

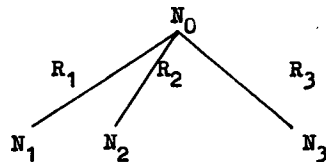
BUNDLES AND SYNTACTIC ANALYSIS

Dorothee Reimann

Zentralinstitut für Sprachwissenschaft AdW DDR, Berlin

Syntactic analysis - in our terms - means to construct a description of the syntactic structures of a given sentence, starting from a formal description of the syntactic features of the wordforms (the result of a preceding wordform analysis). The description of the syntactic structure as well as the formal description of the wordforms have to follow a certain model, in our case the Dependency Grammar by KUNZE (1975).

A central point of the model - in respect to analysis - are the so-called bundles. They represent the necessary syntactic knowledge of a language which is the base for analysing sentences of this language. The bundles are parts of trees in a very generalized form. If



is a bundle, then it means, that there are conditions for the nodes N_0 , N_1 , N_2 , N_3 , that there are relations between the nodes expressed by the edges and their markings R_1 , R_2 , R_3 and that there is some order condition for the nodes, in the example N_1 N_2 N_0 N_3 . Following the model, these bundles give us a high measure of surety for the analysis.

The bundles are stored in the system in a very concentrated form. For example for a noun group with a noun like "table" at the top there is only one bundle with all possible facultative nodes. So we have in a bundle informations about the mode of occurrence of a node, i.e. if a node is obligatory, facultative etc.

The bundles serve also as description for the dominance behaviour of wordforms (one part of the formal description of their features), thus we have a strong connection between a bundle and a wordform which can "stand" at the top of the bundle. For this purpose the bundles have a special external shape for linguistic work. For instance, a finite verb form of the German verb "schreiben" has several bundles in its dominance component, one of these is

$$\text{FIN} (\text{ACT}^{\circ} \wedge \text{OBJ3}^{\text{f}} \wedge (\text{PAT} \triangle \text{DAS})^{\circ})$$

where

FIN	- means, that the top is a finite verb form
ACT ^o	- there <u>must</u> be an actor (subject)
OBJ3 ^f	- there <u>can</u> be an indirect object
(PAT \triangle DAS) ^o	- there <u>must</u> be a direct object, either in the form of a noun group in the accusative case (PAT) or as a sub- ordinate clause with the conjunction "dass" (DAS)

Free modifications like local and temporal circumstances are possible as well.

Of course, in the syntactic analysis system the bundles have another shape and they contain more information.

There are two possibilities to use the bundles for syntactic analysis, the first one is to use them as proper rules, the second one is to use other methods, but to have the bundles as syntactic knowledge in the background. The first possibility is outlined in my dissertation (AROLD 1976) as a strong

bottom-up and multiple-path analysis. But for a working system we intend to follow the second way.

In the paper a short summary of the system will be given. The syntactic analysis of a sentence given by the formal description of the syntactic features of the wordforms is made by three steps. These three steps differ one from the other by different methods.

In the first step an "augmented transition network" (ATN) is used to construct a rough structure of the sentence. The second step uses the bundles which are members of the dominance component of the wordforms to find out so-called subordination possibilities. The result of the second step is a loop-free graph, which is transformed during the third step to a graph containing the dependency trees which describe the syntactic structures of the analysed sentence. In the third step the conditions of the bundles (obligatory and facultative nodes etc.) are used as well as the conditions coming from the shape of the resulting graph.

Up to now we are only at the beginning of the implementation of the system. Some remarks about these problems as well as about the difficulties to include the analysis of coordinated sentences will be made in the paper too.