THE USE OF A SEMANTIC NETWORK IN SPEECH DIALOGUE

G.Th. Niedermaier
ZFE F2 KOM 32, Siemens AG, Otto-Hahn-Ring 6, 8000 Munich 83, FRG

Abstract:
The paper discusses the construction principles and use of a semantic network within the dialogue system for continuous speech input, called SPICOS II. One of the main problems in continuous speech recognition is the very close coupling of syntactic and semantic constraints with the recognition process while maintaining their descriptive independence. The paper will show how the evaluation of the semantic network, that actually models the behaviour of words and their semantically acceptable relations, can be integrated into the linguistic search process and help to reduce combinatorial complexity. The same network is used to support the resolution of anaphoric expressions in various ways.

1. Introduction

Within the framework of the speech dialogue system SPICOS II, which is under development in our research laboratory, a semantic network has been set up which models the semantic relations of the words of our application (office environment) and serves to overcome different problems arising especially in a system for speech dialogue. The paper proposes to use such a network mainly for the early restriction of sentence-hypotheses which are either built up along with the acoustic recognition, or are delivered as a graph of sentence-hypotheses. Additionally the same network can be used to support anaphor-resolution in two ways, first by providing semantic categories for pronouns from the sentential context, and second by licensing anaphoric bindings if there are different terms in anaphor and antecedent.

2. The SPICOS Environment

The prototype of the SPICOS II system, which is currently under development, will be finished in early 1990. It will be capable of leading a limited dialogue on the content of an office database. The system aims at speaker-adaptivity and knows a vocabulary of over 1000 words. An overview of the system is given in (NIE 89).

A dialogue-handler monitors the components of the system and triggers their actions. More important it keeps track of the dialogue-history and the status of the analysis modules to ensure that the system has properly understood the utterance or to resolve ambiguities in the analysis in cooperation with the user. For a more detailed description of the dialogue-handler see (DEV 88).

2.1 The Acoustic Linguistic Interface

In the development of the prototype two different interfaces between acoustic analysis and linguistic processing are being investigated; they are due to differences in acoustic recognition. One is entirely top down oriented and delivers by use of a stochastic language model a lattice of sentence hypotheses from which linguistic analysis filters out the best one. A mainly bottom up oriented approach either delivers sentence hypotheses as above or partially developed word graphs, which are checked for correctness by the linguistic analysis with each step forward in the search through the graph. Both approaches require a simultaneous semantic analysis that during parsing is able to exclude semantically ill-formed hypotheses. In the latter approach the linguistic analysis can ensure that only syntactically and semantically correct hypotheses are pursued any further.

The gain of processing syntactic and semantic constraints together is - because of the complexity of the word-hypotheses-graph - much bigger in the second approach.

The disadvantage is - that for fast processing the syntactic-semantic parsing of each of the hypotheses in the developing word-graph is up to now a rather expensive procedure.

- given the complexity of the word-graph linguistic processing could hardly tolerate any words to be missing in the acoustic hypotheses, which obviously is hardly the case.

To avoid this, acoustic analysis has to be provided with strong syntactic-semantic predictions, which in the future we plan to produce from a grammar which compiles syntactic knowledge from the grammar and semantic knowledge contained in the network.

3. Linguistic Analysis

The linguistic analysis process is based on modified chart-parsing. It works on an extended form of a augmented phrase structure grammar (APSG) (see TRG 89). The APSG allows us to introduce tests which not only check syntactic well-formedness, but also semantic consistency of various types of phrase structures. This ensures maximal correctness of the hypotheses being investigated and helps thus reducing the search space. The semantic network, through which these hypotheses are checked and which will be described in the following section, is naturally tied to the relations of things (and consequently the words) as they are defined for the particular application. They are not necessarily meant to reflect universal semantic properties, although this could be a 'side effect'.

3.1. Use of the Network

In our speech dialogue system this networks serves three different purposes:

- Firstly, it checks the semantic consistency during the syntactic analysis of a sentence.

