

Dialogue Processing in a CALL-System

Veit Reuer

Institut für deutsche Sprache und Linguistik
Humboldt-Universität zu Berlin

Unter den Linden 6

10099 Berlin

GERMANY

Veit.Reuer@compling.hu-berlin.de

Abstract

In a CALL-environment (Computer-assisted language learning) programs should ideally allow the learner to train his/her *communicative competence*, which is one of the main goals of foreign language teaching nowadays. This can be reached by allowing learners to use and train their knowledge of a foreign language in realistic *dialogue-style exercises*. All levels of linguistic and communicational analysis have to be considered to realize such a system. In this paper I will concentrate on the *dialogue component* of the concept, which relies on two main knowledge sources. The *discourse grammar* structures the dialogue elements (or dialogue acts) as possible parts of a dialogue and the *dialogue knowledge base* provides the possible contents of dialogues. Additionally, a framing discourse structure has to be built to provide the specific dialogue-exercise. A FSA (finite state automaton) based on the discourse grammar determines the possible moves which the dialogue might take. On the one hand this concept is restricted enough to allow for (relatively) easy maintenance as well as expansion and on the other hand it is advanced enough to allow for simulated complex dialogues.

1 Introduction

Today the main goal in foreign language teaching is acquiring the so-called *communicative competence* instead of only memorizing the structure of the language. This can be achieved by making active language production one of the main parts of the curriculum. Being an efficient part of the

media to enhance the learning process, computers should present tasks that support the acquisition of communicative competence. In the present concept, a presentation of situations is suggested, in which the learner has to produce language, i.e. produce complete sentences in a simulated dialogue. Various program modules analyse the input linguistically, give feedback in case of errors and present appropriate reactions to continue the dialogue. This gives language learners a chance to use their knowledge of the second language in a meaningful situation apart from the class room setting.

Three goals are of relevance: 1) the language learner should be encouraged to enter free-formed input instead of thinking about the 'expectation' of the program; 2) the program should offer reliable feedback to the learner about his/her performance and 3) the program should be (easily) expandable.

When one uses the program, a situation is presented to the learner in which s/he is required to act in order to solve the particular problem at hand. For example, the learner is asked to buy tickets for a movie and has to engage in a written dialogue with the computer as the seller of the tickets.

The motivation for the development of the system arises from the above mentioned pedagogical considerations and the insight that traditional language learning programs do offer only few or none of the features to reach the above mentioned goals. The main part of this paper, however, deals only with the computational aspects of the problem: a possible way to implement such a dialogue component. In the next section I will focus on the dialogue component, which on the one hand allows the learner to communicate with the computer in various dialogue situations, and on the other hand is restricted enough to be easily expandable and maintainable and gives the possibility for advanced feedback. Finally I will give a

rough sketch of the complete system.

2 Discourse Grammars

One way of realizing such dialogues is to develop discourse grammars which describe the steps through distinct parts of dialogues. This possibility is chosen in the present concept, since it enables the learner to lead written, situation-based dialogues almost as in the class-room situation. The advantage of a discourse grammar over a completely plan-based dialogue structure is the separate representation of possible moves ('dialogue acts' in Alexandersson et al. (1994)) and the content of the discourse. The discourse grammar of an 'information-gathering' dialogue can be used while reporting an accident as well as while ordering a pizza. In the first case the police officer wants to know all about the accident and possible casualties and in the latter case the pizza delivery wants to know the toppings and size of the pizza. On the other hand guidance is needed for the learner in the CALL-scenario. Systems like the one described in Carberry (1990) are much too open to be used for language learning. The system would not be able to give any feedback to the learner in case of erroneous input. Therefore the system uses only restricted knowledge about what types of input to expect and how to react to them since the general intentions of the learner are known to the system through the situation presented to the learner. In other NLP-based systems like 'Herr Kommissar' (deSmedt, 1995) and 'LINGO' (Murray, 1995), the dialogue with the system either allows only single question-answer exchanges or is strongly embedded into the respective scenario. In the first case the structure of a complete dialogue does not become clear to the learner and the initiative is with the learner who might not know what to do. In the second case it is difficult to include new scenarios since not only the content of the new dialogue has to be coded but also the various dialogue structures. Moreover the design of a system might not allow for different types of dialogues.

The dialogue component contains two main knowledge bases: The first one contains the discourse grammars, which structure so-called 'goal-driven dialogues' or 'task-oriented dialogues'.¹ The idea of discourse grammars as a means to handle dialogue situations is for instance presented in Fawcett and Taylor (1989). The second knowledge base contains knowledge about the content of the dialogue itself. This data is used to infer a mean-

¹For a discussion about discourse grammars in general see e.g. Taylor et al. (1989).

ingful reaction to the input sentence. Additionally this base contains slots in which the information given by the learner is stored.

