

The BrailleMathCodes Repository

Paraskevi Riga¹, Theodora Antonakopoulou¹, David Kouvaras¹,
Serafim Lentas¹, and Georgios Kouroupetroglou¹

¹National and Kapodistrian University of Athens, Greece
Speech and Accessibility Laboratory,

Department of Informatics and Telecommunications
p.riga@di.uoa.gr, ic1190004@di.uoa.gr, sdi1000196@di.uoa.gr,
sdi170007@di.uoa.gr, koupe@di.uoa.gr

Abstract

Math notation for the sighted is a global language, but this is not the case with braille math, as different codes are in use worldwide. In this work, we present the design and development of a math braille-codes' repository named BrailleMathCodes. It aims to constitute a knowledge base as well as a search engine for both students who need to find a specific symbol code and the editors who produce accessible STEM educational content or, in general, the learner of math braille notation. After compiling a set of mathematical braille codes used worldwide in a database, we assigned the corresponding Unicode representation, when applicable, matched each math braille code with its LaTeX equivalent, and forwarded with Presentation MathML. Every math symbol is accompanied with a characteristic example in MathML and Nemeth. The BrailleMathCodes repository was designed following the Web Content Accessibility Guidelines. Users or learners of any code, both sighted and blind, can search for a term and read how it is rendered in various codes. The repository was implemented as a dynamic e-commerce website using Joomla! and VirtueMart.

1 Introduction

Braille constitutes a tactile writing system used by people who are visually impaired. It employs embossed dots evenly arranged in quadrangular cells, with each cell being three dots high and two dots wide. Braille codes are systems for transcribing printed material using a braille alphabet. To increase the number of characters that can be represented with the six dots, a code either implements a special braille character called prefix or a composition sign or makes the meaning of a braille character dependent on the context. Braille can be extended to 8-dot cells to support STEM notation in assistive technologies, such as braille displays (Kacorri & Kouroupetroglou, 2013).

Early braille education of the blind is crucial to literacy, education, and employment. While in the 60s, braille literacy of legally blind school-age children in the USA was 50 percent in the United States, only 10 percent use braille as their primary reading medium (American Printing House for the Blind, 1996). Braille education provides visually impaired students with the opportunity to excel in several school areas such as vocabulary, comprehension, and even STEM, but also enhances their employment rates drastically (Riles, 2004).

One of the causes for the decline in braille usage was the move of children with blindness from specialized schools into mainstream public schools, where only a small percentage could afford to train and hire Braille-qualified teachers. To surpass this problem, all involved in the learning process, teaching-staff, visually impaired students, and possibly parents/caretakers, should be trained in the braille codes according to the national standards.

Braille readers use different codes to read and write different types of documents. Visually impaired students are generally taught three main codes: a) literary, b) math, and c) music braille, with math being a lot more difficult as it is not linear, it uses a lot of special characters and must be correctly interpreted at all times (unambiguous).

Some codes integrate literary and math support, while others are math specific. Each country, and even each educational institution within the country, follows its teaching system both for the sighted people seeking qualification and the visually impaired students. The educational material used (user guides with examples and exercises) is divided into courses related to different aspects of math. For blind students, code symbols are introduced as they occur in the print versions of the material being used in the curriculum. For sighted people, on the other hand, code symbols are thematically organized. For example, in Greece, although the Nemeth braille code was officially adopted in 2004 (Kouroupetroglou & Florias, 2003), there is still no preparation course to offer instructions for the Nemeth code to the sighted, and the braille competency certification only refers to literary braille. In developing the educational material to teach such a course, we found a gap in web content. There was no index or repository to search for braille math symbols, Unicode, and LaTeX encodings.

2 Printed math for the blind

2.1 Braille Math Codes

Math notation for the sighted is a global language, but this is not the case with braille math, as different codes are in use worldwide (Library of Congress, 1990). Since rules for the encoding of mathematics are not compatible from one system to the next, braille users who wish to read or write a range of material coming from different countries need to learn different sets of rules. The rules of each code include: a) the representation of math symbols, b) the use of "formatting" information, and c) other indicators.

