

# GELLMU

## Introductory Survey

A Bridge for Authors from  $\LaTeX$  to XML

*William F. Hammond*

Last updated: April 27, 2011

### Recent

- MATHML Demos of past articles from *The New York Journal of Mathematics*<sup>1</sup> made with GELLMU. (15 Oct. 2006: These were rebuilt on 27 Apr. 2011 to add an HTML output with MathML output using *MathJax*<sup>2</sup>).
- **News:** As of March, 2011 the current versions of the “big four” web browsers all support HTML with *MathJax*.
- **Download**<sup>3</sup> (v. 0.8.4.1, 6 July 2007)  
Since the version 0.8.0.5 release in October 2004, “regular” GELLMU has provided a reliable dual presentation route for math articles written in  $\LaTeX$ -like source language to (a) print output and (b) online display in XHTML extended by Mathematical Markup Language (MATHML)<sup>4</sup>.
- (14 Oct 2010) Video, slides, and the write-up for my talk  **$\LaTeX$  Profiles as Objects in the “Category” of Markup Languages**<sup>5</sup> on June 29, 2010 at the 32nd annual meeting of TUG<sup>6</sup> in San Francisco.
- (14 Oct 2009) Bleeding edge materials<sup>7</sup> for making XHTML+MathML **slides** (rather than using PDF made with, say, “Beamer”), using W3C’s *slidy*, with an add-on to the didactic production system. (Sorry, no docs for this yet. See the example for TUG 2007, linked below.)
- (30 May 2009) For those who lack a local-platform HTML/XHTML validation set-up, the small package *gvalidhtml*<sup>8</sup> should suffice. (Simply open the tarball in `/usr/local/share` and make a symbolic link from the name “`validhtml`” somewhere in your command path to the script named “`gvalidhtml`” in the untarred directory.)

---

<sup>1</sup>URI: <http://www.albany.edu/~hammond/demos/nyj/>

<sup>2</sup>URI: <http://www.mathjax.org/>

<sup>3</sup>URI: <http://www.albany.edu/~hammond/gellmu/tarball.html>

<sup>4</sup>URI: <http://www.w3.org/Math/>

<sup>5</sup>URI: <http://www.albany.edu/~hammond/presentations/Tug2010/>

<sup>6</sup>URI: <http://www.tug.org/tug2010/>

<sup>7</sup>URI: <http://www.albany.edu/~hammond/gellmu/gslidy.tgz>

<sup>8</sup>URI: <http://www.albany.edu/~hammond/gellmu/gvalidhtml-0.9.tgz>

- (22 Jul 2007) Video, slides, and pre-meeting notes from my talk *Dual Presentation with Math from one Source*<sup>9</sup> on July 19, 2007 at TUG 2007<sup>10</sup> in San Diego.
- Latest online version of the *Introductory User's Guide to Regular GELLMU*:
  - HTML with MathJax<sup>11</sup>
  - XHTML+MathML<sup>12</sup>
  - terminal window HTML<sup>13</sup>
  - PDF: formatted for 8.5 x 11 in.<sup>14</sup>
  - DVI: formatted for 8.5 x 11 in.<sup>15</sup>
  - GELLMU source<sup>16</sup>
- Latest online version of the *GELLMU Manual*:
  - HTML with MathJax<sup>17</sup>
  - XHTML+MathML<sup>18</sup>
  - terminal window HTML<sup>19</sup>
  - PDF: formatted for 8.5 x 11 in.<sup>20</sup>
  - DVI: formatted for 8.5 x 11 in.<sup>21</sup>
  - Plain text<sup>22</sup> (made from classical HTML using *lynx*)
  - GELLMU source<sup>23</sup>
- Author-level XML document type definition<sup>24</sup> for the markup used in *Regular GELLMU*. It's the "side entrance" in the Regular GELLMU Flow Chart<sup>25</sup>.



- Understand the system by **studying examples**<sup>26</sup>.

