

LaTeX2Tri: Physics and Mathematics for the Blind or Visually Impaired

David M. Thompson

Department of Physics, Harvard University, thompson@physics.harvard.edu

Proceedings from the 20th Conference on Technology and Persons with Disabilities
March 19, 2005

Abstract

A new conversion tool, LaTeX2Tri, makes LaTeX files of physics and mathematics accessible to blind or visually impaired users of WinTriangle. Through a variety of pathways, common file types, such as TeX, Word, and PDF, may be converted to Triangle, the working language of many blind or visually impaired students and researchers. Textbooks, arXiv preprints, class notes, and problem sets are now readily accessible to WinTriangle users, completing the loop of mathematical communication between the blind and sighted communities.

1 Introduction

Computers have revolutionized how both sighted and blind researchers access papers in physics and mathematics. For the sighted community of theoretical physicists, the development of a preprint repository, known as “the arXiv,” has revolutionized how research is shared among workers [1]. For blind or visually impaired researchers, a host of new technologies has enabled them to read and write mathematics in ways unimaginable before the advent of the personal computer [2–9]. In an effort to bridge the gap between blind researchers and the sighted community, this paper introduces the LaTeX to Triangle converter (LaTeX2Tri). This tool takes LaTeX files, such as the files stored on the arXiv, and converts them into Triangle files, the file format used by many blind and visually impaired researchers.

The second section reviews the history of LaTeX, especially as used in the arXiv, and the Triangle language developed at Oregon State University. Section three introduces the LaTeX2Tri converter and section four discusses some of its uses, especially in converting other file formats, such as PDF and Word.

2 Brief History of LaTeX and Triangle

In the late 1970’s, Donald Knuth began work on a typesetting language known as TeX [10]. The TeX format was designed to typeset mathematical equations and, as such, gained wide popularity among those who had to write technical papers. In the early 1980’s, Leslie Lamport improved on the TeX format with the development of LaTeX [11]. LaTeX has since overtaken plain TeX as the most popular typesetting language among those in the technical sciences. Among theoretical high-energy physics preprints on the arXiv, LaTeX submissions currently outnumber plain TeX submissions by roughly ten to one.

The arXiv¹ (pronounced “archive”) began in August of 1991 as a database of preprints in string theory, a branch of theoretical high-energy physics [1]. Beginning with a community of less than 200 string theorists, this free web site now hosts preprint archives for various fields of physics, mathematics, computer science, and biology. For theoretical physicists, the arXiv has become an indispensable part of their daily routine, as a dozen or so new papers are posted to the site every 24 hours. The arXiv has had a dramatic impact on the entire community of researchers, reducing “publishing” times to days and providing free access to anyone in the world who has an internet connection. The arXiv now holds over 300,000 preprints and receives roughly 4,000 papers per month, a rate which is increasing at about 300/month/year, ensuring its continued and growing relevance to researchers worldwide²³. The dominant role of the arXiv, and of the LaTeX format used for arXiv submissions, highlights the importance of making LaTeX files readily accessible to blind students and researchers.

In a separate development, the past decade has seen the development of computer-based technologies aimed at making physics and mathematics accessible to blind students and researchers [2–9]. In the 1990’s, researchers at

¹<http://arxiv.org>

²http://arxiv.org/show_monthly_submissions

³<http://arxiv.org/Stats/hcamonthly.html>

$$\int dx x = \frac{1}{2} x^2 \quad \xrightarrow{\text{LaTeX2Tri}} \quad \int dx x = [F1\langle/\rangle2F]x^2$$

Figure 1: LaTeX equations, displayed on the left-hand side as compiled into PDF, may be converted into Triangle, displayed on the right-hand side, using LaTeX2Tri. The intuitive fonts used in the Triangle language may be read by a sighted person. Here, the [F symbol denotes the beginning of a fraction, the \langle/\rangle symbol separates numerator from denominator, and the F] ends the fraction.

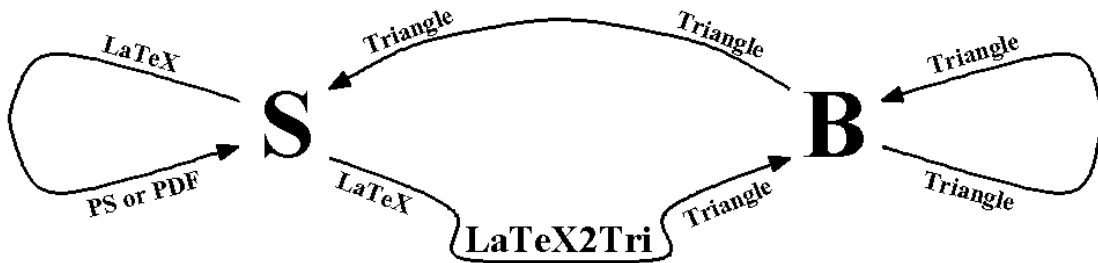


Figure 2: The development of LaTeX2Tri has completed the loop of mathematical communication between (S)ighted and (B)lind. Sighted individuals write in LaTeX and may read postscript, PDF, and Triangle files. Blind users of WinTriangle read and write in Triangle and, now, may convert LaTeX files into Triangle for reading and editing.

