Clarification Dialogues in VERBMOBIL

Elisabeth Maier *

DFKI GmbH
Stuhlsatzenhausweg 3
D-66123 Saarbrücken, Germany
E-Mail: maier@dfki.de

ABSTRACT

In this paper we describe the possibility to carry out clarification dialogues in the framework of the face-to-face translation system VERBMOBIL. We focus on a special subtype of clarification dialogues which occur when the system has insufficient information to continue processing. The clarification dialogues currently incorporated in our system concern three aspects: (i) phonological ambiguities, (ii) unknown words, and (iii) semantic inconsistencies. We describe each of these subdialogues in detail and discuss the extensions and changes that had to be made to the overall system in order to allow for clarification dialogues. An outlook on future developments concludes the paper.

1. Introduction

In this paper we present an approach for the treatment of clarification dialogues in the framework of the face-to-face translation system VERBMOBIL [2]. Currently, this system translates German and a small set of Japanese dialogue contributions into English. The task under consideration is appointment scheduling. In our system, two types of clarification dialogues occur:

- **human-human subdialogues** where a dialogue participant elicits unclear or missing information from his dialogue partner:

  User 1: *... aber Sie konnten erst ab achtundzwanzigsten, wenn ich 's recht erinnere.*
  ... but you only have time after the twentyeighth, if I remember correctly.
  User 2: *Ich könnte nur bis zum achtundzwanzigsten.*
  I only have time until the twentyeighth.

- **human-machine subdialogues** where the machine engages in a dialogue with the user to elicit information necessary for correct processing.

  User: *Ich würde Sonntag, den 19. April vorschlagen.*
  I would suggest Sunday, April 19.
  VM: *Die Angabe Sonntag, 19. April existiert nicht. Meinen Sie Samstag, 19. April ?*
  The date 'Sunday, April 19' does not exist
  Do you mean Saturday, April 19?
  User: *Ja.* Yes.

In contrast to other dialogue systems which actively engage in an interaction with the user the VERBMOBIL system plays a rather passive role, since it only mediates between dialogue participants, producing translations on demand. Only in some well-defined cases the VERBMOBIL system changes from a passive into a rather active mode. This happens, for example, in the case of clarification dialogues. These clarification dialogues are carried out by the dialogue component, which is described in [1].

2. Types of Clarification Dialogues

Wizard-of-Oz studies which accompanied the development of the VERBMOBIL system (see [5]) have shown that the frequent occurrence of system-initiated clarification dialogues disturbs the naturalness of the interaction and has a negative influence in the users' acceptance of such a system. Therefore, we restricted the possibility to carry out clarification dialogues to the implementation of only a few cases, additionally ensuring that the user is able to provide the necessary information; where the problems presented to the user require too much linguistic expertise we consider different recovery strategies (e.g. the use of defaults, the integration of statistical models which determine the most appropriate way to proceed). The following types of clarification dialogues are incorporated in our system:

1. dialogues about phonological similarities (`similar_words`[1]) which cope with possible confusions of words like *Juni* vs. *Juli* (engl.: June vs. July);

*This work was funded by the German Federal Ministry for Education, Science, Research and Technology (BMBF) in the framework of the VERBMOBIL Project under Grant 01IV101K/1. The responsibility for the contents of this study lies with the author. We would like to thank the members of the VERBMOBIL software integration team – in particular Thomas Bub, Andreas Klüter, Stefan Mertens, and Johannes Schwinn – for their help in the design of the component. We are also grateful to Jan Alexandersson for comments on a previous draft of this paper.

[1]Words printed in `teletype font` refer to system-internal messages.

2. dialogues about words unknown to the system (`unknown_words`);
3. dialogues about inconsistent or inexistent dates (`inconsistent_date`), e.g. *um 16 Uhr am Vormittag* (engl.: at 16 hours in the morning) or *am 30. Februar* (engl.: on February 30).

Since the complexity of our clarification dialogues is rather low and since no complex inference mechanisms are required our clarification dialogues are modeled as Augmented Transition Network (ATN).

The possibility to carry out such clarification dialogues can be selectively enabled or disabled in order to reduce the number of subdialogues occurring during the interaction.

In the remainder of this paper we describe the processing of clarification dialogues and discuss each of the clarification types in detail.

