

## DIVIDED AND VALENCY-ORIENTED PARSING IN SPEECH UNDERSTANDING

Gerh.Th.Niedermaier  
ZT ZTI INF

Siemens AG  
Otto-Hahn-Ring 6  
8000 München 83

### **Abstract**

A parsing scheme for spoken utterances is proposed that deviates from traditional 'one go' left to right sentence parsing in that it divides the parsing process first into two separate parallel processes. Verbal constituents and nominal phrases (including prepositional phrases) are treated separately and only brought together in an utterance parser. This allows especially the utterance parser to draw on valency information right from beginning when amalgamating the nominal constituents to the verbal core by means of binary sentence rules. The paper also discusses problems of representing the valency information in case-frames arising in a spoken language environment.

### **0. Setup**

In the framework of a speech understanding system SPICOS (Siemens IPO Philips Continuous Speech Recognition, Understanding and Dialog Project) which is supported by the German Federal Ministry of Technology and Research, new parsing strategies have been investigated. The whole system is designed as an interface for spoken language in German and Dutch to a relational database, that contains office information on the project itself like letters, publications, dates, and persons involved, etc. It should be able to answer all kind of questions and imperatives concerning the subject matter. The vocabulary comprises about 1000 word-forms.

### **1. Goals and Problems**

Whether or not one argues in favour of an interface between the acoustic and linguistic modules that allows for passing information in both directions is to be kept separate of the discussion on what kind of knowledge is available to the linguistic analysis. Only if it is able to reduce the number of possible sentences to a considerable extent it makes sense to organize this knowledge most effectively. In order to reduce the flood of hypotheses one has to make sure that the linguistic module uses as much knowledge as can be made available at the most early time of processing. Whether this knowledge is then used to make predictions for the acoustic modules or if the whole process works sequentially is rather a question of efficiency than of principle. (See Briscoe (1984))

### **1.1. Interface and System Architecture**

To study the effects of different techniques independent of each other we have for the first version decided on a simple sequential interface. The acoustic module delivers a list of word hypotheses with begin, end, and score. But not every word, that starts physically at the same point of time where the previous hypothesis ends, is a possible successor. One can limit the number of successors to those which are phonetically justified. The consequences of this are that the interface is a network of nodes and edges, where the nodes represent the words, rather than a list with beginnings and ends, which a chart-parser normally expects.

The interface also contains scores of the different

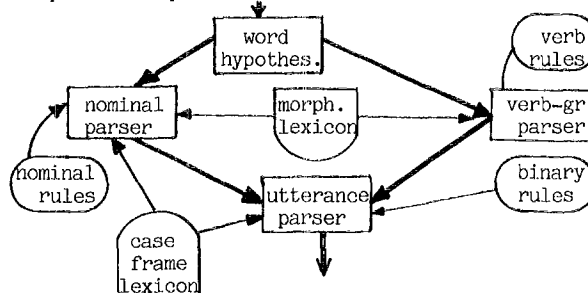
word-hypotheses. Nevertheless we do not use them yet, first because the analysis up to now works exhaustively and non-deterministic. This allows us to see how and where linguistic knowledge can be brought to bear most effectively. (See also Thompson (1984), who also argues for keeping the sources separate during the 'try-out-phase'). Weighing different syntactic structures implies that they have a weight multiplying factor which is inherent to them. (See Woods (1982)). Yet there is no agreement as to how to add up the scores in syntactic analysis. We believe that there is no general procedure. It is highly dependent on the domain and the influence of the domain on the syntactic structures.

### **1.2. Flow of Analysis Components**

Knowledge of relations between objects or objects and processes can be expressed in terms of caseframes. Our parsing strategy is mainly guided by the hypothesis that one of the major sources of restrictive power on the sentence level is to be found through caseframe restrictions. In order to take care of the restrictions that are carried out by the verb simple left to right parsers seem rather inappropriate. The caseframe restrictions can only be applied when the respective verb is encountered, which in German, badly enough, is mostly at the end of a sentence. The nominal and prepositional phrases are then grouped around the verb (see also M. Johnson, arguing that way in a DCG frame-work). To cope with the huge number of hypotheses an attempt is made here to further cut them down through generative power in caseframes and necessarily early verb-recognition.

### **2. Divided Parsing**

This has led us to a parsing strategy, that first splits up the parsing of the word-hypotheses into two different channels. One is the Nominal-Parser that takes care of all terminal elements that belong to a nominal group. The other part is the verb-group parser that is initialised with all verbal categories. They could work in parallel. They are brought together again in the utterance parser, that deals with one verb-hypothesis at a time. This enables us to bring to bear the caseframe restrictions of that particular verb at this early point. One verb-hypothesis is done after the other. Since the type of rules is also different in both cases the parsers can be tuned to the respective requirements.



