A Large Vocabulary Semantic Network for Computerised Speech Recognition

George C. Demetriou and Eric S. Atwell
Artificial Intelligence Division
School of Computer Studies
University of Leeds
e-mail: `george@scs.leeds.ac.uk` `eric@scs.leeds.ac.uk`

Abstract

The work presented in this paper deals with the construction of a large-vocabulary semantic network to assist computerised speech or text recognition. The semantic network is systematically constructed with semantic information about nouns and verbs from the Longman Dictionary of Contemporary English by the application of pattern matching rules. It is represented in the form of a directed graph where nodes represent word senses and links represent the types of conceptual relationships. A semantic score, for pairwise combinations of word candidates in a speech recognition lattice, can be derived by traversing the network and calculating the conceptual distance between the senses of these candidates.

1 Introduction

Search is a major problem for the development of applications that belong to the 'noisy' channel concept, such as computerized speech, handwriting and optical character recognition. Semantics is often used as a higher level knowledge source to restrict search and reduce the number of word sequence hypotheses by providing compatibility constraints between lexical items in the same sentence. To date, nearly all systems in the literature have concentrated on utilising semantic information for small vocabulary recognition (1000 words or less) assuming restricted environments or specific task domains (i.e. e-mail task, resource management task, air travel information task, etc.). To support a reasonable degree of vocabulary and domain independence for speech or text recognition, a semantic knowledge base should provide rich and accurate semantic information about the relationships of more than 20000 words (or more than 100000 words if we include inflected and derived variants). Statistical language models often fail to capture long distance dependencies between words whereas hand-crafted lexica such as Wordnet (Miller et al 1990) have the disadvantage of being costly and time consuming to build and utilize. There is no standard technique to automatically acquire, represent and implement semantic constraints for general purpose, task independent speech recognition for applications like machine dictation or general inquiries through the telephone or speech-to-speech machine translation (one possible exception is Rose and Evett (1992) for large vocabulary text recognition).

The objectives of the work discussed in this paper are:

- to investigate and develop an automatic technique for extracting conceptual relationships between words from the Longman Dictionary of Contemporary English (LDOCE)

• to represent the derived semantic associations in a semantic network or directed graph and to implement appropriate score functions for them

• to utilise the semantic network for large vocabulary lattice disambiguation (i.e. to apply it for the selection of the correct sequence of word candidates during search)

• to compare the performance of the derived model with definition-overlap Wordnet-based models for speech recognition.

2 Semantic Information and Machine Readable Dictionaries

The motivation of our research lies in the use of Machine Readable Dictionaries (MRDs) for the extraction of semantic information for broad-coverage NLP. Several researchers have experimented with the identification of semantic relations between words implicitly expressed in sense definitions and examples in MRDs. For example, a program should be able to identify the 'is-a', 'purpose' and 'object' relations from analysis of:

carpet(n,1)*: heavy woven often woollen material for covering floors or stairs*

```
carpet:
      is-a: [material]
      purpose:  [cover]
            object: [floor,stairs]              (1)
```

Similarly, for ***cover(v,1)****: to place or spread something upon, over,or in front of (something) in order to protect, hide, etc.* we may have:

```
cover:
      is-a: [place,spread]
            object: [something]
      purpose: [protect,hide]                   (2)
```

In the literature, Calzolari (1984) examined the hyponymy and restriction relations by applying lexical pattern matching procedures to an Italian dictionary. Along the same line are the works of Chodorow et al (1985) to specify genus terms for nouns and verbs from the Webster's 7th, Markowitz et al (1986) who attempted to discover the defining generic patterns used for the construction of dictionary definitions and Nakamura and Nagao (1988).

General purpose syntactic grammars to parse the dictionary definitions has been used by Jensen and Binot (1987), Montemagni and Vanderwende(1992) and Dolan et al (1993), whereas tailor-made parsers have been constructed by Alshawi(1988) and Wilks et al (1989) to acquire the semantic information from LDOCE sense definitions.

All the above approaches aim not only at the verification that a relationship between the headword and a word in the definition holds, but also, at the identification of the nature of that semantic relationship. In contrast, approaches using simple word overlap techniques for sense disambiguation or information retrieval (Lesk 1987, Guthrie et al 1991, Demetriou 1993 among others) or text recognition (Rose and Evett 1992) cannot generally distinguish between the defining function patterns (such as 'a', 'part of', 'used for', etc.) and defining key concepts (such as 'person', 'vehicle', 'book', etc.).

