym in a given definition text does not necessarily imply that no hypernym exists. For example, "this is a corkscrew" clearly entails (and is not entailed by) "this is a device" (the hypernym given by the *COBUILD* and the *CED*). In many cases, the lack of a hypernym seems to be the result of concern over space and/or readability. We can imagine, for example, that the definition for *corkscrew* could be more fully specified as "a *device* consisting of a pointed spiral piece of metal..." In such cases, lexicographers rely on the reader's ability to deduce that something made of metal, with a handle, used for pulling corks, can be called a device. However, for some terms, such as *cutlery* or *dinner service*, it is not clear that a hypernym exists. Note that we have voluntarily excluded problematic terms of this kind from our corpus, in order to restrict our evaluation to the best case.

3.1.3. Missing overlaps : 8-14%

Another problem results from the necessary choices that lexicographers must make in an attempt to specify a single superordinate, when concepts in the real world overlap freely. For instance, a *saucepan* can be said to be a *pot* as well as a *pan* . "This is a saucepan" entails both "this is a pot" (the hypernym given by the *CED* and *W9*) as well as "this is a pan" (the hypernym given by the *LDOCE, OALD, and COBUILD*). On the other hand, "this is a pot" does not entail and is not entailed by "this is a pan", which is to say that *pot* and *pan* are not synonyms, nor is one the hypernym of the other. In terms of classes, *pan* and *pot* are distinct but overlapping, and *saucepan* is a subset of their intersection (figure 2.a). This is no longer a strict hierarchy since it includes merging branches (figure 2.b). We will call it an "overlapping hierarchy". Although a tree representation of such a hierarchy is impossible, it presents no problems on either logical or computational grounds.

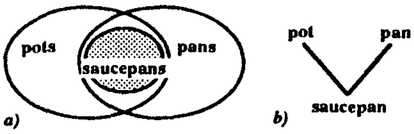

Figure 2. Overlapping hierarchy

Assuming the above relations, it would be more logically correct to phrase the definition of *saucepan* as "a *pan* AND a *pot*...". However, lexicographers never use "and" in this way, but usually give only one of the alternatives. For example, each of the five dictionaries in our study chooses either *pot* or *pan* as the genus term for *saucepan*. When this occurs, one of the hypernyms is missing. This problem arises in our sample corpus relatively frequently, 8-14% of the time depending on the dictionary (figure 8).

3.2. Difficulties at higher levels

At the higher levels of the hierarchy, terms necessarily become more general, and they often become less clearly defined. For example, most people will agree on whether some object falls into the category *fork* or *spoon*, but there is much less agreement on what objects are *implements* or *utensils*. In addition, at the higher levels some concepts simply lack a term to designate them exactly. As a result, there is confusion at the higher levels of hierarchies implicit in dictionary definitions.

3.2.1. OR-conjoined heads : 7-10%

For 7-10% of the terms in our corpus, definitions give a list of head nouns separated by the conjunction *or*, as in the following:

utensil an *implement, tool* or *container*... [CED]

In this case, none of the three alternatives is a hypernym of *utensil*. First, it is clearly not true that "this is a utensil" entails "this is a container". For the other two, it is not clear whether or not "this is a utensil" entails "this is a tool" and "this is an implement", and it is even less clear that the reverse entailments do not apply. Regarding the three terms as hypernyms of *utensil* would produce the hierarchy in figure 3. However, by enumerating the paths upwards from *spatula* (defined as a *utensil*), we see that *spatula* is a kind of *container*, which is obviously incorrect.

This solution amounts to regarding the class of *utensils* as the intersection of the classes of *implements, tools,* and *containers*. Regarding the conjunction *or* as denoting the union of these classes would be more correct on logical grounds, since if X is included in A or X is included in B, then X is included in $A \cup B$. This relation cannot be fitted into a tree, but it can be pictured as in figure 4. However, this does not help to determine whether *spatula* is an *implement, tool,* or *container*, or some subset of the three. In any case, lexicographers do not use *or* with a consistent, mathematical meaning. *Or*-conjoined heads appear not to be usable in constructing hierarchical trees without considerable manipulation and addition of information.