The notations that are currently in use include the Antoine Notation (French Braille Code), Nemeth Code, Unified English Braille Code (UEB), British Mathematics Notation (BAUK), Spanish Unified Mathematics Code, Marburg Mathematics (German Code), Woluwe Code (Notaert Code), Italian Braille Code, Swedish Braille Code, Finnish Braille Code, Russian Code, and Arabic Code. Some of them are solely and others partially mathematics-related.

French Braille, created by Louis Braille in 1837, is the original braille alphabet and the basis of all others. It was meant for literary braille. In 1922, the braille code was first adapted to Mathematics by Louis-Auguste Antoine, a mathematician who became blind at war. The last revision of Antoine Notation was in 2007 (Croisetièrre, D'Amour, Ferland, & Rainville, 2008).

Abraham Nemeth first wrote The Nemeth Braille Code for Mathematics (Nemeth, 1972) in 1952. "The code is intended to convey as accurate an impression as is possible to the braille reader of the

corresponding printed text” and the result is a compact human-readable markup language where an ink-print math symbol with two different meanings is brailled the same in both instances. Nemeth Code is the standard for teaching and doing mathematics in braille in the US. 8-dot braille code has been introduced for complex Nemeth symbols (Martos, Kouroupetoglou, & Argyropoulos, 2015)

UEB was initially specified in 1991 (Cranmer & Nemeth, 1991). It is intended to form one set of rules, the same worldwide, which could be applied across various types of English-language material, e.g., for literary and technical material. Even though at the beginning of 1992, UEB integrated the Nemeth Code, it fails to handle mathematics as compactly. Besides requiring more space to represent and more time to read and write, the verbosity of UEB can make learning mathematics more difficult.

British Mathematics Notation was first designed in 1970, and its last revision was in 2005 (Kingdom, 2005). It was used in the UK until UEB replaced it, but it is still used in many places in Africa.

Spanish Unified Mathematics Code, known in Spanish as “Codigo Matematico Unificado” (CMU), is widely used in Spanish and Portuguese speaking countries. The Spanish National Organization of the Blind (ONCE) proposed CMU in 1970 to unify the math symbology but, it was only accepted in 1987 by all parties.

In 1955, the “Internationale Mathematikschrift für Blinde” notation was designed at the Marburg school for the Blind in Germany. Marburg notation, as it is known, was last updated in 1992 (Britz, Epheser, & Pograniczna., 1992) and is used by German-speaking countries.

In 1975 Flanders agreed on a code to represent mathematical formulae in braille, the Woluwe Code (Notaert, Suij, & Vandekerkhove, 1975) (after the Woluwe School for the blind), also known as Notaert Code (after Gilbert Notaert). It was created based on the Marburg Code and is used in Dutch-speaking parts of Belgium.

The last specification of the Italian Braille code, which incorporates math, was published in 2003 (Biblioteca Italiana, 2003).

In 1992, a report was published in Sweden to present the Swedish braille system with a set of symbols and writing principles for modern mathematics (Becker, Stenberg, Lindqvist, & Trowald, 1992). It was adopted by the Swedish Braille Council in June 1972, was also used in Iceland till 1990, and was last revised in 1912.

The Finnish mathematical notation was published along with physics and chemistry notation in 1979 (Central Union of the Blind, 1979). In the latest edition of 2014 ASCIIMath is used instead of braille, as students have moved to electronic books.

The formal braille code for math, physics, chemistry, and astronomy was published in a second edition in 1982 in Russia (Bykov, Egorov, Morozova, & Proskurjakov, 1982). It is also used in Belarus.

Since 1951, with the introduction of braille machines, braille in Arabic is read from left to right to be in line with the braille reading format of all languages. The braille codes in the Arabic nations include Arabic pre-2002, Arabic, Dari, Dhivehi, Farsi, Somali, and Urdu Pakistani. The Unified Arabic Braille Project started in 2015, is an effort to reduce the limitations of the current systems (Mada, 2021). The math code of the Unified Arabic Braille is supported by the Liblouis translator.