<sup>9</sup>URI: <http://math.albany.edu/math/pers/hammond/Presen/Tug2007/>

<sup>10</sup>URI: <http://www.tug.org/tug2007/>

<sup>11</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/userdoc5.html>

<sup>12</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/userdoc.xhtml>

<sup>13</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/userdoc.html>

<sup>14</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/userdoc.pdf>

<sup>15</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/userdoc.dvi>

<sup>16</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/userdoc.glm>

<sup>17</sup>URI: <http://www.albany.edu/~hammond/gellmu/glman/glman5.html>

<sup>18</sup>URI: <http://www.albany.edu/~hammond/gellmu/glman/glman.xhtml>

<sup>19</sup>URI: <http://www.albany.edu/~hammond/gellmu/glman/glman.html>

<sup>20</sup>URI: <http://www.albany.edu/~hammond/gellmu/glman/glman.pdf>

<sup>21</sup>URI: <http://www.albany.edu/~hammond/gellmu/glman/glman.dvi>

<sup>22</sup>URI: <http://www.albany.edu/~hammond/gellmu/glman/glman.txt>

<sup>23</sup>URI: <http://www.albany.edu/~hammond/gellmu/glman/glman.glm>

<sup>24</sup>URI: <http://www.albany.edu/~hammond/gellmu/xml/axgellmu.dtd>

<sup>25</sup>URI: <http://www.albany.edu/~hammond/gellmu/glman/glman.html#flow>

<sup>26</sup>URI: <http://www.albany.edu/~hammond/gellmu/examples/>

- The latest *fully tested* version of the **syntactic translator** `gellmu.el`<sup>27</sup>, which is all that is required for *basic* GELLMU.

## Table of Contents

1	Introduction.....	3
2	First Demonstrations.....	4
3	Can Content-Level MATHML be a Derived Format?.....	5
4	Brief Introductions.....	6
	4.1 Basic GELLMU.....	6
	4.2 Advanced GELLMU.....	6
	4.3 <i>Regular</i> GELLMU: The Didactic Production System.....	6
	4.4 Other Production Systems.....	8
5	Materials.....	8
6	Relevant Public Discussion and Comment.....	8
7	Pointers to a Few Related Things.....	9
8	About this Document.....	12

## 1 Introduction

Generalized Extensible L<sup>A</sup>T<sub>E</sub>X-Like Markup (GELLMU) is my concept for using L<sup>A</sup>T<sub>E</sub>X-like markup to create documents in an easy plain text format that may be faithfully converted to high-powered documents marked up under SGML. T<sub>E</sub>X is the classical typesetting markup language (with robust handling of mathematics) that was created by Donald E. Knuth of Stanford University around 1980. The *L<sup>A</sup>T<sub>E</sub>X document preparation system* was created shortly thereafter by Leslie Lamport of Digital Equipment Corporation. L<sup>A</sup>T<sub>E</sub>X is a simplified markup interface to T<sub>E</sub>X designed to let “the user concentrate on the structure of the text” rather than on typesetting. SGML, an abbreviation for *Standard Generalized Markup Language* (ISO 8879:1986), is the name of a family of markup languages, unspecified in number, designed for efficient automatic text processing with shared tools of a certain type.

During the period 1993-1998 the most familiar example of a markup language in the SGML family was *Hypertext Markup Language* (HTML), the now familiar language of the World Wide Web<sup>28</sup>. HTML is a rather low-powered member of the SGML family. The notion of “power” for a language under the umbrella of SGML has to do with the number of available translations to other document languages, both within and without SGML.

---

<sup>27</sup>URI: <http://www.albany.edu/~hammond/gellmu/gellmu.el>

<sup>28</sup>URI: <http://www.w3.org/>

One of the ideas in my design for GELLMU is that with existing stable freely available SGML tools one may go to almost any presentation format. For the community of mathematicians and scientists, who have become accustomed to using  $\TeX$  to create finely typeset documents for printing, this design provides a way automatically to create other carefully crafted forms from a single source document without over-burdening Donald Knuth's program  $\TeX$ .

