

MULTIMODAL INFORMATION SEEKING DIALOGUES ON THE WORLD WIDE WEB

José Rouillard, Jean Caelen

Laboratoire CLIPS-IMAG, équipe GEOD
Université Joseph Fourier, Campus Scientifique, BP 53 38041 Grenoble Cedex 9 - France
Tél.: 33 + (0)4.76.63.56.51 and 33+ (0)4.76.51.46.27 - Fax: 33+ (0)4.76.63.55.52
E-mail: { Jose.Rouillard, Jean.Caelen }@imag.fr

ABSTRACT

We have developed the Halpin system to implement our multimodal conversational model for information retrieval. This dialogue-oriented interface allows the access to a digital library database, on the internet, in a natural language way, and gives its oral responses via usual browsers. The results of the first experiments show that the Halpin system provides some interesting dialogues (in particular with the beginners), according to the user's goals and skills, that leads to information retrieval success, while searches with the original user interface (traditional web form) failed.

1. INTRODUCTION

While seeking a document or information, "some are looking for the ocean and some others for a grain of sand" [6]. In fact, seeking relevant information in a large database is not an easy task. There are many research works about information retrieval, but very few in which the natural language (NL) plays an important role. Most of the classical user interfaces and search engines try to improve the efficiency of the search task with better indexing and retrieval methods. In such models, the human-machine interaction is limited to an exchange of the type : query/database_access/reply. As information seeking and retrieval are interactive processes, we believe that providing a flexible and cooperative human-machine dialogue is a complementary way to improve information retrieval systems [1]. An intelligent conversational system must be capable of adapting itself to the user's goals and skills, interpreting speech acts within the context and

negotiating ambiguous information using a NL interface [12]. In our previous papers, we have also shown that it is possible to gather interesting human-machines dialogues on the Web without using the Wizard of Oz strategy [9]. Having in mind these important observations in order to show that NL dialogue systems may improve interaction quality, we proceeded to create an interactive search and navigation environment to incorporate adaptability and conversational capabilities to an existing digital library information research system (the Inria's database counts 83297 documents available). In the following we will present the Halpin system and its architecture.

2. THE HALPIN SYSTEM

Our work is based upon results of the ORION project [7] which is about new multimodal technologies for Web based navigation and information research [8]. The Halpin system uses Xerox's morphological tools [5] to convert the sentence given by the user to a canonical form which may be analysed more easily by the Halpin concept detection module. It also uses the Elan Informatique speech engine [4] to synthesise its answers which will be sent to a Java applet in the user's browser which will produce audio output of the given answer. Inspired from the works of Brun [2] our dialog manager uses a two steps algorithm of concept recognition which leads to understand the user queries.

2.1. The Cooperative Model

Our goal was to propose a system not only responding to the users sentences, but also proposing related information (similar authors or keywords) depending on needs of the user.