Information about each code is found in the original publications, in dedicated teaching books, and online web pages made for learning. Until today, there was no available electronic repository to use as an electronic index or search tool for these codes.

2.2 Math pseudo-codes

Pseudo-codes are written using a qwerty-keyboard or WYIWYG applications and have a visual nature. They are not considered braille codes because they still require translation to braille (Whapples, 2007). Examples include LaTeX (Knuth, 1984) as well as variations of LaTeX such as Human Readable Tex (HrTex) (Suzuki, Kanahori, Ohtake, & Yamaguchi, 2004). Contrary to dedicated braille codes, they are not developed purely for the blind, but they are used when braille math notation is not an

option. The LaTeX notation, which provides a more uniform format than braille, is commonly used to allow for easier communication with sighted users.

When it comes to higher education, LaTeX seems the most flexible notation; the user can expand it, and expansions are provided in packages by a supporting community. Although all braille math codes can express the most common mathematical structures in secondary education, most of them require enhancements to cover higher education mathematics.

2.3 Math XML

Presentation MathML (W3C Math Working Group, 1998), is an XML application for encoding mathematics on the Web based on their visual structure, an integral part of HTML5. MathML is used in braille translators to produce braille math codes and in AT applications to produce an acoustic rendering of math expressions, sometimes along with the haptic rendering.

LaTeX and MathML were found to be too complicated for use in primary and lower secondary education.

3 Assistive technologies and math

When math is in an electronic form, not graphically presented, then it can be rendered both haptically and acoustically. Screen readers, e.g., NVDA (Access, n.d.) and JAWS (Freedom Scientific, n.d.) are software applications that convert text displayed on screen into synthesized speech or braille. In the case of mathematics, the synthetic speech is valuable when math is in a format that can be rendered acoustically either straight from the screen reader or by another software application (e.g., MathPlayer). When math is expressed using a braille notation or some pseudo-code, then a refreshable braille display offers direct access to information, allows the user to check the format, spacing, and spelling in the text.

Students with visual impairments face challenges in learning from digital libraries and the web as the content is not expressed in an accessible way. The least the editors of such content can do is avoid graphic rendering of math expressions and use MathML or Unicode instead. This allows users to listen to the equations and even translate them to their braille code.

4 Math Braille Translators

At the moment, the best way for a blind person to read and write mathematical expressions is by using a mathematical braille notation. Communication of math between sighted people and braille users is done by math translation. Sighted people translate braille math to printed math, either by themselves or by using electronic braille translators. 100% correct braille translation can only be done by a human, as it usually depends on the context and requires text understanding.

Printed-to-braille math translators translate sighted users' code to braille code(s). UMCL provides a transcription of mathematical expressions from several mainstream formats to several mathematical braille codes (Archambault & Guyon, 2011), and math2braille translates MathML to the braille code used in the Netherlands (Crombie, Lenoir, McKenzie, & Barker, 2004). Duxbury DBT is a commercial program that handles LaTeX files and translates them to 9 math braille codes (Duxbury Systems, 2020).

Tools that backtranslate the braille math back to a format that can be read by a sighted individual. SBT is a tool that translates Spanish math braille to MathML (Alonso, Fuertes, González, & Martínez, 2006).

Translation and backtranslation programs translate some code used for the authoring of math by sighted users to some braille code(s) and vice versa. MAVIS developed the first Nemeth braille code to LaTeX backtranslator (Karshmer, Gupta, & Geiger, 1998). The Lambda Mathematical Code (Schweikhardt, Bernareggi, Jessel, Encelle, & Gut, 2006) was directly derived from MathML. It was designed to be used with braille peripherals and speech synthesis. It is automatically (back)convertible, in real-time and error-free, into an equivalent MathML version and, through it, into the most popular mathematical formats. LaBraDoor started with translating LaTeX to Marburg notation and subsequently backtranslated Marburg to LaTeX (Murillo-Morales, Miesenberger, & Ruemer, 2016). Liblouis is an open-source literary and math braille translator and back-translator of MathML and supports Nemeth, Marburg, and Arabic Braille (Egli, 2009).