For typeset printed presentation, SGML-based processing to the language  $\TeX$  should be optimal, while SGML-based processing to Lamport (v.2)  $\LaTeX$  is used in didactic examples found below. (See also "`jadetex`" at The Comprehensive  $\TeX$  Archive Network (CTAN); brief comment on "`jade`" may be found below.) Most of the magic is due to Charles Goldfarb, the inventor of SGML, James Clark, the author of "`nsgmls`", and David Megginson, the author of "`sgmlsp1`". The GELLMU to SGML transliterator that I am still writing could have been done in many languages, but ELISP, the language of GNU Emacs, probably the best-documented of all languages, and probably also the most easy-to-debug general purpose language, seemed to be just right for this. Beyond that I am grateful to Richard Stallman for encouragement and answers. Of course, when things do not work, the problems should in no way be attributed even in part to the antecedent work.

## 2 First Demonstrations

For a quick look, intended for those who know  $\LaTeX$ , there is *A Silly Little GELLMU Article*<sup>29</sup> of about three printed pages. Alongside the HTML form of this article are other versions:

- the original GELLMU source markup<sup>30</sup>.
- its syntactic translation to an SGML document<sup>31</sup>, involving only minimal knowledge of the document type.
- its translation to an XML dialect<sup>32</sup>, that is essentially equivalent to the source.
- $\LaTeX$  source that was generated from the XML version<sup>33</sup>
- DVI<sup>34</sup> made from the  $\LaTeX$  version.
- PDF<sup>35</sup> made with *pdflatex*.
- XHTML+MATHML<sup>36</sup>.

And yes, of course, both HTML versions were generated from the XML version.

---

<sup>29</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/silly.html>

<sup>30</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/silly.glm>

<sup>31</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/silly.sgml>

<sup>32</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/silly.xml>

<sup>33</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/silly.ltx>

<sup>34</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/silly.dvi>

<sup>35</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/silly.pdf>

<sup>36</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/silly.xhtml>

### 3 Can Content-Level MathML be a Derived Format?

Mathematical Markup Language (MATHML) is a language under development by the World Wide Web Consortium (W3C)<sup>37</sup> for (1) the display of mathematics in ordinary web pages and (2) automated interchange of mathematical segments among web-compatible software applications.

Corresponding to (1) and (2) above the W3C has provided presentation and content-level versions of MATHML.

While MATHML, which is an XML language (formally “application”), is verbose to a point that makes its writing by human authors almost impossible, the W3C project has not undertaken to provide a language suitable for authors. Moreover, one cannot robustly translate well-structured standard L<sup>A</sup>T<sub>E</sub>X or T<sub>E</sub>X math segments into MATHML without the discipline of rules that are difficult both to formulate and to enforce.

The concept of *generalized L<sup>A</sup>T<sub>E</sub>X* in the GELLMU Project provides such discipline.

The version of “Regular” (see §4.3) GELLMU in the **tarball**<sup>38</sup>, has, since August 2004, provided translation of generalized L<sup>A</sup>T<sub>E</sub>X source markup under the *article* document type to HTML with presentation-level MATHML as well as translation to ordinary L<sup>A</sup>T<sub>E</sub>X.

The key question in designing a system sufficient for generation of mathematics under an umbrella like content-level MATHML either using highly specialized L<sup>A</sup>T<sub>E</sub>X or using an SGML or XML language for authors is how far authors will be willing to diverge from past habits.

The Math Benchmark Document<sup>39</sup> offers an example of various mathematical segments that one might want to have automatically translated to a language with relative semantics such as content-level MATHML.

There is something of an explanation (now in early draft stage<sup>40</sup>), familiar to many research mathematicians but perhaps not to so many computer scientists, of why most legacy T<sub>E</sub>X/L<sup>A</sup>T<sub>E</sub>X markup of mathematics is **not** ambiguous for robots when augmented by adequate “type” information. Legacy practice has been to include “type” information in paper documents as part of an article’s descriptive text. In a few words, mathematicians are usually careful and fussy about notation. GELLMU will eventually provide for “declared symbols” and optional associated alpha-numeric “type” information. Ultimately there should emerge a public formal object, the “mathematical expression” (**mathexpr**) that is something like the “regular expression” (**regex**) that is familiar to users of “ed”, ELISP, “Perl”, etc. One will want a separate, probably simpler syntax for the specification of the *type* of a **mathexpr**.