From the speech recognition perspective, the identification of the nature of semantic relationships between words offers two opportunities. First, it reduces the number of semantically insignificant words (ie. function words) in dictionary definitions and second, because concepts can be organised in semantic network structures, compatibility checks of relations can be facilititated by traversing the correct links through the network. Each link in the network specifies a semantic score so that the NL component should be able to provide feedback in the form of a total semantic distance which indicates the certainty factor about the semantic relationship for each pair of word candidates in context.

3 Identifying and extracting semantic relations from LDOCE

LDOCE exploits the idea of considering the language of dictionary definitions as a particular sublanguage within natural language (Calzolari 1984). Each definition is written with words from a restricted vocabulary of about 2200 words (except for certain cases of cross-referencing in definitions). LDOCE also provides a semantic 'box' system that encodes restrictions between nouns and verbs and nouns and adjectives, as well as a thematic code system that classifies the senses of the words in domain categories.

Our work involves in analysing senses of nouns (total number: 23586, total number of noun senses: 41270) and verbs (total number: 7922, total number of verb senses: 16731)[1]. After filtering all special characters in the dictionary definitions (being there for typesetting purposes), the process of identifying the kind of relationships between a headword and a word in the definition makes use of the defining function patterns in the 'genus' and 'differentia' parts[2]. Generally, the method is similar to the ones by Calzolari (1984), Calzolari and Picchi (1988) and Nakamura and Nagao (1988). First, the defining function patterns are identified statistically and classified into semantic categories each one specifying a semantic relation. For example, the statistical analysis reveals that most noun definitions start by specifying an 'is-a', 'part-of', 'act-of', etc. relation between the headword and the semantic head in the definition.

The morphological analysis and identification of the parts of speech of words in the definitions is done by a simple program that accesses the entries of the CELEX lexical database for the stems of the words. Although the coverage of CELEX is not 100% for all LDOCE entries (since the vocabulary of CELEX represents the intersection of LDOCE and OALD) it was found adequate to process the words of the defining vocabulary of LDOCE and their variants. When a word is not in CELEX (in most cases compound words), the program accesses the corresponding LDOCE entry. For the inflected forms of verbs and plurals of nouns it is possible to retrieve the part of speech directly from CELEX. Syntactically ambiguous words are looked up in LDOCE. The assumption is that the most common entry of a word is listed first in LDOCE (for example, 'pause' is retrieved as a noun

[1] At this stage we avoided dealing with adjectives which we considered semantically 'null' words although their contribution in filtering speech recognition lattices would be as significant as those of nouns and verbs.

[2] Dictionary definitions generally include a "genus" part placing the word in a semantic hierarchy, and a "differentiae" part giving further general information about meaning - see (Calzolari 1984).

and not a verb if it appears in this form in a definition).

Structures of the form <*determiner*> *noun* <*relative clause*> (where relative clause = prepositional phrase, adjective phrase, etc.) directly map an entry to the semantic head of the definition with an isa relation. The semantic relation of the entry and the head of structures that begin with <*determiner*> *noun1 of noun2* <*relative clause*> is usually denoted by the semantics of *noun1*, for example 'part of' denotes a 'part-of' relation, 'act of' denotes an 'action' relation and so on. Two questions are encountered for the computational analysis of these structures: 'how many different relations can be specified from these different function nouns' and 'how can this classification be done automatically'. One extreme in this process would be to label each relation with its corresponding functional pattern from which it is derived. For example, 'a group of' structure labels a 'group' relation, a 'state of' structure a 'state' relation, etc. Although this classification has the advantage that it can be automated, it cannot probably realise semantic relations that are generally the same (ie. 'part-of', 'top-of') but are expressed in a different way (due to lexicographic preferences). The other extreme would be to try a very general approach such as the one proposed by Sinclair (Renouf and Sinclair 1991) that classifies the nouns of collocates in the 'a/an noun1 of' framework in classes of 'measurement', 'focus' and 'support'. This classification is also not very useful since it cannot provide clear links to inherited properties in the network (i.e. how to specify that 'salmon' as a 'fish' has 'bones'?). We take an intermediate approach and classify all function nouns taking part in 'noun of' structures into seven general semantic categories (is-a, part-of, class-of, action, condition, measure, form) roughly similar to the ones in Nakamura and Nagao (1988).

