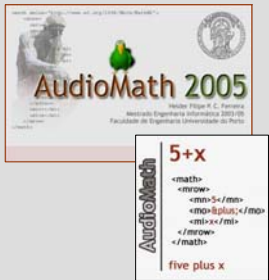


AudioMath – Speaking MathML for the Blind

Helder Ferreira
helder.filipe@sapo.pt
FEUP, Portugal

Diamantino Freitas
dfreitas@fe.up.pt
FEUP, Portugal



Introduction

The AudioMath project (<http://lpf-esi.fe.up.pt/~audiomath>) aim is to build a tool to operate either as standalone or integrated in a speech interface (TTS – text-to-speech), capable of:

- Mathematics Audio Rendering**
 - Parsing, interpretation and conversion of MathML into plain text format
 - Generation of the appropriate prosodic contour for reading of the math formula's text
 - An intra-formula browsing device (navigation)
- Recognition and conversion of any other text or markup elements not directly "understandable" by the TTS engine**, such as: numerals, abbreviations, acronyms and network references.

State of the Art

Publishing Scientific e-Contents

- Use of images to represent math expressions
- Use of alternative formats: PDF, TeX, Postscript, Microsoft Word, RTF
- Use of Java applets that create math graphical representations
- Use of plug-ins to represent graphical math expressions
- Use of HTML and Symbol Fonts
- Use of XHTML and CSS
- Use of markup languages: MINSE, **MathML**, SVG and others

Speaking Mathematical Expressions

- ASTER from T.V. Raman
- MathTalk from Robert Stevens
- MathSpeak from Abraham Nemeth
- MathGenie from Arthur Karshmer
- MathPlayer from Design Science
- LAMBDA Project
- **AudioMath Project**

MathML – Mathematical Markup Language

$$(a + b)^2$$

Presentation Markup	<pre><math> <msup> <math> <mi>a</mi><mi>b</mi> </math> </msup> </math></pre>
Content Markup	<pre><math> <msup> <math> <mi>a</mi><mi>b</mi> </math> </msup> </math></pre>

AudioMathENGINE – Analysis, Interpretation and Conversion

- Developed in PERL (available for Windows and Linux/Unix)
- Available in 4 different ways:
 - ActiveX DLL
 - **.NET component**
 - CGI interface
 - Executable EXE
- Modular, flexible and extensible architecture
- Several logic modules for:
 - numerals, abbreviations, acronyms, network references
 - **mathematical expressions** and their **intra-navigation**
 - automatic recognition
- Currently only in *European Portuguese* (new languages under construction)
- Usage of a internal markup language – **AKML** (AudioMath Knowledge Markup Language)
- Native support for **MathML Presentation Markup**
- Support for **MathML Content Markup**
- Partial support for **Unicode**

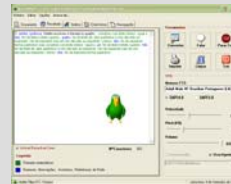
AudioMathGUI – Graphical Interface for Demo and Simulate Navigation



Document View to open or edit XHTML + MathML



Graphical View to visualize the document



Result after conversion



Speaking Math Expression



Stats regarding the conversion and AKML tree view



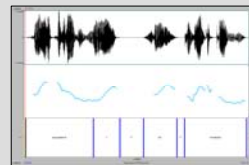
Navigation on a Math Expression

Mathematical Prosody

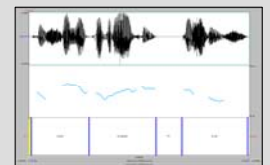
A mathematical corpus has been defined to study mathematical prosody, including: algebra, fractions, roots, powers, trigonometry, matrices, integrals, derivatives and sums.

This corpus is composed by 74 chosen expressions (out of possible 200).

This study focused more on the pauses. Several conclusions were extracted from the recorded waves (like the existence of 2 types of pauses - short pauses and long pauses – and the influence of operators in these pauses).



"square root of" + long pause + "a" + long pause + "end" + short pause + "of radicand"



"cosine squared" + long pause + "of x"

Test and Validation of Results

-Objective tests:

- AudioMathENGINE had some errors in the identification and conversion of the text element to be converted. This was caused by the lack of more dictionary entries and the need of some enhancements on the automatic recognition engine.
- Very good results converting large documents (online and offline).
- Converted documents are bigger (since they are marked with XML).

- Subjective tests:

- About 14 users were tested (2 of them PhDs in Mathematics and other 4 were blind users).
- 3 types of tests:
 - Written interpretation
 - Oral interpretation
 - Math Navigation
- Very good results on the first two tests. All the users understood the math expressions.
- Navigation worked but needs some enhancements (mainly on keyboard choices).

Future Work

- Complete the support for MathML Content Markup
- Study in more detail mathematical prosody
- Implement a proper blind tool
- Add more languages
- Enhancements on XHTML support
- Implement SAPI, SSML support for TTS technologies