Error handling in a stochastic dialog system through confidence measures ☆

F. Torres *, L.F. Hurtado, F. García, E. Sanchis, E. Segarra

Departamento de Sistemas Informáticos y Computación (DSIC), Universidad Politécnica de Valencia (UPV), Camino de Vera s/n, 46022 Valencia, Spain

Abstract

In this work, we present an approach to take advantage of confidence measures obtained during the recognition and understanding processes of a dialog system, in order to guide the behavior of the dialog manager. Our approach allows the system to ask the user for confirmation about the data which have low confidence values associated to them, after the recognition or understanding processes. This technique could help to protect the system from recognition or understanding errors. Although the number of confirmation turns could increase, it would be less probable for the system to consider data with a low confidence value as correct. The understanding module and the dialog manager that we have used are modelled by stochastic automata, and some confidence measures are proposed for the understanding module. An evaluation of the behavior of the dialog system is also presented.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Stochastic dialog system; Stochastic language understanding; Speech acts; Confidence measures

1. Introduction

Spoken dialog systems for semantically restricted domains are generally characterized by the following features: telephone access, tasks that are restricted to specific domains, and mixed initiative. Descriptions of some of these systems, which have been developed in the last few years, can be found in (CMU, 1999; Pieraccini et al.,

1997; Lamel et al., 2000; Glass and Weinstein, 2001; San-Segundo et al., 2001a; López-Cózar et al., 2000).

The development of spoken dialog systems, such as natural language queries to database systems, involves several knowledge sources which must cooperate in order to answer user queries. Generally, a spoken dialog system consists of several modules which represent these different knowledge sources: the recognition module, the understanding module, the dialog manager, the answer generator, etc. One of the main problems which must be considered by the dialog manager is the propagation of errors through the different modules. The recognition module must deal with the effects of spontaneous speech and with noisy environments; consequently, the sentence provided by this module could incorporate some errors. The understanding module could also add its own errors (which are mainly due to the lack of coverage of the semantic domain). Finally, the semantic representation provided to the dialog manager might also contain certain errors.

The dialog manager can apply different strategies to deal with these possible errors. It can use implicit confirmations, but this strategy is only effective with highly reliable data, because implicit confirmation of erroneous data generates confusion in the user. It can use explicit confirmations on all the main items of the task: ''origin'', ''destination'' and ''date'', for example, in the case of queries to an information system with information about train timetables. However, this kind of dialog system strategy leads to very long dialogs because the system uses many turns in explicit confirmations. To avoid this drawback, the dialog manager can be designed so that it makes use of explicit confirmations only when relevant uncertainties in the recognition and/or understanding processes are detected. Therefore, it would be desirable to provide the dialog manager with information about what parts of the user utterance have been clearly recognized and understood and what parts have not. From this information, the dialog manager could direct its strategy to confirming or asking only about the uncertain parts of the user utterance. The need to detect errors or data with low reliability makes the association of confidence scores to each component of the decoding (acoustic, lexical or semantic) relevant.

In the literature, several approaches have been proposed to detect and handle the errors generated in the recognition and understanding processes of the input utterances. In the case of the recognition process, supplying several sentence hypotheses (N-best, graphs of words) instead of a single hypothesis, or associating confidence scores with words, allows the following modules of the dialog system to work with some alternatives and to take the reliability of words into account (Wessel et al., 1998; Zhang and Rudnicky, 2001; San-Segundo et al., 2001b; Hazen et al., 2002). Although the confidence scores associated to words in the recognition process are the most frequently used measure of the reliability of the information provided to the dialog manager, it is also interesting to use confidence scores obtained in the understanding process. In (Sturm et al., 1999; Bouwman et al., 1999), the use of confidence measures associated to concepts or attributes in the understanding process has proven to be appropriate to guide the implicit and explicit confirmation strategies in the ARISE information system. There are also other similar contributions such as Hacioglu and Ward (2002) and Raymond et al. (2003). Recently, in (García et al., 2003), we have proposed the definition of two confidence measures in the understanding module. One of them is related to the association of words and concepts, and the other one is related to the probability of sequences of concepts. In these cases, the understanding module provides the dialog manager with the semantic representation (generally one frame or several frames) associated to the user input together with its confidence scores.

In this paper, we study the interest of the use of confidence measures in minimizing the errors in the behavior of a dialog system. We focus our interest on the study of some confidence measures obtained during the understanding process and on their use by the dialog manager to establish its strategy: asking for explicit confirmations, using implicit confirmations, accessing the database system, etc. The novelty consists of the capability of our stochastic dialog system to automatically