5 BrailleMathCodes Repository

5.1 Content

The first step to the creation of the repository was the gathering of the information. Materials and content were gathered from the web, and symbol tables were created for each mathematical braille code. For codes with an official book reference, the online book was used as a source. For codes with no book reference, we relied on the content provided by national blind associations and libraries for the blind. No changes were made to the written forms of the symbols other than their transcription from braille patterns or graphics to text. This was needed to make the content accessible to those using adaptive technologies (refreshable braille displays, speech synthesizers, etc.) and allow error fixes.

Each symbol or indicator was matched with an image, a Unicode, and a name. Unicode is the most standard multilingual character encoding and is descriptive of the symbol form (The Unicode Consortium, 2000). To match Unicode with a braille formed symbol, we did both textual and visual search of Unicode tables when there was no reference of it in the original source. The name was descriptive, as it is in the Nemeth book, and it was also accompanied by its mathematical use(s) in parentheses, when different, as the verbal rendering in math changes based on the context. For example, the Unicode character "circled plus", U+2295 is referred to as: a) direct sum (an operation from abstract algebra), b) dilation (mathematical morphology), and c) exclusive or (a logical operation). The name of the symbol changes with the spoken language of each country; currently we provide the verbal rendering of symbols in English and Greek. However, we can extend the verbal rendering translations to cover more languages

The code tables were then joined to one table, where we have all the existing forms for each symbol. The join process was semi-automatic. As symbol names were not expressed in the same way for each code, unicodes were not always present, and some notations presented their own braille indicators with no equivalency.

To organize the whole table in sections, we followed the classification of symbols and indicators made by Nemeth, where symbols are grouped based on their use or the rule that governs them. However, we changed the order and made a distinction between symbols and indicators (Table 1). We did not classify the symbols as they occur in the print versions of the mathematic texts, as the repository's scope is not to teach mathematical concepts.

Then, we matched each math braille code with its LaTeX analogous and forwarded it with Presentation MathML. For each math symbol, we selected a characteristic example in Nemeth code, translated in MathML, when applicable (Table 2).

Most math braille codes cover the symbols used in pre-university curricula (secondary education). To study maths in high-level education, a braille user has to extend the given notations. There are works, some continuing, reviewed and published in conferences and journals of special education, that extend

some math braille codes, e.g., Nemeth (Martos, Kouroupetroglou, Argyropoulos, & Deligiorgi, Towards the 8-dot Nemeth braille code, 2014). When these extensions are adopted by official codes, we will include them in our repository, so that braille learners will not have to reinvent them.

Math Symbols	Math Braille Indicators 7
. Signs and symbols	. Alphabetic indicators 39
.. Punctuation signs and symbols 17	. Capitalization indicators 2
.. Numeric signs and symbols 13	. Fraction indicators 14
.. Signs and symbols of operation 25	. Level indicators (superscripts and subscripts) 15
.. Radicals 1	. Radical indicators 2
.. Modifiers	.. Order of radical 3
... Arc 2	. Shape indicators 6
... Bar 2	. Arrow direction indicators 4
... Caret 4	. Cancellation indicators 2
... Arrow 48	. Type-form indicators for words, phrases, and mathematical statements 4
... Other modifiers 6	. Type-form indicators for letters, numerals, and compound expressions 4
.. Signs and symbols of comparison 116	. Spatial arrangements 16
.. Signs and symbols of grouping 42	. Modification indicators 7
.. Reference signs and symbols 7	
.. Miscellaneous signs and symbols 37	
. Shapes	
.. Basic shapes 25	
.. Shapes with interior modification	
... Angle 4	
... Circle 12	
... Square 6	
.. Shapes with structural modification	
... Angle 11	
... Triangle 5	
. Function names and their abbreviations 38	

Table 1: Classification of math symbols and braille indicators. Each category is followed by the number of its entries.

