FI MU

Control of the Hypertext System Audis Using the Dialogue

by

Pavel Gaura

Control of the Hypertext System Audis Using the Dialogue

Pavel Gaura

Department of Information Technology, Faculty of Informatics
Masaryk University Brno, Czech Republic
gazi@fi.muni.cz

Abstract. The brief description of the blind user oriented hypertext system AUDIS and the principals of the dialogue approach to the system control are described in the paper. The emphasis is placed on the improvements to the control of the system using the dialogues. The system is developed at the Faculty of Informatics, Masaryk University, Brno. AUDIS is developed primarily as a support that would help visually impaired students to study various materials (see [3], [1] and [2]).
Key words: hypertext, multimodality, dialogue, VoiceXML, visually impaired

1 Introduction

Large amount of electronical educational textbooks contain special structures, e.g. tables, graphs, pictures or mathematical expressions. Reading such structures by the screen reading software, impaired users have often problems to understand and to interpret the content correctly. Studying a mathematical textbook is not easy even for sighted student, but for blind person without a sighted co-worker is it almost impossible. Therefore, we develop the hypertext system that will be used by the handicapped people to get information quickly and comfortably. We have built the prototype of the hypertext system that uses a keyboard, voice commands and a mouse as an input and the speech synthesis, earcons, background sounds and additional graphics as an output. The prototype that is still in the development, serves as an textbook for visually impaired students. An overview of the implemented functions and features are described in the next chapter.

The principles of the system input has been designed in the way that would allow the quick system control and an interactive document navigation. Generally, the dialogue access to the

system can slow down the interaction, but sometimes it can be more comfortable. Therefore, we are trying to find the possibilities of controlling the system using the natural language dialogue while keeping the original approach or to combine these two approaches.

2 Brief Description of the System

Main goal of the system is to afford the access to the technical texts and electronical documents to the blind user. Additionally, the system allows easy navigation of the user in the document and facilitates his/her orientation in the text.

AUDIS can be controlled via keyboard, voice commands and mouse; an additional graphical output of the spoken text is also implemented in the system. It means that the system can be used by the shortsighted, visually or manually impaired users.

The document in AUDIS is divided into chapters that makes the passing of the whole document more synoptically. There are several types of structures that appears in chapters: common text, links, lists, tables, pictures and mathematical expressions. When presenting the document, AUDIS reads the spoken form of the text. This text is created from the content of the structures contained in the particular chapters (see [2] for the details on document representation in AUDIS).

The orientation in the document is facilitated by earcons - short environmental sounds with a special meaning. Earcons help the user to determine the following part of the document or a special structure (e.g. the beginning and the end of a link in a hypertext document). All sounds used in AUDIS are fully customizable according to the user needs.

There is a possibility to spell difficult words or repeat the said text on the user's demand for better understanding of the synthesized speech. More information about the features of the system can be found in [1] or in [2].

3 Current State of the Control of the System

Generally, there are three possible ways for controlling the system.

- The control buttons and the other control elements of AUDIS are large and contrast so that partially sighted user can control the system using mouse or any other pointing device.
- Faster and more comfortable is the control of the system using keyboard. This should be very effective method for advanced users. We tried to define the default definition of keys to be intuitive, but the user can redefine it at any time.
- The third possibility of the control is voice commands recognition that can be used by the manually impaired users as well.

There are the user definable properties for each type of the control that improve the user interaction with the system.

The navigation in the AUDIS document is similar to the movement control in the text editors. In spite of their handicap, visually impaired users work with these applications. Therefore, the same combinations of keys that are used in the applications for the sighted has been chosen for the control of the spoken output in AUDIS.

The exact form of navigation depends on the type of the structure that is currently processed. In the common text, the intuitive navigation is used (e.g. left and right arrows skip the sentence backward/forward and page up, page down keys skip to the previous/next chapter). There are also the commands for skipping or repeating words.

When processing a table, some keyboard commands change its functionality. Pressing up/down arrow moves the position in actual column to the previous/next table row. Left/right arrows do not skip the sentences, but move the position to the previous/next column in the same row. Moreover, the keyboard combinations for accessing the header row or the header column of the table are available from any place inside the table. The behavior of the system changes similarly for the other types of structures.

4   Dialogue Access to the System Functions

Generally, the form of the dialogue for the control of the system Audis will be similar to the access to the menu of a menu driven application, with all the possibilities and advantages of a dialogue

approach. Dialogue starts as a computer initiative one and it can contain the default and user defined shortcuts. Then, the user can follow proposed options (e.g. User shortcuts, Search, Filters, Navigation, Access to menu, Help) or take over an initiative so that he or she can directly say the requested function with its parameters (e.g. "Find me the chapter about VoiceXML" will result in finding the word VoiceXML, terminating the dialogue and starting the document processing at the beginning of the found chapter).

