Dialogue control in the ALPARON system

L.J.M.Rothkrantz, R.J. van Vark, A. Peters, A.C. Andeweg

Knowledge Based System Group
Delft University of Technology,
Zuidplantsoen 4
2628 BZ Delft, The Netherlands
{L.J.M.Rothkrantz,R.J.vanVark}@cs.tudelft.nl

Abstract. The main topic of this paper is on modelling a human operator in the dialogue manager of the Alparon system. First, a corpus of 200 human-human dialogues have been analysed applying an approach surpassing a finite state automation approach. The corpus analysis resulted in a set of common strategies applied by professional human operators in similar situations. Secondly, a prototype system has been built based on the Alparon dialogue manager. This has been done by translating the strategies into knowledge rules and heuristics as these are used by the dialogue control modules in the Alparon dialogue manager.

1 Introduction

Dialogue management is the key factor in automated speech processing systems for information retrieval and transaction services for it can overcome the shortcomings of a one-shot approach. In a one-shot approach the user specifies his requirements to the system in one turn. Then, the system tries to understand the user's requirements using the information provided in this turn and the system gives an appropriate response, also in one turn. If this approach fails, the user has to start again from scratch. In a one-shot approach one cannot speak of a dialogue-like interaction as there is simply an iteration of actions followed by reactions.

To behave in a more user-friendly fashion the system has to be able to cope a wide variety of phenomena like speech recognition errors, misunderstandings, partial specification of user requirements, conflicting requirements, etc. To cope with these phenomena, the system engages in a dialogue with the user using information collected in previous turns to improve the current interpretation of the user's requirements. Therefore dialogue management is an essential element for a user-friendly automated speech processing system.

Researchers at the Delft University of Technology have developed a natural language system aimed at information retrieval and transaction services. The focus of the Alparon system is on dialogue management, especially for speech-enabled applications. Other components of the automated speech processing system, like a speech recogniser, parser and speech generation, are only incorporated to make a complete test bed for testing dialogue strategies. The goal of the Alparon project is to investigate the application of human-human strategies in human-computer dialogues.

This paper describes an approach to model a human operator using the dialogue manager of the Alparon system. This paper especially focusses on retrieving the user's requirements and not on presenting the information itself. Modelling the information of presentation has been presented earlier by researchers of the Delft University of Technology (see [1]). First, applicable human-human strategies have to be analysed. This was done by analysing a corpus of over 200 human-human dialogues between customers requiring travel information and professional operators. The analysis of this corpus is described in the second section. Secondly, the human-human strategies have to be translated to knowledge rules and heuristics as they are applied in the dialogue control components of the Alparon dialogue manager. A description of the Alparon prototype modelling a human operator can be found in section 3. The final section presents conclusions and future work.

2 Computational dialogue management

One way to design an effective, efficient and user-friendly dialogue manager is to apply a corpus-based approach. In such an approach, a corpus of human-human dialogues is studied to derive applicable strategies. In the Alparon project, a corpus was available consisting of 5.000 recorded and annotated human-human dialogues on timetable information in public transport. A selection of 500 was coded using the Alparon coding scheme [2]. This coding scheme applies dialogue acts to code the influence of an utterance to the dialogue context.

A dialogue can be represented as a tree. Every layer in the tree represents a turn in the dialogue and the nodes correspond to individual dialogue acts. As many dialogue acts can be found in a dialogue, the tree grows exponentially as the dialogue progresses. Applying such methods would call for a huge corpus.

Therefore, another way of representing the dialogue discourse was needed. In this paper, a dialogue is considered from the operator's point of view. This corresponds to the dialogue manager's point of view as the the operator will be modelled in the Alparon system. A computational model of the operator combines thoughts about the hypotheses, the client's intentions and the prompt choices. The model is constructed using goal-directed principles. The operator bases his/her behavior on a comparison of a representation of the goal-state and the current state. The operator can use the planning strategy means-end analysis. Means-end analysis requires a measure of distance between the current state and the goal-state. The next step is then chosen based on an evaluation of how much this step will reduce the distance to the goal state. A useful reaction of the operator provides new information, reduces the uncertainty or provides clearness, whereas an inappropriate reaction takes extra time or causes confusion.

As such, the quality of an operator's prompt is defined by the extent the prompt will get the dialogue nearer to the goal, being defined as providing the information the client desires. One way to assess the quality of the prompts is

