Math transcribing with Braille2000 V3

Presented by Robert Stepp
Braille2000, LLC
1:45-2:45 February 28, 2023

7th Biennial Principals of Schools for the Blind
Math & Science Institute for Instruction
Kentucky Science Center
Louisville Kentucky

Synopsis

Braille2000 V3 is just being released, with new mechanisms for importing, editing, proofreading, and converting math, in UEB, UEB-with-Nemeth, and Nemeth notations. The presentation will demonstrate and explain the new math features and answer questions.

Learning Objectives

- You will learn how to use Braille2000 to import math from Word.
- You will learn how to use Braille2000 to proofread math braille in print notation.
- You will learn how to use Braille2000 to convert UEB math to UEB-with-Nemeth.

Agenda

I. Brief Introduction: “Where does math come from?”
   A. digital math: MathML as in DAISY/NIMAS files
   B. digital math: OfficeMath as in Word 2016, 2019, 2021, 365
   C. digital math: other notations, e.g., LaTeX
   D. fingertip math: via an equation typesetter
   E. fingertip math: via six-key braille input

Braille2000 V3 handles a basic-math[1] subset of MathML and OfficeMath.

II. How is braille math notated?
   A. once standardized as EBAE with Nemeth notations
   B. now various alternatives
      1. UEB-technical (worldwide exclusivity except in the US)
      2. UEB-with-Nemeth (rather elaborate semantic-sensitive transcription)
      3. Nemeth (i.e., EBAE+Nemeth, beautifully old-fashioned, ambiguous encoding)
      4. NUBS (Dr. Nemeth’s final offering, largely abandoned)

Braille2000 V3 can encode basic-math[1] in all of the above codes.

III. Strategies for math-braille in education
   A. As an educator you want a braille experience that stimulates learning
   B. As an educator you need materials in a timely manner
   C. Not every learner can cope with today's braille complexities
      1. Grade-school math is OK in any of the codes
      2. Algebra and beyond bog down with upper-digits (UEB)
      3. UEB-with-Nemeth is self-inconsistent
      4. Nemeth is not UEB

IV. Braille2000 V3
   A. Is already popular; for most license-holders, V3 is available without update fee
   B. Is the only WYSIWYG braille tool in the world i.e., the only tool with visual braille-page layout for braille ***and*** print
   C. Can import basic-math[1] from the Clipboard and from RTF, DOCX[2], NIMAS files
   D. Can interpret (as print) UEB, UEB-with-Nemeth, Nemeth[3] materials
   E. Can convert basic-math[1] between math codes

Notes:

[1] At this time, V3 handles basic math forms: math operators, math symbols, Greek letters, subscripts, superscripts, fractions, radicals, and combinations of these. At this time V3 does not handle underbars, overbars, other modifiers, arrows, and shapes.

[2] Importation from DOCX files requires the "Math Tools" option (otherwise you need MS Word so you can open the DOCX file,

copy it to the Clipboard, and paste it into V3). This option also enables the Math toolbar.

[3] Interpretation (for print proofreading and conversion) of Nemeth braille is possible, but only after the user has identified all the math (as opposed to narrative) passages. This tedious task is necessary because a Nemeth transcription is fundamentally ambiguous.

Narrative

I. Brief Introduction: “Where does math come from?”

Modern mathematics and math notation have been around for more than two centuries. Math is easy to write by hand but impossible to type on a typewriter and continues to be awkward even when using the computer. This is because there are non-Roman symbols and non-linear arrangements, such as superscripts, subscripts, and fractions. Fifty years ago, computer automation was brought to bear on newspaper and document printing by the use of "markup" statements intermingled with the text of the document. The UNIX operating system was invented at Bell Labs precisely to apply markup automation to document publishing, well before computer display technology had fonts or the ability to show what the published document would actually look like. Markup statements controlled the layout and style of the document, allowing plain text input to describe how to generate stylistically complex printing. It was thus entirely logical to use markup to describe how math should be shown.

