

Copyright © 2012 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Mathspeak: An Audio Method for Presenting Mathematical Formulae to Blind Students

Azadeh Nazemi, Iain Murray and Nazanin Mohammadi
*Electrical and Computer Engineering Department
Curtin University, Perth, Western Australia*

Azadeh.nazemi@postgrad.curtin.edu.au, i.murray@exchange.curtin.edu.au and n.mohammadi@curtin.edu.au

Abstract - This paper describes the problems involved with learning and understanding math for vision impaired students and developing a computer system approach for rendering mathematical formulae into audio form. Access to mathematics is an obstacle for blind students. The lack of easy access to mathematical resources is a barrier to higher education for many blind students and puts them at an unfair disadvantage in school, academia, and industry [1]. Results from the National Assessment of Educational Progress show that there is great disparity between the math skills of students with disabilities and students without disabilities [2]. A methodology for rendering technical documents, in particular, complex mathematical formula, in an audio descriptive form (Mathspeak) is presented in this paper.

Keywords-component; Vision-impaired; Mathematics formulae; Braille; DotPlus; MathML; ASTER; LaTeX; Amsmath

I. INTRODUCTION

Reading and writing mathematics is inherently different from reading and writing text. Mathematics can even be considered a language of its own [3].

The task of presentation of mathematical formula in an accessible form is very complex. The established techniques to help visually impaired students deal with mathematical equations typically fall into the following general categories.

1) *Tactile methods*: for example 6-dots (Braille) and 8dots (Dots Plus) which are physically raised representations of characters on a flat surface. Additionally, the Pictures In A Flash (PIAF) machine produces high quality tactile graphics. The raised images are easy to produce and ideal for publishers/printers.

2) *Audio methods*: Audio aids read equations to blind students, using tools which synthesize speech as output.

II. TACTILE METHODS

Conveying mathematical equations in Braille is done through the use of specialized math Braille codes. The Braille notations for mathematical characters have no international standards so printing math in Braille can only be done by local specialists [4].

This manual process can take a lot of time, causing significant delays for the person requiring access to the math. For example, Nemeth Braille code is a generally accepted method for conveying mathematics in the United States. The Nemeth Braille Code is used to encode mathematical and scientific

notation linearly, using standard six-dot Braille cells for tactile reading by the visually impaired. The Nemeth Braille Code is an example of a compact human-readable markup language [5].

While Braille is suitable for text representation, which is linear in nature, mathematical formulae are not entirely linear.

Linearity and limitation of character sets are problematic in the use of Braille for mathematical formula representation. In comparison with linear-nature text, mathematical equations are multi-dimensional; they may contain fractions, superscripts and subscripts, simple expressions, algebra, square roots, logarithms, series, sums, products, matrices, integrals, limits, trigonometry, and exponentiations. Thus by increasing complexity of mathematical equations, Braille representation complexity is increased. Additionally, Braille text is represented by a limited number of characters which include upper and lowercase letters, the 10 digits, various punctuation marks and a small set of special characters, represented by the $64=2^6$ combinations of possible dot placements, meaning normal 6-dot Braille can only represent 64 unique characters. Equations may contain any of the normal text characters, plus a large number of special characters. It is obvious that Braille math representation with a 64 character set is very difficult.

This limitation causes Braille difficulties in dealing with mathematical characters and distinguishing between numerous meanings for a single cell [6].

Another tactile method is 8-dot Braille or Dots Plus representation. By using 8 dots, the basic character set is extended from 64 to 256 characters [7], additionally, most 8-dot systems attempt to have some sort of graphics associated with the dot pattern [8], allowing for clearer representation. However, learning and remembering this character set may be problematic for users. Dots Plus is a two-dimensional spatial tactile mathematics notation, which is a composite of standard Braille with raised lines and symbols. Creating a Dots Plus document requires embossing, which can be done on a Tiger Braille Printer [14].