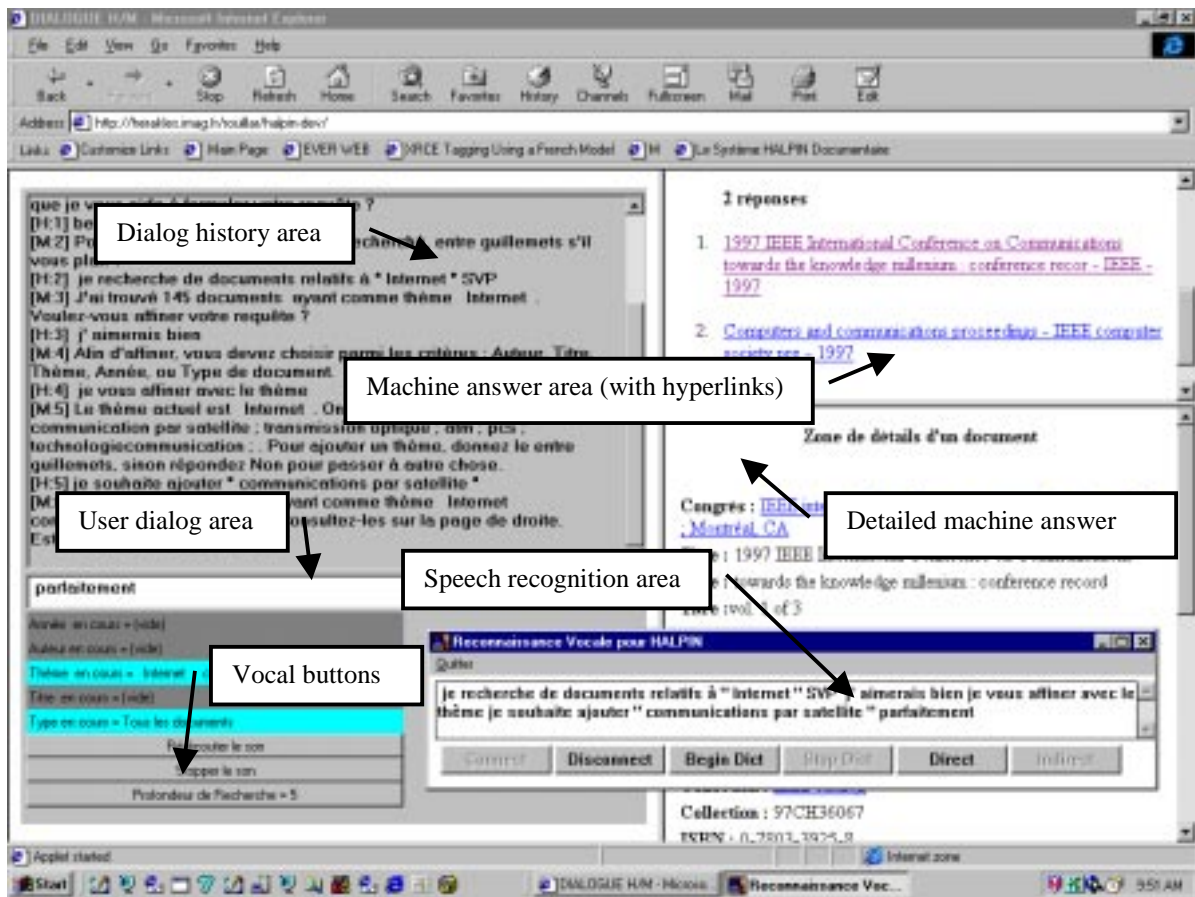


Figure 1 : The HALPIN system interface in a World Wide Web browser

This is why at the beginning of the interaction with the system, we have to determine the user profile (novice or expert) and her aim (finding an already known paper, searching an unknown set of books, discovering the site, etc.). The Figure 1 shows the multimodal interface of the Halpin system.

The COR (Conversational Roles) model of [11] proposed typical *Ideal* and *Alternative* dialogue sequences (cycles). For example, a dialogue between A (information seeker) and B (information provider) can be formalised as : Dialogue (A,B) => request (A,B) promise (B, A) inform (B,A) be-contented(A,B) or also Dialogue (A,B) => offer (B,A) accept (A, B) inform (B,A) be-contented(A,B).

In the same way, our model is a kind of conversational roles and tactics (COR) model augmented with the knowledge about the user and her aims, so that the model can react according to the user profile and the task in progress. We propose, for a finalized and co-operative defined task to follow the rule:

[Profile]. [Goal]. [Speech Act]. [Concepts]. [Task] => [Reply]. [Justification]. [Suggestion].

The concepts database is divided into different files, according to the type of concept which they contain. Indeed, certain concepts are common to all the possible tasks (acceptance, refusal ...) and others are specific to the task (searching information in a digital library for instance). If a sentence is ambiguous, even when the goal of the user is known, the system ask for choices. For example, this French sentence : “je veux un livre de Boole“ (I want a book of Boole) can be interpreted in two different way : (a) The user wants a book talking about Boole. (b) The user wants a book written by Boole. The first interpretation gives 100 responses, while the second gives 3 responses. So, we think that, rather than asking the database with an uncertain query, it’s better to resolve the ambiguity in a co-operative way.

