gefördert vom Bundesministerium für Forschung und Technologie

Verbmobil
Verbundvorhaben

Semantic and Dialogue Processing in the VERBMOBIL Spoken Dialogue Translation System

Elisabeth Maier
Scott McGlashan

DFKI GmbH & FB Computerlinguistik,
Universität des Saarlandes

Report 51
September 1994

September 1994

Elisabeth Maier
Scott McGlashan

DFKI
Stuhlsatzenhausweg 3
66123 Saarbrücken

Tel.: (0681) 302 - 5347
e-mail: {maier}@dfki.uni-sb.de

Das diesem Bericht zugrundeliegende Forschungsvorhaben wurde mit Mitteln des Bundesministers für Forschung und Technologie unter dem Förderkennzeichen 01 IV 101 K/1 und 01 IV 101 R gefördert. Die Verantwortung für den Inhalt dieser Arbeit liegt bei den Autoren.

Semantic and Dialogue Processing in the VERBMOBIL Spoken Dialogue Translation System*

Elisabeth Maier
DFKI GmbH
Stuhlsatzenhausweg 3
D-66123 Saarbrücken, Germany
E-mail: maier@dfki.uni-sb.de

Scott McGlashan
Department of Computational Linguistics
University of Saarbrücken
D-66041 Saarbrücken, Germany
E-mail: Scott.McGlashan@coli.uni-sb.de

Abstract

In this paper we sketch the role of semantic and dialogue processing in the VERBMOBIL spoken dialogue translation system. A full demonstration system is currently under development although a 'mini' demonstration system has been implemented.

1 Introduction

The VERBMOBIL project combines speech technology with machine translation techniques in order to develop a system for translation in face-to-face dialogues. The VERBMOBIL prototype will provide translation for negotiation of business appointments between German and Japanese users who have only a passive knowledge of English. The major requirement is to provide translation as and when users need it, and do so in real-time. In order to meet this requirement, the system is composed of time-limited processing

*This paper also appeared in Heinrich Niemann, Renato de Mori und Gerhard Hanrieder (eds.), *Progress and Prospects of Speech Research and Technology, CRIM/FORWISS Workshop*, p. 270–273, Infix Verlag, München, September 1994.

modules which perform acoustic, syntactic, semantic and dialogue processing. A transfer module exploits semantic and pragmatic information in order to map between source and target languages.

2 Semantic Interpretation

The semantic component of the VERBMOBIL system performs two functions. First, it constructs a context-independent semantic representation from an HPSG-style syntactic representation of user input ([PS94]). The semantic representation is based upon Discourse Representation Structures (DRSs), but augmented with thematic role and semantic sort information, as well as links to domain-oriented concepts ([BMM+94]). In many cases, these representations abstract away from language-specific semantics, so that the same DRS underlies expressions in different languages; for example, *Montag ist mir lieber* and *I prefer Monday* have the same DRS representation. In addition, underspecification is used as a technique for dealing with anaphora and ellipsis, phenomena frequently occurring in appointment dialogues; for example,

(1) [[*die erste Hälfte*][*das ist schlecht*]]
(*the first half. that is bad*)

where *die erste Hälfte* refers back to a proposed month for the appointment (the first half of November), *das* refers not simply to this date, but to the proposition that the participants meet on this date. The representations for these expressions are explicitly marked as anaphoric and the conditions on their antecedents are stated.

The second function is to further specify the semantic representation for transfer. This is achieved by evaluating DRS representations within a conceptual model of the dialogue, implemented in the BACK system ([HKQ+93]).

With elements identified as anaphoric, the conceptual model is searched for appropriate antecedents and the representation is updated. One advantage of this approach is in the treatment of some types of fragmentary input. When an utterance is parsed as a sequence of syntactic phrases, represented in (1) with bracketing, a unitary semantic representation can be constructed as a result of anaphora and ellipsis resolution. This treatment, in turn, contributes to a 'reduction-oriented' translation strategy, where the translator reduces the linguistic content while preserving (the essence of) its semantic and pragmatic content. Thus the utterance in (1) can be translated as *the first half is bad.*