My philosophy, and I think the only realistic philosophy, is that such types for mathexprs should involve *relative*, rather than *absolute*, semantics.

One of the most basic types is categorical “morphism”, which is a generalization of a calculus student’s notion of “function”; for much that is of interest to many, the notion of function will suffice, provided that each function symbol is understood to imply “domain” and “target” with “target” not always the same as “image” or “range”. Regardless, users may conceptualize “morphisms” as “functions”.

---

<sup>37</sup>URI: <http://www.w3.org/>

<sup>38</sup>URI: <http://www.albany.edu/~hammond/gellmu/tarball.html>

<sup>39</sup>URI: <http://www.albany.edu/~hammond/gellmu/authordtd/>

<sup>40</sup>URI: <http://www.albany.edu/~hammond/gellmu/notation>

## 4 Brief Introductions

To summarize there are two concepts in this project.

### 4.1 Basic GELLMU

This may be useful for some authors familiar with  $\text{\LaTeX}$  who wish to write directly for an SGML or XML document type. It provides rudimentary  $\text{\LaTeX}$ -like commands with single argument syntax. SGML attribute strings may be entered using a single  $\text{\LaTeX}$ -like option.

It also offers a  $\text{\LaTeX}$ -like meta-command `\newcommand`, which provides for macros with arguments. See Using the GELLMU Syntactic Translator to Write HTML<sup>41</sup>. For example, the previous anchor would be marked up in HTML as

```
<a href="/~hammond/gellmu/ghtml.html"
  >Using ... <kbd>HTML</kbd></a> ,
```

and this is marked up somewhat more succinctly in GELLMU source as

```
\a[href="/~hammond/gellmu/ghtml.html"
  ]{Using ... \kbd{HTML}} .
```

With the `\newcommand` definition for `\href`

```
\newcommand{\href}[2]{\a[href="#1"]{#2}}
```

the even more succinct markup

```
\href{/~hammond/gellmu/ghtml.html}{Using ... \kbd{HTML}}
```

suffices.

### 4.2 Advanced GELLMU

This goes beyond basic  $\text{\LaTeX}$ -like command / argument syntax to provide  $\text{\LaTeX}$ -like multiple argument / option syntax and also what might be called  $\text{\LaTeX}$ -like grammar including `\begin{...}\end` and, if desired, blank lines to initiate paragraphs.

When desired, advanced GELLMU has knowledge of a few command names, but the author must know the SGML or XML document type.

### 4.3 *Regular* GELLMU: The Didactic Production System

The didactic production system is a beginning at emulating  $\text{\LaTeX}$  with an XML document type. In fact,  $\text{\LaTeX}$  can be modeled more precisely with SGML than with XML.

The didactic production system consists of

---

<sup>41</sup>URI: /~hammond/gellmu/ghtml.html

The GELLMU Syntactic Translator.

An SGML document type.

An XML document type.

The following translators, coded in Perl under David Megginson's *sgmlspl* framework:

1. from the SGML document type (GELLMU Syntactic Translator output) to the XML document type.
2. from the XML document type to HTML. For this there are two routes:
  - (a) The up to date XML form of HTML extended by MATHML.
  - (b) Classical HTML with pseudo-TeX ASCII for math, still useful with terminal window browsers.
3. from the XML document type to  $\LaTeX$

There is validation of each stage of output. Indeed, validation of the GELLMU Syntactic Translator's SGML output is very useful for catching author errors. To assist with this there is line number alignment between the source and GELLMU Syntactic Translator output. If necessary<sup>42</sup> one may intervene at any stage of the processing since the output of each stage is quite readable by humans.

The two document types are parallel; the XML version is intended to be the nearest XML approximation of the SGML version. The SGML version should be regarded as "in-house", while the XML version is suitable for export. (Usable, though not identical, source may be recovered from the XML document type.)