5.2 Design and implementation

The user interface of (BrailleMathCodes repository, 2021) was designed following the Web Content Accessibility Guidelines (WCAG) (World Wide Web Consortium), so the content is perceivable (all three modalities are present in the content) and understandable. The repository was checked to be both operable and robust. The user can either A) search for a symbol by a) name (text input), b) Unicode (dropdown), or c) LaTeX (dropdown) or B) browse through the website's math symbol/braille indicator categories to find what he is looking for (Figure 1).

The repository was implemented as a dynamic e-commerce website using Joomla! (Cao & Yu, 2010), an open-source content management system (CMS), and VirtueMart (VirtueMart, n.d.), one of the oldest and more established eCommerce extensions for Joomla, where each symbol is "a product" and the corresponding codes are its fields. The source code of the CMS and the extension had to be altered to support braille codes and make the website fully accessible. The free template Horme 3, which does mobile detection, was used in full-width layout, with some customizations in the CSS and PHP codes.

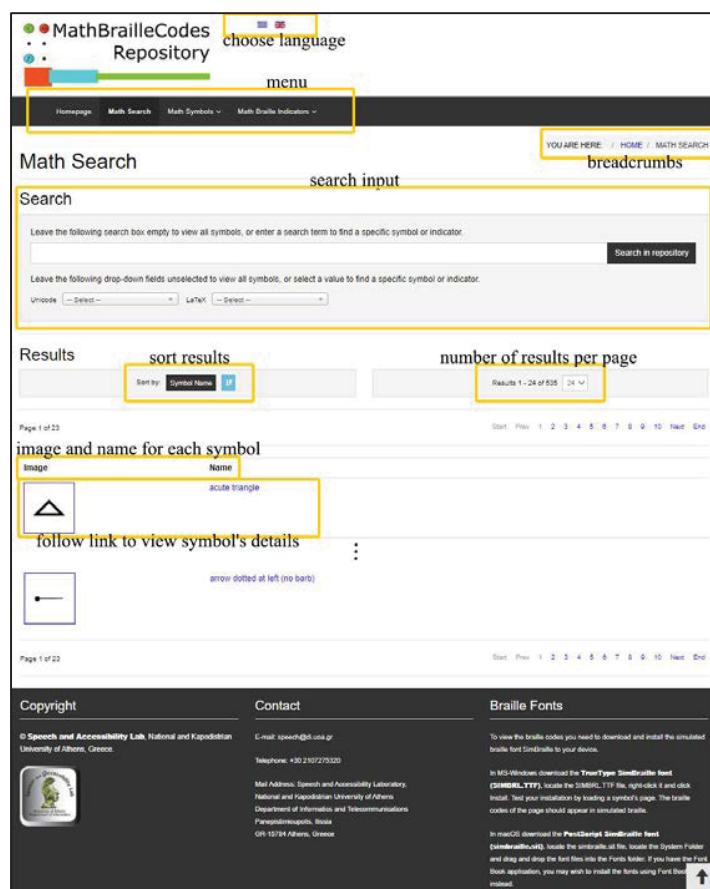


Figure 1: Screenshot of the BrailleMathCodes search page

5.3 Use

As stated previously, the repository is accessible, so visually impaired users can use their computer or smartphone along with a screen-reader and a refreshable braille display to access the repository. No login is required.

Users can either search for a specific code or browse through the math braille symbols and indicators. No conversion tool is used in the background, but learners of any code, both sighted and blind, can search for a term by its name, Unicode or LaTeX form, and then read how it is rendered in all the existing codes.

Search terms in BrailleMathCodes include math symbols, Unicodes, and braille math codes, but they cannot be expressed as mathematical formulae, as the examples in our collection are not large scale and the search algorithms used are string-based.