3. Information Flow in Clarification Dialogues

If the possibility to carry out clarification dialogues is disabled spoken input is sent through components for speech recognition, syntactic, semantic construction, semantic evaluation, transfer, tactical generation and speech synthesis. In figure 1 this information flow is represented by means of small arrows.

In case clarification dialogues are enabled the processing results of the syntactic and semantic components are continuously monitored by the dialogue component: among other intermediate results the following information is passed on to the dialogue component (indicated by bold arrows in figure 1):

- **segment descriptions:** during speech recognition, syntactic analysis and semantic construction each incoming turn is segmented into chunks which ideally correspond to individual sentences; for each of these segments a description is constructed which includes a unique segment identifier, information about the turn-finality of the segment, the recognized word string for the speech input in the source language and, if available, the word string of its translation into the target language.
- **signals concerning the consistency of intermediate results:** for every segment the semantic evaluation component checks the semantic consistency of the results being produced by the preceding components; if the analysis does not yield a semantically consistent result, this fact is signaled to the dialogue component, thereby initiating a clarification dialogue.

For every segment `segment_id` and for each type of clarification dialogue the dialogue component sends a message to the central control component of the VERBMOBIL system indicating whether a clarification dialogue has to be executed or not (<X segment_id> or <no_X segment_id>, where X may be any of {`similar_words, unknown_words, inconsistent_date`}). If a subdialogue has to be carried out, the clarification mode is switched on (`clarification_dialogue on`) and the processing flow of the system is changed. Depending on the clarification type `X`, a synthesized message is sent to the user, informing him of the necessity and reason for a clarification dialogue. These system messages are generated and synthesized from templates which have been filled in with information specific to the clarification dialogue at hand. By way of such messages a list of options for recovery is presented. In order to minimize processing errors the options the user can choose from are formulated as yes/no questions; a yes-/no recognizer with a recognition rate of approximately 100% developed specifically for this purpose processes the user's response. If the user chooses an option that allows a continuation of the dialogue it is used to modify the system's intermediate results; the utterance `segment_id` and the updated message are sent to the control module (`clarification_dialogue_succeeded segment_id` <modified-message>), the system switches back into the normal processing mode (`clarification_dialogue off`), and computation is resumed using the modified data. If the user finds none of the presented options appropriate, the user is requested to reformulate his original utterance, the control component is informed of a failure of the subdialogue (`clarification_dialogue_failed segment_id`) and the clarification dialogue is switched off (`clarification_dialogue off`).

All three subdialogue types realized in the current VERBMOBIL prototype follow this uniform processing scheme.

The clarification dialogues are adaptive to the situation insofar as the system utterances are elaborate in case a clarification type is carried out for the first time and rather compact if it has been carried out already. All clarification types mentioned in this abstract are fully implemented.

In our architecture, clarification dialogues can be triggered by various sources:

- **the dialogue component:** the dialogue component checks the intermediate results produced by various VERBMOBIL modules and examines them for possible problems and inconsistencies. Subsequently, a clarification dialogue is initiated which addresses the source of the problem. This method has been chosen for cases, where simple heuristics for consistency checks could be implemented and where these heuristics do not lead to a duplication of knowledge already available in the system.
- **other subcomponents of** VERBMOBIL**:** more complex consistency checks are resolved in the components that contain the necessary information; this is the case, for example, for the detection of semantic inconsistencies, for which complex reasoning processes are required.