The document types have been designed for translation to many output formats. I have the intention ultimately to write or find others to write translators from the XML document type to other formats.

Finally the *article* document type may have value as a layout vehicle that is useful as an intermediate formatting stage for structure-rich document types such as *DocBook* and *TEI*, and I would encourage those who might be so inclined to think about writing translators from such document types to GELLMU *article*.

The *Brief Introduction to Regular GELLMU*<sup>43</sup> (in XHTML+MATHML) is itself a GELLMU document. It deals mainly with the language. Various other versions are also available here:

- the input source<sup>44</sup>.
- the syntactic translation to SGML<sup>45</sup>.
- the translation to XML<sup>46</sup>.
- translation to classical HTML<sup>47</sup>.

---

<sup>42</sup>But only in very exceptional situations

<sup>43</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/iglm.xhtml>

<sup>44</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/iglm.glm>

<sup>45</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/iglm.sgml>

<sup>46</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/iglm.xml>

<sup>47</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/iglm.html>

- the L<sup>A</sup>T<sub>E</sub>X version<sup>48</sup>.
- DVI made from the L<sup>A</sup>T<sub>E</sub>X<sup>49</sup>.
- portrait PDF made from the L<sup>A</sup>T<sub>E</sub>X<sup>50</sup>

## 4.4 Other Production Systems

An author may use *advanced* GELLMU as a front end to many other SGML or XML production systems with appropriate setting of variables for the GELLMU Syntactic Translator.

## 5 Materials

All that one should need to get started is in the current tarball<sup>51</sup>. One should look at the user guide, the manual, both listed as “Quick Anchors” above, and the examples<sup>52</sup>. Note that the driver scripts found in the `bin` directory of the unpacked tarball may need editing for location names.

Note also that the tarball may be installed in a “Windows” system equipped with *Cygwin*, enhanced by a sufficient array of *Cygwin*-provided packages, using the Linux driver scripts.

In principle, it should also work on MacOS X, but I have no reports, and I have no idea what might be required to port it to earlier versions of MacOS.

Although the project was begun in June 1998, its alpha release was in July, 2001. It will not be considered to have reached beta stage until I have more knowledge about use experience of others.

Some older odds and ends may be found on the GELLMU veterans page<sup>53</sup>, and the very old page for early preview of materials<sup>54</sup> is still available.

## 6 Relevant Public Discussion and Comment

My annotations allude, though not entirely precisely, to the article *The Cathedral and the Bazaar*<sup>55</sup> by Eric Raymond.

**Electronic Math Journals** <sup>56</sup> Use “subscribe EMJ” in the BODY of a message.

There is an archive at the host site<sup>57</sup>.

This is a bazaar. Sometimes technical, sometimes economic or legal, sometimes other.

---

<sup>48</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/iglm.ltx>

<sup>49</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/iglm.dvi>

<sup>50</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/iglm.pdf>

<sup>51</sup>URI: <http://www.albany.edu/~hammond/gellmu/tarball.html>

<sup>52</sup>URI: <http://www.albany.edu/~hammond/gellmu/examples/>

<sup>53</sup>URI: <http://www.albany.edu/~hammond/gellmu/veterans.html>

<sup>54</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/gellmum.html>

<sup>55</sup>URI: [http://www.redhat.com/knowledgebase/otherwhitepapers/whitepaper\\_cathedral.html](http://www.redhat.com/knowledgebase/otherwhitepapers/whitepaper_cathedral.html)

<sup>56</sup>URI: <mailto:listserv@listserv.albany.edu>

<sup>57</sup>URI: <http://listserv.albany.edu:8080/archives/emj.html>

**L<sup>A</sup>T<sub>E</sub>X Development** <sup>58</sup> Use “subscribe L<sup>A</sup>T<sub>E</sub>X-L” in the BODY of a message.

Archive location, if any, unknown.

Neither a bazaar, nor a cathedral. Very sophisticated and technical. User questions are not wanted.