The following figure shows a simplified part of a discourse grammar, which models an information gathering dialogue such as is necessary in the case of collecting information about an accident. Additional items of discourse grammars are of course needed, for example, to start and end a telephone call, etc.

The same type of structures is also used in the analysis of dialogues, e.g. (Carletta et al., 1997). Here dialogues are analysed with the help of a 'Dialogue Structure Coding Scheme', which in particular contains only a limited number of possible moves between dialogue partners. A similar analysis was done in the preparational phase of the Verbmobil project (Alexandersson et al., 1994). In a dialogue system where the intentions of the dialogue partners are known and the fixed structures serve to assess the performance of the language learner, the restrictions will probably not make the overall behaviour of the system worse than more flexible dialogue systems.

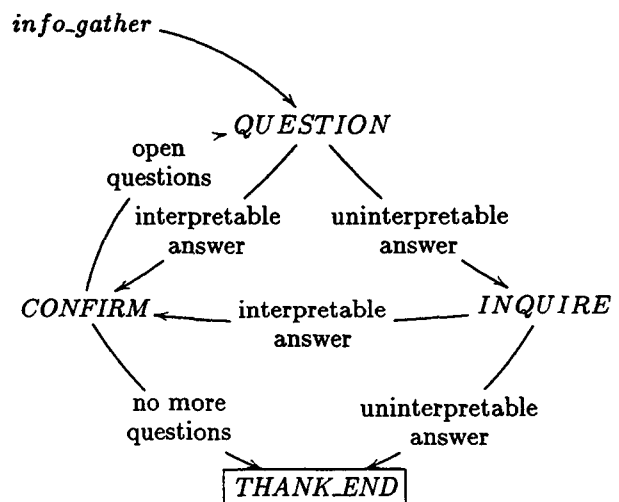


Figure 1: Simplified discourse grammar

The dialogue module uses a surrounding discourse grammar, which includes the grammar parts for starting and ending a telephone call etc. From here the information gathering structure is called to try to fill the variables in the dialogue knowledge base (see below) by asking the learner a question. This process is continued as long as there are open questions (open questions) or until the learner does not provide interpretable input even after a repeated question (*INQUIRE - THANK_END*).

The dialogue knowledge base contains the data

necessary to lead a dialogue with a certain content. The data is organized in a hierarchical structure. In the 'police call' example the root-node consists of a slot with a first reaction of the officer (greeting) to be presented to the learner. The daughter nodes (e.g. accident, theft) contain some slots which are used for the actual presentation of reactions on the screen or for information storage and retrieval. Some slots are:

- question for pieces of information: This includes canned text, which is presented to the learner. For example the police officer might ask 'Are there any injured people?'
- information about expected answer: The semantic structure of the learner's input is checked against the content of this slot and in case of variables it is stored.
- keywords to match the learner's input: In case the parser was not able to produce a semantic representation, the system retreats to keyword matching in order to provide at least some reaction.
- text as answer: A sentence is passed to the learner to acknowledge or confirm the processing of the input ('So, there has been an accident.').

In case the system chooses to ask a question based on the discourse grammar, the question from the appropriate slot in a daughter node (top-down left-right) is passed to the learner. After the grammatical processing of the answer, the content is checked against the expected one. If they match, a confirmation may be passed to the learner and the next step in the discourse grammar is taken. If the answer was considered not appropriate for the question the system tries to find a response in a hierarchy of steps from world knowledge checking to simple keyword analysis. The final output can thus be from the same node, a subnode or from a more general independent source of possible reactions. Some mechanism has to manage the matching-procedure of the sentence. Possible mechanisms thus include:

- the content matches completely: The system was able to recognize the input sentence as some meaningful reaction to the previous question or statement.
- the content fits only partly (too general): There are subnodes which in-

clude variables for more specific information.

- the content fits only partly (only one aspect): A general keyword-based mechanism recognizes only parts of the expected input. If possible the learner is asked for further clarification.
- the content does not fit: A question for rephrasal will be displayed to the learner. Additionally the learner might consult a helpfile with information about how to proceed in the current situation.

A difficulty that might arise is the change of control (or initiative) between the dialogue partners. Allowing the learner to take the initiative has several consequences, which are difficult to realize. In contrast to the present concept the dialogue module should include a language generation device to generate natural language output to database-inquiries. From this follows that the dialogue knowledge base should not contain any contradictions etc. to allow for easy inference of possible answers to the input question. Finally in case the learner keeps on asking questions the system might fail to continue the dialogue in a meaningful way. Thus a system designed for the use by pupils must be rigid enough to deal with this kind of input.

