Axessibility: a LaTeX Package for Mathematical Formulae Accessibility in PDF Documents

Dragan Ahmetovic[1], Tiziana Armano[1], Cristian Bernareggi[2], Michele Berra[1], Anna Capietto[1], Sandro Coriasco[1], Nadir Murru[1], Alice Ruighi[1], Eugenia Taranto[1]

[1]Università degli Studi di Torino
Dipartimento di Matematica
{firstname.lastname}@unito.it

[2]Università degli Studi di Milano
Dipartimento di Informatica
{firstname.lastname}@unimi.it

ABSTRACT

Accessing mathematical formulae within digital documents is challenging for blind people. In particular, document formats designed for printing, such as PDF, structure math content for visual access only. While accessibility features exist to present PDF content non-visually, formulae support is limited to providing replacement text that can be read by a screen reader or displayed on a braille bar. However, the operation of inserting replacement text is left to document authors, who rarely provide such content. Furthermore, at best, description of the formulae are provided. Thus, conveying detailed understanding of complex formulae is nearly impossible.

In this contribution we report our ongoing research on *Axessibility*, a LaTeX package framework that automates the process of making mathematical formulae accessible by providing the formulae LaTeX code as PDF replacement text. *Axessibility* is coupled with external scripts to automate its integration in existing documents, expand user shorthand macros to standard LaTeX representation, and custom screen reader dictionaries that improve formulae reading on screen readers.

CCS Concepts

• **Human-centered computing** → Accessibility systems and tools; • **Social and professional topics** → **Assistive technologies**, **People with disabilities** • **Applied computing** → **Annotation**

Author Keywords

Visual impairment; Document and Math accessibility;

INTRODUCTION

Accessibility of digital learning material is crucial for people with visual impairments [14, 3]. In particular, accessing mathematical content [9], such as notation [4] and graphs [17] is fundamental at any education level from grade school [8, 1] to secondary and postsecondary education [6].

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ASSETS'18, October 22–24, 2018, Galway, Ireland

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5650-3/18/10.

DOI: https://doi.org/10.1145/3234695.3241029

There is a great ongoing effort to make both web pages [5] and digital documents [10] more accessible to people with visual impairments. Thus, a number of assistive technologies has been created for both math content access [15] and authoring [16, 13, 4] for visually impaired users. However studies show that didactic material, and in particular PDF documents are often inaccessible. This is common with university level course material [6], as well as research papers in published journals and conference proceedings, even in the field of education and accessibility [12].

In this paper, we outline our effort towards automated generation of PDF documents with accessible mathematical formulae. We present *Axessibility*, a LaTeX package devised to seamlessly include mathematical formulae typed in LaTeX as replacement text within the created PDF document. We also outline additional extensions to our software to ease authoring effort and provide screen reader integration. We evaluate *Axessibility* with 4 blind users achieving positive results, and we discuss future research and goals.

AXESSIBILITY LATEX PACKAGE

A PDF document that is generated starting from LaTeX code does not provide inherent accessibility of formulae by screen readers and braille bars. This can be achieved with external editors (Adobe Acrobat) or additional LaTeX packages (e.g., *pdfcomment.sty*) that allow to insert hidden comments describing formulae. This task must be manually performed by the author, and the provided tools are limited in expressiveness. For example, *pdfcomment.sty* does not allow to insert special characters like backslash (\), brace ({), etc, in the comment. Furthermore, these solutions often record content in a way that is incompatible with common screen readers, causing comprehension difficulties.

Axessibility is a LaTeX package that enables the creation of PDF documents containing accessible mathematical formulae [2]. It uses *accsupp.sty* and *accessibility_meta.sty* packages that allow to generate tagged PDF documents from LaTeX. To achieve this, *Axessibility* inserts hidden replacement text for each formula using the */ActualText* PDF attribute, which makes it visible to screen readers and braille bars. Additionally, the package enables to copy the LaTeX code of a formula from the PDF reader and paste it elsewhere.

Besides including the package among those used by the document, no other user intervention is needed. The package uses the command '\BeginAccSupp' defined in the existing package *accsupp.sty*. This command has been modified in order to include replacement text containing formulae LaTeX code in correspondence to their graphical presentation, using */ActualText* PDF attribute. We have provided integration with the most used mathematical environments, such as: *equation*, *equation**, \[, \(. Hence, any formula inserted using one of these environments is accessible in the corresponding PDF document.