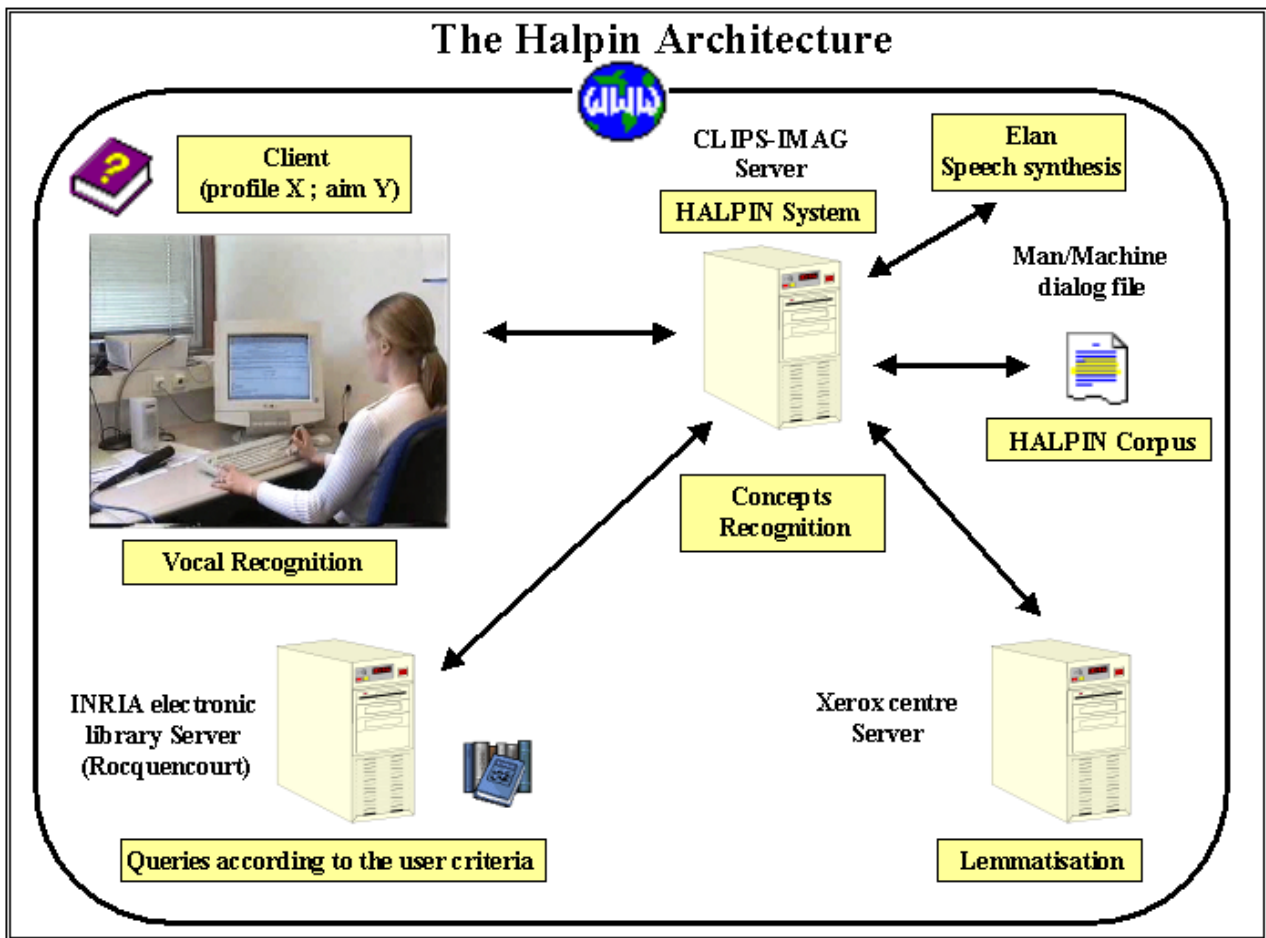


Figure 2 : The Halpin system architecture for a multimodal information retrieval on the Web

2.2. The Halpin Architecture And Functionalities

The Halpin system was developed upon C, Java and Perl and HTML languages. The users can hear the system responding to their questions thanks to a software installed on our Web server which synthesises textual responses to an audio file. This audio file is sent to the browser, then played by a Java applet. The dialogue manager allows not only entries related to the current task, but also about the interface (screen, sound, speech synthesis) and system responses, called meta-information. The system tries to understand, according to the context and found concepts, if the user is speaking about the task (ex "The author is Turing"), about the interface (ex "stop the speech please") or about the meta-information (ex "Why do you ask me that?"). The Figure 2 represents the Halpin architecture. Each module is described in the following.

2.3. Modules Description

- *Speech recognition* : It's the IBM ViaVoice system. This module must be installed on the client machine. It delivers in output a string that the user can still modify. And, if he does not have this software, the user can still enter all the requests by the keyboard.
- *The morphological analyser*: it's the Xerox one [5]. The Halpin system access it via Internet. This module delivers the lemmatised form of each word of the input. So the canonical sentence is more easy to be analysed by the concept recognition module.
- *Comprehension by concepts* is a module developed within the framework of Halpin [10]. It provides a structure from a conceptual analysis (lemmatised) source. This module is carried out on a server of the laboratory.
- *The dialogue controller* was also developed within the framework of Halpin. It makes

possible to manage the dialogue and the current task, according to the goal, skills of the user and dialogue history.