For continuous speech understanding systems it has in the meantime become an accepted fact that all available knowledge in the linguistic part of the
process should be brought to bear as soon as possible. This does not only support acoustic recognition, but also limits the search space of the linguistic analysis, which can very easily become much to large, if the syntactic constraints are processed without the inclusion of the semantic constraints. Since semantic restrictions are usually much more efficient in reducing the amount of combinatorial possibilities than syntactic constraints these restrictions are to be applied at the earliest possible point in time. (NIE 86, THU 88) Especially in speech systems this has often been achieved by using semantic grammar variants (LOW 80, SCHIMA 86, ME-PAE 87) or caseframe oriented approaches (SAG-KU 88, POE 87, SHI 86, YOU 88) with the known effects of either strong application dependence in the grammar or loss of syntactic descriptivity in the case of caseframes, if they are not used in combination with a syntactic grammar. The caseframes have also clear limitations in expressing semantic relations other than 'verb + sentence-function' or 'noun + attribute'. The semantic network however not only allows to express the semantic restrictions for role fillers in the most general way by means of a semantic hierarchy, but also allows to describe specification-relations (nominal or adjectival or other), attribute-value relations, term-classes (for resolution purposes), etc.

The goal of early application of semantic restriction we achieve by checking semantic compatibility between content words within the framework of the APSG. The semantic test forms part of a set of constraints which have to be met immediately each time the rule proceeds with a new constituent part. The test accesses the semantic network and retrieves from it the semantic acceptability of a pair of content-words together with additional information concerning the type of relation, e.g. the respective deepcase in a noun-verb relation.

- Secondly the semantic network is used in anaphor-resolution for attaching semantic categories to semantically empty pronouns (see chapter 7 of this paper) and
- Thirdly it is used for licencing anaphoric bindings, which are indirectly established through synonymy-hypernym or other relations. An account of the use of these relations in dialogues is given in (GEH 89).

4. The Structure of the Semantic Network

The network is a static model, not of the entities in the area of our application, but of the words describing these entities and their respective semantic relationships, i.e. the possible semantic behaviour of these words within an utterance. To avoid any confusion: the network does not represent the data of the application as instances of its concepts nor is it used as a means for meaning representation; this is achieved through other formalisms (see NIE 89).

A partial view of the semantic network is given in fig. 1.

Each content word is associated to a node (concept), of which it is an instance, in a hierarchy of nodes. A node can of course have several instances, if the words behave in the same way in their relation to other words. (This can be regarded as a very narrow definition of synonymity). The nodes are such roughly equivalent to 'semantic classes'.

The backbone of this network is a hierarchy of labelled nodes. The nodes in this taxonomy of classes are linked by is-a-relations. Words can of course also be instances of 'higher' classes, if language provides words for such more general concepts, like 'etwas' (something), denoting all objects except for human beings, or 'dokumente (documents)' for all types of letters, protocols, etc. The requirement is that the word behaves like all the classes (and with it words) that it subsumes.

This unrestricted inheritance has also disadvantages, such as in cases like e.g. the relation between 'document' and 'contract'. One would clearly say that 'contract' is a subclass of 'document', yet it should actually not inherit the property of 'contract' (not in the partial network in fig 1) to be specifiable by 'individuum' as the contractors, expressed by 'contract' between X and Y or 'contract with X', speaking of a * 'document with X ' is somewhat awkward. This suggests to be able to distinguish between inheritable and non-inheritable properties of concepts, which is not possible yet.
This allows a rather general description of, for instance, the constraints of verb frames, since a case relation description of the desired and semantically permissible transitions rules for semantically correct relations differ with the sentence surface structure in which the two words in question appear, like verb + noun, noun + specification (adjectival or prepositional), noun + noun + noun-genitive etc. Take for example a transition rule for two nodes, a noun and a verb. A semantic relation 'semfit' between any verb 'verbterm' and any noun 'nounterm' is assumed to be correct, if there is a link 'case-relation' between 'verbterm' and a concept, whose instance or subclass-instance is 'nounterm', or if 'nounterm' is in a role of relation to this concept or to any of its subclasses. Since the network is implemented in Prolog this can easily be stated as Prolog rules:

```
semfit( nounterm, verbterm):= 
  case-relation( _n-concept, verbterm, _a, b, .......),
  is-a( _n-lower, _n-concept),
  role-of( _n-lower),
  value-of( _n-lower),
```

