Necessities for Math-Access Tools with Speech

Katsuhito Yamaguchi

Nihon University

eugene@gaea.jcn.nihon-u.ac.jp

Toshihiro Kanahori

Tsukuba University of Technology

kanahori@k.tsukuba-tech.ac.jp

Masakazu Suzuki

Institute of Systems, Information Technologies and Nanotechnologies (ISIT)

suzuki@isit.or.jp

Abstract

Some new points that should be required in assistive tools for print disabled people to access math/science with speech output are discussed. The manner of “aloud reading” should be switched to the appropriate one according to the characteristics of a reader’s disability; blind users need a different manner than dyslexic users. Furthermore, the ambiguity of math expressions requires assistive tools to offer a method to assign an appropriate reading flexibly to those expressions according to context. In our accessible math-document editor, those features are really implemented. Combining it with our math OCR software, one can easily convert a math/scientific document in print or PDF into accessible e-books in DAISY or EPUB.

Keywords

MathML, Print Disability, Dyslexia, DAISY, Speech, Math

Introduction

"Knowledge" forms the foundation of contemporary society. In particular, abilities to make full use of scientific information including math expressions must be an essential skill for people to play active roles in society. However, such information was not necessarily accessible for print disabled people until recently. Fortunately, the situation has become remarkably improved these days. For instance, since MathML has been officially adopted in the current version of "the Digital Accessible Information System (DAISY)," we have become able to produce text-based/multi-media DAISY math content, in which one could access a math formula character by character, symbol by symbol with speech synthesis (DAISY Consortium). The International Digital Publishing Forum (IDPF) decided to adopt DAISY XML as a part of EPUB3 standards in 2011 (International Digital Publishing Forum).

However, there still remain several unsatisfactory points even in the current assistive tools to access math with speech. First, the fact that there is certain ambiguity in reading out math expressions is not necessarily taken account of in those tools. Many mathematical symbols or formulas have various names. For instance, "x with a raised 2, x^2" is usually used for x squared; however, it sometimes represents another quantity such as the second component of a vector. Such formulas have to be read in an appropriate manner according to their context. We need a method to define how to read those symbols or formulas with speech synthesis; however, it is difficult to change the manner of “aloud reading” in many math-access tools. Furthermore, “how a math formula should be read aloud” depends on the characteristics of the reader’s disability. We cannot provide such choices in most cases, either.

Math-Access tools

Math-Access tools with Speech

In the last two decades, various approaches have been tried to improve math accessibility. In 1994, T. V. Raman reported on his "Audio System for Technical Readings" which could read out math documents in LaTeX format with DEC Talk synthesizer, which might be the very first trial in this field (Raman). While LaTeX still plays an important role in math accessibility, recently MathML has become a key technology in assistive tools for print-disabled people to access "science, technology, engineering and math" (STEM) with speech. For instance, "MathPlayer" is a plug-in for a web browser such as Internet Explorer (Design Science). With this plug-in, the browser not only can display but also read out math expressions described in MathML embedded in a web page. "ReadHear" is one type of DAISY playback software, which can read aloud math expressions described in MathML, while highlighting each character/symbol as it is read aloud (gh, LLC.). There have been several other challenging activities based on MathML such as "Math Genie" (Gillan, Barraza, Karshmer, and Pazuchanics).

MathML

MathML is one of the XML (extended markup language) family, which was originally used for embedding math expressions in a web page; however, now it is used in various e-book formats such as EPUB3.

Although it is not well known, since the very beginning of its history, MathML has 2 types of notation for describing math expressions: "Presentation Markup" and "Content Markup." Presentation Markup aims only at formulating each symbol in a print math style. For instance, x with a raised 2, x^2, is always described as follows:

<msup>

<mi>x</mi> <mn>2</mn>

</msup>

regardless of its meaning, "x squared" or "the second component of a vector **x**."

On the other hand, Content Markup focuses on the semantics or meaning of math expressions. For instance, when x with a raised 2 means "x squared," it is described as follows:

<apply>

<power/> <ci>x</ci> <cn>2</cn>

</apply>.

When it means "the second component of a vector **x**", it is described as follows:

<apply>

<selector/> <ci type=”vector”>x</ci> <cn>2</cn>

</apply>.

That is, a description of Content Markup for a math expression depends on its semantic meaning. Using it, we can distinguish different math expressions having the same print style from each other. However, since this notation is rather complicated, there are practically no used web browsers that can support Content Markup.