**MathML and the HTML Math WG** <sup>59</sup> Make your message SUBJECT “subscribe”. Message BODY should be blank.

An archive will be found behind the W3C Math<sup>60</sup> web site.

This is a small bazaar in the nave of a cathedral. The cathedral “chapter” has its own private list. Many chapter members, not all, who speak in the nave seem to feel constrained to representation of the chapter.

**UseNet news on SGML (if you get “news”)** <sup>61</sup> A bazaar with many, many important people. Sophisticated and technical, questions about SGML (but not HTML, nor http, nor “the web”, ...) are usually answered well.

**UseNet news on XML (if you get “news”)** <sup>62</sup> A recent spin-off from the SGML discussion. Eventually it should operate at much higher volume than the SGML discussion.

## 7 Pointers to a Few Related Things

**Slides from 2001** <sup>63</sup> A presentation given at The University of Delaware during the 2001 annual meeting of TUG<sup>64</sup>.

**Blahtex** <sup>65</sup> *Blahtex* converts L<sup>A</sup>T<sub>E</sub>X-like math markup to MATHML for use with *MediaWiki*<sup>66</sup>, which is wiki implementation software for *Wikipedia*<sup>67</sup>.

**TBook** <sup>68</sup> The *TBook* System for XML Authoring by Torsten Bronger.

**MathML, Version 2.0, Second Edition** <sup>69</sup> A W3C recommendation (October 21, 2003). In the fall of 2009 <em>MathML, Version 3</em> and the <em>MathML for CSS Profile</em> became candidate recommendations at W3C. See <http://www.w3.org/Math/>.

**OMDoc: Open Mathematical Documents** <sup>70</sup> A content based XML markup format by Michael Kohlhase of Universität Saarlandes and Carnegie Mellon University for mathematics on the Internet that extends OpenMath<sup>71</sup> to the document level. Released November 1, 2000.

---

<sup>58</sup>URI: <mailto:listserv@relay.urz.uni-heidelberg.de>

<sup>59</sup>URI: <mailto:www-math-request@w3.org>

<sup>60</sup>URI: <http://www.w3.org/Math/>

<sup>61</sup>URI: <news:comp.text.sgml>

<sup>62</sup>URI: <news:comp.text.xml>

<sup>63</sup>URI: <http://math.albany.edu/math/pers/hammond/Presen/tug2001/>

<sup>64</sup>URI: <http://www.tug.org/tug2001/>

<sup>65</sup>URI: <http://www.mediawiki.org/wiki/Extension:Blahtex>

<sup>66</sup>URI: <http://en.wikipedia.org/wiki/MediaWiki>

<sup>67</sup>URI: <http://en.wikipedia.org/>

<sup>68</sup>URI: <http://tbookdtd.sourceforge.net/>

<sup>69</sup>URI: <http://www.w3.org/TR/MathML/>

<sup>70</sup>URI: <http://www.mathweb.org/omdoc/>

<sup>71</sup>URI: <http://www.openmath.org/>

**Daniele Giacomini's Sgmltexi** <sup>72</sup> SGMLTEXI provided the first SGML model of *Texinfo*, the language of the GNU Documentation System. Since its first release in the year 2000 *Texinfo* itself has incorporated an XML model.

**itex2MML** <sup>73</sup> *itex2MML* is the T<sub>E</sub>X-math to MATHML converter that at one time had been featured at Paul Gartside's *MathZilla* site. It is now used with Jacques Distler's very active mathematical physics blog *Musings*<sup>74</sup>.

**David Carlisle's xmltex** <sup>75</sup> *xmltex* uses *TeX*, the program, to parse (without validation) an XML document and then set it in T<sub>E</sub>X, according to user rules written in code for T<sub>E</sub>X, that govern what is done for each of the tags in the corresponding XML document type definition. The same items are also available at CTAN in "`macros/xmltex`".

### Sebastian Rahtz's "PassiveTeX"

Uses T<sub>E</sub>X as a formatting back end for documents prepared under an XML language according to an XSL stylesheet. It is available through CTAN<sup>76</sup>.