In the future, we plan to add more examples to make the repository a full learning environment for math braille codes, as well as to apply it in a mathematical information search and retrieval engine (Pattaniyil & Zanibbi, 2014) (Novotný, 2019) (Aizawa & Kohlhase, 2021).

6 Conclusions

The BrailleMathCodes repository meets the need of math braille users to search and find the codes for a symbol or indicator without the need to look it up in scattered sources. It contains the latest edition of each code and can be easily expanded to more codes and languages. We believe that as a dictionary, it will be a helpful tool for every braille math user, sighted, or visually impaired. Once all the codes and examples are gathered and imported, we plan to extend the repository's content to include math and code tutorials. At that point, we will need to enhance the search capabilities of the repository.

References

- Access, N. (n.d.). Retrieved from NonVisual Desktop Access (NVDA) screen reader: <http://www.nvda-project.org/>
- Aizawa, A., & Kohlhase, M. (2021). Mathematical Information Retrieval. *Information Retrieval and Access Tasks*, 169-185.
- Alonso, F., Fuertes, J., González, Á., & Martínez, L. (2006). SBT: A Translator from Spanish Mathematical Braille to MathML. *ICCHP* (pp. 1207-1214). Linz: Springer.
- American Printing House for the Blind. (1996). Estimated Number of Adult Braille Readers in the United States. *Journal of Visual Impairment & Blindness*, 90(3), 287.
- Archambault, D., & Guyon, F. (2011). UMCL Transcription Tools - Universal Maths Conversion Library. *AAATE Conference*, (pp. 416-423). Maastricht.
- Becker, S., Stenberg, G., Lindqvist, B., & Trowald, N. (1992). *Matematiska Symboler i Punktskrift*. Uppsala: Department of Educational Research, School of Education.
- Biblioteca Italiana. (2003). *Codice Braille Italiano*. Ciechi "Regina Margherita".
- BrailleMathCodes repository. (2021). *Speech and Accessibility Lab*. University of Athens. Retrieved from <https://access.uoa.gr/mathbraillecodes>
- Britz, K., Epheser, H., & Pograniczna, D. (1992). *Internationalen Mathematikschrift für Blinde*. Marburg: Deutsche Blindenstudienanstalt e.V.
- Bykov, A., Egorov, M., Morozova, F., & Proskurjakov, I. (1982). *System of Symbols for Math, Physics, Chemistry, and Astronomy: A Learning Aid*. Moscow: The All Russia Association for the blind.
- Cao, X., & Yu, W. (2010). Using Content Management System Joomla! to Build a Website for Research Institute Needs. *International Conference on Management and Service Science* (pp. 1-3). Wuhan: IEEE.
- Central Union of the Blind. (1979). *Finnish Braille Committee's Approved Mathematics, Physics and Chemistry Notation*. Helsinki: Central Union of the Blind.
- Cranmer, T., & Nemeth, A. (1991). *A Uniform Braille Code*.
- Croisetière, P., D'Amour, M., Ferland, P., & Rainville, G. (2008). *Code Braille Français Uniformisé Pour la Transcription des Textes Imprimés (CBFU)*. Québec: Drummondville: Office des personnes handicapées du Québec.
- Crombie, D., Lenoir, R., McKenzie, N., & Barker, A. (2004). math2braille: Opening Access to Mathematics. *ICCHP* (pp. 670-677). Berlin, Heidelberg: Springer.
- Duxbury Systems, I. (2020, 12 31). Retrieved from Braille Translation Software from Duxbury Systems: <https://www.duxburysystems.com>
- Egli, C. (2009). Liblouis - a uniiversal solution for Braille transcription services. *The Daisy 2009 Conference*. Leipzig.
- Freedom Scientific, I. (n.d.). Retrieved from Job Access With Speech (JAWS): <http://freedomscientific.com/Products/Blindness/JAWS>