The identification of the relations between a noun entry and the differentiae part is a much more difficult task. This is because the words in the differentiae part are often used not for describing a direct relation with the word entry but with other words in the definition, as for example, the words 'gases', 'enter', 'escape', 'engine' in:

manifold (n,1)*: a pipe with holes connecting it to a number of smaller pipes, to allow gases to enter or escape from an engine, such as that of a car.*

The use of pattern matching rules can be also used to identify a number of certain relations by analysing the regularities of these patterns. For example, for the 'purpose' relation, recurring patterns such as 'used for', 'for', 'for the purpose of', 'in order to', 'to' can be used to associate their arguments with the word entry. Obviously, this process becomes complicated and laborious when someone wants to declare all possible semantic associations between words and because of considerable variation in the LDOCE defining descriptions the number of pattern matching rules is high[3]. In addition, the issue of classifying the recurring patterns in certain classes (each one specifying a particular relation) without intuitive knowledge has not been solved yet.

One way to develop a specially designed grammar for LDOCE definitions would be to follow Vossen et al (1989)'s approach (that attempts a classification of the structure of noun

[3]Much controversy surrounds the adequacy of string patterns for extracting semantic relations from dictionary definitions. Montemagni and Vanderwende (1992) argue that only patterns based on structural information provide reliable relations for the differentiae part.

definitions with the use of statistical analysis), with the expense of a rather large set of manually coded rules. Taking into account that the distinction between senses is often delicate and fine-grained we limit ourselves to the investigation of the extraction of six basic relations between a noun entry and nouns and verbs in its definition ('has-part', 'purpose', 'location', 'time', 'subject-of', 'object-of') and two selectional relations between verbs and nouns in the definition ('subject', 'object'). Some of these relations are not too difficult to identify since they are usually introduced by prepositions. For example, a prepositional clause introduced by 'with' just after the semantic head noun of the genus term may specify a 'has-part' relation between the noun entry and the head noun the prepositional clause, whereas a relative clause introduced by 'for' specifies a 'purpose' relation between the noun entry and the head verb of the relative clause (ie. *car(n,1): a vehicle with 3 or 4 wheels and driven by a motor, esp. one for carrying people*).

The parsing rules we use are rather simple and general. The procedure tries to locate all recurring function patterns and to associate them with the following nouns and verbs in a left-to-right basis. At the moment of writing this paper, no claim can be made about the accuracy of the extracted information since our general purpose parsing method cannot help in the complete semantic analysis of all definitions. By randomly inspecting the performance of the parser for a small number entries, it appears that the identification of the semantic head (genus term) is quite satisfactory while relations identified for other words still get poor results. Most errors are produced from the inability of parsing rules to correctly analyse all different patterns of the same semantic relation (an issue we will work on) and the presence of conjuctive ('and') and disjunctive ('or') elements in the definitions that need further handling. Many words (especially nouns) are left semantically unspecified. This is partially due to the fact that our set of semantic relations cannot cover all possibilities (for example, 'instrument' relations in noun definitions cannot be identified) and partially due to parsing (at the moment, all such words get an 'unspecified' relation).

The same approach has been adopted for verbs. Most verbs usually start with a 'to verb' pattern, so that it is fairly easy to identify an 'is-a' relation by directly mapping a verb entry to the first verb following 'to' in its definition. The types of relations distinguished between a verb entry and other words in its definition are: 'object', 'subject', 'use-of', 'manner', 'purpose', 'location' and 'time'. Main problems encountered here are to try to locate the typical subjects and objects for a verb (though not as difficult as in noun definitions) and the correct disambiguation between 'use-of' and 'manner' relations which generally use the same phrasal patterns (i.e. in most cases 'with'). For the latter we seek assistance from the 'box' coding system in order to identify abstract and concrete nouns. For example, the entry of ***frighten(v,1)****: to fill with fear* specifies a 'manner' relation between 'frighten' and 'fear' because 'fear' belongs to the abstract class of nouns.

4 A semantic metric for speech recognition

For speech recognition and understanding, an appropriate semantic metric should be developed to take advantage of the inheritance capabilities provided by the network. Generally, while 'is-a' relations can be used for inheriting properties from the very general to the very specific, the same is not always true for the 'part-of', 'location', or other relations. For the purpose of this paper, however, we will not try to explore the possibilities of deriving plausible inference rules that could be used to indicate which edges of the network are to be preferred in order to link (or not) two nodes[4]. Instead, we use a strategy that assigns a semantic distance score (weight) to each relation and then traverses all paths through the network for all pairs of content word candidates in a sentence hypothesis. For our preliminary experiments, the 'is-a' relation had a zero weight whereas all other relations were attached weights equal to 1. It is one of our intentions to experiment with various weighting schemes in order to optimise the contribution of each semantic relation.