Infty Software

Our group, the Infty Project and the not-for-profit organization, the Science Accessibility Net, have been developing math-access tools for more-than-15 years (Infty Project, sAccessNet). Our math-OCR system, "InftyReader," can recognize STEM documents in print or PDF and convert them into various accessible format such as LaTeX, MathML and Microsoft Word XML. "ChattyInfty" is an accessible math-document editor with speech. Using it, blind users not only

can read but also author a STEM document for themselves.

We worked on upgrading ChattyInfty thoroughly so that a recognized result with InftyReader or an edited file with ChattyInfty can be converted into DAISY XML format. Thus, the new version of ChattyInfty, "ChattyInfty3," becomes an accessible authoring tool for text-based/multi-media DAISY (Yamaguchi and Suzuki, “Accessible;” “On Necessity”). It is also useful for people with low vision or dyslexia as well as blind users. In the software, both the technical and non-technical content can be read aloud. Furthermore, since a result recognized by InftyReader can be imported directly, both sighted and print-disabled people can produce a DAISY book easily from printed or PDF material by making use of Infty software only.

In terms of multilingual support in ChattyInfty3, users can author/change not only how to read mathematical content but also captions in menu items and dialogs as they like since those things are all stored in independent files on a main program. Users, therefore, could customize ChattyInfty3 for each local language if necessary. For the present, although only Japanese and English versions are available, we are working on developing French and some other-language versions. Incidentally, the foreign-language versions other than Japanese use Microsoft Speech API, Ver.5 as a speech engine. Users can produce Text-based/multimedia DAISY content (DAISY3) in which all the math expressions are represented in MathML. In addition, output capability to EPUB3 is also implemented.

ABBYY "FineReader" is known as one of most powerful OCR software applications in the world (ABBYY). Recently, we released FineReader plug-in for InftyReader. By combining it with InftyReader, recognition rate for European languages including extended Latin characters becomes remarkably improved.

What Is Required of Math-Access Tools

Three Kinds of Aloud-Reading for Math Expressions

By making use of Infty software, we started a collaborative project with the Japan Braille Library two years ago to provide DAISY math/science textbooks to various print-disabled students. Through this activity, we noticed that "how a math formula should be read aloud" does depend on the characteristics of the reader's disability. For blind users, extra messages for which there are no corresponding print symbols explicitly in a math formula should be provided to show the structure of that formula exactly. For instance, they do need messages such as "Begin Fraction" and "End Fraction." The manner of "aloud reading" originally defined in ChattyInfty was essentially of this style. However, for other print-disabled people who can "see" the math expression up to a certain level, dyslexic users for instance, those extra messages actually disturb their understanding. They are often confused with the messages that do not have corresponding print symbols. Furthermore, even for blind users, when reading material repeatedly, they tend to want to skip those messages as well.

Thus, we implemented a new function in ChattyInfty3 so that three different types of "aloud reading" for math formulas can be selected. "Plain-Reading mode" is based on one which may be most widely used in English-speaking countries. It is natural; however, a spoken math expression is often ambiguous. We assume that people with low vision and dyslexia use it. In "Smooth-Reading mode," minimum-necessary speech guides for blind users to grasp the structure of a math formula correctly are added. "Detailed-Reading mode" is assumed to be used when a blind user wants to know the math-formula structure in the most detail.

"Yomi" Function in ChattyInfty3

As was mentioned previously, there are many ambiguities in reading math expressions. However, most math-access tools with speech do not seem to take this point into account. A fact that they are based on MathML in the presentation-markup notation may be one reason for that. As far as blind users are concerned, they could take down spoken math contents once in Braille and read it (a Braille document) over again. In this case, "presentation-based aloud-reading" is probably enough for their demands. Actually, "MathSpeak," that is a system of “aloud reading” for blind users given by A. Nemeth, seems to be of this type (MathMonkeys, LLC.).

However, people with other print disabilities such as those with severe low vision or dyslexia do need to directly understand the spoken math content, itself. If MathML in the content-markup notation could be used, it might be possible for us to assign appropriate “aloud-reading” automatically to each math expression. However, since it is impossible now, we need a new way to handle how to read out each symbol or math formula locally according to their context. We refer to this new concept of assigning a pronunciation as "Yomi" (a Japanese word that means "a manner of aloud reading"). We actually give a concrete method to assign Yomi to math expressions in DAISY3/EPUB3, which is implemented in ChattyInfty3.

Discussion

Here, at first, we list some samples to show the difference among three modes of “aloud reading”: Detailed Reading, Smooth Reading and Plain Reading in ChattyInfty3. The 10th terms of a sequence, c_{10} is represented in each of them as follows:

Detailed Reading: "c sub ten sub-end,"

Smooth Reading: "c sub ten (a short pose),"

Plain Reading: "c ten."