- *The answer generation* is a module which is based on a model of “holes” generation.
- *The voice synthesis* is the system of the French company “Elan Informatique” [4]. This module runs on a server of our laboratory. Contrary to the recognition, the user does not need to install it on his own machine.
- *The graphic interface* runs under a browser like Internet Explorer or Netscape. A Java Applet displays all the buttons, textarea and windows, while a Perl Module displays the hyperlinks.
- *The access to the INRIA’s database* is done via Internet.

2.4. Results

Some interesting dialogues gathered with this system show that a cooperative and relevant multimodal man-machine dialogue can be a solution to the users problems encountered on the Web. For example, a user believes the name of the author she’s searching is Krakoviak, but this name gives no answer. In a traditional Web information retrieval task, the search stops here. But, with Halpin, after a negotiating phase, the computer finally gives the relevant name of the author (Krakowiak with a W). The user finds 15 documents, and tell the machine that it was exactly what she was looking for.

3. CONCLUSION AND FUTURE WORK

The Halpin system is currently used by many people on the Web. The first results show that the users readily co-operate with the machine. This kind of multimodal natural language interaction, is a valid answer to the scientific problems, known in information seeking interfaces and particularly in hypertext environments, which are confusion, cognitive overload, and relevance of the answers. To go further, we wish to integrate a voice recognition module to the system, in order to allow the users to freely dialogue with the

machine, in a more natural and more effective way. The integration of a powerful thesaurus is also expected, for a broader cover of the vocabulary used, as well in input as at output.

4. REFERENCES

[1] BATEMAN, J. A.; HAGEN, E., & STEIN, A. *Dialogue modeling for speech generation in multimodal information systems*, in P. Dalsgaard, et al. (Ed.), *Proceedings of the ESCA Workshop on Spoken Dialogue Systems - Theories and Applications* (pp. 225-228). Aalborg, Denmark: ESCA/Aalborg University, 1995.

[2] BRUN, C., *A Terminology Finite-State Preprocessing for Computational LFG*. 36th International meeting of the Association for Computational Linguistics & 17th International Conference on Computational Linguistics, Montreal, Quebec, Canada, August 1998.

[3] CONKLIN, J., *Hypertext: an introduction and survey*, IEEE Computer, pp. 17-41, September 1987.

[4] <http://www.elan.fr>

[5] GAUSSIER, E., GREFFENSTETTE, G., SCHULZE, M., *Traitement du langage naturel et recherche d’informations : quelques expériences sur le français*. Premières Journées Scientifiques et Techniques du Réseau Francophone de l’Ingénierie de la Langue de l’AUPELF-UREF, Avignon, Avril 1997.

[6] HARDIE, E., *A grain of sand or the ocean ; User aims in search engine interactions*. Fifth International WWW Conference - Poster Proceedings, INRIA/CNIT, Paris La Défense, May 1996.

[7] <http://www.gate.cnrs.fr/~zeiliger/Orion99.doc>

[8] ROUILLARD, J. et CAELEN, J. *A multimodal browser to navigate and search information on the Web*. Fourteenth International Conference on Speech Processing (ICSP97), IEEE Korea Council, IEEE Korea signal processing society. August 1997, Seoul, Korea.

[9] ROUILLARD, J., *Hyperdialogue Homme-Machine sur le World Wide Web : Le système HALPIN*, ERGO’IA 98, Biarritz, Novembre 98.

[10] ROUILLARD, J. et CAELEN, J., *Etude du dialogue Homme-Machine en langue naturelle sur le Web pour une recherche documentaire*, Deuxième Colloque International sur l’Apprentissage Personne-Système, CAPS’98, Caen, Juillet 98.

[11] STEIN, A. & MAIER, E., *Structuring collaborative information-seeking dialogues*, *Knowledge-Based Systems*, 8(2-3, Special Issue on Human-Computer Collaboration): 82-93., 1995.

[12] STEIN, A., GULLA, J. A., MÜLLER, A. & THIEL, U., *Conversational interaction for semantic access to multimedia information*, in M.T. Maybury (Ed.), *Intelligent Multimedia Information Retrieval* (pp. 399-421). Menlo Park, CA: AAAI/The MIT Press, 1997.