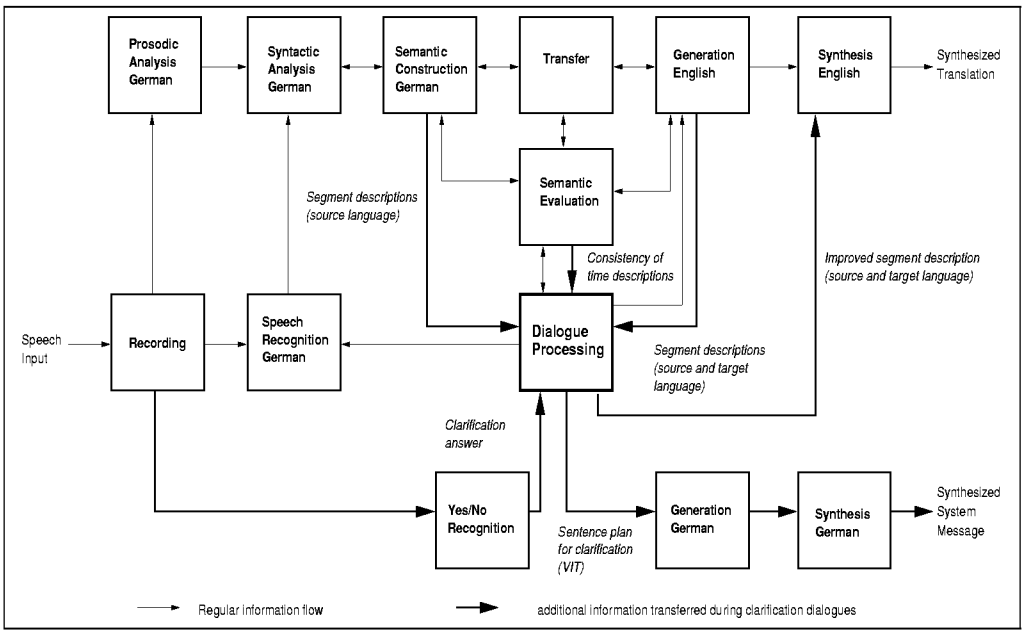

Figure 1: The information flow in VERBMOBIL

In the following we describe the various types of clarification dialogues as they are implemented in our current system.

4. Dialogue-Initiated Clarifications

Phonological Similarities

The dialogue system has access to a list of words that are often confused on the basis of a high degree of phonological similarity. Not all of the word pairs included in this list are intuitive candidates for an average VERBMOBIL user. Examples are e.g. the German word pairs *Halle - fahren* or *Modus - Morgen*. We compiled a subset of this list that contains only word pairs whose similarity is plausible for a user who has no phonological expertise. This list includes word pairs like e.g. *Sonntag - sonntags* (engl.: *Sunday - sundays*) or *fünfzehn - fünfzig* (engl.: *fifteen - fifty*).

If the word string processed by the syntactic/semantic components contains a member of this word list the dialogue initializes the generation of a system message that points out the potential confusion to the user. If for example the original input sentence is *Wie wär's Sonntag?* (engl.: *How about Sunday?*) the system triggers the message *VERBMOBIL hat eine mögliche Verwechslung erkannt. Meinen Sie die Angabe 'Sonntag'?* (engl.: *VERBMOBIL encountered a possible ambiguity. Do you mean the word 'Sunday'?*). Depending on the answer of the user either the proposed word is accepted or the remaining other candidate (in this case *sonntags*, engl: sundays) is used. The chosen word is then inserted into the intermediate processing result, so that the translation later contains the word selected by the user.

Unknown Words

The speech recognizers of the VERBMOBIL system are able to recognize input as unknown to the system; if such a fragment is encountered the symbol `unk_` followed by the SAMPA-transcription of the fragment (e.g.<`UNK_maI6`> for the unknown spoken input *Maier*) is inserted into the output of the recognizers. In our domain, unknown words often refer to names, e.g. of locations or persons. The user is asked to confirm this assumption. A message including a synthesized version of the SAMPA transcription is presented to the user, e.g. *Handelt es sich bei* `maI6` *um einen Namen?* (engl.: Is `maI6` a name?). If this assumption is confirmed, syntactic processing is continued treating the fragment as a name. The SAMPA transcription is later included in the output of the English generator and pronounced accordingly.

5. Clarifications Initiated by Other VERBMOBIL Components

Semantic Inconsistencies

If a user tries to propose nonexistent or inconsistent dates, this is signaled to the dialogue component by the semantic module. If possible, this module also proposes alternative dates. The message `clarify_date([dom:31,moy:apr],[dom:30,moy:apr])`, for instance, indicates both that April 31 is an inconsistent date and that the user might have meant April 30. The dialogue component transforms this information into a message presented to the user: *Die Angabe 31. April existiert nicht. Meinen Sie die Angabe 30. April?* (engl: The date 'April 31' does not exist. Do You mean April 30?). If the user chooses the alternative date, it is passed on to the relevant components and the resulting translation includes the correct date.