Braille math materials can often be created from electronic files with Braille translation software. However, since many of the symbols and spatial typesetting conventions used in mathematical equations cannot be replicated with the American Standard Code for Information Interchange (ASCII) character set, nor recognized by standard Optical Character Recognition (OCR) systems, additional processing is often required. Tactile output is also very helpful for blind math

students in representing the graphical output of a function as shown by Haptic technology [9], but this technology is still too expensive to be widely implemented.

III. AUDIO METHODS

Most publishers do not generally produce specialized alternative-format materials such as Braille. Many textbook publishers create electronic versions that can be used with many types of assistive technologies [10]. In this method, once a submission has been accepted for publication, text and math are converted to a markup language which is then used to create the article for printing. For example, Mathematical Markup Language (MathML) is a format for universally designed math. MathML can be used for math equations within standard of extensible Hyper Text Markup Language (XHTML) or extensible Markup Language (XML) content. Unlike image formats, using MathML enables mathematics to be served, received, and processed in digital environments such as the World Wide Web. MathML also provides standard tagging content with an information structure that can be navigated by assistive technology in a comparable manner to disabilities. MathML equations can be accessed via standard assistive technology applications, such as synthetic speech technologies. Text to Speech (TTS) technology can generate the audio form of an equation provided by MathML. This type of audio has been one of the most popular and successful output media for blind students in understanding the process of mathematics. Audio System for Technical Readings (ASTER) and Talking Emacs (EmacSpeak) are two projects aiming to represent math to blind students through audio output. ASTER reads complex equations with special intonation to emphasize parts of the equation [11]. EmacSpeak reads C code to the user, providing descriptive information about the code in addition to the basic program material [12]. Since audio output forms of complicated maths formulae are difficult to comprehend for vision impaired students a comprehensive system of notation, capable of expressing all mathematical relationships clearly, is essential for intelligibility.

Making a number of professional mathematical publications available in Digital Accessible Information System (DAISY) XML format, which makes all text, math, and figures fully accessible to people with vision disability, will encourage many other publishers to make their publications available in accessible DAISY XML format as well [13]. However the original DAISY XML specification does not include any capability for mathematics or accessible graphics beyond displaying the relevant characters as images with "alt text" labels. An extension of MathML permits MathML to be included in DAISY XML.

In creating math equations, letters and numbers are typed as usual. Special math characters such as Greek letters can be input by typing the LaTeX equivalent. Special functions such as fractions, radicals and integrals may also be input by using their LaTeX equivalents.

IV. MATHSPEAK AN APPLICATION FOR READING MATHS FORMULA WITH TEXT TO SPEECH

In development of the software application to provide accessible math formula, the attempt has been made to preserve conceptual information of formulae and discard the visual description. Since the math is complicated and technical, it needs practical ways to represent itself to the visually impaired. This application contains two main steps, as follow:

1) Converting the math equation to a format of markup language in the TEX family.

Audio renderings and navigation of the math formula (rehearing parts of the expression to understand them) have been considered. Finding the way to communicate math content to blind people and avoiding, where possible, graphical representations of the mathematical notation (using images of math) were the first steps in this process. These graphical representations create a barrier for the software in accessing the math in the accessible documents.

Using an authoring tool for the mathematical document to render math formula will make it readable. Free and open-source tools include MathJax for web pages, EPUB3 specifications for standard eBook format, and Math type plugin for Microsoft Word. The written mathematical formulae are calculated by various functions and written as output into files in the notation of MathML or LaTeX. These standard formats can generate the text math formula which is convertible to audio format by TTS function but not necessarily understandable by users. If documents contain only a few simple mathematical formulae, plain LaTeX may be sufficient. Amsmath is a LaTeX package that provides miscellaneous enhancements for improving the information structure and printed output of scientific documents containing mathematical formulae.

The Amsmath package introduces several new commands that are more powerful and flexible than the ones provided by LaTeX. The input to the Mathspeak system should be a document written in the TEX family of markup languages: TEX, LaTeX and AMS-TEX. The result of audio renderings produced by the Mathspeak system depends on how much of the document's logical structure is recognized and converted to markup language. Though currently this application requires input to be in Amsmath format, it can be reconfigured to deal with other markup languages in the TEX family.