4.1 Information Retrieval

Very important part of the system is navigation inside the document that allows quick information retrieval. For navigation inside the opened document there are not many possibilities how to improve the control commands or how to replace them by using the dialogue approach. The control commands directly accessed via keyboard or voice are faster than use of any type of the dialogue strategy. Therefore, the commands for navigating inside the document will be represented as a group of options in the navigation subdialogue. This fact allows the system to be controlled using the telephone or various remote control devices that are supported in dialogue approach. Finally, all the navigation options will have the same keyboard shortcuts as now the navigation functions of the system have, so that nothing changes for keyboard oriented user.

Unlike the navigation, the dialogue approach can be successfully utilized for the search of the relevant parts in the document. The use of the fluent speech recognition allows the user just to say the words that is he looking for. Then, user can choose some of the attributes of search (e.g. starting position, type of the structure to search in).

The dialogue control of the system can also directly handle the speech synthesis properties. This action will replace current dialogue screen for setting the voice properties for the document processing. Moreover, the user can set the voice properties without interrupting the document processing (e.g. by saying "Slow down, please." or "Volume up!").

4.2   Comfortable Configuration

The dialogue approach can handle the configuration of the system settings efficiently and comfortably. There are three possible dialogue strategies for setting the attribute values in the system. First one is user initiative. The user can directly say the attribute name and its desired value. This will set the system behavior without the need of interrupting the document processing. Second dialogue strategy is mixed initiative. The user wants to set the specified attribute, but he or she does not include the attribute value in his utterance. Then, the system interrupts the document processing and executes the subdialogue for finding the desired value. Third dialogue strategy is computer initiative dialogue that asks the user gradually for attribute names and their values. The dialogue approach to the configuration of the system includes the synoptical listing of the system properties to be set. Usually, the number of system property names is very high. The form of presenting them to the user has to be solved yet, because to read them all is too inefficient, even in some structured form.

4.3   Dialogues for Working with Documents

For working with the document content and the document structure we develop the supporting modules in Audis. The functions contained in these modules can be efficiently controlled using the dialogue approach. The automatic hypertext links generation function can serve as an demonstration example. The system is able to divide the actual document into several parts and to create the hierarchical tree structure from these parts by the occurrence of the words, entered by the user. Then, the user can proceed only into the parts of the document that he or she is interested in (e.g. that contain the subset of entered words and does not contain another subset of entered words). The efficiency of this function decreases with the number of words that has to be entered. Conversely, using the fluent speech recognition for entering words and the dialogue control for navigating in the created structure the user can obtain a very quick access to the relevant parts of the document.

5 Conclusions

Audis is the hypertext system, that allows visually impaired students to get information quickly and comfortably. The use of a dialogue based approach to control the hypertext system seems to be a very efficient way. There are several important advantages of the dialogue based approach:

- The communication based on the speech recognition is the natural one (especially recognition of the fluent speech). User can concentrate on the information he or she obtains instead of concentrating on control.
- The user can be driven by a computer initiative dialogue to control the application using the natural speech input, no matter how experienced he or she is.
- Using the dialogue, the user can obtain information faster.
- Number of various input devices can be used to obtain the information from the document, also remote control devices can be used.

6 Acknowledgment

References

1. Gaura, P., Nygryn, P., "Hypertext System Audis for Visually Impaired Users", proceedings of ISM'2000, pp. 69-73, Roznov pod Radhostem, 2000
2. Gaura, P., "Internal Document Structure of the Blind User Oriented Hypertext AUDIS", proceedings of ICCHP'2000, pp. 237-244, Karlsruhe, 2000
3. Kopecek, I., "The Architecture of Blind User Oriented Hypertext AUDIS", proceedings of ICCHP'98, p.215-223
4. Smith, R.W., Hipp, D., "Spoken Natural Language Dialog Systems - A Practical Approach"
5. Laboratory of Speech and Dialogue - http://www.fi.muni.cz/lsd/
6. VoiceXML Interpreter Elvira - http://www.fi.muni.cz/lsd/elvira

Publications in the FI MU Report Series are in general accessible
via WWW and anonymous FTP:

```
http://www.fi.muni.cz/informatics/reports/
ftp  ftp.fi.muni.cz (cd pub/reports)
```

Copies may be also obtained by contacting:

Faculty of Informatics
Masaryk University
Botanická 68a
602 00 Brno
Czech Republic