SYNTAX, SEMANTICS AND PRAGMATICS IN CONCERT: AN INCREMENTAL, MULTILEVEL APPROACH IN RECONSTRUCTING TASK-ORIENTED DIALOGUES

Manfred Gehrke

Project "Prozedurale Dialogmodelle"*
Department of Linguistics and Literature
University of Bielefeld, Bielefeld, FRG

ABSTRACT

This paper gives an overview of a model for the reconstruction of task-oriented dialogues based on an interactive, multilevel parsing formalism. It is applied to route description dialogues. It will be shown, how the pragmatic aspects of such dialogues *are* taken into account on different levels of processing. The approach described is based on an extension of the concept of cascaded ATNs. Furthermore this approach uses knowledge sources (KSs) for ev*ery* participant in the dialogue in which knowledge about the world and a partner model is build up during the analysis of a dialogue. These KSs are supplied to the parsing process, as well. In this paper special importance is laid on the description of the interaction and cooperation of the different processing components of this formalism.

INTRODUCTION

In the project "Prozedurale Dialogmodelle" we have concerned ourselves with solving the following problems by reconstructing task-oriented dialogues:

1. What dependencies exist between the different linguistic processing components (LPCs) in realistic task-oriented dialogues** of the special type 'route description'?
2. Which knowledge sources (KSs) participate in such dialogues?
3. What informations do these KSs contain?
4. At which point and with which information do the LPCs interact?

Furthermore, we assume that the cooperation of the various LPCs starts at a very early stage in processing an utterance in a dialogue. In our model the LPCs are organized as coroutines, and not, as suggested by the results of Marslen-Wilson and Tyler [6] , as parallel processes. So that, e. g. as soon as a partial syntactic description of an utterance is derived, the semantic components are triggered.

The linguistic basis of our approach is a scheme of dialogue interaction in a task domain[7], in which the purpose of an utterance in the dialogue context is determined with respect not only to the

maintainance of the interaction.

THE MULTILEVEL PARSING FORMALISM

The reconstruction of dialogues is carried out by a multilevel parsing formalism [2, 3] in which the different components are independent, but closely cooperating. Under this viewpoint "parsing" is taken to mean 'application of recognition procedures' and not 'recognition of syntactic sentence structures' .

The architecture of the formalism is based on the concept of cascaded ATN (CATN) [10] with the following important differences:

(1) the different levels of CATN are not necessarily ATNs, and
(2) levels of an CATN can be 'by-passed'.

In addition to the usual means of communication in a CATN such as TRANSMIT there is a 'blackboard' (the common KS) in order to store the results of the various cascade levels. While the RUS Parser [l] is a 2-stage cascade, our formalism allows an arbitrary number of stages, in this application 6.

The syntactic component (SYNC) is an ATN. Due to our dialogues, which were uttered in a rather colloquial language and therefore overwhelmed with incomplete utterances, this component must be v*ery* robust. Therefore, a complete syntactic analysis is of minor importance. This is, e.g., obtained by enabling SYNC to reach a permissible final state, if its analysis of an utterance is not rejected by the semantic or pragmatic components. So the main purpose of SYNC is to detect phrases in a utterance (noun-, prepositional-phrase, adverb-preposition-groups, etc.) and to make suggestions on their role in the utterance. Additionally it should detect wh- and imperative structures. As soon as a phrase or even a part of it is analysed, the semantic component is triggered in order to see how it can process the results of SYNC. This includes an early detection of wrong pathes of analysis.

The semantic component (SEMC) is a case oriented production rule system. Because of the incremental manner of parsing, the semantic interpretation has to be tentative. The determination of a syntactic construct is - until the main verb is detected - more a suggestion rather than a categorization. Thus the case slots are defined in two ways. Firstly, there are very general definitions of case slots applying before tne main verb is detected which build up a proposed interpretation of the utterance.

Secondly, more specific definitions of these slots are connected to the respective verb frame, thus meeting its special demands. When the verb is transmitted, the hitherto constituted case slots are tested and the definitions of the required case slots, which have not yet been constituted, are supplemented by the special requirements of the verb. Every constituted case slot is transmitted to the task-communication component.

