LaTeX, MATHML, AND TEX4HT: TOOLS FOR CREATING ACCESSIBLE DOCUMENTS

(A BRIEF TUTORIAL)

JACEK POLEWCZAK

***Note**: This tutorial can be found at* http://www.csun.edu/~hcmth008/mathml/acc_tutorial.pdf *(pdf format) or at* http://www.csun.edu/~hcmth008/mathml/acc_tutorial.html *(html format).*

Contents

1. Accessible documents

The aim of this tutorial is to present a selection of already available tools for creating accessible documents. The term *accessible* is understood here in the way W3C Accessibility Initiative understands it (see also, [1] for more information on ADA/508 compliance). While LaTeX provides a powerful desktop publishing tool for creating scientific documents, Mathematical Markup Language (MathML) facilitates the use of mathematical and scientific content on the Web. And TeX4ht is a tool for converting LaTeX input into hypertext document, including MathML. There is elegance and efficiency when the same LaTeX (ASCII) source file can produce different outputs; dvi, postscript, and pdf for printing/viewing, or XML/MathML for accessible viewing in browsers.

Note: This tutorial does not address a separate process of creating accessible personal webpages; at the same time, the techniques provided here produce hypertext documents that constitute standalone accessible content on the Web. An accesible front webpage, without accessible documents (subject lessons, essays, tests, homeworks, etc), will reduce itself to perhaps stylish, though empty shell.

2. What is LaTeX ?

LaTeX is a document markup language (as groff/troff and html languages are) for representing structured documents. LaTeX, initially designed and implemented by Leslie Lamport [2] in 1994, is based on Donald E. Knuth's work (1984) *The TeXbook* [3] and is essentially a collection of TeX macros.

TeX is a high quality typesetting program offering extensive desktop publishing features and automation, such as numbering and cross-referencing, tables and figures, detailed page layout, bibliographies, and indexing. Also, TeX/LaTeX is the only **VIABLE** tool for creating high quality documents that contain math/physics/chemistry/biology/engineering notations.

In contrast to most word processors, where one sees the document more or less as it will look when printed, LaTeX focuses on the meaning of what is being written without distractions by the visual presentation of the information.

Finally, Open Source TeX/LaTeX is a professional typesetting and publishing tool (used by major publishing houses) that is free to use and/or to modify.

3. What is MathML ?

MathML is an application of XML for describing mathematical notations, and capturing both their structure and content. It aims at integrating mathematical notation into World Wide Web documents so they can be accessible to the visually impaired.

As LaTeX, XML is a markup language for representing structured documents. However, in contrast to LaTeX, XML is **NOT** page layout language. Also, XML is an interchange and manipulation interface designed for machine, and not to be edited by humans.

3.1. **Presentation and Content MathML.** From Wikipedia entry for MathML:

> *MathML deals not only with the presentation but also the meaning of formula components (the latter part of MathML is known as "Content MathML"). Because the meaning of the equation is preserved separate from the presentation, how the content is communicated can be left up to the user. For example, web pages with MathML embedded in them can be viewed as normal web pages with many browsers but visually impaired users can also have the same MathML read to them through the use of screen readers (e.g. using the MathPlayer plugin for Internet Explorer, Opera 9.50 build 9656+ or the Fire Vox extension for Firefox).*
>
> ...
>
> *Presentation MathML focuses on the display of an equation, and has about 30 elements, and 50 attributes. The elements all begin with* `m` *and include token element:* `<mi>x</mi>` *- identifiers;* `<mo>+</mo>` *- operators;* `<mn>2</mn>` *- number. Tokens are combined using layout elements which include:* `<mrow>` *- a row;* `<msup>` *- superscripts;* `<mfrac>` *- fractions. The attributes mainly control fine details of the presentation. A large number of entities are available which represent letters* `&pi` (π, my addition); *symbols* `&RightArrow`; *and some non-visible character such as* `&InvisibleTimes`; *representing multiplication.*

This tutorial focuses only on *Presentation MathML*. (see, [4] for further information on MathML)

3.2. **The important qualifications.**

- TeX/LaTeX provides extremely detailed page layout. HTML/XML/MathML formats do not! They are *functional* mark-up languages and *NOT* page layout languages. Their exact rendering is not given by the document but decided by a browser, by a window's size, resolution, and font selection. The results are good for browsing but not for printing.
- The only way to produce precise page layouts is to represent documents in a page layout languages such as PDF, Postscript, or DVI formats. By the way, these are all open file formats.

4. What is TeX4ht ?

TeX4ht is a system that converts TeX/LaTeX inputs into various hypertext documents: HTML or XML/MathML:

$$\text{LaTeX input} \implies \text{TeX4ht} \implies \text{HTML/XML/MathML output}$$

TeX4ht has been designed and maintained by Eitan M. Gurari [5] and [6] (see also [7]). First, a LaTeX source code is compiled by TeX/LaTeX program together with loading of the additional macros for creating hooks in the output. Next, this output is post-processed by the program `tex4ht` to produce hypertext. Additional files, such as `.css` and, if needed, image files are created by the program `t4ht`.

5. How to do it ?

This tutorial is supplemented with ready to download and use complete TEX/LATEX/TeX4ht packages for Mac and Windows platforms (Sections 9 and 10, respectively). Linux packages are not included since the vast majority of Linux users have them already installed on their systems; however, just in case, I also provide Linux installation instructions (Section 11). In addition to the full TeX/LaTeX system, the packages also include additional tools like Ghostscript, Ghostview, dvips, image converters, as well as Firefox browser extensions: fonts package for better MathML rendering and Open Source Fire Vox screen reader (all platforms).