### 3. The Nominal and the Verb Parser

The nominal-parser is in essence a chart parser (see Winograd 1983), working with augmented context-free grammar rules. It also triggers actions to pop features up to the dominating node. Prepositional and nominal phrases specifying NP's get also attached here according to the caseframe information of the head of the NP-constituent. This is only true for immediately neighbouring PP's. Focused and therefore moved PP's have to be treated differently. The parsing of the verbal groups, whose parts may be scattered all over the sentence like:

' Wer hat am 17. Mai einen Brief geschrieben'  
(who has written a letter on the 17th of May)

is carried out by a modified chart-parser, which is also able to take care of discontinuous elements in the grammar, like:

VG ← VRB (finite part) +: VNF (non-finite part)

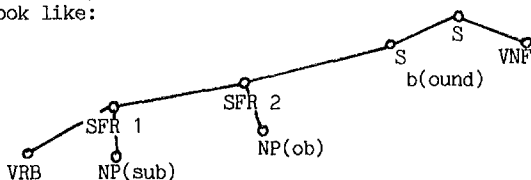
where ++ indicates that the next constituent is somewhere to the right. The output of this parser is a complete list of possible verb groups. Which caseframe they point to is a feature accumulated during the parse. In the case of verbal adjuncts the feature is a result of both components.

### 4. The Utterance Parser

This again is a chart parser, which for one go is initialized with the NP's, PP's, and AP's as terminal categories and the parts constituting the first verb-group hypothesis. It selects the right constituents according to the information given in the caseframe of the current verb-hypothesis. Since our semantic representation is a kind of predicate calculus formalism (see Bunt 1985 and v. Deemter 1985) at this level almost every constituent (except for focused PP's) can become arguments of the predicate. For this reason there is no point whatsoever in generating nodes that combine constituents into any other than S(sentence)-nodes, as it is quite common in traditional grammars. It does not contribute any additional meaning to a syntax-tree that has to be transformed into a predicate-argument structure. The only purpose is for us to restrict the linear precedence. (They serve very much the same purpose as the LP rules in the GPSG formalism (Gazdar 1985). For linear precedence in German see also Russell(1985).

Rules like these would yield very flat trees and lead to quite a number of rules. In order to avoid that we have set up a set of binary rules that is a lot smaller. The deepest level nodes always take a verb and one of the surrounding constituents. They create an artificial node that in turn can take another constituent and build a new artificial node.

The binary dependency trees generated by these rules look like:



That this grammar also demands rules of the kind:

S ← S + VNF  
b

may be surprising at first sight. But taking into consideration that no additional information is conveyed by the nodes higher up in the hierarchy this does not seem so bad any longer. These rules can be indexed according to how many NP-arguments they contain. This is a lexical feature of a verb. Only those rules will be invoked whose index does not exceed the maximum number of NP-valencies of the verb.

### 5. The Valency Lexicon

When adding new constituents to a node, their features and, if necessary, also strings are tested. Although the algorithm is not based on pattern-matching like some other frame based approaches (Hayes 1981 and Hayes 1985), the entries in the case-frame lexicon sometimes do have to come very close to it, in order to be of a power that not only describes case indicators and fillers that may occur, but at the same time excludes the wrong ones; a feature that is generally referred to as strong generative power.

As one can see from the rules, each constituent carries an index, that is passed on to the test procedure as one of its parameters. It tells which function this constituent has in the surface structure. This function is also an entry in the caseframes since some case-roles behave differently depending on the function they have to fulfill in the surface structure. It is at the same time a means to restrict the ordering of the constituents on the surface, which even in German is not as liberal as to make this kind of information redundant. (See Russell 1985) Case-roles look of the following kind:

name: TIME-POINT						
func-tion	prep	case	role-attribute	filler-descript.	filler-value	slot
PNM	vom	DAT		+intervall		1
PNM	vom	DAT			month-value	2
PNM	von	DAT			year-value	3
PNM	mit	DAT	datum			4
PNM	aus	DAT		jahr		5
...	..	..				

POB = prepositional object

PNM = prepositional phrase as noun modification

The test procedure checks whether a certain slot can be realised according to the feature parameters of the constituent. If the test is successful, it returns a number which indicates the number of the case-role, that the slot belonged to. In order to prevent doubling of case-roles in a sentence, which with this kind of input can easily happen, it is checked whether this case-role is not yet a member of a set of case-roles, already accumulated. This set is kept as a feature of the nodes in the binary tree. If not, it is made a new member of this set and passed on to the next level node.

Each caseframe comprises a selection of case-roles. There are frames for verbs as well as for nouns.

The noun frames become crucial when attaching the proper PP's to NP's. The caseframes are, unlike in other systems ( Brietzmann 1984, Hayes 1985), static data structures that are not instantiated, nor do they trigger any actions. Since it is very unclear as to the criterion of whether roles, especially prepositional objects and the like are obligatory, no distinction is made in the caseframes. The fact that certain NP's are obligatory for a verb is taken care of by the argument-number of the verb. There is also no distinction made between immediately verb dependent and free prepositional complements. First one does not really know where to draw the borderline between the two, (see Vater 1978, Jacobs 1985) and second, from the point of view of semantic interpretation, they all have to be treated in the same way, namely as arguments of the verb.