Early on, Professor Donald Knuth became fascinated with the mechanics of automated typography and invented a system he called T_EX to specify and automate document style and layout. The letters T_EX are Greek, tau epsilon chi, the abbreviation (in Greek) for techne ("technical"). A popular document layout system built on his tool is known as LaT_EX (the name has nothing to do with rubber). About this same time other researchers were studying generalized markup languages, notably SGML (Structured General Markup Language) and its simplified spin-offs HTML (used to layout web pages), and XML (used for documents and data, and for DAISY and NIMAS publisher file standards). Specialized markup strategies have been created for math and music and other non-Roman things. MathML notation is popular and is used in DAISY and NIMAS publisher files. Although second-generation Word files (.docx) use XML, Microsoft chose to invent their own XML-based math markup called OMML (Office Math Markup Language). OMML notation is also available in the latest versions of Rich Text (RTF).

Braille2000 V3 understands a basic-math[1] subset of MathML and OMML.

In the 80's, Xerox Corporation introduced the non-markup strategy of WYSIWUG (what you see is what you get) based on the advancement of computer displays to show various fonts and graphics. Yet there were symbols and forms needed that were not found on the keyboard, and those would be entered through menu-based input. The Apple MacIntosh followed this lead, as did myriad

versions of Microsoft Windows. In a WYSIWUG tool, the menus to input math are known as an equation typesetter. But we should not forget that braille is itself a math markup language, and so another way to get math is to input six-key braille and have the computer interpret it as printed math.

MS Word is a WYSIWYG tool. It contains an equation typesetter. Braille2000 is a WYSIWYG tool. Braille2000 contains an equation typesetter, and can display print math entered as braille.

II. How is braille math notated?

Braille itself came about because of math: patterns of dots denoted numbers in the azimuth (firing angle) tables for French gunneries wanting to aim and fire their canons in the dark. That use is said to have given Louis Braille the idea for text braille. Braille later made its way to England. Henry Martyn Taylor (1842-1927) who taught math at Trinity College created a novel braille code that used a mixture of both upper- and lower-cell digits. In those days, college math didn't go beyond algebra. In the recent past, Abraham Nemeth (1918-2013) taught math at the University of Detroit. As a math student, he found the Taylor code unsuitable (plus nobody knew how to transcribe it), and so he invented his own system that bears his name. He later revised the system, calling it NUBS. With the adoption of UEB, NUBS is largely forgotten.

It turns out that braille is a rather concise notation (i.e., markup) for both math and music. But in the modern era, the work product needs to be producible by automation. And for that, the markup interpretation needs to be unambiguous (i.e., without any question as to what it says). Traditional Nemeth braille is English Braille American Edition with math notations. It is fundamentally ambiguous, for example "⠽⠕⠥ ⠅⠝⠕⠺ ⠊⠞" is both "you know it" and "y < x". A human reader easily figures out which interpretation is useful, but a computer tool cannot know which utterance to print out. And for that reason, traditional Nemeth is difficult to deal with by computer.

Newer braille codes have been designed to be essentially unambiguous. This is true for UEB and NUBS. And although Nemeth code is itself totally ambiguous, when used uncontracted and set off by code-switch indicators, the UEB-with-Nemeth combination is also essentially unambiguous. Being unambiguous is important when it comes to ink printing from braille, to speech output from braille, and to code conversion.

Braille2000 can ink-print braille for the unambiguous codes, and it can convert from one such code to another.

III. Strategies for math-braille in education

At this time, not every braille reader has a refreshable braille device, and such devices are, by and large, limited to one line of braille. There is still much literature suitably delivered only on paper and that braille is obviously

static—it is not adjustable to the occasion. There is thus a tension/uncertainty in the transcription process about how much auxiliary information (font styles, font colors, etc.) to inject into the prose, as well as what constitutes the best page layouts for reading and/or reference—the transcriber can only guess what tasks the reader will undertake, besides linear reading. Obviously, the braille code must be known to the reader, to the degree that reading is a fluid and flowing process. It is this last aspect that has some readers dissatisfied with UEB-technical, where upper-cell digits in a math formula involving the letters a-j lead to a great many injected indicators, enough to slow reading and/or lead to misreading and tediousness.