6. First steps with LATEX

LATEX file contains both the text and the instructions (the markup commands). The instructions tell LATEX how it is to appear. This file is usually created with system's text editor; the name of the file should end with `.tex` to identify the file's content. Let's say we call it `foo.tex`. When LATEX processes `foo.tex`, it creates a new file of typesetting commands, `foo.dvi`. `dvi` stands for `Device Independent` and `foo.dvi` is used to create output on printers; it is also used for viewing.

Typographical design is a craft and it is here where LATEX shines. In contrast to most WYSIWYG word processors, such as MS Word, LATEX concentrates on the logical structure rather than on the appearance of the document. Document design should make the document easier to read, not prettier. A basic set of standard `document classes` comes with LATEX: `article`, `book`, `report`, `letter`, and `slides`. These classes determine exactly how documents will be formatted: Additional document classes can be created by a user, although one should know basic principles of typographical design before starting to create a new document class.

The following simple LATEX file together with the interspersed comments, provides a good first look at the structure of a LATEX file (see also Figure 1 below).

In Figure 1 there are a number of words that start with \ (for example, see lines 11 and 12). These are LATEX commands that describe the structure of the document. All LATEX commands start with \ followed by one or more characters. LATEX commands are case sensitive: `\Begin` and `\begin` are not the same. There are also commands like `\command{text}`: e.g., `\emph{this is emphasized}` (line 26) or `\textbf{this is bold}` (line 27) in Figure 1. The actual text of the document always starts with `\begin{document}` and ends with an `\end{document}` command (see lines 12 and 40). Any text that comes after `\end{document}` command is ignored. At least one command must appear in the preamble, `\documentclass` command. In Figure 1, it is `\documentclass{article}` (line 11), which specifies that `article` class is use in the document. As mentioned above, there are other document classes, as well as there are many options in each class. There are also different environments, type styles, sectioning commands, tables of contents, tabular material, cross-referencing, citations, and indexing commands. For these and more, I refer the reader to a number of tutorials and books on how to start using LATEX; they are listed in Section 6.4.

6.1. Mathematical typesetting in LATEX.

Mathematical typesetting is different from text typesetting. There are two modes for mathematical expressions: *math mode* and *display math mode*.

Math mode commands are surrounded by `\(...\)` or by `$...$`, and thus `\(a^2+b^2=c^2\)` or `$a^2+b^2=c^2$` produce $a^2 + b^2 = c^2$.

```
% This is a small sample LaTeX input file (Version of 10 April 1994)
%
% Use this file as a model for making your own LaTeX input file.
% Everything to the right of a  %  is a remark to you and is ignored by LaTeX.

% The Local Guide tells how to run LaTeX.

% WARNING!  Do not type any of the following 10 characters except as directed:
%                &   $   #   %   _   {   }   ^   ~   \

\documentclass{article}        % Your input file must contain these two lines
\begin{document}               % plus the \end{document} command at the end.


\section{Simple Text}          % This command makes a section title.

Words are separated by one or more spaces.  Paragraphs are separated by
one or more blank lines.  The output is not affected by adding extra
spaces or extra blank lines to the input file.

Double quotes are typed like this: ``quoted text''.
Single quotes are typed like this: `single-quoted text'.

Long dashes are typed as three dash characters---like this.

Emphasized text is typed like this: \emph{this is emphasized}.
Bold       text is typed like this: \textbf{this is bold}.

\subsection{A Warning or Two}  % This command makes a subsection title.

If you get too much space after a mid-sentence period---abbreviations
like etc.\ are the common culprits)---then type a backslash followed by
a space after the period, as in this sentence.

Remember, don't type the 10 special characters (such as dollar sign and
backslash) except as directed!  The following seven are printed by
typing a backslash in front of them:  \$  \&  \#  \%  \_  \{  and  \}.
The manual tells how to make other symbols.

\end{document}                 % The input file ends with this command.
```

FIGURE 1. A sample LaTeX file

Display math mode commands are surrounded by `\[...\]`; and thus `\[a^2+b^2=c^2\]` produces a displayed equation

$$a^2 + b^2 = c^2$$

And here is another variant of *display math mode* that produces an equation number:

```
\begin{equation}
f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x)}{n!}
\end{equation}
```

generates

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} \tag{1}$$

6.2. **A typical command line session with LaTeX.**

`latex foo.tex` produces `foo.dvi` (dvi file)
`pdflatex foo.tex` produces `foo.pdf` (pdf file)
`dvips -o foo.ps foo.dvi` produces `foo.ps` (postscript file)

There is another variant (sometimes preferred) for producing pdf file from LaTeX file:

```
latex foo.tex
dvips -o foo.ps foo.dvi
ps2pdf foo.ps
```

produces `foo.pdf` (pdf file)

where `ps2pdf` is postscript to pdf converter included in most distributions of LaTeX.

Here is a pdf file produced by typesetting sample LaTeX file shown in Figure 1. The following LaTeX file is also worth looking into, if you are new to LaTeX. And here is its pdf output.

6.3. **Front-ends for LaTeX.** With the use of graphical front-ends there is no need to know many commands or technical details of LaTeX, or even type-in the above command lines. They also provide templates for most styles, macros for commands, and viewers for dvi files. Output pdf files can be viewed by standard pdf viewers, e.g., *Acrobat Reader*. Two Open Source front-ends, *TeXnicCenter* (Windows platform) and *TeXShop* (Mac platform) are included with the packages described in this tutorial (see Sections 9 and 10). Below, I provide the links to five Open Source front-ends examples and one shareware example that are easy to install and use. LaTeX distribution is required for typesetting LaTeX files with these editors.