- Kacorri, H., & Kouroupetroglou, G. (2013). Design and Developing Methodology for 8-dot Braille Code Systems. *Lecture Notes in Computer Science (LNCS), Volume 8011*, pp. 331-340, 2013., 8011, 331-340.
- Karshmer, A., Gupta, G., & Geiger, S. (1998). Reading and writing mathematics: the MAVIS project. *ASSETS* (pp. 136-143). Marina del Rey, CA: ACM Press.
- Kingdom, B. A. (2005). *Braille Mathematics Notation*. Peterborough: RNIB.
- Knuth, D. E. (1984). *Computers & Typesetting, Volume A: The TeXbook*. Boston, USA: Addison-Wesley.
- Kouroupetroglou, G., & Florias, E. (2003). *Greek Braille Scientific Notation – Application in Information Systems for the Blind*. Athens, Greece: Education and Rehabilitation Center for the Blind. Retrieved from <https://speech.di.uoa.gr/sppages/spppdf/BG1a.pdf>
- Library of Congress. (1990). *World Braille usage*. National Library Service for the Blind, & Physically Handicapped. Paris: UNESCO.
- Mada, Q. A. (2021, 1 3). Retrieved from The Unified Arabic Braille Project: <https://mip.qa/braille>
- Martos, A., Kouroupetroglou, G., & Argyropoulos, V. (2015). *International Conference on Enabling Access for Persons with Visual Impairment (ICEAPVI)* (pp. 115-119). Athens: University of Athens. Retrieved from <https://access.uoa.gr/ICEAPVI-2015/proceedingsFiles/117-121.pdf>
- Martos, A., Kouroupetroglou, G., Argyropoulos, V., & Deligiorgi, D. (2014). Towards the 8-dot Nemeth braille code. *Lecture Notes in Computer Science, 8547*, 533-536. doi:10.1007/978-3-319-08596-8_83
- Murillo-Morales, T., Miesenberger, K., & Ruemer, R. (2016). A LaTeX to Braille Conversion Tool for Creating Accessible Schoolbooks in Austria. *ICCHP* (pp. 397-400). Linz: Springer.
- Nemeth, A. (1972). *The Nemeth Braille Code for Mathematics and Science Notation*. Louisville, Kentucky: AAWB-AEVH Braille Authority.
- Notaert, G., Suij, M., & Vandekerckhove, E. (1975). *Handleiding Braillesymbolen wiskunde (Manual for braille mathematics)*. Woluwe.
- Novotný, V. (2019). *Document Representations for Fast and Accurate Retrieval of Mathematical Information*. Brno.
- Pattaniyil, N., & Zanibbi, R. (2014). Combining TF-IDF text retrieval with an inverted index over symbol pairs in math expressions: the tangent math search engine at NTCIR. *NTCIR* (pp. 135-142). Tokyo, Japan: online.
- Riles, R. (2004). *Research Study: Early Braille Education Vital*. Future Reflections.
- Schweikhardt, W., Bernareggi, C., Jessel, N., Encelle, B., & Gut, M. (2006). LAMBDA: A European System to Access Mathematics with Braille and Audio Synthesis. *ICCHP* (pp. 1223-1230). Linz: Springer, Berlin, Heidelberg.
- Suzuki, M., Kanahori, T., Ohtake, N., & Yamaguchi, K. (2004). An Integrated OCR Software for Mathematical Documents and Its Output with Accessibility. *ICCHP* (pp. 648-655). Linz: Springer.
- The Unicode Consortium. (2000). *The Unicode Standard, Version 3.0*. Addison-Wesley Longman.
- VirtueMart. (n.d.). *an Open Source E-Commerce solution*. Retrieved from <http://virtuemart.net/>
- W3C Math Working Group, S. B. (1998). Retrieved from Mathematical Markup Language (MathML) 1.0 Specification: <http://www.w3.org/TR/REC-MathML/>
- Whapples, M. (2007). Obtaining Braille mathematical documents. *MSOR Connections* 7.3, 18-21.
- World Wide Web Consortium, W. (n.d.). *Web Content Accessibility Guidelines (WCAG) Recommendation*. Retrieved from W3C: <https://www.w3.org/WAI/standards-guidelines/wcag/>