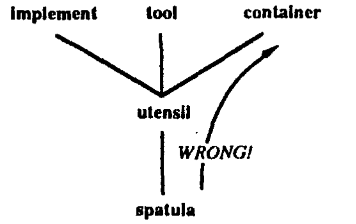

Figure 3 : problematic hierarchy

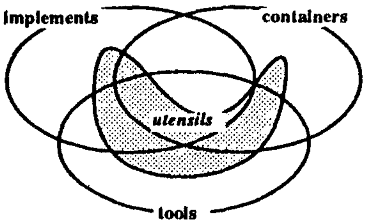

Figure 4. OR as class union

3.2.2. Circularity : 7-11%

It is well known that circularity exists in dictionary definitions, especially when concepts are high up in the hierarchy. For instance, consider the definitions below:

tool	an *implement*, such as a hammer... [CED]
implement	a piece of *equipment*; *tool* or *utensil*. [CED]
utensil	an *implement, tool* or *container*... [CED]

Circular definitions yield hierarchies containing loops (figure 5.a). Unlike merging branches, loops have no interpretation in terms of classes. A loop asserts both that *A is a sub-class of B* and *B is a sub-class of A*, which yields $A = B$. This is why Amsler (1980) suggests merging circularly-defined concepts and regarding them as synonyms (figure 5.b).

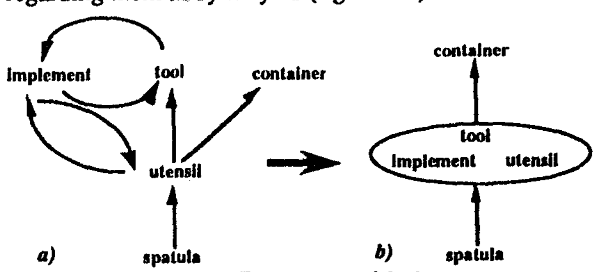

Figure 5. Taxonomy with loops

However, in most cases this solution leads to erroneous results; it is clear, for example, that many *implements, tools,* and *utensils* (e.g., *spatula*) are not *containers*. This problem is similar to the one cited above in section 3.2.1. If dictionary definitions are to be interpreted in terms of set theoretical relations, a more complex mathematical treatment is required. The definitions above can be represented by the following relations:

$$\textbf{tool} \subset \textbf{implement}$$
$$\textbf{implement} \subset (\textbf{equipment} \cup \textbf{tool} \cup \textbf{utensil})$$
$$\textbf{utensil} \subset (\textbf{implement} \cup \textbf{tool} \cup \textbf{container})$$

which, once solved, do not equate *tool, implement,* and *utensil*, but instead define the overlapping classes in figure 6. This representation is clearly more sound on logical grounds. It still does not indicate exactly where *spatula* should appear (since we have no indication that it is not a *container*), but at least it shows that there may be some utensils which are ***not*** containers.

Although this representation is more intuitively accurate than the representation in figure 5.b, ultimately it goes too far in delineating the relations among terms. In actual use, the distinctions among terms are much less clear-cut than figure 6 implies. For instance, the figure indicates that all tools that are containers are also implements, but it is certainly not clear that humans would agree to this or use the terms in a manner consistent with this specification. Dictionaries themselves do not agree, and when taken formally they yield very different diagrams for higher level concepts.

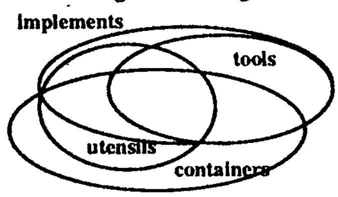

Figure 6. Solving "loops"

Figure 8 shows that 7-11% of the definitions use a hypernym that is itself defined circularly.