- Kile – an integrated LaTeX editor for the KDE desktop environment. KDE is available for many architectures such as PC, PowerPC (Mac for example) and SPARC;
- Texmaker, available on all platforms;
- XEmacs, available for all platforms;
- GNU TeX, WYSIWYW (What You See Is What You Want, and not WYSIWYG) TeXmacs editor for scientists, available for all platforms;
- LyX - The Document Processor, available for all platforms;
- Winedit, a popular Windows only editor (shareware).

In the case of *LyX* and *GNU TeX* one does not have to know LaTeX at all; although when using LyX, one suffers from lack of portability with LaTeX users who do not use LyX. Just in case you wanted to know, I use Emacs for all my TeX/LaTeX work.

6.4. **Tutorials and books on LaTeX.**

- TUG (TeX Users Group) website, and especially http://www.tug.org/begin.html for many online LaTeX tutorials and documentation;
- LaTeX project's website;
- The list of many books on LaTeX and TeX from TUG website;
- G. Grätzer, Math into LaTeX, or its fourth edition, More Math into LaTeX;
- M. Goossens, F. Mittelbach, and A. Samarin, The LaTeX Companion;
- M. Goossens, S. Rahtz, and F. Mittelbach, The LaTeX Graphics Companion;
- M. Goossens, S. Rahtz, E. M. Gurari, R. Moore, and R. S. Sutor, The LaTeX Web Companion: Integrating TeX, HTML, and XML;
- Tobias Oetiker, The Not So Short Introduction to LaTeX2e;
- Jon Warbrick, Essential LaTeX ++;
- LaTeX cheat sheet;
- Tables of LaTeX commands for a large number of mathematical and other symbols, from G. Grätzer's, *More Math into LaTeX*.

There are so many advantages of using LaTeX (see, e.g., the list given in Tobias Oetiker, *The Not So Short Introduction to LaTeX2e*, Section 1.2.3). Furthermore, there are hundreds of free add-on packages for typesetting tasks ranging from applications in physics, chemistry, biology, engineering to typesetting music. Many of them are described in the above mentioned two books: M. Goossens, F. Mittelbach, and A. Samarin, *The LaTeX Companion* and M. Goossens, S. Rahtz, and F. Mittelbach, *The LaTeX Graphics Companion.*

There are also disadvantages: *LaTeX does not work well for people who have sold their souls...*

7. Adjusting your browser for MathML and for screen reader

You can download *Firefox* browser from the site:

http://www.mozilla.com/en-US/firefox (all platforms)

or *Opera* browser from the following site:

http://www.opera.com/download (all platforms)

Firefox (at least 2.0 and up, and perhaps, even older versions than 2.0) and Opera (at least 9.52 and up) should properly render MathML through native presentation on all three platforms (Linux/Mac/Windows). Actually, I only tested Firefox (2.0.0.16, 3.0.2-3) and Opera (9.52) browsers on Linux/Mac/Windows. For better rendering of MathML in Firefox, one should install additional STIX Beta fonts. They can be downloaded from:

http://www.csun.edu/~hcmth008/mathml/browsers/STIXBeta.zip

The fonts can be installed by extracting the files from the above Zip archive, then following the instructions for Microsoft Windows from:

http://www.microsoft.com/typography/TrueTypeInstallWin95.mspx,

or copying the files to a `~/Library/Fonts` folder on Mac OS X, or to a `~/.fonts` directory (which may need to be created) in Linux. More information can be found at:

http://www.mozilla.org/projects/mathml/fonts

The Internet Explorer (5.5 and up) browser renders MathML properly after installation of *MathPlayer* plugin from the *Design Science* website:

http://www.dessci.com/en/dl/MathPlayerSetup.asp

The link will download a file called `MathPlayerSetup.exe`. Make a note of where you download it to your computer. Once the download is completed, run MathPlayerSetup.exe.

I also suggest installing Adobe SVGview 3.03 that translates simple graphics into SVG images SVG images. SVG stands for Scalable Vector Graphics. Download the following file and follow the instructions:

http://download.adobe.com/pub/adobe/magic/svgviewer/win/3.x/3.03/en/SVGView.exe

I could not adjust *Safari* browser (used on Mac desktops) for proper MathML rendering.

7.1. **Testing your browser.** You can test for proper MathML rendering by opening the following two files in your browser:

- http://www.csun.edu/~hcmth008/mathml/browsers/test1/test1.xml
- http://www.csun.edu/~hcmth008/mathml/browsers/test2/test2.xml

For comparison, Figure 1 in `test1.XML` and `test2.xml` represents MathML rendering by your browser, while Figure 2 in `test1.xml` and `test2.xml` represents rendering of the same expression via a graphics image.

By right-clicking on Figure 1 in `test1.xml` and `test2.xml` and selecting *View MathML source*, Firefox browser opens a new window containing the MathML source of the corresponding math expression. In Internet Explorer, by right-clicking on Figure 1 and selecting *Copy MathML*, you can paste the corresponding MathML source code, for example, to Notepad and view it there.

Finally, here is a rather complicated and real life document; the original (unchanged source code) solutions to Assignment 2 (Applied Differential Equations, Math 280), given to my students in the Spring of 2008:

http://www.csun.edu/~hcmth008/mathml/browsers/test3/test3.xml, and for comparison, a pdf version of the same assignment:

http://www.csun.edu/~hcmth008/mathml/browsers/test3/test3.pdf

Note: *Internet Explorer without installed* MathPlayer *will not properly render MathML code, or may even refuse to open a file and may display an error message.*

7.2. **Enabling screen reader in Firefox.** Fire Vox is an Open Source, freely available talking browser extension for Firefox web browser created and maintained by Charles Chen. You can think of it as a screen reader that is designed especially for Firefox. The installation page provides instructions for all platforms. For Mac/Windows platforms download `.xpi` extension using the link:

http://firevox.clcworld.net/downloads.html.