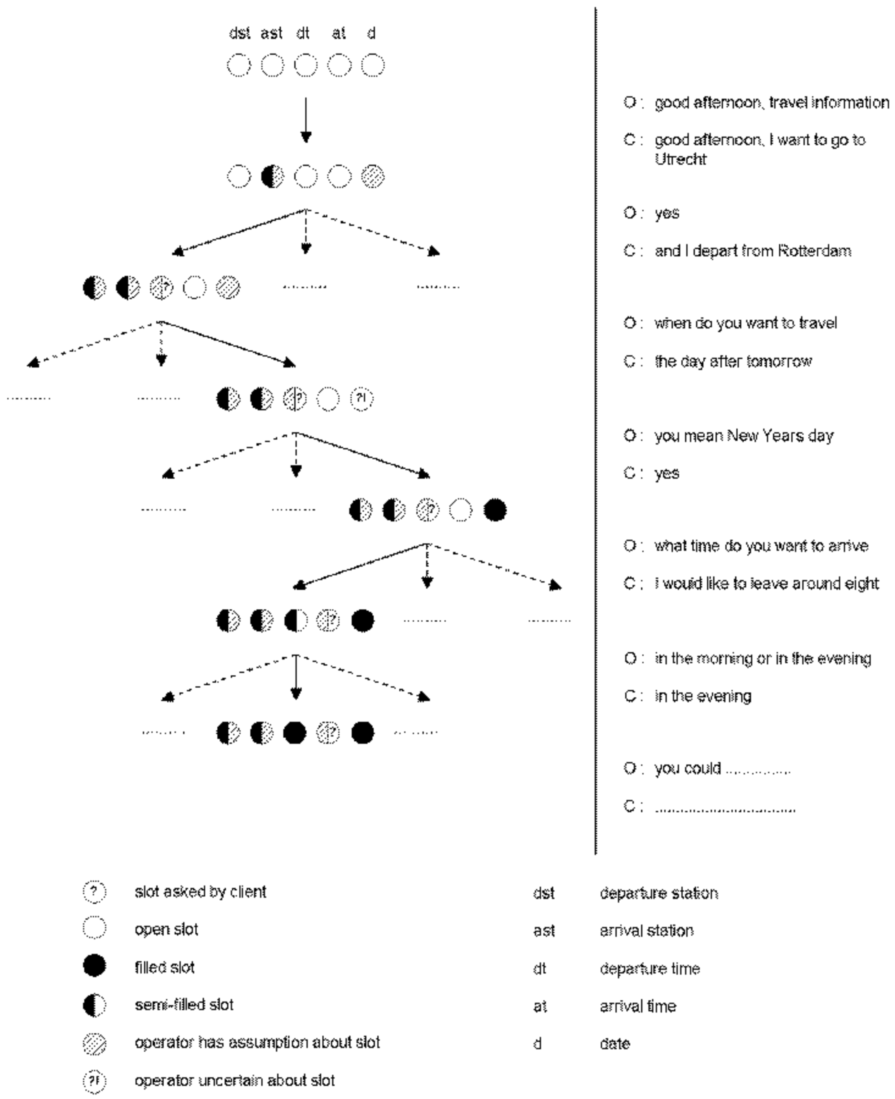

Fig. 1. Dialogue model from the operator's perspective

by applying an evaluation function that provides a score for each prompt. At each step the prompt with the highest score is chosen. It is possible to think several steps ahead and combine the weights that belong to these steps. This way all the possible paths through a dialogue are scored and the path with the highest score can be chosen. Thinking several steps ahead is effective in finding the best path through a dialogue. The shortest path without detours produces the most efficient dialogue but it is questionable whether the client appreciates such a dialogue. The factors that influence the professional operator in deciding on which prompt to choose have to be combined in these weights.

In figure 1 the operator's mental state during a dialogue is displayed. The operator's thoughts are illustrated in the picture by using slots, represented by five circles. The goal is to fill the slots. All slots are empty at the beginning of the dialogue. Every state represents the operator's thoughts at that time. A new state is displayed after an operators prompt and the client's reaction to it. The next state indicates if there are differences in the operator's thoughts about the slots as compared to the previous state. At every state there are different ways to continue the dialogue. This is indicated in the figure by multiple arrows on a level. Using an optimal strategy implies that in the next state a maximal slot-filling is achieved. This slot-filling is done by reducing uncertainty about the slots or asking for new information.

Dialogue graphs were constructed for 200 dialogues selected from the Alparon corpus. In a one-step approach those prompts were selected providing a maximal filling of the slots. In about half of the cases it was possible to generate knowledge rules and heuristics how to select the appropriate prompt. In the other cases the history has to be taken into account or a multiple step approach, or the client took the freedom to take the initiative or could not provide the information. Some examples of heuristics derived using this strategy, are:

- if there is a (non-empty) subset of open slots and a (non-empty) subset of filled slots, ask for information covering as much of the open slots as possible;
- if one slot in the current state is not complety filled, immediately ask for the missing information to solve the ambiguity concerning the slot;
- if the subset of not complete filled slots contains more than one slot, handle the individual slots one after the other;
- as long as new information can be provided, assumptions of the operator are not verified;

It should be stressed that in the current application the operator needs information from the client to define a query for a database. This might cause the operator's behaviour to be rather slot-oriented, as was found for the current domain. There is a tendency from the operator to formulate open questions. In case of automated speech-driven systems there is a tendency to generate closed questions to prevent misunderstanding. In the next section an approach is presented which compromises both extreme viewpoints.