Limitations

Since *Axessibility* creates PDF tags to make formulae accessible, it is currently incompatible with Adobe Acrobat DC PDF tagging functionality. To address this concern we are investigating the tagging capabilities of *accessibility* LaTeX package. The PDF produced using *Axessibility* and *accessibility* jointly is tagged correctly and contains accessible formulae. However, using *accessibility*, the text is not correctly read by screen readers. We are currently working on improving *accessibility* tagging to enable screen reader usage.

Axessibility is currently unavailable for *align* and *multiline* environments. It is possible to typeset multiline formulas using *aligned* environment of `amsmath` package. Display mathematical mode[1] and inline math ($$, $) are also not supported in this version of the package. Additionally, to preserve compatibility with Acrobat Reader, which does not correctly display the underscore (_) character inside replacement text, *Axessibility* suggests to use the equivalent command **sb**. However, external scripts provided as companion software, and described in the following section can also address these use cases.

EXTERNAL SCRIPTS

In addition to the core LaTeX package, we developed additional software to address three use cases: 1) Applying *Axessibility* on existing documents, 2) Expanding user defined commands within formulae, and 3) screen reader integration.

Preprocessing scripts

We provide an external package through which it is possible to apply *Axessibility* on existing LaTeXdocuments that include unsupported commands and environments. The external script uses regular expressions parsing to substitute unsupported commands and environments with suitable replacements.

Expansion of user macros

In LaTeX, it is possible to create macros, that is sequences of LaTeX instructions mapped to custom commands. This is a common practice to avoid code repetitions, simplify refactoring, and so on. *Axessibility* is transparent to commands used in math environments, which means that it will include standard LaTeX as well as custom macros within PDF replacement text. However, custom commands used by an author may bear no meaning for other readers. We therefore provide a script that complements *Axessibility* enabling to parse LaTeX documents and expand user macros before compiling the PDF.

[1] the use of display math mode is deprecated by TUG: https://www.tug.org/~hvoss/PDF/mathmode.pdf

Screen reader dictionaries

While *Axessibility* replacement text is easy to read by a LaTeX proficient user with a braille bar, LaTeX code read by a screen reader may appear awkward and hamper comprehension.

To address this issue, we provide additional dictionaries for *JAWS* and *NVDA* screen readers to convert LaTeX commands into their natural language counterparts (e.g., '\dfrac{a}{b}' into "a divided by b"). As a future work we will explore more sophisticated natural language processing techniques to also consider formula complexity, user's math proficiency and context during translation.

USER EVALUATION

We evaluated *Axessibility* with four blind users using Acrobat DC PDF reader. Two of them (**A** and **B**) are blind from birth, proficient in braille, and expert LaTeX users. The others are late-onset blind [7], do not use braille and have beginner (**C**) to intermediate (**D**) LaTeX proficiency. Braille users tested *Axessibility* replacement text with a braille bar, and all users tested the package with NVDA (**A** and **D**) or JAWS (**B** and **C**) screen readers.

The participants were able to correctly read formulae inside PDF documents produced using *Axessibility*. Participants **B** and **C** also tested the screen reader dictionary, reporting consistency with LaTeX commands. Finally, participants **A** and **C** successfully used the package for generating PDF documents with accessible formulae.

CONCLUSION AND FUTURE WORK

We have developed *Axessibility*, a LaTeX package that generates PDF documents with braille bar and screen reader-accessible mathematical formulae. Our package is complemented with additional external scripts to assists authors during content creation and readers during document access via screen reader. Through a preliminary evaluation with 4 blind users we uncover that *Axessibility* is effective in making mathematical formulae accessible.

As future work, we intend to explore different representations for the embedded accessible formulae, such as *mathml*. This would enable the integration with Mathplayer [15], a software that allows navigation of mathematical formulae content in details. We will also explore different braille rendering representations for braille bars, such as Nemeth Math [11]. Another key future work will consist in natural language representation of formulae that considers contextual information to verbalize the semantic meaning of mathematical symbols, as well as natural language adaptation based on user's mathematical proficiency.