### *TeX4ht*, *htlatex*, ... : Work of Eitan Gurari at Ohio State University. <sup>77</sup>

An important way to make HTML and XML versions of T<sub>E</sub>X and L<sup>A</sup>T<sub>E</sub>X documents. This is based on a C program *TeX4ht*<sup>78</sup>, and on a related macro package for T<sub>E</sub>X. The macro package causes "T<sub>E</sub>X, the program", to add specials to its DVI output. The program *TeX4ht* operates on a DVI that has been so prepared and makes HTML or XML. (The DVI format has the abstract structure of a classical assembly language. There are several "special" instructions that serve as wildcards. These "specials" are of use only to processors that know about them on a case-by-case basis. They *should*, in theory, be ignored by processors that do not recognize them.)

In recent editions of TUG's TeXLive a convenient default interface for using *TeX4ht* to make classical HTML from L<sup>A</sup>T<sub>E</sub>X is the command *htlatex*, while the interface for making HTML with MATHML is the command *mzlatex*. Aside from the standard *TeX4ht* docs, those interested in this approach might want to consult <http://facweb.knowlton.ohio-state.edu/pviton/support/swphtpa4.html>.

Sadly, Eitan Gurari, the author of *TeX4ht*, died in 2009.

### HyperLaTeX

An early (mid 90's) package (unfortunately not on CTAN) for the production of L<sup>A</sup>T<sub>E</sub>X and HTML from a single specialized L<sup>A</sup>T<sub>E</sub>X source document. *Hyperlatex*<sup>79</sup> is somewhat similar to GELLMU in its use of an Emacs Lisp program for generating HTML though it seems not to provide a method for conscious writing under other SGML or XML document types.

**The L<sup>A</sup>T<sub>E</sub>X3 Project** <sup>80</sup> Information is available in the document section of the current L<sup>A</sup>T<sub>E</sub>X2E base distribution under the filename "`ltx3info.tex`" (with DVI and PostScript version nearby). On the web one may consult the PDF version<sup>81</sup>. Plans for SGML are

---

<sup>72</sup>URI: <http://www.archive.org/details/sgmltexi>

<sup>73</sup>URI: <http://golem.ph.utexas.edu/~distler/blog/itex2MML.html>

<sup>74</sup>URI: <http://golem.ph.utexas.edu/~distler/blog/>

<sup>75</sup>URI: <http://www.dcarlisle.demon.co.uk/xmltex/>

<sup>76</sup>URI: <http://ctan.tug.org/>

<sup>77</sup>URI: <http://www.cse.ohio-state.edu/~gurari/>

<sup>78</sup>URI: <http://www.cse.ohio-state.edu/~gurari/TeX4ht/mn.html>

<sup>79</sup>URI: <http://www.cs.uu.nl/~otfried/Hyperlatex/>

<sup>80</sup>URI: <http://www.latex-project.org/>

<sup>81</sup>URI: <http://ctan.tug.org/tex-archive/macros/latex/doc/ltx3info.pdf>

mentioned in this document. There is a mailing list on the topic of L<sup>A</sup>T<sub>E</sub>X3 development at the address `LATEX-L@relay.urz.uni-heidelberg.de`.

**TeXML**<sup>82</sup> TeXML is an XML vocabulary for describing T<sub>E</sub>X syntax that has evolved from Doug Lovell's TeXML<sup>83</sup>, which became available in the late 1990s.

It's useful for converting XML documents to T<sub>E</sub>X, L<sup>A</sup>T<sub>E</sub>X, or Context, but it's not useful for translating T<sub>E</sub>X documents to XML document types. One writes an XSL style sheet to translate an XML document type into TeXML. Another program then translates TeXML to T<sub>E</sub>X.

**Bruce Miller's LaTeXXML**<sup>84</sup> *LaTeXML* is a Perl program for converting L<sup>A</sup>T<sub>E</sub>X documents to the *LaTeXML* XML document type. A separate program is provided for translating the *LaTeXML* XML document type to XHTML+MathML. While *LaTeXML* tries to mimic the actions of L<sup>A</sup>T<sub>E</sub>X, the program, in typesetting L<sup>A</sup>T<sub>E</sub>X documents as DVI or PDF, it does not employ a T<sub>E</sub>X engine.