The Mathspeak system can read equations input as text files. At this stage, the system can recognize all the alphanumeric characters, almost all Greek letters, and other frequently used symbols in mathematical formulae. It can analyse formulae, including fractions, square roots, subscripts, superscripts, integrals, limits, summation, and matrices.

One of the most significant tools of this system is the ability to navigate through different sections of a math formula.

To aid in the navigation of sections and blocks of mathematical expressions, "sections" and "subsections" of the equation must be marked out as separate components of the overall expression. Sections may be separated from each other

by the “\” character, which is recognizable as a keyword separator, for example: “\frac, \sum, \prod, \log, \int, \log, \sin”.

Each subsection contains all of the elements which are surrounded by open “{“ and closed “}” braces.

Two arrow keys (left and right) on keyboard have been implemented as the navigation commands. During the playing of an audio format formula (which is generated by the Text to Speech (TTS) function), the user can change play position to the previous subsection in the formula by pressing the left arrow key. If there is no subsection before the current position, play position is changed to the previous main section. Similarly, by pressing the right arrow, the user is able to set the play position to the next subsection, or in the absence of one, to the following main section.

2) Preparation of Amsmath format:

The Mathspeak system uses a modified form of Amsmath, whereby the Amsmath format of a mathematical equation is transmitted to TTS to produce an audio form. This technique must be preceded by some pre-processing steps to avoid issues such that may result in ambiguous audio production.

In some cases of mathematical formula, recognition of written notation can be confusing and hard to understand. The absence of brackets can cause this problem, so before sending Amsmath linear text to TTS, further formatting conversions are necessary to make the audio representation closer to the equivalent of the visual data. Additional formatting also fixes some Amsmath conflicts and adds some useful settings, symbols, and environments to the Amsmath format.

Distinguishing among several different meanings for a symbol in code is the basic task of the preparation stage. For example in Amsmath form, the character ‘_’ is usually used for presenting “index” meaning, however if this character is appeared in the formula containing “\sum”, “\prod” and “\int” it would mean “from”, with “\log” it would symbolize “to the base of...”, and finally with “\lime” it means “as approaches”. Also the ‘^’ character, which generally means “to the power of”, after “\sum”, “\prod”, and “\int” means “to”.

In some cases, recognizing the exact concept of a formula can be difficult. As shown in Table I, the output of simple linear TTS could be ambiguous. There is no difference between the outputs of two different inputs, but after using Amsmath and the pre-processing detailed, the outputs could be differentiated.

TABLE I.

Math formula	Text format Before preparation	Amsmath	Text format After preparation
$\frac{x}{y}+2$	x over y plus2	<code>\frac{x}{y}+2</code>	x over y plus 2
$\frac{x}{y+2}$	x over y plus2	<code>\frac{x}{y+2}</code>	x over open bracket y plus 2 close bracket

In Amsmath format, a fraction has two elements which are separated from each other by enveloping each element in braces ‘{ }’ and this is the main key to solving the ambiguity in some formulae that contain fractions, however sometimes,

this causes perturbation in audio output. As shown in Table II. if a fraction numerator or denominator includes just one

Math formula	Amsmath	Text format Before preparation	Text format After preparation
$\frac{1}{\sqrt{2} + \frac{1}{\sqrt{2} + \dots}}$	<code>\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{2}}+\dots</code>	open bracket 1 close bracket ,over open bracket square root of 2 plus ,open bracket 1 close bracket over open bracket square root of 2 plus open bracket 1 close bracket over open bracket square root of 2 plus ,until to ,close bracket ,close bracket square root of 2 plus ,until to ,close bracket ,close bracket	1 over open bracket square root of 2 plus 1 over open bracket square root of 2 plus 1 over open bracket square root of 2 plus ,until to ,close bracket ,close bracket

element, brackets around this element must be eliminated after the formula comes out from Amsmath, to reduce ambiguity and perturbation.

TABLE II.