The injected indicators of UEB-technical get in the way, as do the code-switch indicators of UEB-with-Nemeth. And thus, there is still some transcription being done using traditional Nemeth. Note that automated conversion from an unambiguous code to traditional Nemeth is possible, but going the other way fails in those utterances that get interpreted "wrong" due to ambiguity. (Braille2000 has a feature to mark a Nemeth document to disambiguate the utterances, but it is a tedious human process.)

IV. Braille2000 V3

A WYSIWYG braille tool for math is an elegant concept but it is very complicated software. And, alas, Braille2000 V2 has proven not to be up to the task. (Braille2000 V1 did a good job handling Nemeth, but it was designed before UEB.) Braille2000 V3 is a major revision with an

improved mechanism for handling math. Although it was supposed to be (desired) ready months ago, it is just now becoming ready for first release. Most licensed users of V2 will be able to use V3 without any fee. And until V3 is proven to be reliable, installing V3 will not replace V2. At this time, V3 handles basic[1] math only—additional capabilities such as arrows and modifiers will come later (they can be done now via six-key input).

Braille2000 V3 is designed to handle UEB, UEB-with-Nemeth, Nemeth[3], and NUBS. Like V2, it can import print-text from RTF and NIMAS files. Unlike V2, it can also import print-text from DOCX[2] files.

When using Braille2000 V3 you can copy print-text from Word or Wordpad or a browser or spreadsheet to the Clipboard and then paste it into Braille2000. This includes print-text containing math created by the equation tool in Word (but not pictures of equations from an external tool).

Braille2000 can interpret math braille into print notation for on-screen proofreading and the print notation can be printed to paper or copied to the Clipboard and pasted into a Word document.

V. Demonstration of V3

Image of a page shown by Office 2021 with equations created in Office.

Math in WORD can be imported into Braille2000 V3.

- With the Math Feature, you can use Open, Source File to import a .docx file
- Without the Math Feature, you can Copy to the Clipboard and then Paste into Braille2000.

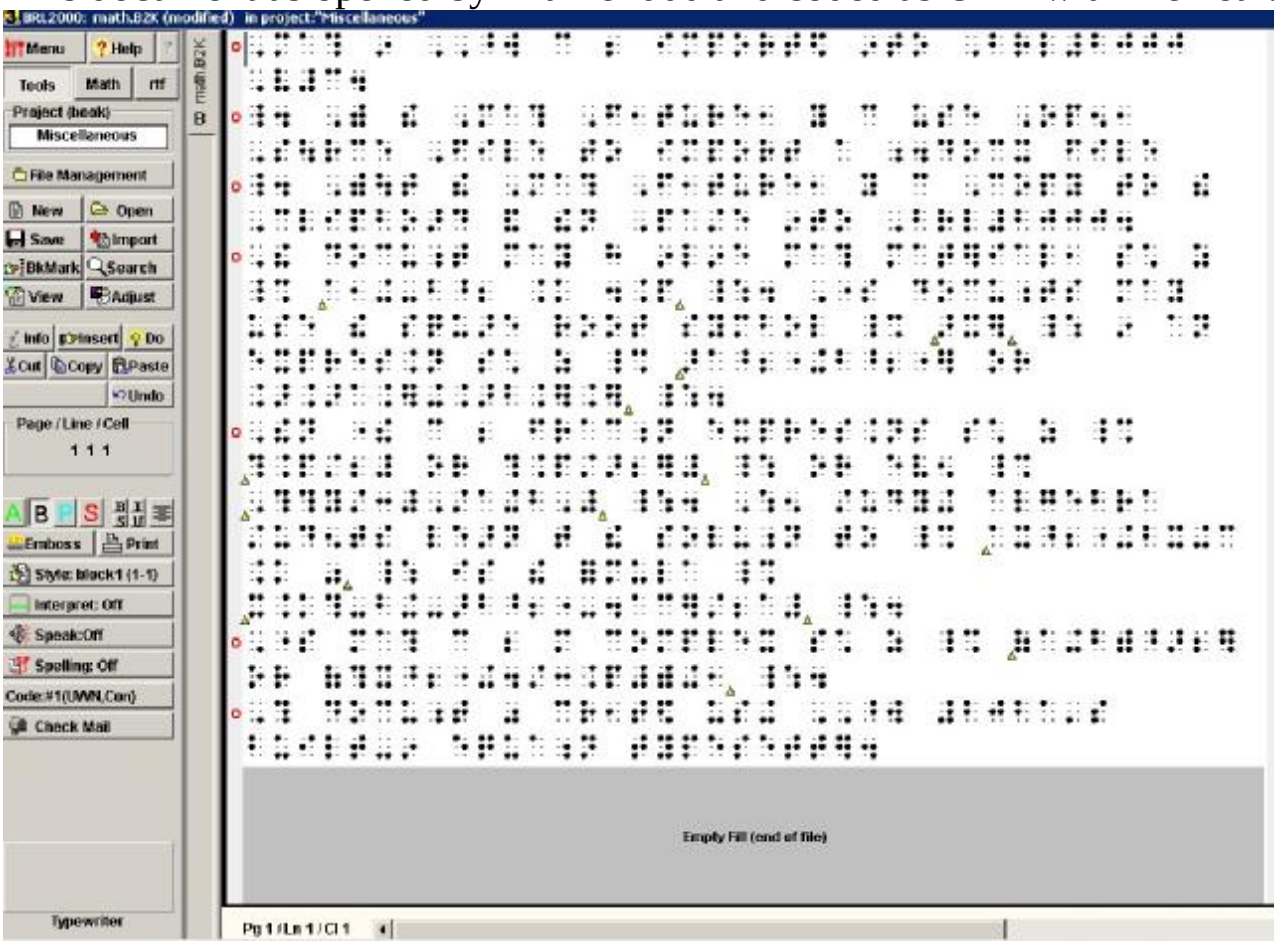
The document may have inline math material, such as $a_1 \pm b^2 = 4\pi$. Some documents may use the square root symbol $\sqrt{x}$ in an expression such as $\sqrt{a^2 + b^2}$ or $\sqrt{\sqrt{a} + \sqrt{b}}$.