When checking a phrase structure in the syntactic grammar these rules are applied to the relevant terms in the phrase by executing a test operation. For this purpose a test operator, like 'semfit' is introduced into the grammar, which takes the terms as its parameters. Depending on the syntactic categories of the words to be checked, different sets of transition rules are activated to prove compatibility of the terms. The test can in some cases, like the above mentioned verb-noun relations, also return the deepcase as a parameter, which then is stored as a grammatical feature and used to avoid multiple realisations of the same deepcase in a sentence.

`schrieb er an den ... (did he write to the ....)`

The possible semantic categories for 'er (he)' can be retrieved from the network by proper application of the transition rules, (i.e. regarding the pronoun as a 'variable' whose possible values have to be computed). In the above example we would retrieve the semantic class, that the concept-node of 'schreiben (write)' immediately points to its filler in subject position, which is 'individuum'. In our office domain, and probably elsewhere as well, we would only allow something of the class 'individuum' but not e.g. 'event'.

Providing these categories is especially helpful in a speech system where short words like pronouns turn up frequently in the hypotheses list. Deciding for each of them in the resolution module whether or not it can be bound to some antecedent and therefore inserted is very important in the antecedent's semantic category for the current phrase hypotheses during the parse will easily become too time consuming. Additionally, the cases in which pronoun-hypotheses can be rejected on these grounds may turn out to be too few to make it worth while spending so much effort at this stage.

Instead we go the opposite way. We assign the possible
semantic categories to a pronoun from its available sentential context. The possible categories for a pronoun are those which in turn would pass the semantic test for two content words as described above. However this cannot always be done so easily with composed proforms for prepositional phrases, like 'daran', which contain the preposition 'an' and a pronoun 'da(r)'. The special treatment of those will be discussed elsewhere.

In this way the class of the pronoun can be deduced from the network and be returned to the syntactic analysis as the hypothetical semantic class of the semantically unmarked pronoun in this particular context. The anaphor-resolution can use the instantiated semantic class(es) of pronouns, which have survived as part of correctly parsed sentence-hypotheses, in order to find correct antecedents within the range of these classes. This not only speeds up but also semantically constrains the search for possible antecedents.

The third purpose of the network applies when the resolution module tries to bind anaphoric expressions to their antecedents. In certain cases, where related terms like 'brief (letter)' and 'document (document)' are referentially bound, the resolution module needs to be provided with information on the semantic agreement between two terms with respect to synonymy or hypernymy and other relations like part/whole or property relations. The already mentioned hierarchy-relations and others like 'property' and 'role-of' in the semantic network can provide such information. For the correct analysis and the correct binding of meronymic relations an extension beyond the now existing 'has-property'-relation is certainly required. See also (FRE-GE 88, GEH 89). Similar to the semantic tests in the syntax several sets of transition rules for various types of relations are used to state admissible relations between anaphors and antecedents.

8.Conclusion

Preliminary test with the network covering the terms of our current applications have shown that in 95% the semantic compatibility as well as incompatibility is assigned correctly. This has been done by generating pairs of terms through application of the transition rules and checking their semantic validity. Only 5% were illformed, due to data-mistakes as well as overgeneralization in the network.

The general difficulty with the construction of a network for the above mentioned purposes is, as usual in speech understanding, that it not only has to depict all possible correct semantic relations but to block incorrect ones; this is our main purpose, else incorrect paths are further pursued in the parse and may prevail even over the correct ones. The strenght of the network in blocking the incorrect relations can easily and rightly be judged by the above mentioned generative method.

References:

[CHA-HER 84] R.Chaffin, D.J.Herrmann: The similarity and diversity of semantic Relations. in: Memory and Cognition, 1984, 12(2), 134-141