Another mathematical construct which requires preparation is the matrix. In the Amsmath form of this matrix the character ‘\’ denotes ‘next row’ and the character ‘&’ denotes ‘next element’. In some matrices, where all values are displayed for at least one row and at least one column, the number of ‘\’ characters determines the number of rows and the number of ‘&’ characters in a row determines the number of columns. The size of the matrix is the number of rows times the number of columns. Therefore at the beginning of matrix reading, it is useful and straightforward to give a brief overview to the user by presenting the number of rows and the number of columns. As shown in Figure I, if the matrix is presented in an abbreviated form, and contains dots to denote absent elements, extracting the dimensions of the matrix is done by accessing the two subscript indices of the last element of the matrix (which lies at the bottom right corner). Navigation through a matrix is done in three ways: by specific row, by specific column, by specific element.

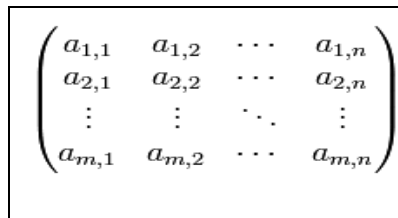


Figure 1. Sample of a matrix

`\begin{pmatrix}`

$$\begin{matrix} a_{1,1} & \& a_{1,2} & \& \dots & \& a_{1,n} \\ a_{2,1} & \& a_{2,2} & \& \dots & \& a_{2,n} \\ \dots & \& \dots & \& \dots & \& \dots \\ a_{m,1} & \& a_{m,2} & \& \dots & \& a_{m,n} \end{matrix}$$

Distinguishing between math and statistical symbols is another part of the preparation stage, for example the '|' character in statistics symbolizes "conditional probability" however, an expression enveloped in the same '|' character in math symbolizes, "the absolute value of the expression".

Table III shows the process which is implemented in this step.

TABLE III
CONVERSION OF MATH FORMULAE TO READABLE TEXT FORMAT

Math Formulae	Amsmath format	Text format before transmit to TTS
$\frac{x}{y}$	<code>\frac{numerator}{denominator}</code>	numerator over denominator
$\binom{n}{k}$	<code>\binom{n}{k}</code>	binomial n choose k
$\sum_{i=n}^k t_i$	<code>\sum_{i=n}^k t_i</code>	Sum from I equal n to k of t index i
$\prod_{i=n}^k t_i$	<code>\prod_{i=n}^k t_i</code>	Product from I equal n to k of t index i
x_k^n	<code>x_k^n</code> (Powers and indices are equivalent to superscripts and subscript in text mode)	X index n power k
$\sqrt[n]{k}$	<code>\sqrt[n]{k}</code>	root n of k
$\lim_{x \rightarrow \infty} x$	<code>\lim_{x \rightarrow \infty} x</code>	Lim x to infinity of x
$\log_n x$	<code>\log_n x</code>	Logarithm based n for x
$\ln x$	<code>\ln x</code>	Natural Logarithm for x
$\int_{-x}^x y$	<code>\int_{-x}^x y</code>	Integral from -x to x, y
$\cos^2 \theta$	<code>\cos^2 \theta</code>	Cosine power two theta
$P(A \cap B)$	<code>\P(A \bigcap B)</code>	Probability of A and B
$P(A \cup B)$	<code>\P(A \bigcup B)</code>	Probability of A or B

V. FEATURES EXTRACTION IN MATHSPEAK

Traditionally, an experienced mathematics reader might view a complex formula at high-level and then study its various details or components. In contrast, by a linear audio rendering, Mathspeak Speech system allows a listener access to such a high-level view by replacing sub-expressions in complex mathematics.

This option is implemented by pre-counting the number of occurrences of each keyword, distinguished by the "\" character, the most repeated keyword in formula makes the title of the high level formula view.

VI. FUTURE DEVELOPMENT

The problems associated with mathematical access for blind students are not only related to reading formulae, but also involved mathematical graphical information (such as line graphs, pie charts, etc.) which generally conveys the intended message of a mathematical document. The graphical information is common and pervasive and access to graphs, drawing, and figures is an important component in the math learning process. However, this visual representation of numerical data limits access to information for students with visual impairment. The output of a function can also be a multi-dimensional structure. Presenting these forms of mathematical representation audibly is more complicated. Hearing the graphed math formulae alone could not be sufficient to transfer and describe complex mathematical constructs. Therefore, finding a solution to make graphs and charts accessible is necessary. This solution must include a graphical content processing part, which detects and categorizes important information from the graphical resources and extracts information from them.