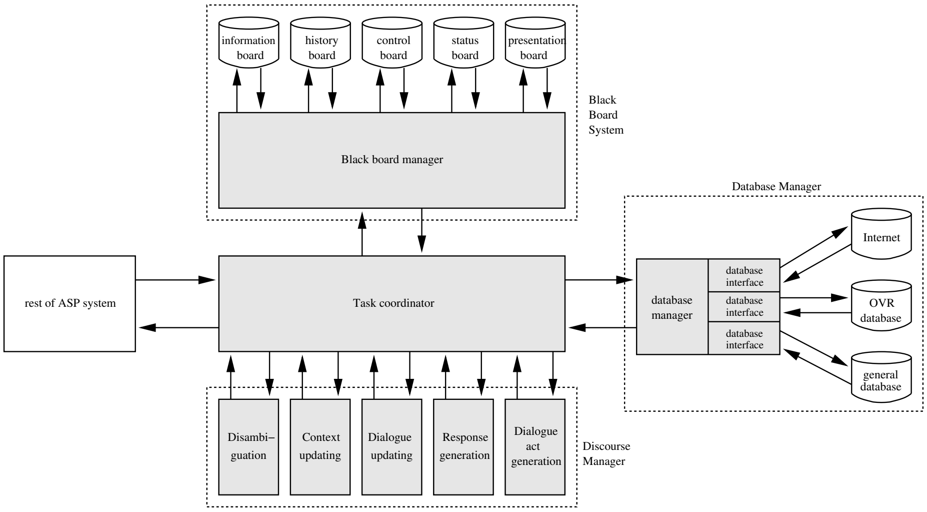

Fig. 2. Alparon dialogue manager overview

3 The Alparon natural language system

The Alparon natural language system focusses on applying human-human strategies in dialogue management for information and transaction services. The Alparon dialogue manager consists of several modules (see figure 2). This modular design was chosen to create a flexible and extensible system. When implementing information services, the dialogue manager should be able to interpret the user's utterance in the context of the ongoing dialogue. This interpretation is then used to generate a response to the user's utterance. Information on the current state of the dialogue and domain knowledge can be used to perform these tasks.

The main topic of this paper is on using human operator strategies to implement the dialogue control mechanism in the Alparon dialogue manager. This mechanism is to reason out a system response after the user's utterance is interpreted. These responses are the system's only means of influencing the user and thus controlling the dialogue. An approach surpassing a finite-state automation approach should be applied.

In the original design of the Alparon system, the task of generating system responses was placed with the response generation module and the control board. All the processing necessary to generate the system response should be done by the response generation module. This would include processing the user's turn on a dialogue level, since the context-updating module only takes care of the user's turn on an information level. In the actual design for dialogue control as described in this paper, another module was added, the dialogue updating module, to determine the consequences of the user's turn on a dialogue level.

The dialogue updating module is charged with updating of goalstack according to the state of the dialogue in each turn. The goalstack is used to structure the goals that define what must be achieved during the dialogue. In information retrieval dialogues the dialogue structure highly resembles the task structure. So a plan based approach is used in which the goals embody the plan to accomplish the task and subtasks in such a dialogue. When disambiguation and context updating modules have interpreted the user's turn and extracted the factual information from it, the dialogue updating module processes the consequences of the new information and the user's turn on a dialogue level. The goalstack is updated to reflect this new situation.

The task of the response generation module is to generate the actual system response. Given the goals that still have to be achieved, the response generation module places communicative actions linked to those goals and selects communicative actions to be combined in a system response. The communicative actions are linked to goals and they are designed to help achieve the goals.

The control board is used to store the goal structure and the communicative actions, both the possible ones and the ones selected for the system response. This goal structure and the communicative actions implement the rules and strategies extracted by our corpus analysis. Other information related to dialogue control is also placed on blackboards, e.g. conflicts in the system beliefs of the dialogue. All the information used for dialogue control is present on the blackboards.

4 Conclusions

In this paper we presented a prototype of a dialogue management system. The dialogue control strategies implement a human operator applying rules and strategies extracted from a corpus of human-human dialogues. In the research environment of the Alparon system, dialogue control is designed as a generic framework. Within this framework various rulesets can be defined to specify the behaviour of dialogue control. By changing these rulesets various approaches to dialogue control can be implemented and refined. The rulesets specify one engine, consisting of all components for dialogue control. As a result, the actual strategy for dialogue control is defined by these rulesets.

References

1. M.M.M. Rats, R.J. van Vark, and J.P.M. de Vreught. Corpus–based information presentation for a spoken public transport information system. In *ACL/EACL Workshop on Interactive Spoken Dialog Systems: Bringing Speech and NLP Together in Real Applications*, pages 106–113, 1997.
2. R.J. van Vark, J.P.M. de Vreught, and L.J.M. Rothkrantz. Classification of public transport information dialogues using an information based coding scheme. *Dialogue Processing in Spoken Language Systems, Lecture Notes in Artificial Intelligence (1236)*, pages 55–69, 1997.