(The current version is http://firevox.clcworld.net/clc-4-tts_bundle_v4.0_release.xpi)

Start Firefox and go to File, Open File. Then choose the file you just downloaded. Click the *Install Now* button on the window that pops up. Restart Firefox. Read the manual at:

http://clc4tts.clcworld.net/clc-firevox_doc.html

Tutorials can be found at:

http://firevox.clcworld.net/tutorial/tutorial.html

and Charles Chen's presentation video (in multimedia container format: avi):

http://firevox.clcworld.net/presentation.avi

For Linux, one has to install an additional `Java jar` package from:

http://firevox.clcworld.net/downloads.html

(The current version is http://firevox.clcworld.net/clc4tts_freetts_installer_1.2.jar

Go to the root directory of your *Java* distribution that is used by Firefox; next run the following command (as root):

```
./bin/java -jar clc4tts_freetts_installer_1.2.jar
```

The above command requires `clc4tts_freetts_installer_1.2.jar` file to be in the root directory of the *Java* distribution. Finally, restart Firefox.

Note: MathPlayer *plugin for Internet Explorer will speak MathML expression when you right-click on it and select* Speak Expression*, however, it is not a screen reader as* Fire Vox *is.*

8. TeX4ht in action

TeX4ht is a system that converts TeX/LaTeX inputs into various hypertext documents: HTML or XML/MathML. Unless specified otherwise the commands apply to Linux/Mac/Windows. The following Tex4ht command acting on `foo.tex`:

```
htlatex  foo.tex
```

produces `foo.html` (HTML 4.01 Transitional) version of `foo.tex`, along with some supplementary files (e.g., `.css` files and image files). Most math expressions are converted into graphics images.

On Linux/Mac platforms usage of Tex4ht is simplified via the Perl script `mk4ht` which can be called directly to combine various options. Thus,

```
mk4ht  oolatex  foo.tex   (Linux/Mac platforms)
oolatex  foo.tex          (Windows platform)
```

produces ODF format that can be read by *OpenOffice*.

The TeX4ht system has possibilities for using Unicode and fonts suited to the Gecko engine of the Mozilla browser. The command

```
mk4ht  mzlatex  foo.tex   (Linux/Mac platforms)
mzlatex  foo.tex          (Windows platform)
```

produces MathML/XML (XHTML 1.1 plus MathML 2.0) file, `foo.xml`, properly rendered by Firefox/Opera browsers, but not Internet Explorer browser.

For XHTML+MathML (XHTML 1.1 plus MathML 2.0) code to be served by *Firefox*, *Opera*, and Internet Explorer (with MathPlayer), one uses the following command line:

```
mk4ht mzlatex foo.tex "html,mathplayer"   (Linux/Mac platforms)
mzlatex foo.tex "html,mathplayer"         (Windows platform)
```

The result is `foo.xht` file.

Note: *Do* ***NOT*** *change the extension* `.xht` *to* `.xml`, *otherwise Internet Explorer will not properly render MathML code, or may even refuse to open a file and display error message. Firefox and Opera browsers are not affected by this change.*

Another versatile command of Tex4ht,

```
mk4ht mzlatex foo.tex "html,pmathml"   (Linux/Mac platforms)
mzlatex foo.tex "html,pmathml"         (Windows platform)
```

produces `foo.xml` file that renders properly Presentation MathML by detecting the rendering possibilities available to the current browser, any preferences specified in the document, and sets up an appropriate transformation. This is done by including additional XSLT stylesheets files: `pmathml.xsl` and `pmathmlcss.xsl` in the directory/folder along with the main document. The stylesheet file, pmathmlcss.xsl, transforms Presentation MathML to XHTML+CSS+JavaScript, so rendering MathML (to somewhat variable quality) in a standard HTML browser without any extra plugins. However, even here, *Safari* browser does not seem to work correctly with MathML.

The files, `pmathml.xsl` and `pmathmlcss.xsl`, can be downloaded from W3C site:

http://www.w3.org/Math/XSL/pmathml.xsl

http://www.w3.org/Math/XSL/pmathmlcss.xsl

Take a look at http://www.w3.org/Math/XSL/Overview-tech.html for further information/help.

Important Note: *Make sure that along with the output files* (`foo.html`, `foo.xml`, *or* `foo.xht`), *you also include (on the server side)* `.css` *and image* (`*.png` *or* `*.gif`) *files. Also, for large LaTeX documents with many graphics images it is convenient to keep images in separate directory/folder. With another option to the above commands,* `imgdir:images/`, *the browser will look for* `*.png` *files in* `images/` *sub-directory of the directory where,* `foo.tex` *is located. However, you need to put the images files to that directory. The corresponding commands are:*

```
mk4ht mzlatex foo.tex "html,mathplayer,imgdir:images/"   (Linux/Mac platforms)
mzlatex foo.tex "html,mathplayer,imgdir:images/"         (Windows platform)
```

or

```
mk4ht mzlatex foo.tex "html,pmathml,imgdir:images/"   (Linux/Mac platforms)
mzlatex foo.tex "html,pmathml,imgdir:images/"         (Windows platform)
```

The references [5], [6], and [7] provide the authoritative documentation on more options of TeX4ht as well as additional help.

9. Download and Installation instructions for Mac platform

The full TeX/LaTeX (TeX Live 2008, Mac edition) you are about to download is provided by TUG (TeX Users Group) with the relevant Mac site at http://www.tug.org/mactex.