Evaluation within a conceptual model also contributes to translation by further specifying the meaning of ambiguous and polysemous expressions whose senses correspond to the different expressions in the target language. In our domain, two classes can be distinguished depending upon whether local or global contextual information is required. Evaluation based on local context is required for verbs and prepositions where the appropriate sense is correlated with the semantic sort of its argument(s). For example, some senses of the preposition *nach* can be translated into English as *according to*, *to* or *after* depending on whether its argument is a thing, location or temporal, respectively. Other cases of disambiguation, however, are more complex requiring inference over non-local context:

domain knowledge *bei mir*, for example, can be translated as *at my place* or *for me* depending whether a date or a location is being negotiated.

dialogue state *Yoroshiku onegaishi masu* would be translated into English as *thank you* in a confirmation phase of the dialogue, but as *good bye* in a closing phase.

3 The Dialogue Component

The dialogue module has the task to increase efficiency and robustness of the VERBMOBIL system and to provide information to support the translation process. In order to do so, it follows the interaction so that it can supply other VERBMOBIL components with information about the current dialogue state and about the interaction pursued so far. Deep processing of dialogue contributions in VERBMOBIL only takes place when the user requests a translation. In cases where the dialogue participants speak English and therefore no translation is necessary the dialogue is monitored only by means of a key word spotter which only reacts when items of a predefined set occur in the input utterance. Finally, the dialogue module carries out clarification dialogues in cases where information necessary for correct processing is incomplete or missing.

To meet these requirements a hybrid architecture has been proposed (see [AMR94]). It is composed of three submodules which differ with respect to processing speed and linguistic sophistication: a statistical module, a finite

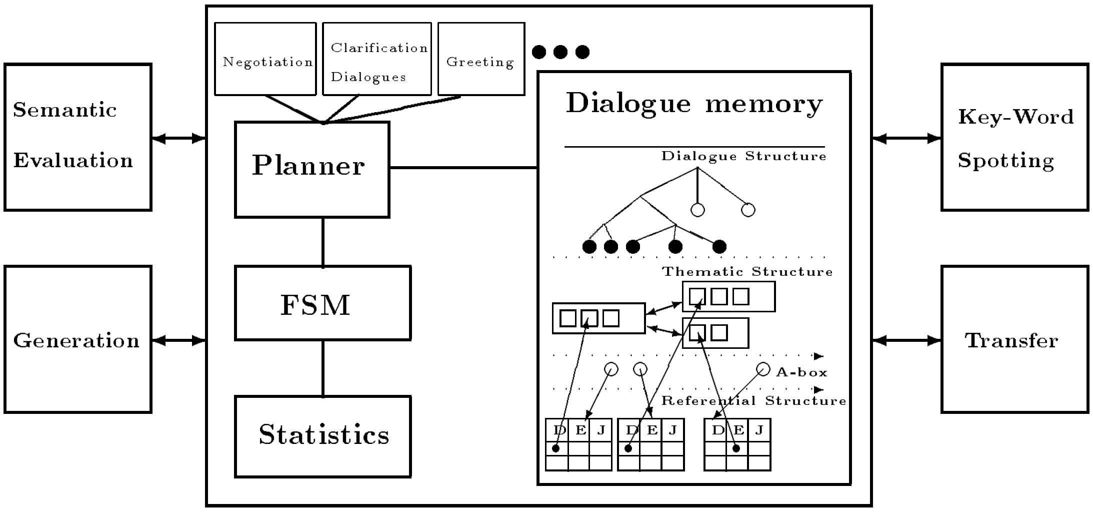

Figure 1: Architecture of the dialogue component in VERBMOBIL

state machine and a dialogue planner (see Figure 1). All three components are based on the assumption that interactions can be modeled by means of dialogue acts. The dialogue model therefore consists of a description of all potential sequences of dialogue acts. Similar models have been already successfully employed for integrated speech systems, like e.g. EVAR ([MKK+92]) and SUNDIAL ([Bil91]). In the following we briefly discuss the three subcomponents:

