ARET – Automatic Reading of Educational Texts for Visually Impaired Students

Martin Grůber[1]*, Jindřich Matoušek*[2]*, Zdeněk Hanzlíček*[1]*, Zdeněk Krňoul*[2]*, Zbyněk Zajíc*[1]

[1]New Technologies for the Information Society (NTIS),

[2]Department of Cybernetics

Faculty of Applied Sciences, University of West Bohemia, Czech Republic

{gruber,zhanzlic,zzajic}@ntis.zcu.cz, {jmatouse,zdkrnoul}@kky.zcu.cz

Abstract

This paper deals with a presentation of an application which was developed to help in education of visually impaired pupils at a secondary school, i.e. at the pupils' age of 12 to 14 years. The web-based application integrates speech and language technologies to make the education easier in several areas, e.g. in mathematics, physics, chemistry or languages (Czech, English, German). TTS system is used for automatic reading of educational texts and it makes use of a special preprocessing of the texts, namely any formulas which may occur therein. The application is used by both teachers to create and manage the teaching material and pupils to view and listen to the prepared material. The application is currently being used by one special school for visually impaired pupils in daily lessons.

Index Terms: speech synthesis, human-computer interaction, education of visually impaired

1. Introduction

Current technologies are very helpful for people with many types of impairments or disabilities. There are various assistive technologies, e.g. for people with cognitive disabilities [1], mobility impairment [2] or hearing loss [3], [4].

For individuals diagnosed with a print disability (learning, visual, or physical disability), the transition from textbooks accessible on tapes to electronic texts read by computers is considered as one of the most effective methods for providing access to textbooks in the last decade [5]. However, this is not always true about automatic reading of e.g. mathematics. Development and implementation of classroom materials for visually impaired and blind pupils attending technical courses such as mathematics, physics or chemistry is a demanding and not comprehensively solved task.

This paper presents a newly developed web-based application ARET (Automatic Reading of Educational Texts for Visually Impaired Students, available online at http://ucebnice.zcu.cz) which aims at innovations and enhancement of schooling of visually impaired students at lower secondary schools. The application is basically composed of two main parts, one to be used by teachers for administration and management of the textbooks, and one to be used by the students for viewing and listening to the educational material.

Compared to conventional TTS systems, the developed application is more specialized and tailored to reading of educational texts. It could be considered as an alternative or a supplement to screen readers. As the educational texts are mainly focused on mathematics, physics and chemistry, a new approach to processing of various formulas was developed within the system framework exploiting the extra information about them. The following school subjects are currently available in the system: mathematics, physics, chemistry, languages (Czech, German, English), biology and work teaching.

The system architecture, i.e. both the administrative interface (backend) and the public interface (frontend), is described in Section 2. In Section 3, the employed TTS system is presented.

2. System architecture

The developed web application is based on a client-server architecture, running on *Apache* HTTP server with *MySQL* database system. The core of the system is based on *Symfony*, an open-source web application framework. The ARET application is composed of two main parts: frontend and backend. Frontend is used by students to view and listen to the textbooks prepared by teachers. Backend is an administrative interface and serves for teachers to create, update and manage the textbooks. The system architecture is depicted in Figure 1. Both backend and frontend are accessible via standard web browsers and thus available to anyone with a compatible device and an internet connection.
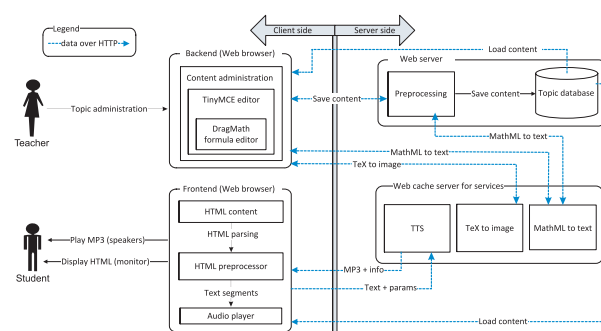

Figure 1: *System architecture.*

2.1. Backend

In the backend interface, teachers are allowed to create and modify the textbooks via a *WYSIWYG* text editor. The textbooks are then stored in an HTML format in a DB. In the ARET application, the text editor is enhanced so that some specific templates and formulas might be inserted in the textbooks. The templates are used to clarify the meaning of particular fragments of the documents and various synthetic voices can be assigned to each template. There are also special templates e.g. for mark-

ing an exercise with a hidden solution which can be shown later or for indicating fragments in languages (English, German) different from the native language (Czech) of the whole textbook. In addition, there are two ways of inserting formulas (not only mathematical but also physical or chemical ones). Simple formulas (called "inline formulas") like $y = x + 1$ are written as a plain text and are marked as a special template to indicate that they must undergo a special preprocessing. More complex formulas are written in the *WIRIS* editor[1] which allows teachers to write such complex formulas easily while providing two representations of these formulas: *MathML* which is used as an input for TTS text preprocessing and TEXwhich is used for converting formulas to images to be displayed in the application frontend.