Oregon State University's Science Access Project, under the direction of John Gardner, began work on what has now become WinTriangle version 2.0.0⁴. Wintriangle is a Rich Text Format (RTF) word processor that uses specialized fonts to represent mathematical symbols and operations which, in turn, are voiced by a speech synthesizer [2, 3]. For the remainder of the paper, this RTF format, together with the Triangle TrueType font, will be referred to as the Triangle language.

WinTriangle allows blind users to hear mathematical text, write mathematical text, and perform calculations. This software serves as the pen and paper that sighted students and researchers take for granted. Users can archive their own work for future reference or communicate with other WinTriangle users by saving their work electronically or printing as GS Braille with an embossing printer. Of course, this would be of little use if the user's work could not be communicated to the larger sighted community. Students must be able to read homework assignments and hand in their solutions. Blind researchers must be able to read the work of their peers and disseminate their own papers. Thanks to the intuitive fonts employed by WinTriangle, communication from the blind user to the sighted reader is easy. For an example of an equation in Triangle, see figure (1). Communication in the other direction, from sighted to blind, has now been addressed by the creation of LaTeX2Tri, the subject of this paper.

3 Introducing LaTeX2Tri

Sighted workers communicate research to each other by writing their papers in LaTeX and then compiling to PS or PDF for other sighted people to read. Blind researchers, on the other hand, can communicate with each other by writing and listening in WinTriangle. Sighted people can read work written in the Triangle language thanks to the intuitive fonts developed for WinTriangle. The one missing path is that from the sighted community to the blind researcher, from LaTeX to Triangle. The LaTeX to Triangle converter bridges this gap as sketched in figure (2).

Perhaps in recognition of LaTeX's growing importance, several projects have been designed to address the issue of LaTeX access. The Aster program, a Ph.D. project by T. V. Raman, processes LaTeX files and reads them out

⁴The most recent version may be downloaded from <http://tap.oregonstate.edu/WinTriangle/WinTriangle.htm>

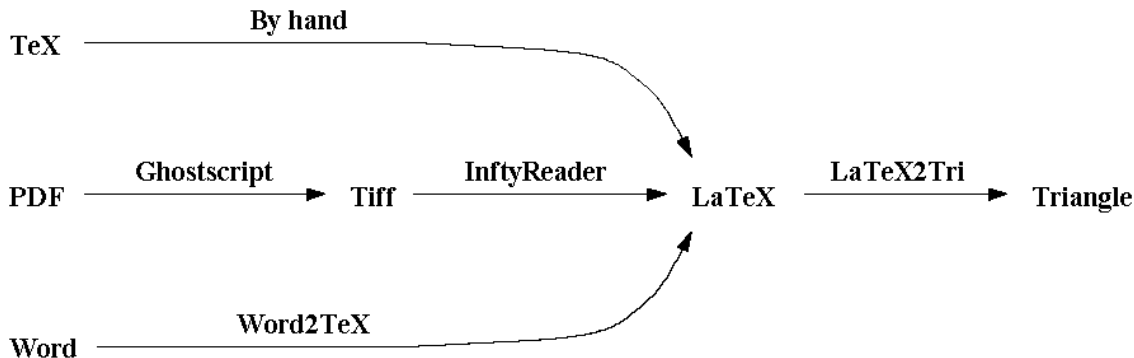


Figure 3: In addition to the conversion of LaTeX files into Triangle, the LaTeX2Tri converter has completed conversion pathways to Triangle from many common formats, such as TeX, PDF, and Word.

loud [5]. The Human readable TeX (HrTeX) format was designed to simplify LaTeX code for direct reading [6]. Some students have even used LaTeX directly as a means of communication [12]. Other projects have included Labrador, which converts LaTeX to Braille [7], Insight, which converts Braille to LaTeX [9], and Techread, which reads TeX out loud [8]. But none of these allow for conversion of LaTeX into the working language of WinTriangle users. With the development of LaTeX2Tri, blind researchers can read LaTeX files in WinTriangle, edit them and add to them at will, and print them for reading by sighted people.

The LaTeX2Tri converter has been built on top of version 1.9.15 of latex2rtf⁵, an open source program first written in 1994 at Vienna University. The outputted format, RTF, is a word-processing language developed by Microsoft and used by many common word-processing programs such as Microsoft Word. LaTeX2Tri is currently in beta version 0.3 and already recognizes most of the more common commands used in LaTeX2e. The converter has been ported to Windows so that anyone using JAWS on their own personal computer can run and operate LaTeX2Tri.