The information encoded in the **Statistical Module** has been derived from a corpus of appointment scheduling dialogues which have been annotated with dialogue acts. For every dialogue state the module can provide the possible follow-up acts together with their likelihood of occurrence (for similar approaches see [And92], [WN92] and [NM93]). The **Finite State Machine** (FSM) parses dialogue acts provided by the semantic component and checks them for compatibility with the underlying dialogue model. Also, the Finite State Machine can provide information which subsequent dialogue acts are possible according to the model. The most knowledge-intensive subcomponent is a hierarchical **Dialogue Planner** that also includes a representation of the dialogue model. Planning proceeds by means of so-called plan operators which are specialized for the treatment of specific dialogue phases like, for example, initialization, negotiation, clarification or repair. In addition to the Finite State Machine the Planner also takes contextual and pragmatic features into account when testing whether a new dialogue act is in line with the underlying model. The Dialogue Planner is also responsible for the incremental construction of the dialogue memory, which contains intentional, thematic and referential information concerning the interaction going on so far.

The results produced by the dialogue module are used by various VERBMOBIL components: the predictions determined by the Statistical Module and the Finite State Machine influence which language model will be used by the Speech Analysis Component or the Key Word Spotter when processing the next utterance. The dialogue memory and the context represented there is used by the semantic component to infer dialogue acts. Also, this information is exploited to support transfer and verb disambiguation.

4 Interaction between Semantics and Dialogue

The semantic and dialogue components interact by means of a sign-based representation where different levels of information about the input utterance are encapsulated in a single, structured representation. In addition to the levels of syntactic and semantic information, the sign includes a representation of the 'pragmatic' level. While the semantic component provides an underspecified assignment of the speech act — as derived, for example, from performative verbs and discourse cue phrase — the dialogue component checks the consistency of the assignment with the underlying dialogue model and adds more specific information to the pragmatic level.

5 Further Development

Integration of these components in the VERBMOBIL prototype is currently under way. A full demonstration system will be completed by Spring 1995, and a 'mini' system, with limited functionality and robustness, will be available by September 1994. By the time of the workshop, we can report further on the development of these systems as well as preliminary results concerning efficiency and robustness.

6 Acknowledgements

This work was partly funded by the German Ministry for Research and Technology (BMFT) under contracts 01 IV 101 K/1 and 01 IV 101 R. We would like to thank our colleagues in the VERBMOBIL project, especially those in the semantics, dialogue and translation workpackages. The authors are solely responsible for the views expressed in this paper.

References

[AMR94] Jan Alexandersson, Elisabeth Maier, and Norbert Reithinger. Robust and efficient dialogue processing using a 3-layered architecture. Technical report, DFKI GmbH, April 1994.

[And92] Francois Andry. Static and dynamic predictions: a method to improve speech understanding in cooperative dialogues. In *Proceedings of the International Conference on Spoken Language Processing*, pages 639–642, Banff, October 1992.

[Bil91] Eric Bilange. A task independent oral dialogue model. In *Proceedings of the Fifth Conference of the European Chapter of the Association for Computational Linguistics (EACL-91)*, pages 83–88, Berlin, Germany, April 1991.

[BMM+94] Johan Bos, Elsbeth Mastenbroek, Scott McGlashan, Sebastian Millies, and Manfred Pinkal. The Verbmobil semantic formalism. Technical report, Department of Computational Linguistics, University of Saarbrücken, February, 1994.

[HKQ+93] Th. Hoppe, C. Kindermann, J.J. Quantz, A. Schmiedel, and M. Fischer. BACK V5 tutorial and manual. KIT Report 100, Technical University of Berlin, March 1993.

[MKK+92] M. Mast, R. Kompe, F. Kummert, H. Niemann, and E. Nöth. The dialog module of the speech recognition and dialog system EVAR. In *Proceedings of the International Conference on Spoken Language Processing, Banff, Canada*, pages 1573–1576, 1992.

[NM93] Masaaki Nagata and Tsuyoshi Morimoto. An experimental statistical dialogue model to predict the Speech Act Type of the next utterance. In *Proceedings of the International Symposium on Spoken Dialogue*

(ISSD-93), pages 83–86, Waseda University, Tokyo, Japan, November 1993.

[PS94] C. Pollard and I. A. Sag. *Head-Driven Phrase Structure Grammar.* Standford: CSLI, 1994.

[WN92] Petra Witschel and Gerhard Th. Niedermair. Experiments in dialogue context dependent language modeling. In Günther Görz, editor, *Proceedings of KONVENS-92.* Springer Verlag, 1992.