Then there can be fraction expressions such as $\frac{\pi}{2}$ or $\frac{\pi}{\sqrt{2}}$ or even $\frac{\frac{y}{a}}{a+b}$. When studying algebra students learn that the solution to $ax^2+bx+c = 0$ is the formula $x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$.

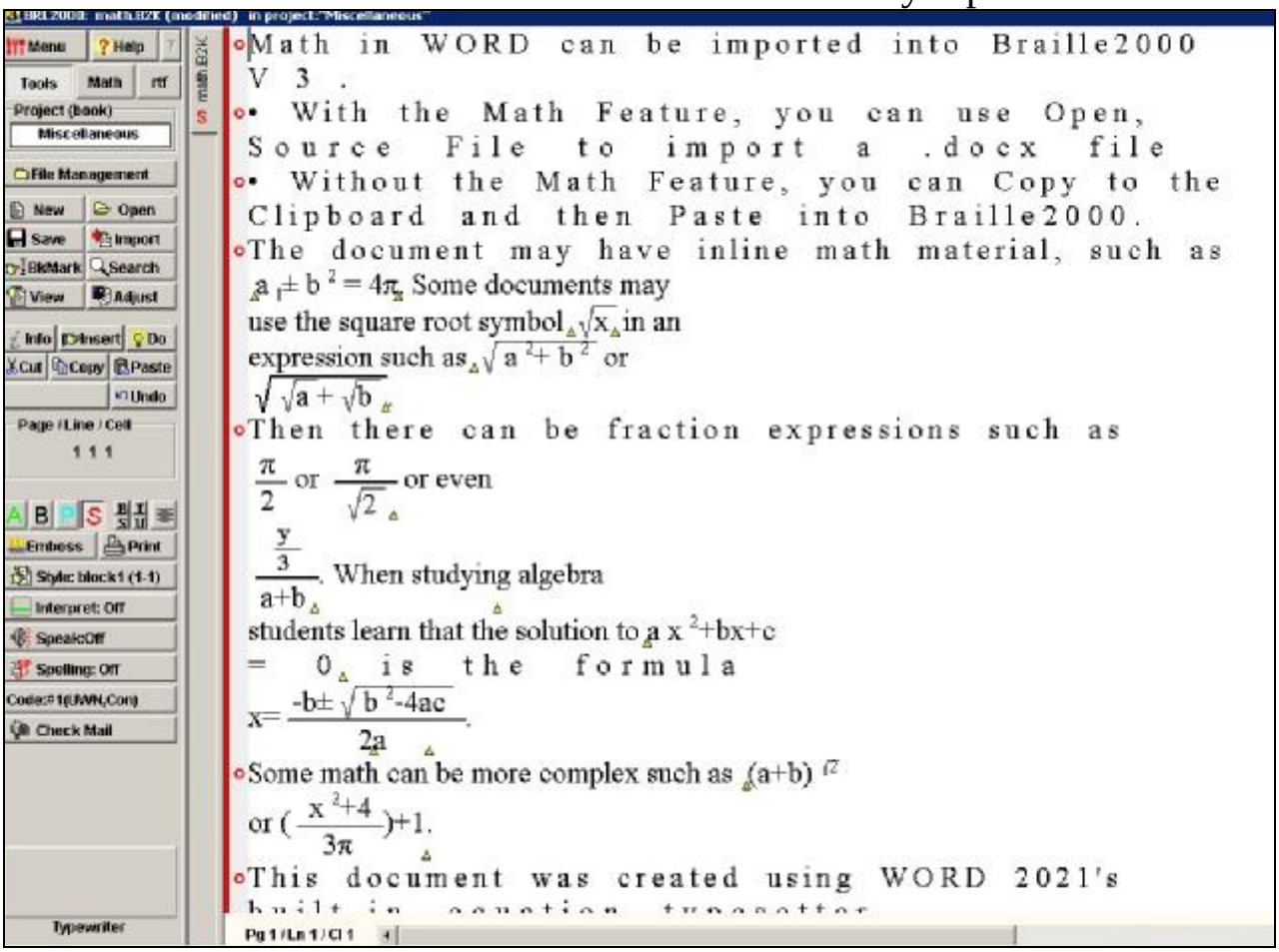
Some math can be more complex such as $(a + b)^{\sqrt{2}}$ or $\left(\frac{x^2+4}{\sqrt[3]{\pi}}\right)+1$.

This document was created using WORD 2021's built-in equation typesetter.