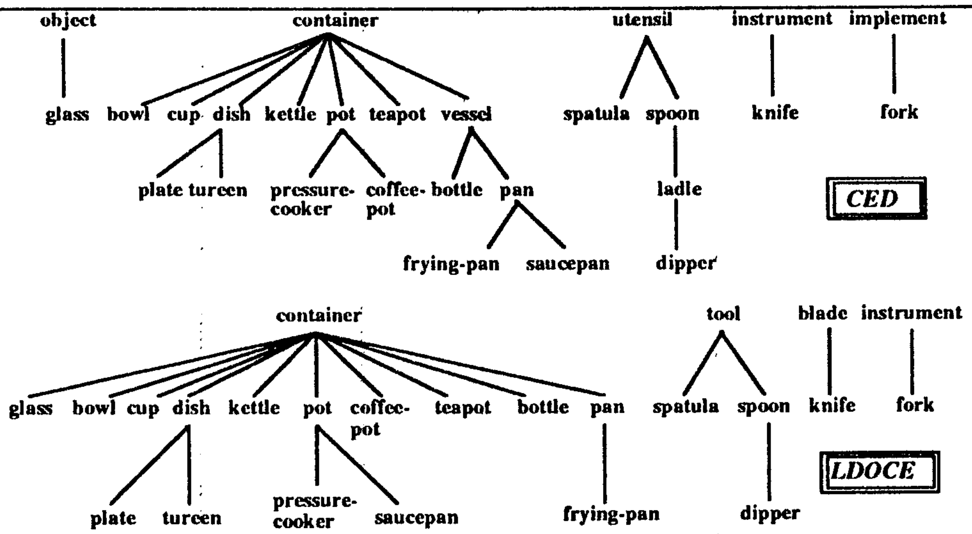

Figure 7. Hierarchies for the CED and LDOCE

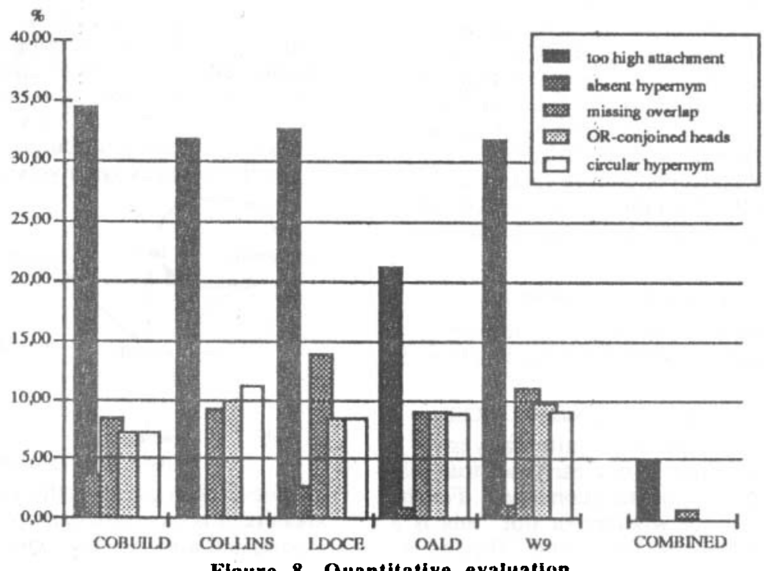

Figure 8. Quantitative evaluation

3.3. Summary

Altogether, the problems described in the sections above yield a 55-70% error rate in automatically extracted hierarchies. Given that we have attempted to consider the most favorable case, it appears that any single dictionary, taken in isolation, is a poor source of automatically extracted semantic information. This is made more evident in figure 7, which demonstrates the marked differences in hierarchies extracted from the *CED* and *LDOCE* for a small subset of our corpus. A summary of our results appears in figure 8.

4. REFINING

We have concluded that hierarchies extracted using strategies such as that of Chodorow, Byrd, and Heidorn are seriously flawed, and are therefore likely to be unusable in NLP systems. However, in this section we discuss various means to refine automatically extracted hierarchies, most of which can be performed automatically.