Download `MacTeX.mpkg.zip` package from

http://www.csun.edu/~hcmth008/mathml/mac/MacTeX.mpkg.zip

The above file is $> 1.1GB$. Downloading it from your CSUN's office is perhaps your best option. Next, unzip it by double-clicking on `MacTeX.mpkg.zip`. `MacTeX.mpkg` package will be created. Double click on it and follow the instructions. If asked for paper size, choose `letter` or `A4` (most likely you want `letter`). You can change this setting later by entering one of the following commands in Apple's Terminal program:

`sudo tlmgr paper letter` (for letter size)

or

`sudo tlmgr paper A4` (for A4 size).

(enter your user password when asked)

The installation process will take a while; at the end, MacTeX-2008 distribution will be installed in `/usr/local/texlive/2008` folder, occupying at least 1.8GB of disk space. If you want to know more details about what was installed read *Read Me First* file in `/Applications/TeX` folder. The same folder contains several applications, among them, front-end to TeX/LaTeX, *TeXShop* application. Double-click on it and open *Preferences* menu. Go to *Typesetting* menu and make sure *LaTeX* (under *Default Command*) and *TeX+Ghostscript* (under *Default Script*) are checked. Restart *TeXShop*. The default *Pdftex* under*Default Script* works with LaTeX files that do NOT have embedded graphics files, while checking *TeX+Ghostscript* in *Default Script* works well in all cases.

For testing purposes, download

http://www.csun.edu/~hcmth008/mathml/mac/test1_testing.tex,

or

http://www.csun.edu/~hcmth008/mathml/mac/test2_testing.tex,

and open `test1_testing.tex` in *TeXShop*. Click on *Typeset* button (located in the upper left corner of the screen). You should see two new windows opened: one is *test1_testing console* that stores log files of compilation process and the other window is Mac's pdf viewer with `test1_testing.pdf`. For comparison, here is what you should see in pdf viewer window:

http://www.csun.edu/~hcmth008/mathml/mac/test1_testing.pdf

You can also try sample LaTeX file and another LaTeX file from Section 6.2. The most frequently asked questions about MacTeX can be found at:

http://www.tug.org/mactex/faq

9.1. **Adjusting TeX4ht.** For proper working of TeX4ht with *Internet Explorer's MathPlayer* download the following file to your Desktop folder:

http://www.csun.edu/~hcmth008/mathml/mac/mathplayer.4ht.

Then, open Apple's Terminal program and enter the following command in its window:

```
mkdir -p Library/texmf/tex/generic/tex4ht
```

Next, move `mathplayer.xht` to `~Library/texmf/tex/generic/tex4ht` directory by executing the command:

```
mv Desktop/mathplayer.4ht Library/texmf/tex/generic/tex4ht/
```

Note: *The above procedure works only for a user that has downloaded the above file. On the other hand, the following procedure works for all users on a computer: use* `sudo` *to copy the file* `mathplayer.4ht` *to the directory:* `/usr/local/texlive/2008/texmf-dist/tex/generic/tex4ht/`. *In this process you overwrite the existing file.*

While Apple's terminal is still open, for testing purposes, create the folder `~tests` folder by executing the command:

```
mkdir tests
```

and move the following file:

http://www.csun.edu/~hcmth008/mathml/mac/test1_testing.tex

to `~tests/`directory. From `~tests` directory enter the following command in Apple's Terminal window:

```
mk4ht mzlatex test1_testing.tex "html,mathplayer"
```

and open the output, `test1_testing.xht`, in *Firefox* or *Opera* browser. Check your browser rendering of `test1_testing.xht` with

http://www.csun.edu/~hcmth008/mathml/mac/test1_testing.pdf

9.2. **Optional Installation items.**

(1) The command `mk4ht mzlatex myfile.tex "html,mathplayer"` is quite long and thus time consuming to type in, especially when entered many times. The following procedure simplifies this process: download the batch (bash) script to Desktop folder:
http://www.csun.edu/~hcmth008/mathml/mac/tomath
and perform the following commands in Apple's Terminal window:

```
chmod a+x Desktop/tomath
sudo mv Desktop/tomath /usr/texbin/
```

(enter your user password when asked)

Now, in order to produce hypertext document with MathML, properly rendered by *Firefox, Opera,* and *Internet Explorer* browsers, enter:

```
tomath myfile.tex
```

in Apple's Terminal window. The output `myfile.xht` will be created. By the way, the name of the file `tomath` can be changed to anything you want, as long its name doesn't match another executable file.

Note 1: *Make sure that along with the output file* `myfile.xht`, *you also include (on the server side)* `myfile.css` *and image* (`*.png` *or* `*.gif`) *files.*

Note 2: *While working in Apple's Terminal window, you can invoke previously entered commands by pressing (repeatedly, if necessary) the* `Upper arrow` *key on your keyboard.*

(2) SimpleTeX4ht is an Open Source Mac OS X graphical user interface for TeX4ht.
Note 3: *SimpleTeX4ht depends on proper setup of TeX4ht and the use of commands from Section 9.1.*
Here is one download site:
http://www.csun.edu/~hcmth008/mathml/mac/SimpleTeX4ht.dmg

After installation, you can create various hypertext documents by simply dragging files into *SimpleTeX4ht* window. There are three menus: *HTML*, *Other formats*, and *Expert*. When you check *MathML* button in *Other formats* menu and drag a LaTeX file, say `myfile.tex`, into *SimpleTeX4ht* window, the XML/MathML output produced (`myfile.xml`) is the same as using `mk4ht mzlatex myfile.tex` command in Apple's Terminal window. Both *Firefox* and *Opera* render correctly included MathML code, but not *Internet Explorer*.