This document as opened by Braille2000 and coded as UEB-with-Nemeth.



The same document in Braille2000 viewed in print notation (by scrolling down the remainder of the braille page is revealed). This image would be the same if the braille had all been entered via six-key input.



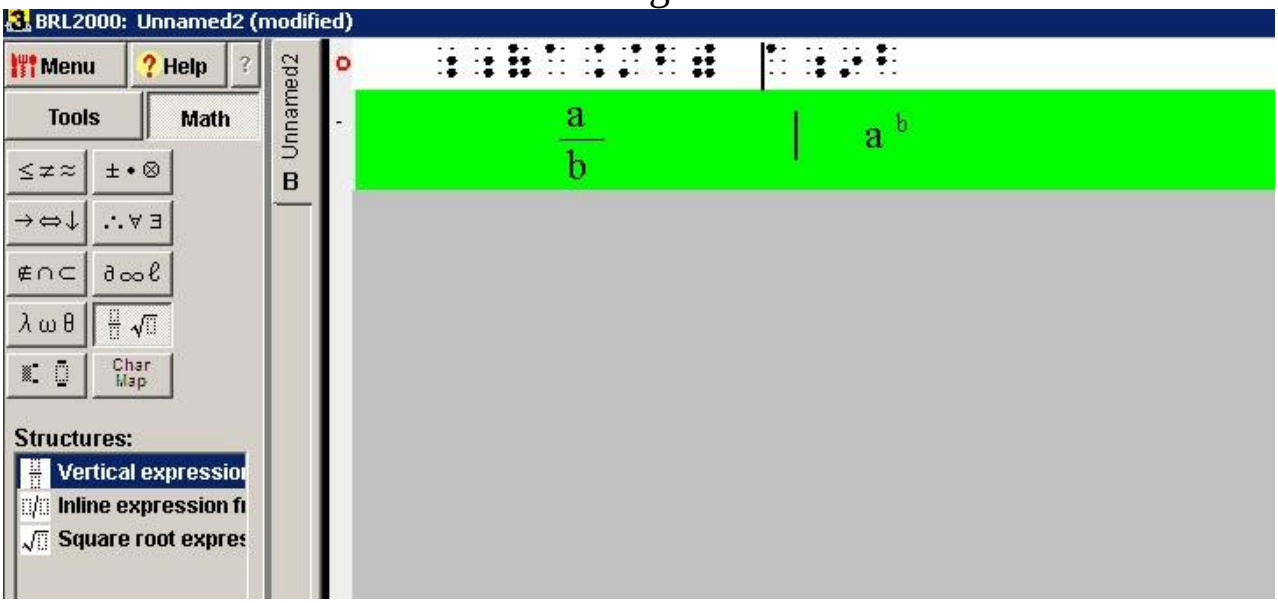
A page capture when the document is printed by Braille2000 to a laser or ink-jet printer. (The layout is WYSIWYG for proofreading Braille. For a "pretty" version, paste it into Word.)