2.2. Frontend

The frontend (screenshot is depicted in Figure 2) is a public interface where the textbooks are displayed and read to the students. Before displaying, the whole textbook is automatically preprocessed, e.g. the *MathML* formulas are converted to word-level description. Thus, raw text fragments are extracted and sent to our *Web TTS server* with proper settings (language, voice, speed rate, etc.) to be synthesized. The synthetic utterances are then played within the frontend via a JavaScript player using *Adobe Flash* or *HTML5* audio player. Paragraph- and section-based navigation is also supported so that the students can skip to a specific location in the document. The text currently being read is also highlighted.
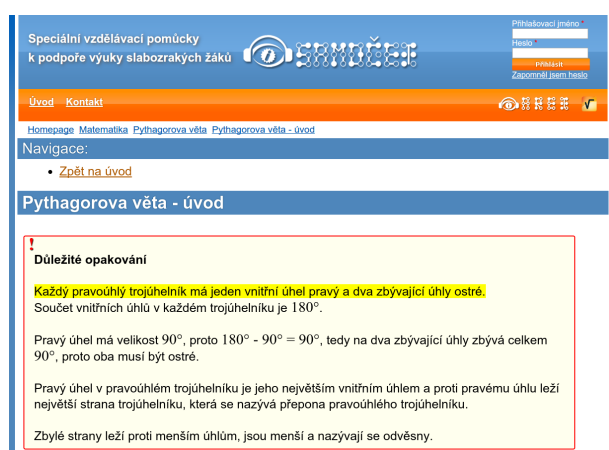

Figure 2: *Screenshot of the frontend.*

3. Text-to-speech

For generating the audio representation of the textbooks in the frontend, a TTS system is employed and available as a web service. Our TTS system ARTIC [6] manages several Czech male and female voices as well as two male voices for English (one American English and one British English), a female voice for the Slovak language and a female voice for the Russian language. Thus, it is used as the main TTS system for the ARET application as the textbooks are mainly in Czech. In textbooks for teaching foreign languages, ARTIC can be easily replaced by another TTS system for the relevant fragments as the communication protocol is very simple and easy to adapt. Currently, for the German language, MaryTTS system[2] [7] is used.

During the speech synthesis, the input text is a subject of a thorough analysis and preprocessing. Within the ARET application, this task also includes preprocessing of various formulas (mathematical, physical and chemical), symbols, tables, and other non-standard notations. Generally, reading formulas is a very complex task, especially if there is no limitation for the complexity of the equation structure. Moreover, Czech is an inflective language, thus all operands in the formula should be converted into the correct grammatical form (which can differ in various mathematical contexts).

4. Conclusions and future work

This paper presents an application which employs TTS technologies for automatic reading of educational textbooks for visually impaired students. As the textbooks contain many non-standard parts like mathematical formulas, the text preprocessing is a more complex task then in standard TTS systems. In addition, fragments in different languages may appear in textbooks for teaching foreign languages and this task was solved by putting these fragments into a special templates which indicate the language to be used. Thus, the proper settings of our ARTIC TTS system or the MaryTTS system can be used to synthesize these fragments.

Future work might be focused on ensuring compatibility with other tools for the visually impaired, e.g. Braille Displays, or on integrating voice control.

5. Acknowledgements

6. References

[1] M. Dawe, "Desperately seeking simplicity: how young adults with cognitive disabilities and their families adopt assistive technologies," in *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, Montreal, Canada, 2006, pp. 1143–1152.

[2] J. Přibil and A. Přibilová, "Czech and Slovak speaking voice communicator based on PDA/smartphone device for handicapped people," in *Proc. of the International Conference on Applied Electronics*, Pilsen, Czech Republic, 2012, pp. 219–222.

[3] A. Pražák, L. Müller, J. V. Psutka, and J. Psutka, "Live TV subtitling – fast 2-pass LVCSR system for online subtitling," in *Proc. of the International Conference on Signal Processing and Multimedia Applications*, Barcelona, Spain, 2007, pp. 139–142.

[4] M. Hrúz, P. Campr, E. Dikici, A. A. Kindiroglu, Z. Krňoul, A. Ronzhin, H. Sak, S. Daniel, Y. Hulya, A. Lale, A. Oya, A. Karpov, M. Saraclar, and M. Železný, "Automatic fingersign-to-speech translation system," *Journal on Multimodal User Interfaces*, vol. 4, pp. 61–79, 2011.

[5] P. Lewis, S. Noble, and N. Soiffer, "Using accessible math textbooks with students who have learning disabilities," in *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2000)*, Orlando, Florida, USA: ACM, 2010, pp. 139–146.

[6] J. Matoušek, D. Tihelka, and J. Romportl, "Current state of Czech text-to-speech system ARTIC," in *Text, Speech and Dialogue*, ser. Lecture Notes in Artificial Intelligence, vol. 4188, Berlin, Heidelberg: Springer, 2006, pp. 439–446.

[7] M. Schröder, M. Charfuelan, S. Pammi, and I. Steiner, "Open source voice creation toolkit for the MARY TTS platform," in *Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech 2011)*, Florence, Italy, 2011, pp. 3253–3256.

---

[1]http://www.wiris.com/en/editor

[2]http://mary.dfki.de