It should be noted that this kind of representation includes two kinds of relationships: those that directly link a word entry to a word in its definition (such as 'carpet'→'is-a'→'material) and indirect, nested relations (for example, 'carpet' is indirectly linked to 'floor' and 'stairs' through 'cover' in (1)). During search, the direct relation that connects the parent node (eg. 'purpose' in 'carpet'→'purpose'→'cover') is traversed up to the child node (ie. 'cover') carrying with it all indirect relations it may contain (e.g. 'cover' is extended not only to relations specified by the words in its definition but also to 'floor' and 'stairs' through the 'object' relation in (1)).

If $R=\{r_1,r_2,...,r_n\}$ is the set of semantic relations, $W=\{w_1,w_2,...,w_n\}$ is the set of the correspon-ding weights and $P=\{p_1,p_2,..,p_k\}$ is the set of scores of all possible paths that link two concepts, then the semantic score for path j (j=1,...,k) can be given by:

$$p_j = \sum_{i=1}^{n} a_i \cdot w_i$$

where a_i the number of times the relation r_i exists in path j. The distance between two concepts c_1 and c_2 is defined as the minimum score derived between c_1 and c_2 i.e.

$$D(c_1,c_2) = \min(p_j) \qquad j=1,2,..,k$$

The strategy for lattice disambiguation follows the same steps as in Demetriou (1993). That is, for each set of word candidates $V=\{v_1,v_2,...,v_m)$ of a sentence hypothesis, the set of all possible sense combinations is specified[5]. For each combination $C_k=\{c_1,c_2,....c_n\}$ (where c_i represents a sense of word candidate w_i) a semantic score function calculates the sum of distances of pairwise combinations of senses:

$$S(C_k) = \Sigma\, D(c_i,c_j) \qquad (i,j=1,...,l \text{ and } i \text{ not equal } j)$$
$$(k=1,...,q \text{ where } q \text{ the number of sense combinations})$$

The sense combination with the lowest score represents the combination with the highest semantic relatedness. This also gives a total score for the particular set of word candidates V:

$$S_V = \min(S(C_k)) \qquad (k=1,...,q)$$

which can be transformed by an appropriate function to combine with its acoustic score.

---

[4]For example, because 'heart' has a 'location' relation 'chest', 'chest' is a 'part-of' body, it seems plausible to infer that 'heart' has a 'location' relation with 'body' and continue traversing through this point.

[5]Assuming that each word v_i has a total of s_i senses (i=1,...,m) then the number of sense combinations is $\pi_{i=1m}\, s_i$.

5 Discussion and further research

At a first look, the above approach is an alternative to the one that replaces words within a definition with their sense definitions[6], where the problem of when the search should stop is unsolved. However, this implementation is more flexible because it offers two opportunities: to activate/deactivate the traversal for certain links (relations) by appropriate adjustment of weights, and to provide an upper threshold beyond which search will be pruned. For example, by limiting search to 'is-a' and 'part-of' relations between nouns and certain relations between verbs and nouns, the system can perform semantic compatibility checks between words fairly easy with rules in case-frame form:

```
test_for_compatibility(Verb,Relation,Noun):-
          Check=..[Verb,Relation,Noun], call(Check),!.

test_for_compatibility(Verb,Relation,Noun):-
          inherit(Noun,ParentNoun),
          test_for_compatibility(Verb,Relation,ParentNoun).

inherit(Noun,ParentNoun):-
          (Check=..[Noun,is-a,ParentNoun]
          ;
          Check=..[Noun,part-of,ParentNoun]),
          call(Check).
```

(where Relation can be constrained to 'subject' or 'object')

Unfortunately, a dictionary definition cannot describe all links a word can have with all other words in the language. For example, no sense of 'house' in LDOCE contains any information about gardens, vegetables or flowers so that combinations of concepts such as 'house'-'vegetable' or 'house'-'flower', even though plausible, are assumed conceptually distant with the above method. However, 'house', 'vegetable' and 'flower' are mentioned in ***garden(n,1)****: a piece of land, often near a house, on which flowers and vegetables may be grown*. The possibility of developing a system that accesses not only the definitions of a word, but also, all other definitions that contain that word, would probably provide a better measure of the strength of conceptual association between word senses (Dolan et al 1993). In that case, the semantic metric function should be revised to account for both 'outward' and 'inward' links to a word sense.