The remaining levels of the cascade are responsible for the interpretation of the dialogue from pragmatic points of view. For that purpose they are supported by KSs that contain knowledge of the world and a partner model (WPKSs). The information, gathered by the initiator of the direction-giving dialogue (addresser) during the dialogue is stored in his/her WPKS as well as the assumptions on how the dialogue may proceed. The content of the addresser's WPKS could be stated after a first piece of route description, such as

 Problem: addresser wants to get to the opera
 Addresser wants to know: solution of the problem
Addresser believes: Addressee knows the solution
Addresser believes: Addressee will tell him/her the
                    solution
Addressee says: go straight ahead to the church
Addressee says: arrive at Rathenauplatz

On the other hand, the addressee's knowledge about the course of dialogue is stored in his/her WPKS. In addition, each partner should have a problem solver which takes into account the appropriate KS for finding the way asked *for.* But this is beyond the scope of our project.

The Task-Communication-Component (TCC) contains the task dependent, pragmatic categories. It is responsible for the detection of utterances bearing solutions to subproblems, while a general solution is the purpose of the dialogue. In the case of direction-giving dialogues the following task dependent categories are used:

(1) route description, i.e. utterances containing a piece of the proposed route, e. g. "You go straight ahead on that street."
(2) place descriptions, i.e. utterances bearing some information about where a change of direction will take place, or confirming that one is on the right path, e. g. "Then there is Woolworth's to the left."
(3) partial goal determination, i.e. when an intermediate goal is reached and where a change of direction often takes place, e. g. "Turn left at Wollworth's."
(4) goal declaration, i. e. that point in which the reaching of the goal is stated, e. g. "Then you arrive at the station."
(5) destination specification, i. e. where the addresser utters the goal he will reach, e. g. "We are looking for the bus station.".

An Alteration of this type of task-oriented dialogues only requires a change in this component at the pragmatic level.

The Question-Answer-Interaction-Component (QAIC)

has to detect the patterns of speech acts used to control the cooperative organisation of mutual understanding. These are, e.g., the exchange of question, answer, assurance, confirmation, etc. The Task-Interaction-Component (TIC) provides a general scheme of task-oriented dialogues with categories like ask-for-information, information-giving (e.g. "There you'd better ask someone else."). The final component is responsible for starting, keeping and ending the social contact (SCC).

All the components on the pragmatic level are ATNs, because task-oriented dialogues consist of rather fixed patterns of pragmatic categories. While it is true that a dialogue may be interrupted by a clarification dialogue, such a clarification has a fixed order, as well. Thus it is only necessary to detect the beginning of an embedded dialogue. Markers for clarification *are* wh-phrases inside the dialogue such as "Where?" or phrases like
"I don't get it! Would you please repeat that?"

INTERACTIONS WITH THE SYNTACTIC COMPONENT

The interaction of an ATN with its neighbouring components is carried out by the TRANSMIT-action. After traversing an arc with this action, processing in the ATN is interrupted and control is passed over to the next component. Such TRANSMITS occur in SYNC for two reasons:

1. at critical points in the analysis where a wrong decision can be rather resource consuming, e.g. when the head noun of a nominal phrase or a preposition after a nominal phrase is detected,
2. at points where important results for SEMC are detected such as verbs, deictic or directional adverbs.

SYNC can receive two types of message from SEMC. On the other hand, if SEMC lacks further information to process, then SYNC starts its analysis at the goal node of that arc on which the TRANSMIT occured. If, on the other hand, something is wrong with the analysis of SYNC, then backtracking is forced.

INTERACTIONS WITH THE SEMANTIC COMPONENT

SEMC can receive various types of message from SYNC. If a full verb is recognized, then it becomes head of the verb frame; previously filled case slots are matched against the specific requirements on case slots of that verb. In the case that further frames are possible for one verb these frames have to be instatiated and processed in a similar way.

All other syntactic constructs (nominal-, prepositional phrases, embedded sentences) are provided with case slots or categorized as modifiers of a verb, e.g. adverbs, auxiliary verbs, to mark tense, modality, etc.

When the end of an utterance is indicated, then the frame is checked for completion (i.e. a verb is recognized and all obligatory case slots are filled), otherwise possibly additional instantiated verb frames are not completed, it is assumed that the

utterance is some sort of cotextual ellipsis [4] and an inference mechanism (see below) based on the pragmatic assumptions is initiated.