Math in WORD can be imported into Braille2000 V 3 .
• With the Math Feature, you can use Open, Source File to import a .docx file
• Without the Math Feature, you can Copy to the Clipboard and then Paste into Braille2000.
The document may have inline math material, such as $a_1 \pm b^2 = 4\pi$. Some documents may use the square root symbol $\sqrt{x}$ in an expression such as $\sqrt{a^2 + b^2}$ or $\sqrt{\sqrt{a} + \sqrt{b}}$.
Then there can be fraction expressions such as $\frac{\pi}{2}$ or $\frac{\pi}{\sqrt{2}}$ or even $\frac{\frac{y}{3}}{a+b}$. When studying algebra students learn that the solution to $ax^2+bx+c = 0$ is the formula $x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$.
Some math can be more complex such as $(a+b)^{\sqrt{2}}$ or $(\frac{x^2+4}{3\pi})+1$.
This document was created using WORD 2021's built-in equation typesetter.

When copied to the clipboard and then pasted into Office, the layout is no longer that of the braille page—it flows according to Word's page size.

Math in WORD can be imported into Braille2000 V3.
• With the Math Feature, you can use Open, Source File to import a .docx file
• Without the Math Feature, you can Copy to the Clipboard and then Paste into Braille2000.
The document may have inline math material, such as $a_1 \pm b^2 = 4\pi$. Some documents may use the square root symbol $\sqrt{x}$ in an expression such as $\sqrt{a^2 + b^2}$ or $\sqrt{\sqrt{a} + \sqrt{b}}$.
Then there can be fraction expressions such as $\frac{\pi}{2}$ or $\frac{\pi}{\sqrt{2}}$ or even $\frac{\frac{y}{3}}{a+b}$. When studying algebra students learn that the solution to $ax^2+bx+c = 0$ is the formula $x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$.
Some math can be more complex such as $(a + b)^{\sqrt{2}}$ or $(\frac{x^2+4}{\sqrt[3]{\pi}})+1$.
This document was created using WORD 2021's built-in equation typesetter.

To enter math expressions through the keyboard, you can turn on the interpreter line (the so-called "green line") and six-key input math braille. The screenshot below shows this being done for UEB.



The dynamically displayed print is positioned (as much as possible) to follow the braille layout (so that the print rendition is WYSIWYG relative to braille page layout). When back-translation is used for proofreading, it gives immediate feedback if, for example, the G1 indicator is missing (the material wouldn't be math at all).

The Math toolbar[2] is shown. The toolbar operates like an equation typesetter. The various categories of expressions offer buttons to generate symbols and math structures to create math in lieu of six-key input. Through repeated use of the toolbar, expressions of arbitrary complexity can be created. The toolbar work in both the braille view of the document as well as in the print view.

Please note that this document was prepared using an early working copy of Braille2000 V3 and that some illustrations (all actual screen-captures) contain a few content flaws.