Efficient Natural Language Access to Databases: the TAMIC-P System

Bernardo Magnini, Alessandro Artale, Mihai Huber, Carlo Strapparava and Massimo Zancanaro
ITC-Irst, Istituto per la Ricerca Scientifica e Tecnologica
Via Sommarive 18, 38050 Povo (Trento), Italy
magnini@itc.irst.it

Abstract

TAMIC-P proposes an interaction style in which direct manipulation and natural language are strictly integrated, each of them providing a communication style complementary to the other. The system has been designed to be used by a Public Administration operator during an information session with a citizen. On field evaluation demonstrated the high usability of the system even with a short training phase.

Introduction

The demonstration shows the TAMIC-P (Transparent Access to Multiple Information for the Citizen - Pensions) prototype system, a natural language based interface to Public Administration databases, developed in the framework of the European Union project LE-4253. Both an Italian and an Austrian version of the system have been designed and implemented.

The system proposes an interaction style in which direct manipulation and natural language are strictly integrated, each of them providing a communication style complementary to the other [Cohen 1992]. TAMIC-P provides a number of tools to limit the typical problems of natural language interfaces to databases, such as the opacity of both the linguistic and the conceptual coverage, and the user false expectations about the "intelligence" of the system [Androstopulos et *al*. 1995]. The following are relevant features implemented in TAMIC-P:

- linguistic flexibility (synonyms, complex terms, temporal expressions, etc.);
- use of abbreviations;
- contextual interpretation;
- flexible presentation of the information;
- integration among the citizen personal data, the general norms and the explanation of technical terms;
- hypertextual navigation through the domain knowledge;
- error administration;
- graphical representation of the history.

The system has been connected to several archives of INPS (Istituto Nazionale per la Previdenza Sociale) the bigger Italian institution for contributions and pensions, serving about 19 million working people and several million retired people. The system is conceived as a tool to support INPS operators in providing information to the citizens.

1 Architecture

The TAMIC-P modules are grouped in four layers, whose relations are sketched in figure 1: (i) the graphical interface layer defines the possibilities of the user to interact with the system and is in charge for displaying the output to the user; (ii) the natural language layer is responsible for the linguistic processing of the user's request, including the management of the context; (iii) the presentation layer, which, depending on the different kind of information, decides the more appropriate presentation to be shown to the user; (iv) the data base access layer is responsible of the data retrieval from user physical databases. Tamic-P information sources are mainly of four types: lexical information, grouped in a lexical knowledge base, used for linguistic inferences; domain oriented information, which is split in two sources for architectural reasons: the Logical Data Model, organised around entity relationship principles, and the Conceptual Data Model, organised around the notion of frame; finally, a

Textual Data Base which contains the relevant norms to be displayed to the user.

Figure 1: Tamic-P main modules.

2 The Demonstration

The demonstration shows a typical working session of an INPS operator who has to address a citizen problem. TAMIC-P is based on the idea of consulting *information cards*. In order to accede to a card through natural language it is sufficient to insert a short description of its content in the suited space of the graphical interface (Figure 2). For example, as answer to an input sentence like “voluntary contributions between 1980 and 1990” a list of contributions of the above kind and for the indicated time period is obtained.
Other information can be found at the TAMIC-P site: http://ecate.itc.it:1024/projects/tamic-p.

Acknowledgements

This work has been partially sponsored by the European Union under the LE-TAMIC-P 4253 project.

References

Androutsopoulos I., Ritchie G.D. and Thanisch P. (1995) *Natural Language Interfaces to databases - an Introduction*. Natural Language Engineering 1(1).

Cohen, P. R. (1992) *The Role of Natural Language in a Multimodal Interface*. In Proceedings of UIST’92, pp. 143-149. ACM Press, New York