Discourse Management in Voice Systems for Accessing Web Services

Marta Gatius   **Meritxell González**

Technical University of Catalonia, Software Department

Jordi Girona 1-3, Campus Nord, 08034 Barcelona, Spain

`{gatius,mgonzalez}@lsi.upc.edu`

Abstract

This paper describes the discourse management component of a dialogue system that supports voice and text access to the web services in different languages: English, Spanish, Catalan and Italian. The dialogue manager follows the information state theory and uses communication plans that are generated when the system is adapted to a new web service. To facilitate the generation of these plans we have defined general communication plans for different types of web services.

1 Introduction

There already exist voice systems for accessing specific web services. However, most of these services only support system-initiative dialogues, in which the system drives the interaction asking the user the information the service needs.

System-initiative dialogues have proved efficient when only the system knows which particular data the service needs from the user, i.e. when inexperienced users wants to perform an online transaction that require particular data. However, there are situations in which the users can take the initiative to give the information the system needs, e.g., in the case of information seeking, only the user knows what (s)he wants to find.

Understanding what the user's intends becomes more difficult in mixed-initiative dialogues, because user's interventions may be difficult to predict (i.e., they may be unrelated to the system's questions previously asked). In this paper we describe the approach we followed in the multilingual dialogue system (DS) to access public administration web services we developed in the context of the European project HOPS (http://www.hops-fp6.org/). In this DS, the dialogue is controlled by a general module, the dia-logue manager (DM), which uses communication plans to recognize the user's intention and decide next actions.

2 The Dialogue System

The DS has been designed to support access to different types of web services through several channels. The final prototype supports access through the telephone and the web in different languages (English, Spanish, Catalan and Italian) to two web services: a transactional service for collecting large objects and an informational service on cultural events.

The voice component of the DS is that of the Loquendo multilingual VoiceXML platform. The speech recognizer uses grammars (following the standard SRGS syntax) modeling user utterances. The transcription of the utterance is passed to a natural language parser and processor, which performs a deep syntactic and semantic analysis. The resulting semantic interpretation is then passed to the DM. The same parser is used to process the text introduced through the web.

3 The Dialogue Manager

The discourse management component we have developed supports mixed-initiative dialogues and other complex phenomena, such as feedback strategies, which are especially needed to deal with misunderstandings. An example of dialogue supported by our system can be seen in Figure 1.

Our DM follows the information state theory (IST) of language modeling (Traum et al., 1999). This model combines the efficiency of structural approaches and the flexibility of plan-based approaches. Although the IST deals with complex aspects of dialogue, such as beliefs, intentions and plans, it is not difficult to implement because it provides the tools to formalize this knowledge.

As in other practical DSs (Larsson, 2002), the plans used by our DM to recognize the user's

intention and determine next communication actions are generated statically, when a new service is incorporated to the service. Because we consider that each task a web service can perform is a possible user's goal, we generate a communication plan for each service task.

S1: Welcome to the automatic platform of Barcelona. May I help you?
U1: I search cultural events
S2: You said I search cultural events. I'm sorry, I don't understand.
U2: I search for information about cultural events
S3: Wait a moment, please. Can you tell me some data about the event or events you are interested in?
U3: Movies in the filmoteca
S4: Ok, you are interested in the title of the event. The event type is cinema. The place is filmoteca.
  *** <database consultation>

....

Figure 1: An example of dialogue

4 The Communication Plans

The objects of the plans are the service tasks and their parameters. The operators (actions) of the plans can be simple and complex actions. Possible simple actions are any of the central dialogue moves (ask, answer and ask for confirmation) and the accesses to the web services. Complex actions consist of a sequence of simple actions. The preconditions of the actions are related to the parameters values of the service tasks. Additional information about the service tasks parameters (i.e., the set of possible values) can also be included in the plans.

In order to facilitate the generation of the plans we have defined general communication plans for different types of services. There already are general descriptions of the interactions that take place when accessing web services. We have represented those general descriptions as top-level plans with abstract operators that can be decomposed into a group of steps. These steps can bee semi-automatically instantiated for each web service. We have considered two types of services: transactional and informational.

4.1 Plans for Transactional Services

We consider transactional services are those performing a transaction. Usually, those services require specific information from the user. The top-level plan to follow when accessing those services consists of the following sequence of general actions:

- *Obtaining the value of the input parameters*. The system asks the user this data.

- *Access the transactional service*.

- *Give information about the transaction*.

4.2 Plans for Informational Services

Seeking-information dialogues have focused several works. We have only considered the communication that usually takes place when the user accesses a simple web service describing a particular item (or a set of items). The top-level plan to access those services is the following:

- *Obtaining the value of the focus parameters*. Focus parameters are those needed to restrict the search. Although usually there is a set of possible parameters to restrict the search, unlike in the case of transactional services, the user may choose to give the value of only a subset of them. Optionally, the system also asks the specific data the user wants about the element searched (i.e., the location).

- *Access the informational service*.

- *Presentation of the results*. If two many elements are obtained the DM determines the new parameters to restrict the search, which are then asked to the user.

5 Conclusion and Future Work

Experiments performed by ten volunteers showed that the task success achieved using the text mode is high (90%), higher in transactional services (95%) than in informational services (85%). More formal tests on the performance of the DM will be done in the future. Future work will also include the adaptation of the DM to access other different types of web contents.

Reference

Araki, M. and Tachibana, K. Multimodal Dialog Description Language for Rapid System Development. In the Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue, (Sidney, 2006).

Larsson, S. Issue-based Dialogue Management. PhD. Thesis, Goteborg University, 2002.

Traum, D., Bos, J., Cooper, R., Larsson, S., Lewin, I., Mathesson, C., Poesio, M. A model of Dialogue Moves and Information State Revision. Trindi Technical Report D2.1, 1999. http://www.ling.gu.se/projekt/trindi/publications.html.