For proper rendering of MathML by all three browsers click on *Expert* menu. Enter the following three options, in the order shown below, to the top three dialog boxes in *Expert* menu:

```
xhtml,mozilla,mathplayer
-cmozhtf
-cvalidate
```

Then drag your LaTeX file, `myfile.tex`, into *SimpleTeX4ht* window and click *Convert* button. The output produced (`myfile.xht`) is the same as using

```
mk4ht mzlatex myfile.tex "html,mathplayer"
```

command in Apple's Terminal window. It is unfortunate that the entries in *Expert* menu are not preserved to the next *SimpleTeX4ht* session. Hopefully it will be changed in the next version of *SimpleTeX4ht*. Time permitting, I may try to recompile the source code of *SimpleTeX4ht* to include this option.

Enjoy!

10. Download and Installation instructions for Windows platform

The full TeX/LaTeX (Windows based MiKteX) you are about to download is provided by TUG (TeX Users Group) at

http://www.tug.org/protext

Download it from

http://www.csun.edu/~hcmth008/mathml/windows/ProTeXt-2.2-071608.exe

It is a large $\approx 700MB$ compressed executable. Downloading it from your CSUN's office is perhaps your best option. In order to proceed further you need to have a pdf reader. You can download free Adobe Acrobat Reader , or you can search the Web for other pdf readers.

Create a new folder `C:\tmp` and move the just downloaded file `ProTeXt-2.2-071608.exe` to `C:\tmp`. Double-click on it and, after a while, a new folder `C:\tmp\ProTeXt-2.2` will be created. Next, double-click on `Setup.exe` (or `Setup`, depending how you setup your file manager). Setup program actually opens the interactive pdf file `protext-install-en.pdf`. (If you want you can look at it in `C:\tmp\ProTeXt-2.2\Install\protext-install-en.pdf` before installation process.) The process consists of installing four application: MiKTeX, TeXnicCenter, Ghostscript, and GSview. Install all of them. If you have other versions of these applications, you will be guided

to remove them, by clicking on the indicated links in the pdf file. Also, if you have another version of TEX/LATEX, **REMOVE IT**, before installing MiKTeX. Now, it is possible to have several versions of TEX/LATEX system, however, taking care of it goes beyond this tutorial. Although not required, it is advisable you install MiKTeX to a folder with the name not containing spaces, for example, `C:\MiKTeX2_7`. You are installing MiKTeX, version 2.7. The instructions in Section 10.2 use `C:\MiKTeX2_7` as the root directory for MiKTeX.

During the installation you will be asked to select paper size, `letter` or `A4`, also choose to install additional packages on-the-fly. You system must have working network connection. Depending on your system it make take from 30 to 45 minutes for the full installation of MiKTeX. Then read section 1.6.3 (*After the installation*) to perform the first run of LATEX. Open one of the following files in *TeXnicCenter*:

http://www.csun.edu/~hcmth008/mathml/windows/test1_testing.tex

http://www.csun.edu/~hcmth008/mathml/windows/test2_testing.tex

next, press `Ctrl+F7` to typeset the file and then `F5` to view the output.

You can also try sample LATEX file and another LATEX file from Section 6.2.

10.1. **Adjusting TeX4ht.** *Below, I assume that you installed MiKTeX to* `C:\MiKTeX2_7` *folder. If this is not the case, adjust the instructions below by entering a proper (root) folder name of MiKTeX distribution.*

Move the content (but not the folder `tex4ht`) of `C:\MiKTeX2_7\tex\generic\tex4ht` to another folder, e.g., `C:\tmp\tex4ht_original` (create it before this operation).

Download the following zip file:

http://www.csun.edu/~hcmth008/mathml/windows/tex4ht_new.zip

to `C:\tmp`, open it and copy the content to `C:\MiKTeX2_7\tex\generic\tex4ht`.

In *Command Prompt* window (found in *Accessories* sub-menu of *All Programs* menu) enter the command:

```
texhash
```

This will update database for MiKTeX system. You can achieve the same by invoking *Setting* menu in *MiKTeX* submenu of *All Programs* menu, and press *Refresh FNDB* button. As you can see by this example, entering a simple command in DOS prompt window sometimes can be done faster than by invoking GUI (Graphical User Interface).

Finally, TeX4ht needs image converter (to `png` or `gif` format), if LATEX file has embedded graphics. Open Source ImageMagic is used for this purpose. There are two versions of the program you can install. Version with Q8 in the name produces 8 bits-per-pixel component (e.g. 8-bit red, 8-bit green, etc.), whereas, Q16 in the file name produces 16 bits-per-pixel component. Q16 version permits you to read or write 16-bit images without losing precision but requires twice as much resources as the Q8 version. You can download them from:

http://www.csun.edu/~hcmth008/mathml/windows/ImageMagick-6.4.4-4-Q8-static.exe

http://www.csun.edu/~hcmth008/mathml/windows/ImageMagick-6.4.4-0-Q16-static.exe

Double click on the downloaded file and follow the instructions for installation. Accept all default choices.