Finally SEMC can be "by-passed", when e.g. an interjection such as "yes, uhh" is detected by SYNC. An application of SEMC to such words doesn't contribute to the whole understanding process. Therefore they are immediately routed to QAIC, in which they may be categorized as an assurance, confirmation, etc.

TCC may signal to SEMC that its result is not correct. This could be caused by a wrong choice of the appropiate frame, e.g. a completed frame was sent as a interpretation, while an uncompleted frame held true for an elliptical utterance. Then the latter frame is transmitted instead, otherwise SYNC has to resume its analysis.

In cases where no verb frame can be completed, results are transmitted, too, but marked as uncomplete. Then several cases are to distinguish:

1. To disambiguate elliptical utterances we had to consider that the missing phrase X is somehow described by the unsatisfied case slot and that X was in some way uttered earlier. Then the WFKS of the addressed participant is looked up, assuming that there is a referent for X. For an illustration, consider the following part of a dialogue:
"da kommen sie auf den Rathenauplatz, den Liberqueren sie auch."("then you come to the Rathenauplatz, you cross it, too.")

The second utterance has two alternative analysis. Either it is a relative clause with an incorrect word order. Or it is an independent utterance where the noun referring to the definite article "den" is missing. The frame of "liberqueren" requires as an object case slot a physical object that could be crossed. Secondly, the addresser has the strong assumption that the addressee will somehow conclude the route description. Despite the missing phrase the utterance strengthens this assumption. Thus the addresser's WFKS is looked up in reverse order (see above) for a route or place description containing a referent to "den". Here it will conclude in finding "Rathenauplatz" as referent.

2. For deictical particles referents are similarly found, because they are often used to denote the last partial goal. A subgoal can be expressed as a named location in the GOAL case slot in the last route description, or as a location in the PLACE case slot in the last place description.

These particles (esp. "there") may have the special meaning of denoting the goal, as in "and then you are there."

3. If a verb is missing, then it is first assumed, that the utterance is a route description. Therefore, this case frame is tested with verbs of the field "MOVE", otherwise the test is continued with verbs referring to location.

FURTHER WORK

Hitherto SYNC and SEMC are implemented together with the procedures to build up the KSs as well as the procedures for the interaction between the components. At present we begin to incorporate the pragmatic components, so we do not have confirmed results on the interaction between them and SYNC and SEMC.

The parsing formalism runs on a TR-440 at the University of Bielefeld. It is written in FLAVORS [8], an object-oriented language, embedded in MACLISP.

REFERENCES

[1] R. J. Bobrow/B. Webber, "Knowledge representation for syntactic/semantic processing." Proc. AAAI, Stanford, 1980.

[2] T. Christaller/D. Metzing, "Parsing Interaction: a multilevel parser formalism based on cascaded ATNs." in: K. Sparck-Jones/Y. Wilks, Automatic Natural Language Parsing. Chichester, 1983.

[3] M. Gehrke, "Rekonstruktion aufgabenorientierter Dialoge mit einem mehrstufigen Parsing-Algorithmus auf der Grundlage kaskadierter ATNs." inf9j.

[4] M. HuBmann, "Ellipsenrekonstruktion im Kontext; Forderungen an eine Ellipsenkomponente und ihre Realisierung in SWYSS." in[9].

[5] W. Klein, "Wegausklinfte." Zeitschrift fur Linguistik und Literaturwissenschaft, 9: 9 -57 (1575).

[6] W. Marslen-Wilson/L. K. Tyler, "The termporal structure of spoken language understanding." cognition, 8: 1 - 72 (1980).

[7] D. Metzing, "Dialogue interactions in a Task Domain: Model and some Problems of Application," in: Hein/Hagglund (ed.), Models of Dialogue. Munchen, 1982.

[8] F. di Primio/T. Christaller, "A poor man's flavor system." (to appear), ISSCO, Geneva, 1983.

[9] W. Wahlster (ed.), "Proc. of 6th German Workshop on AI", Berlin-Heidelberg-New York, 1982.

[10] W. A. Woods, "Cascaded ATN Grammars." Journal of ACL, 6:1 (1980) 1 - 13.