WORD	COBUILD	COLLINS	LDOCE	OALD	W9	Combined
ladle	spoon	spoon	spoon	spoon	spoon	spoon
basin	container	container	container	bowl	vessel	bowl
ewer	jug	jug OR pitcher	container	pitcher	pitcher OR jug	pitcher
saucepan	pot	pan	pot	pot	pan	pot AND pan
grill	*(absent)*	device	*(absent)*	device	utensil	device AND utensil
fork	tool	implement	instrument	implement	implement	tool, implement AND instrument

Figure 9. Merging hierarchies

4.1. Merging dictionaries

It is possible to use information provided in the *differentiae* of definition texts to refine hierarchies; for example, in the definition

vessel any *object* USED AS a *container*... [CED]

the automatically extracted hypernym is *object*. However, some additional processing of the definition text enables the extraction of *container* following the phrase "used as". It is also possible to use other definitions. For example, the *CED* does not specify that *knife* and *spoon* are *implements*, but this information is provided in the definition of *cutlery:*

cutlery implements used for eating SUCH AS *knives, forks,* and *spoons*. [CED]

The extraction of information from *differentiae* demands some extra parsing, which may be difficult for complex definitions. Also, further research is required to determine which phrases function as markers for which kind of information, and to determine how consistent their use is. More importantly, such information is sporadic, and its extraction may require more effort than the results warrant. We therefore seek more "brute force" methods to improve automatically extracted hierarchies.

One of the most promising strategies for refining extracted information is the use of information from several dictionaries. Hierarchies derived from individual dictionaries suffer from incompleteness, but it is extremely unlikely that the same information is consistently missing from all dictionaries. For instance, the *CED* attaches *cup* to *container*, which is too high in the hierarchy, while the *W9* attaches it lower, to *vessel*. It is therefore possible to use taxonomic information from several dictionaries to fill in absent hypernyms, missing links, and to rectify cases of too high attachment.

To investigate this possibility, we merged the information extracted from the five English dictionaries in our database. The individual data for the five dictionaries was organized in a table, as in figure 9. Merging these hierarchies into a single hierarchy was accomplished automatically by applying a simple algorithm, which scans the table line-by-line, as follows:

1) regard cells containing multiple heads conjoined by *or* as null, since, as we saw in section 3.2.1, they do not reliably provide a hypernym.

2) if all the cells agree (as for *ladle*), keep that term as the hypernym. Otherwise:
a) if a term is a hypernym of another term in the line, ignore it.
b) take the remaining cell or cells as the hypernym(s).

This algorithm must be applied recursively, since, for example, it may not yet be known when evaluating *basin* that *container* is a hypernym of *vessel*, and *vessel* is a hypernym of *bowl*, until those terms are themselves processed. Therefore, several passes through the table are required. Note that if after applying the algorithm several terms are left as hypernyms for a given word, we effectively create an overlap in the hierarchy. For example, *saucepen* is attached to both *pot* and *pan*, and *fork* is attached to *tool, implement*, and *instrument*.

We evaluate the quality of the resulting combined hierarchy using the same strategy applied in section 3. It is interesting to note that in the merged hierarchy, all the absent hypernym problems (including absence due to *or*-heads) have been eliminated, since in every case at least one of the five dictionaries gives a valid hypernym. In addition, almost all of the attachments too high in the hierarchy and missing overlaps have disappeared, although a few cases remain (5% and 1%, respectively). None of the dictionaries, for instance, gives *pot* as the hypernym of *teapot*, although three of the five dictionaries give *pot* as the hypernym of *coffeepot*. A larger dictionary database would enable the elimination of many of these remaining imperfections (for example, *New Penguin English Dictionary*, not included in our database, gives *pot* as a hypernym of *teapot*).

Merging dictionaries on a large scale assumes that it is possible to automatically map senses across them. For our small sample, we mapped senses among dictionaries by hand. We describe elsewhere a promising method to automatically accomplish sense mapping, using a spreading activation algorithm (Ide and Véronis, 1990).