For testing TeX4ht, download the file:

http://www.csun.edu/~hcmth008/mathml/windows/test1_testing.tex

to `C:\tmp` directory. Next, enter the following commands in DOS *Command Prompt* window:

```
cd C:\tmp
```

```
htlatex test1_testing.tex
```

Then open `test1_testing.html` in your browser (*Firefox*, *Opera*, or *Internet Explorer*). For comparison, check your browser rendering with

http://www.csun.edu/~hcmth008/mathml/mac/test1_testing.pdf

For pure MathML code, enter the following command:

```
mzlatex test1_testing.tex "html,mathplayer"
```

and open the output, `test1_testing.xht`, in a browser.. As above, check your browser rendering of `test1_testing.xht` with

http://www.csun.edu/~hcmth008/mathml/mac/test1_testing.pdf

Note: *While working in DOS Command Prompt window, you can invoke previously entered commands by pressing (repeatedly, if necessary) the* `Upper arrow` *key on your keyboard.*

10.2. **Optional Installation items.**

(1) The command `mzlatex myfile.tex "html,mathplayer"` is quite long and thus time consuming to type in. The following procedure simplifies this process: download the DOS batch file to Desktop folder:
http://www.csun.edu/~hcmth008/mathml/windows/tomath.bat
and copy it to `C:\MiKTeX2_7\miktex\bin` directory. I assume here that you installed MiKTeX `C:\MiKTeX2_7` directory.
Now, in order to produce hypertext document with MathML, properly rendered by *Firefox, Opera,* and *Internet Explorer* browsers, enter:
`tomath myfile.tex`
in DOS Command Prompt window. The output `myfile.xht` will be created. By the way, the name of the file `tomath` can be changed to anything you want, as long its name doesn't match another executable file.

(2) TeX Converter is an Open Source Windows graphical user interface for TeX4ht.
***TeX Converter* depends on proper working of TeX4ht using the commands in Section 10.1.** Follow the instructions in the order given below:
(a) After downloading the files:
http://www.csun.edu/~hcmth008/mathml/windows/Htrun.exe
http://www.csun.edu/~hcmth008/mathml/windows/HTRUN.INI
to the Desktop folder, copy both of them to `C:\MiKTeX2_7\miktex\bin` directory. As in item (1) above, I assume here that you installed MiKTeX to `C:\MiKTeX2_7` directory.
(b) Download two files:
http://www.csun.edu/~hcmth008/mathml/windows/texconv.zip
http://www.csun.edu/~hcmth008/mathml/windows/TeXConverter.ini

to C:\tmp. Double-clicking on texconv.zip will unzip the content to C:\tmp\texconv folder. Next, double-click C:\tmp\texconv\SETUP.EXE file to install *TeX Converter*. Accept all default choices.

Finally, copy C:\tmp\TeX Converter.ini to C:\Program Files\TeX Converter folder. This will replace the existing TeX Converter.ini file and pre-assigns many choices in *TeX Converter* for quick use. After invoking *TeX Converter* you will see seven buttons at the top of the window: HeVeA, Tth, TeX4ht, LaTeX2HTML, DVI, PDF, HELP. Only TeX4ht, DVI, PDF, HELP are functional for this setup. The other buttons correspond to non-existing modules on your computer and I will not describe them in this tutorial.

TeX4ht setup in *TeX Converter* (press TeX4ht button, if necessary) corresponds to the following command line:
mzlatex myfile.tex "html,mathplayer,imgdir:images/
with myfile.xht as the output. I pre-assigned saving output to C:\tmp, however, you can change this destination as you wish by pressing Change Directory button in the lower right of the window. For testing TeX4ht, download the file:
http://www.csun.edu/~hcmth008/mathml/windows/test1_testing.tex
to C:\tmp and point to it in the dialog box of *TeX4ht* menu that is located in the upper left corner. Then, press Convert button. DOS Command Prompt window will open showing the progress of the conversion process. When the conversion is completed Press any key to continue... to exit DOS Command Prompt window. Firefox browser will open test1_testing.xht file. *Internet Explorer* will be used if you check the button Use own Web Browser. Finally, the output is stored in C:\tmp\test1_testing folder, while possible graphics images are stored in C:\tmp\test1_testing\images folder. In general when you convert myfile.tex, the output is stored in C:\tmp\myfile folder and graphics images are in C:\tmp\myfile\images folder. Both names, C:\tmp and images, can be changed at any time. Finally, press Advanced button to see many different options you can choose. If after you made changes, *TeX Converter* doesn't work properly, replace C:\Program Files\TeX Converter\TeX Converter.ini file with the one you have in C:\tmp, or found at
http://www.csun.edu/~hcmth008/mathml/windows/TeXConverter.ini

You can also Press DVI or PDF buttons to enter windows, where you can typeset your LaTeX file to produce .dvi or .pdf output and view them. You can also edit your LaTeX file by pressing Edit Tex File button. *TeXnicCenter* program will open your file.

Enjoy!

11. Installation instructions for Linux platform

Almost every Linux distribution has TeX/LaTeX and TeX4ht packages. Below, I provide the list of them for Fedora Core 8 and Fedora Core 9 distributions.

Fedora Core 8

tetex-3.0-44.9.fc8.i386.rpm
tetex-afm-3.0-44.9.fc8.i386.rpm
tetex-doc-3.0-44.9.fc8.i386.rpm
tetex-dvipost-1.1-8.fc8.i386.rpm

tetex-dvips-3.0-44.9.fc8.i386.rpm
tetex-eurofont-1.1.3-6.fc6.noarch.rpm
tetex-fonts-3.0-44.9.fc8.i386.rpm
tetex-IEEEtran-1.7.1-1.fc8.noarch.rpm
tetex-latex-3.0-44.9.fc8.i386.rpm
tetex-preview-11.85-1.fc8.noarch.rpm
tetex-prosper-1.5-3.fc6.noarch.rpm
tetex-tex4ht-1.0.2008_02_28_2058-3.fc8.1.i386.rpm
tetex-xdvi-3.0-44.9.fc8.i386.rpm

With Yum installer it is easy to install the above packages. Use the command (as root):

```
yum install tetex
```

to install the first package in the above list. If `tetex` package is already installed, Yum installer will inform you about that and you can continue with the second on the above list, and so on. Also, Yum will automatically install any dependency packages needed for the one you selected.