4 Uses for LaTeX2Tri

At Harvard, LaTeX2Tri has proven to be indispensable in converting physics textbooks, arXiv papers, journal articles, class notes, and problem sets. The conversion of graduate-level physics textbooks into Triangle presents an especially daunting challenge. Since the availability of electronic formats varies from publisher to publisher, there is a need to find conversion routes from several different formats: LaTeX, TeX, PDF, and Word. Recently published textbooks are more likely to be written in LaTeX, making conversion with LaTeX2Tri a fairly simple matter. For textbooks written in the older plain TeX format, it is necessary to convert by hand into LaTeX. Fortunately, the similarities between TeX and LaTeX far outweigh the differences; for many deprecated commands, a judicious use of search and replace can speed the conversion along. Oftentimes, a PDF file is all that a publisher has to offer. In these cases, a helping hand is lent by a program developed by Masakazu Suzuki's group in the Graduate School of Mathematics at Kyushu University [4]. This exceptional program, called InfyReader, is an optical character reader that can recognize complicated mathematical expressions and convert them to LaTeX⁶. The conversion process begins by saving the PDF files as Tiff image files using Ghostscript and then feeding them into InfyReader. The resulting LaTeX file requires a small amount of editing by hand, after which it may be passed through LaTeX2Tri. If the document is available only in Word with Equation Editor, then the commercial program Word2TeX⁷ may be used to convert to LaTeX. Again, a small amount of tidying up by hand is required before passing through LaTeX2Tri. A summary of these conversion paths is presented in figure (3).

When it comes to research, a theoretical physicist must have access to papers on the arXiv. Fortunately, the arXiv makes the original LaTeX files available for download. These files can be converted directly using LaTeX2Tri. For papers published before 1991, it is necessary to scan the original hard copies into InfyReader or to copy them by hand into LaTeX. To aid a blind or visually impaired student enrolled in a physics or mathematics class, a scribe can take notes in class on a laptop. With the use of LaTeX2Tri, these LaTeX notes are then immediately available to the student in Triangle format.

⁵<http://latex2rtf.sourceforge.net>

⁶<http://www.inftyproject.org>

⁷<http://www.chikrii.com>

References

- [1] P. Ginsparg, “Winners and losers in the global research village,” *Serials Librarian* **30** (1997), no. 3/4, 83–95.
- [2] J. A. Gardner, R. Stewart, J. Francioni, and A. Smith, “Tiger, AGC, and WinTriangle, removing the barrier to SEM education,” in *Proceedings of the 2002 CSUN International Conference on Technology and Persons with Disabilities*. 2002.
- [3] J. A. Gardner, “Access by blind students and professionals to mainstream math and science,” in *Proceedings of the 8th International Conference on Computers Helping People with Disabilities*, K. Miesenberger, J. Klaus, and W. Zagler, eds., vol. 2398, pp. 502–507. Springer-Verlag, 2002.
- [4] M. Suzuki, F. Tamari, R. Fukuda, S. Uchida, and T. Kanahori, “INFTY – An integrated OCR system for mathematical documents,” in *Proceedings of ACM Symposium on Document Engineering*, pp. 95–104. 2003.
- [5] T. V. Raman, *Audio system for technical readings*. PhD thesis, Cornell University, 1994.
- [6] M. Batusic, K. Miesenberger, and B. Stöger, “Access to mathematics for the blind – Defining HrTeX standard,” in *Proceedings of the 5th International Conference on Computers Helping People with Special Needs*, pp. 609–616. 1996.
- [7] K. Miesenberger, M. Batusic, and B. Stöger, “LABRADOOR: LaTeX-to-Braille-Door,” in *Proceedings of the Conference on Telematics in the Education of the Visually Handicapped*. Paris, 2–3 June, 1998.
- [8] D. Fitzpatrick and A. I. C. Monaghan, “Browsing technical documents: Document modelling and user interface design,” *Bulletin de Linguistique Appliquee et Generale* **24** (1999) 5–18.
- [9] N. Annamalai, D. Gopal, G. Gupta, A. Karshmer, and H. Guo, “INSIGHT: A comprehensive system for translating Braille based mathematical documents to LaTeX,” in *Proceedings of the International Conference on Human Computer Interaction*, pp. 1245–1249. 2003.
- [10] D. E. Knuth, *The TeXbook*. Addison-Wesley Professional, 1984.
- [11] L. Lamport, *LaTeX: A Document Preparation System*. Addison-Wesley Professional, 1994.
- [12] E. Meyer zu Bexten and M. Jung, “LaTeX at the University of Applied Sciences Giessen-Friedberg – Experiences at the Institute for Visually Impaired Students,” in *Proceedings of the 2004 CSUN International Conference on Technology and Persons with Disabilities*, pp. 508–509. 2004.