Our plans for further research go along three axes:

• to improve the performance of the knowledge extractor from LDOCE definitions. For this, we do not restrict ourselves from using a general purpose parser in the near future, if the results suggest so.

• to compare the efficiency of the model with the method that uses the word overlap technique between dictionary definitions by applying both to speech recognition.

[6]as Rose (1993) has discussed and found that this introduced more 'red herrings' between candidates while the overall performance did not rise significantly.

• to compare the efficiency of the model with a similar one that utilizes semantic information in Wordnet.

REFERENCES

1. Alshawi H. (1987), Processing Dictionary Definitions with Phrasal Pattern Hierarchies. In: Computational Linguistics 13, no. 3-4, 195-202.
2. Calzolari N. (1984), Machine readable dictionaries, lexical databases and the lexical system. In: Proceedings of COLING-84, 170-173.
3. Calzolari N. and Picchi E. (1988), Acquisition of semantic information from an on-line dictionary. In: Proceedings of COLING-88, 87-91.
4. Chodorow M. S., Byrd R. J. and Heidorn G.E. (1985), Extracting semantic hierarchies from a large on-line dictionary. In Proceedings of the Association for Computational Linguistics, 299-304.
5. Demetriou G. (1993), Lexical Disambiguation Using Constraint Handling In Prolog (CHIP). In: Proceedings of the 6th European Conference of the Association for Computational Linguistics, 431-436.
6. Dolan W., Vanderwende L. and Richardson S. D. (1993), Automatically deriving structured knowledge bases from on-line dictionaries. In: Proceedings of the Conference of the Pacific Association for Computational Linguistics (PACLING-93), 5-14.
7. Guthrie J., Guthrie L., Wilks Y. and Aidinejad H. (1991), Subject dependent co-occurence and word sense disambiguation. In: Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics, 146-152.
8. Jensen K. and Binot J-L. (1987), Disambiguating prepositional phrase attachments by using on-line dictionary definitions. In: Computational Linguistics 13, no. 3-4, 251-260.
9. Lesk M. (1987), Automatic sense disambiguation using Machine Readable Dictionaries: how to tell a pine cone from an ice-cream cone. In: Proceedings of the ACM SIG-DOC Conference, Ontario, Canada.
10. Markowitz J., Ahlswede T. and Evans M. (1986), Semantically significant patterns in dictionary definitions. In: Proceedings of the Conference of the Association for Computational Linguistics, 112-119.
11. Montemagni S. and Vanderwende L. (1992), Structural patterns vs. string patterns fro extracting semantic information from dictionaries. In Proceedings of COLING-92, 546-552.
12. Miller G. A., Beckwith R., Fellbaum C., Gross D. and Miller K. J. (1990), Introduction to Wordnet: An on-line lexical Database. In: Journal of Lexicography, 3(4):235-244.
13. Nakamura J. and Nagao M. (1988), Extraction of semantic information from an ordinary English dictionary and its evaluation. In: Proceedings of COLING-88, 459-464.
14.Niedermair G. T. Streit M. and Tropf H. (1990), Linguistic processing related to speech understanting in SPICOS II. In: Speech Communication 9, 565-585.
15. Renouf A. and Sinclair J. (1991), Collocational frameworks in English. In: English Corpus Linguistics: Studies in Honour of Jan Svartvik, Aijmer K. and Altenberg B. (eds), Longman, 128-143.
16. Rose T. G. and Evett L. J. (1992), A large vocabulary semantic analyzer for handwriting recognition. In: AISBQ 80, 34-39.
17. Rose T.G. (1993), Large Vocabulary Analysis for Text Recognition, PhD Thesis, Nottingham Trent University.
18. Sussna M. (1993), Word sense disambiguation for free-text indexing using a massive semantic network. In: Proceedings of the 2nd Int. Conf. on Information and Knowledge Management,

Airlington, Virginia, USA.
19. Vossen P., Meijs W. and den Broeder M. (1989), Meaning and structure in dictionary definitions. In: Computational Lexicography for Natural Language Processing, B. Boguraev and T. Briscoe (eds), Longman, 171-192.
20. Wilks Y., Fass D., Guo C., McDonald J., Plate T., and Slator B. (1989), A tractable machine dictionary as a resource for computational semantics. In: Computational Lexicography for Natural Language Processing, B. Boguraev and T. Briscoe (eds), Longman, 193-228.