Decalog 2007: Proceedings of the 11th Workshop on the Semantics and Pragmatics of Dialogue, pages 171–172.
Trento, Italy, 30 May – 1 June 2007. Edited by Ron Artstein and Laure Vieu.

Adapting a Statistical Dialog Model for a New Domain

Lluís F. Hurtado, David Griol, Encarna Segarra, Emilio Sanchis
Departament de Sistemes Informàtics i Computació
Universitat Politècnica de València, E-46022 València, Spain
{lhurtado,dgriol,esegarra,esanchis}@dsic.upv.es

Abstract

In this paper, we present our current work for adapting a statistical methodology for dialog management within the framework of a new domain. This methodology, that is automatically learned from a data corpus and is based on a classification process, has been previously applied in a spoken dialog system that provides railway information. We summarize this approach and the work that we are currently carrying out to apply it for developing a dialog system for booking sports facilities.

1 Introduction

Within the framework of dialog systems, the application of statistical methodologies to model the behavior of the dialog manager is nowadays a growing research area (Williams and Young, 2007).

In this field, we have recently developed an approach to manage the dialog using a statistical model that is learned from a data corpus (Hurtado et al., 2006). This work has been applied within the domain of a Spanish project called DIHANA (Benedí et al., 2006). The task that we considered is the telephone access to information about train timetables and prices in Spanish. A set of 900 dialogs was acquired in the DIHANA project using the Wizard of Oz technique. This corpus was labeled in terms of dialog acts to train the dialog model.

Currently, we are adapting this methodology in order to develop a dialog manager for a new project called EDECAN. The objective of the ongoing EDECAN project is to increase the robustness of a spontaneous speech dialogue system through the development of technologies for the adaptation and personalization of the system to different acoustic and application contexts. The task that we have selected is the booking of sports facilities in our university. Users can ask for the availability, the booking or cancellation of facilities and the information about his/her current bookings.

2 Dialog management in the DIHANA project

We have developed a Dialog Manager (DM) based on the statistical modelization of the sequences of dialog acts (user and system dialog acts). A detailed explanation of the dialog model can be found in (Hurtado et al., 2006). A formal description of the proposed statistical model is as follows:

We represent a dialog as a sequence of pairs (*system-turn, user-turn*):

$$(A_1, U_1), \cdots, (A_i, U_i), \cdots, (A_n, U_n)$$

where A_1 is the greeting turn of the system, and U_n is the last user turn. We refer to a pair (A_i, U_i) as S_i, the state of the dialog sequence at time i.

The objective of the dialog manager at time i is to generate the best system answer. This selection, that is a local process, takes into account the previous history of the dialog, that is to say, the sequence of states of the dialog preceding time i:

$$\hat{A}_i = \underset{A_i \in \mathcal{A}}{\operatorname{argmax}}\, P(A_i | S_1, \cdots, S_{i-1})$$

where set $\mathcal{A}$ contains all the possible system answers.

As the number of all possible sequences of states is very large, we defined a data structure in order to establish a partition in the space of sequences of states (i.e., in the history of the dialog preceding time i). This data structure, that we call Dialog Register (DR), contains the concepts and attributes provided by the user throughout the previous history of the dialog. Using the DR, the order in which the user provided the information is not taken into account, and the selection of the best system answer is made using this maximization:

$$\hat{A}_i = \operatorname*{argmax}_{A_i \in \mathcal{A}} P(A_i | DR_{i-1}, S_{i-1})$$

The last state (S_{i-1}) is considered for the selection of the system answer due to the fact that a user turn can provide information that is not contained in the DR, but is important to decide the next system answer. This is the case of the task-independent information (*Affirmation*, *Negation* and *Not-Understood* dialog acts).

The selection of the system answer is carried out by means of a classification process, in which a multilayer perceptron (MLP) is used. The input layer holds the codification of the pair (DR_{i-1}, S_{i-1}) and the output of the MLP can be seen as the probability of selecting each one of the 51 different system answers defined for the DIHANA task.

3 Our present work

The task defined for the EDECAN project is to provide an oral interface for booking sports facilities. The main difference with regard to the DIHANA task is that now the dialog manager not only provides information but also modifies the application data (i.e. after booking or cancelling a court). The module that controls the booking application (*Application Manager, AM*) performs two main operations. On the one hand, this module has to perform the queries to the database. On the other hand, it has to verify if the user query follows the regulations defined for the task (a user can book only one court a day, the facilities can not be booked if the user is suspended, etc.).

The result of the queries to the AM has to be considered in order to generate the system answer. For instance, in order to book the facilities (i.e. a tennis court), if there is not any available court, the system can suggest a change in the user restrictions (i.e the AM verifies if it is possible to perform the booking changing the hour). In case of only one court available, the system confirms if everything is correct before making the booking. Finally, if there is more than one available court, the system asks which court has to be booked.

In order to use the information provided by the AM for selecting the system answer, we consider that two phases are needed. In the first phase, the information contained in the DR and the last state S_{i-1} are used to select the best request to be made to the AM ($\hat{A}_{1_i}$):

$$\hat{A}_{1_i} = \operatorname*{argmax}_{A_{1_i} \in \mathcal{A}_1} P(A_i | DR_{i-1}, S_{i-1})$$

where $\mathcal{A}_1$ is the set of possible requests to the AM.

In the second phase, the final system answer ($\hat{A}_{2_i}$) is generated taking into account $\hat{A}_{1_i}$ and the information provided by the AM (AM_i):

$$\hat{A}_{2_i} = \operatorname*{argmax}_{A_{2_i} \in \mathcal{A}_2} P(A_i | AM_i, A_{1_i})$$

where $\mathcal{A}_2$ is the set of possible system answers.

A preliminary evaluation of this approach has been made by labeling the person-to-person dialog corpus and defining a training and test partitions. Currently we are working in the development of the different modules in the system in order to carry out a supervised acquisition with real users, using the Wizard of Oz technique but also evaluating the automatic answers provided by the DM. A user simulator has also been developed to test and improve the behavior of the DM.

Acknowledgements

Work partially supported by the Spanish MEC and FEDER under contract TIN2005-08660-C04-02.

References

J.M. Benedí, E. Lleida, A. Varona, M.J. Castro, I. Galiano, R. Justo, I. López, and A. Miguel. 2006. Design and acquisition of a telephone spontaneous speech dialogue corpus in Spanish: DIHANA. In *Proc. of LREC'06*, Genove.

L.F. Hurtado, D. Griol, E. Segarra, and E. Sanchis. 2006. A stochastic approach for dialog management based on neural networks. In *Proc. of Interspeech'06-ICSLP*, pages 49–52, Pittsburgh.

J. Williams and S. Young. 2007. Partially Observable Markov Decision Processes for Spoken Dialog Systems. In *Computer Speech and Language 21(2)*, pages 231–422.