A fraction, two over three, 2/3 is read as:

Detailed Reading: "Frac two over three frac-end,"

Smooth Reading: "Frac two over three (a short pose),"

Plain Reading: "two over three."

Incidentally, we could assign "two thirds" to it with the Yomi function. Contrary to our expectation, we realized that, except for complicated math expressions, blind users seem to prefer Smooth Reading to Detailed Reading. Although most people with severe low-vision and dyslexia like Plain Reading, some of them tend to choose Smooth Reading since they do not look at a display when reading a DAISY book. Thus, we now assume Smooth Reading as a default in ChattyInfty. These three modes of “aloud reading” could be switched to another at each line if necessary. Unfortunately, since most of the other math-access tools with speech do not have such a feature, they are not necessarily useful for print-disabled people other than blind users.

We made a brief survey on the ambiguity in “aloud reading” of math formulas in several math textbooks. Although it is difficult to evaluate our result quantitatively, if we may say so, we confirmed that such ambiguity often appears in those books. Here are some examples we listed.

In fraction-type formulas:

A mixed fraction, 2 1/3: "two and a third" can be read as "two times a third."

A derivative, dy/dx: "(derivative) dy by dx" can be read as "(fraction) dy over dx."

In Super/Subscript type Formulas (in addition to the sample listed before):

f with a raised -1, f^{-1}: "f inverse" can be read as "f to the -1 power."

In math formulas with enclosing symbols:

A function A of an argument -x, A(-x): "A of -x" can be read as "A times a quantity, -x."

A combination, C(5, 3): "combination of 5 things taken 3 at a time" can be read as "a point C of coordinates 5 and 3."

Since each of them has the same form in print, access tools with speech cannot choose a proper “aloud-reading” automatically. In ChattyInfty3, one can assign appropriate manners to them with the Yomi function. Furthermore, although it is not an ambiguity issue, as situations require, we often have to choose the suitable manner of “aloud reading” for a math expression such as "2 and 3" in arithmetic and "a plus b" in algebra. The Yomi function is also useful in handling those cases.

Conclusion

Persons who verse themselves in mathematics probably could understand content of a STEM document according to its context even if some math formulas were read out in a wrong manner. However, for low-skilled students, it is clear that a textbook should be read aloud as correctly as possible. Three types of “aloud reading” and the Yomi function could give a certain contribution for those cases. We believe the other math-access tools with speech should take into account these points as well to support various types of print-disabled people.

Works Cited

ABBYY. *OCR, ICR, OMR - Optical Character Recognition and Linguistic Solutions.* ABBYY, n.d. Web. 10 Aug. 2013.

DAISY Consortium. *Making Information Accessible for All.* The DAISY Consortium. n.d. Web. 10 Aug. 2013.

Design Science. *How Science Communicates.* Design Science. n.d. Web. 10 Aug. 2013.

gh, LLC. *Innovation. Accessibility. Opportunity. Results.* gh, LLC. n.d. Web. 10 Aug. 2013.

Gillan, Douglas, Paula Barraza, Arthur I. Karshmer, and Skye Pazuchanics. "Cognitive Analysis of Equation Reading: Application to the Development of the Math Genie." *Proc. 9th International Conference on Computers Helping People with Special Needs, LNCS*. vol. 3118, Springer 2004: 630-37. Print.

Infty Project. Infty Project. n.d. Web. 10 Aug. 2013.

International Digital Publishing Forum. *The Trade and Standards Organization for the Digital Publishing Industry.* The International Digital Publishing Forum. n.d. Web. 10 Aug. 2013.

MathMonkeys, LLC. *MathSpeak!*. MathMonkeys, LLC. n.d. Web. 10 Aug. 2013.

Raman, T. V. *Audio System for Technical Readings.* Diss. Cornell University, 1994. Berlin: Springer, 1998. Print.

sAccessNet. s*AccessNet – Not-for-Profit Organization Science Accessibility Net.* NPO: Science Accessibility Net. n.d. Web. 10 Aug. 2013.

Yamaguchi, Katsuhito, and Masakazu Suzuki. "Accessible Authoring Tool for DAISY Ranging from Mathematics to Others." *Proc. 13th International Conference on Computers Helping People with Special Needs, LNCS*. vol. 7382, Springer 2012: 130-37. Print.

Yamaguchi, Katsuhito and Masakazu Suzuki. "On Necessity of a New Method to Read Out Math Contents Properly in DAISY." *Proc. 12th International Conference on Computers Helping People with Special Needs, LNCS*. vol. 6180, Springer, 2010: 415-22. Print.