Fedora Core 9

tetex-IEEEtran-1.7.1-1.fc8.noarch.rpm
tetex-bytefield-1.2a-3.fc6.noarch.rpm
tetex-dvipost-1.1-9.fc9.i386.rpm
tetex-elsevier-0.1.20071024-1.fc9.noarch.rpm
tetex-eurofont-1.1.3-6.fc6.noarch.rpm
tetex-font-cm-lgc-0.5-9.fc9.noarch.rpm
tetex-font-kerkis-2.0-15.fc9.noarch.rpm
tetex-fonts-hebrew-0.1-8.fc9.noarch.rpm
tetex-perltex-1.3-1.fc6.noarch.rpm
tetex-prosper-1.5-3.fc6.noarch.rpm
tetex-tex4ht-1.0.2008_02_28_2058-3.fc9.i386.rpm
tetex-unicode-0-7.20041017.fc6.noarch.rpm
tex-preview-11.85-7.fc9.noarch.rpm
texi2html-1.78-3.fc8.noarch.rpm
texlive-2007-30.fc9.i386.rpm
texlive-afm-2007-30.fc9.i386.rpm
texlive-context-2007-30.fc9.i386.rpm
texlive-doc-2007-30.fc9.i386.rpm
texlive-dvips-2007-30.fc9.i386.rpm
texlive-dviutils-2007-30.fc9.i386.rpm
texlive-east-asian-2007-30.fc9.i386.rpm
texlive-latex-2007-30.fc9.i386.rpm
texlive-texmf-2007-22.fc9.noarch.rpm
texlive-texmf-afm-2007-22.fc9.noarch.rpm
texlive-texmf-context-2007-22.fc9.noarch.rpm
texlive-texmf-doc-2007-22.fc9.noarch.rpm
texlive-texmf-dvips-2007-22.fc9.noarch.rpm
texlive-texmf-east-asian-2007-22.fc9.noarch.rpm
texlive-texmf-errata-2007-4.fc9.noarch.rpm

texlive-texmf-errata-afm-2007-4.fc9.noarch.rpm
texlive-texmf-errata-context-2007-4.fc9.noarch.rpm
texlive-texmf-errata-doc-2007-4.fc9.noarch.rpm
texlive-texmf-errata-dvips-2007-4.fc9.noarch.rpm
texlive-texmf-errata-east-asian-2007-4.fc9.noarch.rpm
texlive-texmf-errata-fonts-2007-4.fc9.noarch.rpm
texlive-texmf-errata-latex-2007-4.fc9.noarch.rpm
texlive-texmf-errata-xetex-2007-4.fc9.noarch.rpm
texlive-texmf-fonts-2007-22.fc9.noarch.rpm
texlive-texmf-latex-2007-22.fc9.noarch.rpm
texlive-texmf-xetex-2007-22.fc9.noarch.rpm
texlive-utils-2007-30.fc9.i386.rpm
texlive-xetex-2007-30.fc9.i386.rpm

12. Final Remarks

(1) When you create a new document and mistype or incorrectly enter markup command(s), LaTeX will not typeset your document and displays error message. After typing `?`, LaTeX stops and waits for you to tell it what to do. You can type `x` and return to your editor and correct the mistake(s). The best thing is to hit `Return` key, several times if needed, and allow LaTeX to run as if nothing had happened and go back to your editor to correct your LaTeX file. The books and tutorials in 6.4, provide additional details on handling typesetting errors. For quick introduction on this topic go to Tobias Oetiker, The Not So Short Introduction to LaTeX2e; or Jon Warbrick, Essential LaTeX ++. Finally, TeX4ht will not convert your document to hypertext if LaTeX reports typesetting error(s).

(2) Use of front-ends for LaTeX is not always useful. They often introduce their own styles that make your LaTeX file difficult, if not impossible, to typeset by LaTeX on other computers without this particular front-end editor. Using standard styles is the best option to have your LaTeX file portable.

(3) **Warning:** upgrading TeX4ht using MikTeX package manager can make it unusable.

(4) The current setting for the CSUN web server do not allow for proper rendering of XHTML documents with `.xht` and `.xhtml` extensions. Creating MathML Web sites from Design Science describes in some details those issues. One way to bypass the present CSUN web server settings is to create XML/MathML documents using (see Section 8) the following command:

```
mk4ht  mzlatex  foo.tex  "html,pmathml"   (Linux/Mac platforms)
mzlatex  foo.tex  "html,pmathml"          (Windows platform)
```

This will create `foo.xml` output. However, two files, `pmathml.xsl` and `pmathmlcss.xsl` must reside in the same directory/folder as `foo.xml`. They can be downloaded from W3C site:
`http://www.w3.org/Math/XSL/pmathml.xsl`
`http://www.w3.org/Math/XSL/pmathmlcss.xsl`

(5) My slide presentation on creating accessible documents is at
`http://www.csun.edu/~hcmth008/mathml/converting_to_mathml.pdf`

(6) XML/MathML version of this document can be found at
`http://www.csun.edu/~hcmth008/mathml/tutorial/acc_tutorial.xml`

References

[1] Section 508 software accessibility standards, see also http://www.section508.gov
[2] Leslie Lamport, *A document Preparation System*
[3] D. E. Knuth, *The TEXbook*
[4] MathML Frequently Asked Questions: http://www.w3.org/Math/mathml-faq.html
[5] E. M. Gurari, The authoritative documentation for TeX4ht
[6] E. M. Gurari, MathML via TeX4ht and other tools
[7] E. M. Gurari and S. Rathz, From LaTeX to MathML and Back with TeX4ht and PassiveTeX
[8] The W3C MathML software list

Mathematics, CSUN, Northridge

E-mail address: jacek.polewczak@csun.edu