VII. CONCLUSION

The research undertaken has shown that it is possible to present mathematical structures to people with little or no vision in a logical and comprehensive manner if structured correctly. Issues remain with the communication of graphical information, for example, plots of trigonometric functions and the comprehension and recalling of long strings of symbols. Mathspeak has been implemented within an embedded system as part of a proposed "complete reading system" for people with print disabilities. At this stage, Mathspeak takes only LaTeX equations as input but does reliably and automatically produce comprehensible audio renderings of mathematic expressions. Work is currently being conducted in the automated conversion of line graphs, bar/pie charts and simple flow charts that may aid the intelligibility of a range of mathematical expressions.

REFERENCES

- [1] C. Jayant , (2006), "A Survey of Math Accessibility For Blind Persons and An Investigation on Text/Math Separation, Available at <http://metrc.uoregon.edu/index.php/related-questions/122-how-can-mathematical-text-be-made-accessible-to-students-with-print-disabilities.html>, retrieved 17 May 2012
- [2] S. Noble, (2008), "Universal Design for Math Learning: Bridging the Technology and Policy Divide." Available at <http://www.dessci.com/en/reference/articles/lwd07-noble.htm>, retrieved 17 May 2012
- [3] A.I. Karshmer, G.Gupta, S.Geiger, and C.Weaver, (1999), Reading and Writing Mathematics": The MAVIS Project, BIT (Behaviour & Information Technology).
- [4] F. Larsson, (2008), "Braille Representation and Mathematics Notation comparison different electronic formats for mathematics and their implications for Braille production.", Available at http://www.punktskriftsnamnden.se/filer/pn/Braille_Representation_and_Mathematics_Notation.pdf, Retrieved 17 May 2012
- [5] A. Nemeth, (1996)." Teaching Mathematics as a Blind Person, Speech given to the Mathematical Association of America, Orlando, Florida",

- Available <http://people.rit.edu/easi/easisem/nemeth1.htm>, Retrieved 17 May 2012
- [6] Dotless Braille for Nemeth,2004,Available at:
<http://www.dotlessbraille.org>, Retrieved 17 May 2012
- [7] J. Dixon,(2007), "Eight-dot Braille", Available at
<http://www.brailleauthority.org/eightdot/eightdot.pdf>, Retrieved 17 May 2012
- [8] J. Gardner, (2003), "DotsPlus Braille Tutorial: simplifying communication between sighted and blind people. ", CSUN 2003, Available at <http://www.csun.edu/cod/conf/2003/proceedings/284.htm>, Retrieved 17 May 2012
- [9] S. Sahyun and J. Gardner "Audio and Haptic Access to Math and Science – Audio Graphs, Triangle, MathPlus Toolbox and Tiger Printer", 15th IFIP World Congress, Vienna, 1998 2004, pp.6-9.
- [10] University of Washington, "How can publishers create accessible math textbooks?", January 27, 2010, Available at:
<http://www.http://www.washington.edu/doi/Faculty/articles?384>, Retrieved 2 May 2012
- [11] T.V. Raman, (1994), "Audio System for Technical Readings", Available at <http://emacspeak.sourceforge.net/raman/phd-thesis/aster-thesis.pdf>, Retrieved 17 May 2012
- [12] A.I. Karshmer, G. Gupta, E. Pontelli and K. Miesenberger, "The Development of a Tool to Enhance Communications Between Blind and Sighted Mathematicians, Students and Teachers", A Global Translation Appliance, Proceedings of the HCI Conference, New Orleans, 2001
- [13] DAISY Consortium Press Release: "Mathematics Now Added to the DAISY Standard", 2007, Available at: http://www.daisy.org/news/news_detail.asp?NewsId=296, Retrieved 1 May 2012
- [14] J. Gardner, (2001),"Introduction to DotsPlus Braille", available at <http://dots.physics.orst.edu/dotsplus.html>, Retrieved 17 May 2012