Because of the requested power of the caseframes the information given in the slots has to be as general as possible but also as explicit as necessary to prevent the attachment of those hypotheses which are produced but we rather would not want to fit. In fact each slot represents a combination of casemarkers and categories or particular strings of casefillers. The latter are usually to be found as head of the noun phrase.

Taking a closer look at the heads of the phrases one will notice that they can play very different roles with respect to their function in realising a certain case-role. They are either explicit designators of the role in cases like:

'Brief mit Datum 17.1.86'  
 ("letter with date ....")

where 'mit' works as an empty rolemarker, i.e it is not role-specific, and the head 'datum' takes the role of the preposition. Those strings we call role-attributes. They can also be descriptions of the value, that is meant to fill the slot. On the syntactic surface however they appear very much the same, namely as :

Prep + Nom + Propername

like in: 'von Monat Mai' (from (the month of) May)

In these cases one can use the semantic categories of the heads in order to identify them as proper fillers.

There are also cases where there are specific value descriptions, whose status is usually somewhat inbetween the two mentioned above. They demand a particular preposition, like 'jahr' in the slots above. For example using 'Jahr' as a value description demands 'aus' as a preposition, which in turn cannot be used with a possible value of 'Jahr', which would for instance be '1985'. If the head of the constituent is only a value like in:

'der Brief von 1985'

The semantic category of this value has to appear in the slot restrictions. That this is heavily application dependent is clear enough. This kind of information can be kept in a separate network. Leaving it to the database to decide whether a value is appropriate means that the database can never decide whether this value has not been stored or whether it looked for a non-existent relation. This again clearly influences the possible response of the system. With this information stored in the case-frames the system can answer maybe that it has no letters from this year, whereas in the

alternative case it could just answer 'no result retrieved'.

They also behave differently in terms of case-role restrictions, as in the above example you could say: 'aus dem jahr 1985' (from the year 1985) but not: \* 'aus 1985' (from 1985) Therefore we have decided to demand semantic categories in the caseframes for values too.

## 7. Conclusion

We have introduced a parsing strategy that heavily relies on case-frame and therefore also on semantic labelling information. In order to detect the verbs, that set up the appropriate case-frames has caused us to split the parsing process first into two parallel processes. One parses the nominals and the prepositional phrases, the other one the verb groups. The two processes are brought together and a sentence-parse is tried on the basis of the hypothesised verb-frame. The parsers work with augmented context-free grammars, that also percolate features to the higher nodes. The nodes do not have to convey any additional information. The also trigger tests to check case-role restrictions.

## Literature

- Barr A., Feigenbaum E.A.:** Understanding Spoken Language. in: The Handbook of Artificial Intelligence. Vol I, London 1981
- Brietzmann A.:** Semantische und Pragmatische Analyse im Erlanger Spracherkennungsprojekt. in: Arbeitsberichte des Instituts f. Mathem. Maschinen und Datenverarbeitung Bd.17 Nr.5, Erlangen 1984
- Briscoe, Boguraev:** Control Structures and Theories of Interaction in Speech Understanding Systems. Proceedings of the Coling 1984
- Bunt H.:** Mass Nouns and Model Theoretic Semantics. 1985
- v.Deemter K.:** The Logical Languages of TENDUM and SPICOS. 1985
- Gazdar G., Klein E., Pullum G., Sag I.:** Generalized Phrase Structure Grammar. Cambridge, Mass. 1985
- Hayes P.:** Semantic Caseframe Parsing and Syntactic Generality. in: Proceedings of the ACL 1985
- Hayes, Carbonell:** Multi Strategy Parsing. 1981
- Jacobs J.:** Thesen zur Valenz. Unpubl. MS, 1985
- Lea W.(ed):** Trends in Speech Recognition. N.York 1980
- Pereira F.:** A New Characterisation of Attachment Preferences. in: Dowty, Karttunen, Zwicky (ed): Nat. Language Processing: Psycholinguistic, Computational and Theoretical Perspectives, Cambridge 1984
- Proudian D., Pollard G.:** Parsing Head Driven Phrase Structure Grammar. in: Proc. of the ACL 1985
- Russell G.:** A GPS-Grammar for German Word Order. in: Klemk U.(ed):, Kontextfreie Syntaxen. Tübingen 1985, p.19-32
- Thompson H.:** Speech Transcription: An Incremental Interactive Approach. in: Proc. of the ECAI 1984
- Vater H.:** Probleme der Verbvalenz. in: KLAG 1, 1978
- Winograd T.:** Language as a Cognitive Process. Vol I, 1983
- Woods W.A.:** Optimal Search Strategies for Speech Understanding Control. in: Artificial Intelligence 18, 1982, p.295-326