The seemingly limited flexibility in this system is not really a disadvantage, because 1) the learner is supposed to act in a foreseeable way and 2) the system should give feedback in case of deviating action. Especially the latter seems only possible if a discourse grammar structures the moves which dialogue partners might take.

3 System Overview

The idea behind the system is to extend the types of training which the student gets in a class room setting into a computer. One important kind of training is the practising of dialogues. Therefore the program realizes small written dialogues for the learner to train her/his 'communicative competence', as explained above.

The system consists of four main modules. The *dialogue control module* mainly functions as an interface. It organizes the flow of the input data between the user-interface and the various processing modules. Every input sentence is first passed to the *linguistic module*, which checks it for orthographic and syntactic errors. The orthographic check is done in the spirit of Oflazer (1996). With

the help of a finite state recognizer mildly deviating strings are identified and correct versions are presented to the learner if necessary. The syntactic check follows a rather traditional path. The main work is done by a LFG-parser (Bresnan and Kaplan, 1982), originally implemented by Avery Andrews (Australian National University) and now modified to suit the needs of error detection with the help of modified grammar processing including error rules (Krüger et al., 1997). As a next step the analysis of the sentence is checked against a *world knowledge base*, from which feedback follows to the learner if the sentence does not match the internal model of the world. In contrast to the dialogue knowledge base this model of the world cannot be altered by the learner because of its usage for inference and the absence of a consistency-checking module to prevent contradictions etc. If the student has made an error, the system provides feedback to support the learner in typing a syntactically correct or semantically more plausible sentence. After this step the *dialogue component* tries to find a reaction to continue the dialogue, as described above.

The main focus in all the analyses is to continue the dialogue but without ignoring the errors made by the learner. Only the orthographic check will actually interrupt the dialogue with a suggestion of correct words for the misspelled items. In all other cases the dialogue partner will react to the erroneous input depending on the type of error.

4 Conclusion

The aim of this concept is to provide a foreign language learner with exercises that enhance his/her communicative ability. One important module in the system is the dialogue component itself which organizes e.g. the turntaking in a dialogue. At least two knowledge sources are necessary for limited flexibility and reusability. The structure of a dialogue can be handled by a discourse grammar whereas the content of a dialogue is stored into an (expandable) knowledge base. To add new dialogues, only the dialogue knowledge base and the surrounding discourse grammar have to be updated whereas the other specialized discourse grammars can be reused. The structure of the dialogue knowledge base could also allow for the handling of questions and multi-sentence answers given by the learner. Nodes in the tree do not only represent variables to be instantiated by the learner, but might also include knowledge the system can provide to the learner.

The kind of dialogue component presented here allows for 1) easy maintenance and expansion

with new dialogues, 2) advanced feedback to the learner and 3) flexible and pedagogically sound exercises which enhance the process of acquiring 'communicative competence'.

References

- Jan Alexandersson, Elisabeth Maier, and Norbert Reithinger. 1994. A robust and efficient three-layered dialog component for a speech-to-speech translation system. *Verbmobil-report 50*, DFKI, Saarbrücken.
- Joan Bresnan and Ronald M. Kaplan, 1982. *Lexical-Functional Grammar: A Formal System for Grammatical Representation*. In Bresnan (Bresnan, 1982).
- Joan Bresnan, editor. 1982. *The mental representation of grammatical relations*. MIT Press, Cambridge (MA).
- Sandra Carberry. 1990. *Plan Recognition in Natural Language Dialogue*. MIT Press, Cambridge, MA.
- J. Carletta, A. Isard, S. Isard, J. Kowtko, G. Doherty-Sneddon, and A. Anderson. 1997. The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23(1).
- William H. deSmedt, 1995. *Herr Kommissar: An ICALL Conversation Simulator for Intermediate German*. In Holland et al. (Holland et al., 1995).
- R.P. Fawcett and M.M. Taylor, 1989. *A Generative Grammar for Local Discourse Structure*. In Taylor et al. (Taylor et al., 1989).
- V. Melissa Holland, Jonathan D. Kaplan, and Michelle R. Sams, editors. 1995. *Intelligent Language Tutors*. Erlbaum, Mahwah (NJ).
- Anja Krüger, Hendrik Dittman, and Maureen Murphy. 1997. Grammar based error diagnosis in CALL. Informatics research reports, University of Ulster.
- Janet H. Murray, 1995. *Lessons Learned from the Athena Language Learning Project*. In Holland et al. (Holland et al., 1995).
- Kemal Oflazer. 1996. Error-tolerant finite-state recognition with applications to morphological analysis and spelling correction. *Computational Linguistics*, 22(1).
- M.M. Taylor, F. Neel, and D.G. Bouwhuis, editors. 1989. *The structure of Multimodal Dialog*. Elsevier, North-Holland.