ACKNOWLEDGEMENTS

The authors wish to thank the bank foundation 'Fondazione Cassa di Risparmio di Torino', LeoClub (Biella, Italy) and the several volunteers with visual impairment who provided their fundamental contribution.

REFERENCES

1. Dragan Ahmetovic, Valeria Alampi, Cristian Bernareggi, Andrea Gerino, and Sergio Mascetti. 2017. Math Melodies: Supporting Visually Impaired Primary School Students in Learning Math. In *Proceedings of the 14th Web for All Conference on The Future of Accessible Work*. ACM, 26.
2. Tiziana Armano, Anna Capietto, Sandro Coriasco, Nadir Murru, Alice Ruighi, and Eugenia Taranto. 2018. An automatized method based on LaTeX for the realization of accessible PDF documents containing formulae. In *International Conference on Computers Helping People With Special Needs*. Springer.
3. Tiziana Armano, Anna Capietto, Marco Illengo, Nadir Murru, and Rosaria Rossini. 2015. An overview on ICT for the accessibility of scientific texts by visually impaired students. In *Congresso Nazionale SIREM 2014*. Sie-L Editore, 119–122.
4. Cristian Bernareggi. 2010. Non-sequential mathematical notations in the LAMBDA system. In *International Conference on Computers for Handicapped Persons*. Springer, 389–395.
5. Cristian Bernareggi and Dominique Archambault. 2007. Mathematics on the web: emerging opportunities for visually impaired people. In *Proceedings of the 2007 international cross-disciplinary conference on Web accessibility (W4A)*. ACM, 108–111.
6. Catherine S Fichten, Jennison V Asuncion, Maria Barile, Vittoria Ferraro, and Joan Wolforth. 2009. Accessibility of e-learning and computer and information technologies for students with visual impairments in postsecondary education. *Journal of Visual Impairment & Blindness* 103, 9 (2009), 543.
7. Madeleine Fortin, Patrice Voss, Catherine Lord, Maryse Lassonde, Jens Pruessner, Dave Saint-Amour, Constant Rainville, and Franco Lepore. 2008. Wayfinding in the blind: larger hippocampal volume and supranormal spatial navigation. *Brain* (2008).
8. Andrea Gerino, Nicolo Alabastro, Cristian Bernareggi, Dragan Ahmetovic, and Sergio Mascetti. 2014. Mathmelodies: inclusive design of a didactic game to practice mathematics. In *International Conference on Computers Helping People With Special Needs*. Springer, 564–571.
9. Chandrika Jayant. 2006. A survey of math accessibility for blind persons and an investigation on text/math separation. *Seattle: University of Washington* (2006).
10. Ross Moore. 2014. PDF/A-3u as an archival format for accessible mathematics. In *Intelligent Computer Mathematics*. Springer, 184–199.
11. Abraham Nemeth. 1972. *The Nemeth Braille Code for mathematics and science notation*. American Printing House for the Blind.
12. Julius T Nganji. 2015. The Portable Document Format (PDF) accessibility practice of four journal publishers. *Library & Information Science Research* 37, 3 (2015), 254–262.
13. Alessandro Pepino, Corinna Freda, Fiorentino Ferraro, S Pagliara, and Francesco Zanfardino. 2006. "BlindMath": a new scientific editor for blind students. In *International Conference on Computers for Handicapped Persons*. Springer, 1171–1174.
14. Marshall Raskind, Bob Hoffman, Kendall Hartley, and Randall Boone. 2005. Reaching accessibility: Guidelines for creating and refining digital learning materials. *Intervention in School and Clinic* 40, 3 (2005), 171–176.
15. Neil Soiffer. 2007. MathPlayer v2. 1: web-based math accessibility. In *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 257–258.
16. Volker Sorge. 2016. Supporting visual impaired learners in editing mathematics. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 323–324.
17. Marzia Taibbi, Cristian Bernareggi, Andrea Gerino, Dragan Ahmetovic, and Sergio Mascetti. 2014. Audiofunctions: Eyes-free exploration of mathematical functions on tablets. In *International Conference on Computers Helping People With Special Needs*. Springer, Cham, 537–544.