*LaTeXML* is the converter that was used in the project called *arXMLiv*<sup>85</sup> for converting L<sup>A</sup>T<sub>E</sub>X documents at The arXiv<sup>86</sup> to XML.

### Smart Documents.

There are various forms of "smartness". SGML will provide easily for all of them. See Richard Fateman's material on *More Versatile Scientific Documents ...*<sup>87</sup>.

### Linux Documentation.

The "How To" documents for Linux systems are based on an SGML language with ancestry in the L<sup>A</sup>T<sub>E</sub>X-like language of the QWERTZ document formatting system from the University of Exeter (U.K.) in the early 1990's. The SGMLtools-Lite Project<sup>88</sup> is a recent effort to bring Linux documents under the DocBook<sup>89</sup> language.

### Luc Maranget's Hevea

*Hevea*<sup>90</sup> is a L<sup>A</sup>T<sub>E</sub>X to HTML translator, said to produce correct HTML 4.0.

### Latex2html and Latex2html-with-MathML.

The familiar Perl package *latex2html*<sup>91</sup> gained popularity in math departments during the early days of the web not only by translating the L<sup>A</sup>T<sub>E</sub>X commands that could be marked up into HTML but also by automatically putting out mathematics in graphic objects housed in "<img>" tags; the graphics were created with subprocesses that used T<sub>E</sub>X, *dvips*, and some *netpbm* utilities. Many features have been added.

A 1998 variant<sup>92</sup> at The Geometry Center offers the option of replacing the graphic objects with MATHML<sup>93</sup> objects.

---

<sup>82</sup>URI: <http://getfo.org/texml/>

<sup>83</sup>URI: <http://www.alphaworks.ibm.com/formula/texml>

<sup>84</sup>URI: <http://dlmf.nist.gov/LaTeXML/>

<sup>85</sup>URI: <http://kwarc.info/projects/arXMLiv/>

<sup>86</sup>URI: <http://www.arxiv.org>

<sup>87</sup>URI: <http://http.cs.berkeley.edu/~fateman/MVSD.html>

<sup>88</sup>URI: <http://sgmltools-lite.sourceforge.net/>

<sup>89</sup>URI: <http://www.oasis-open.org/docbook/>

<sup>90</sup>URI: <http://cristal.inria.fr/~maranget/hevea/>

<sup>91</sup>URI: <http://www.ctan.org/tex-archive/support/latex2html/>

<sup>92</sup>URI: <http://www.geom.umn.edu/~ross/webtex/webtex/>

<sup>93</sup>URI: <http://www.w3.org/Math/>

**The philosophy of Kernighan and Pike.** <sup>94</sup> If you have never looked at their classic 1984 book, here are a few quoted paragraphs. Don't let their use of a trademark get in your way.

## 8 About this Document

This document, which is primarily a web page, is itself a *regular* GELLMU document (see §4.3). Versions of this document other than the HTML version include the original GELLMU source<sup>95</sup>, its translation to XML<sup>96</sup> (from which the HTML version is derived), the derived translation to XHTML+MathML<sup>97</sup>, and the derived L<sup>A</sup>T<sub>E</sub>X source<sup>98</sup> from which a device independent (DVI)<sup>99</sup> file and a file in Adobe's portable document format (PDF)<sup>100</sup> were compiled. The PDF copy, which was generated using the free program *pdflatex*, is tuned for printing on 8.5 x 11 inch paper by those who have yet to equip themselves (freely) for printing DVI.

---

<sup>94</sup>URI: <http://math.albany.edu/math/pers/hammond/unixphil.html>

<sup>95</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/igl.glm>

<sup>96</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/igl.xml>

<sup>97</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/igl.xhtml>

<sup>98</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/igl.ltx>

<sup>99</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/igl.dvi>

<sup>100</sup>URI: <http://www.albany.edu/~hammond/gellmu/igl/igl.pdf>