Currently, the following types of semantic inconsistencies are handled:

- inexistent dates, as e.g. *September 31*;
- a mismatch between weekdays specified by the user and the calendar information, as e.g. *Monday, April 20, 1997*;
- inconsistencies between exact time and period of day, as e.g. *in the afternoon at 4pm*;
- inconsistencies between public holidays and their exact date in the calendar, as e.g. *Good Friday, April 27*.

6. Related Work

In contrast to systems where dialogue components play an active role (like e.g. in flight booking systems, in train information systems) clarification dialogues have so far not been integrated into systems where the dialogue component remains rather passive, i.e. only monitors the dialogue. Comparable speech translation systems as e.g. developed at ATR and CMU (see [7] and [6]) do not provide the possibility to augment the system's information by an interaction with the user.

Among the spoken dialogue systems that actively guide the dialogue and that allow for clarification subdialogues is the system described in [3]. The authors describe a system giving information about public transport which uses clarification dialogues to elicit missing information. In the SPEAK! system for information retrieval from a full-text database [4] user-initiated clarifications can be carried out, requesting additional information about previous system utterances.

7. Conclusion and Future Work

In this paper we described three types of clarification dialogues that have been fully integrated in the current prototype of VERBMOBIL. The possibility to carry out clarification dialogues increases the robustness of the overall system insofar as missing information is being elicited, inconsistent knowledge is being clarified and ambiguous information is being resolved.

For future VERBMOBIL prototypes the following improvements are planned:

- **implementation of more clarification types:** we will emphasize on subdialogues which lead to the resolution of ambiguities necessary to determine correct translational equivalents for expressions in the source and the target language.
- **design of an improved communication concept:** in the current prototype, clarifications have been carried out using word strings as underlying data type. The material acquired during the clarification dialogue is integrated into the word string included in the segment description. This updated segment description is then handed over to the English synthesis component. In the future, clarifications will be carried out with respect to the so-called VERBMOBIL *interface terms* (VITS), which serve as input for the linguistic components of the VERBMOBIL system. This method allows us to re-process the clarified material by some of the system's subcomponents, thereby improving the quality of the final result.

8. REFERENCES

1. Jan Alexandersson, Norbert Reithinger, and Elisabeth Maier. Insights into the Dialogue Processing of VERBMOBIL. In *Proceedings of the Fifth Conference on Applied Natural Language Processing, ANLP '97*, pages 33–40, Washington, DC, 1997.
2. Thomas Bub and Johannes Schwinn. Verbmobil: The evolution of a complex large speech-to-speech translation system. In *Proceedings of ICSLP-96*, pages 2371–2374, Philadelphia, PA., 1996.
3. Wieland Eckert and Scott McGlashan. Managing spoken dialogues for information services. In *Proceedings of EUROSPEECH-93*, Madrid, Spain, September 1993.
4. Brigitte Grote, Eli Hagen, Elke Teich, and Adelheit Stein. Speech Production in Human-Machine Dialogue: A Natural Language Generation Perspective. In Elisabeth Maier, Marion Mast, and Susann LuperFoy, editors, *Dialogue Processing in Spoken Language Systems*. Lecture Notes in Artificial Intelligence, Springer-Verlag, Heidelberg, 1997.
5. Detlev Krause. Using an Interpretation System - Some Observations in Hidden Operator Simulations of VERBMOBIL. In Elisabeth Maier, Marion Mast, and Susann LuperFoy, editors, *Dialogue Processing in Spoken Language Systems*. Lecture Notes in Artificial Intelligence, Springer-Verlag, Heidelberg, 1997.
6. Alon Lavie, Lori Levin, Yan Qu, Alex Waibel, Donna Gates, Marsal Gavalda, Laura Mayfield, and Maite Taboada. Dialogue processing in a conversational speech translation system. In *Proceedings of ICSLP-96*, pages 554–557, Philadelphia, PA, 1996.
7. Yasuhiro Sobashima, Osamu Akamine, Jun Kawai, and Hitoshi Iida. A Bidirectional Transfer-Driven Machine Translation System for Spoken Dialogues. In *Proceedings of the 15th International Conference on Computational Linguistics (COLING 94), August 5–9, 1994, Kyoto, Japan*, volume 1, pages 64–68, 1994.