4.2. Covert categories

There remain a number of circularly-defined hypernyms in the combined taxonomy, which demand additional consideration on theoretical grounds. Circularly-defined terms tend to appear when lexicographers lack terms to designate certain concepts. The fact that "it is not impossible for what is intuitively recognized as a conceptual category to be without a label" has already been noted (Cruse, 1986, p. 147). The lack of a specific term for a recognizable concept tends to occur more frequently at the higher levels of the hierarchy (and at the very lowest and most specific levels as well--e.g., there is no term to designate forks with two prongs). This is probably because any language includes the most terms at the *generic level* (Brown, 1958), that is, the level of everyday, ordinary terms for objects and living things (*dog, pencil, house,* etc.).

Circularity, as well as the use of *or*-conjoined terms at the high levels of the hierarchy, results largely from the lexicographers' efforts to approximate the terms they lack. For example, there is no clear term to denote that category of objects which fall under any of the terms *utensil, tool, implement, instrument*, although this concept seems to exist. Clearly, these terms are not strictly synonymous--there are, for example, *utensils* that one would not call *tools* (e.g., a *colander*). If a term, let us say X, for the concept existed, then the definitions for *utensil, tool, implement*, and *instrument* could simply read "an X that...". Since this is not the case, lexicographers define each term with a list including the others, which enables the delineation of a concept which encompasses all of them.

One way to resolve difficulties at the higher levels of extracted hierarchies is to introduce "covert categories", that is, concepts which do not correspond to any particular word. We therefore do not merge circular terms into a single concept, but instead create a common "covert" hypernym for all of them. In this way, *tool, utensil, implement*, and *instrument* each appear in the hierarchy as kinds of INSTRUMENTAL-OBJECT (covert categories names are capitalized).

We need a means to determine when and where covert categories are necessary. Circularities in dictionary definitions clearly indicate the presence of covert categories. However, we obviously cannot use a single dictionary to determine them, because the loops contained in one dictionary rarely include all of the terms that may be involved in the "constellation" representing a given covert category. For instance, the CED contains the loop *tool-implement-utensil*, while the *COBUILD* contains a loop for *tool-instrument;* this provides strong evidence that all four terms should be involved in a constellation. Supporting information can be derived by looking at the hyponyms for each of the candidate terms in different dictionaries. The word *fork*, for example, is defined as *tool (COBUILD), implement (CED, OALD, W9)*, and *instrument (LDOCE)*, while spoon is defined as *object (COBUILD), utensil (CED, OALD), tool (LDOCE)* and *implement (W9)*, which adds further support to the idea that *tool, utensil, instrument*, and *implement* belong to the same constellation.

Even if it is relatively easy to automatically detect circularities, the final determination of which covert categories to create and the terms that are involved in them must be done manually. However, this task is not as daunting as it may first appear, since it involves only the higher levels of the hierarchy, and likely involves a relatively small number of covert categories.

4.3. Summary

By merging five dictionaries, all but 6% of the problems found in individual dictionaries were eliminated (figure 8). This result is made clear in figure 10, which includes the same small subset of the sample corpus as in the individual hierarchies given in figure 7. Although there remain a few imperfections, the combined hierarchy is much more accurate and complete, and therefore more useful, than the hierarchy derived from any one of the dictionaries alone.

5. CONCLUSION

The results of our study show that dictionaries can be a reliable source of automatically extracted semantic information. Merging information from several dictionaries improved the quality of extracted information to an acceptable level. However, these results were obtained for a selected corpus representing a best case situation. It is likely that different results will be obtained for larger, less restricted cases. Our results suggest that this is an encouraging line of research to pursue for refining automatically extracted information.

REFERENCES

AKKERMAN, E., MASEREEUW, P. C., MEIJS, W. J. (1985). Designing a computerized lexicon for linguistic purposes. *ASCOT Report No. 1*, Rodopi, Amsterdam.

AMSLER, R. A. (1980). *The structure of the Merriam-Webster Pocket Dictionary.* Ph. D. Diss., U. Texas at Austin.

BROWN, R. W. (1958) How shall a thing be called? *Psychological Review*, 65, 14-21.

BYRD, R. J. (1989) Discovering relationships among word senses. *Proc. 5th Conf. UW Centre for the New OED*, Oxford, 67-79.

BYRD, R. J., CALZOLARI, N., CHODOROW, M. S., KLAVANS, J. L., NEFF, M. S., RIZK, O. (1987) Tools and methods for computational linguistics. *Computational Linguistics*, 13, 3/4, 219-240.

CALZOLARI, N.(1984). Detecting patterns in a lexical data base. *COLING'84*, 170-173.

CHODOROW, M. S., BYRD. R. J., HEIDORN, G. E. (1985). Extracting semantic hierarchies from a large on-line dictionary. *Proc. 23rd Annual Conf. of the ACL*, Chicago, 299-304.

CRUSE, D. A. (1986). *Lexical semantics*. Cambridge University Press, Cambridge.

IDE, N., M., VÉRONIS, J. (1990). Mapping Dictionaries: A Spreading Activation Approach, *Proc. 6th Conf. UW Centre for the New OED*, Waterloo, 52-64.

KLAVANS, J., CHODOROW, M., WACHOLDER, N (1990). From dictionary to knowledge base via taxonomy. *Proc. 6th Conf. UW Centre for the New OED*, Waterloo, 110-132.

LYONS, J. (1963) *Structural semantics*. Blackwell, Oxford.

MARKOWITZ, J., AHLSWEDE, T., EVENS, M. (1986). Semantically significant patterns in dictionary definitions. *Proc. 24rd Annual Conf. of the ACL*, New York, 112-119.

NAKAMURA, J., NAGAO, M. (1988). Extraction of semantic information from an ordinary English dictionary and its evaluation. *COLING'88*, 459-464.

VÉRONIS, J., IDE, N., M. (1990). Word Sense Disambiguation with Very Large Neural Networks Extracted from Machine Readable Dictionaries. *COLING'90*, Helsinki.

WILKS, Y., D. FASS, C. GUO, J. MACDONALD, T. PLATE, B. SLATOR (1990). Providing Machine Tractable Dictionary Tools. *Machine Translation*, 5, 99-154.

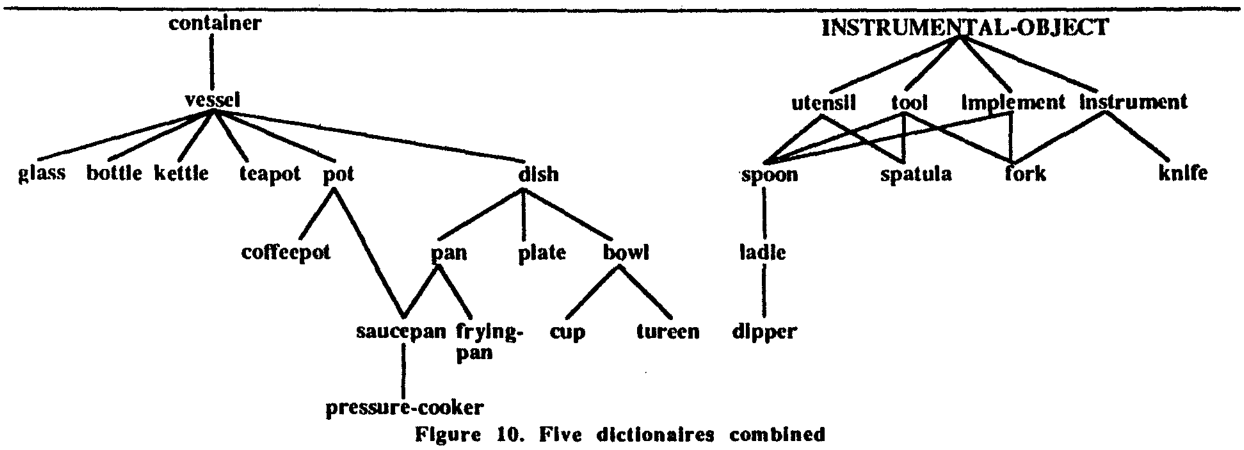

Figure 10. Five dictionaires combined