

Introductory User's Guide to *Regular* GELLMU

Version 0.8.4 release, July 2007

William F. Hammond

Dept. of Mathematics & Statistics
University at Albany
Albany, New York 12222 (USA)

last corrections: April 23, 2011

Table of Contents

1	Introduction.....	1
2	First Steps	2
2.1	Acquiring and Installing	2
2.2	Writing an Article	3
3	Basic Markup Within Sentences	4
4	Sectioning.....	5
5	Lists	6
6	<i>Tabular</i> Environment Emulation.....	7
7	Labels, Cross References, and Anchors.....	7
8	Mathematics	8
9	Beyond this Introductory Guide	14

1 Introduction

Regular GELLMU is the part of the GELLMU Project under which source language markup most closely resembles actual \LaTeX source markup. Much of the markup vocabulary is the same as that of actual \LaTeX . The differences — and, indeed, the whole reason for the GELLMU Project — arises from the author's conclusion in 1998 that it would not be possible in a robust way to write rules for actual \LaTeX source markup so that the collection of documents in

the *article* document class prepared under such rules would admit fully reliable automatic translation to other formats¹.

The version 0.8 release of the GELLMU Project was the first release providing an output stream for the XML form of HTML extended by the World Wide Web Consortium's Mathematical Markup Language (MATHML)². The driver script `mmkg` provides translations of regular GELLMU articles to (1) actual L^AT_EX (and from there to DVI and PDF), (2) classic HTML, and (3) XHTML+MATHML, while the driver script `mkg` omits XHTML+MATHML. Though the PDF output may be read on a screen, it is intended primarily for printing by end consumers.

While one could provide a landscape version of the PDF output with live anchors for screen consumption, the XHTML+MATHML version would still be superior for screen reading since, as with ordinary HTML, the text is re-formatted to conform naturally both to magnification and window-resizing actions by the end user. The quality of the rendering of mathematics in the XHTML+MATHML version is very nearly as good as that in the PDF version.

Indeed the author, as an author, finds that the best venue for proof reading what he writes is XHTML+MATHML when displayed in a suitable web browser³ with larger than normal font size and narrower than normal line length.

Moreover, the XHTML+MATHML version is fully compliant with the Accessibility Guidelines⁴ of the World Wide Web Consortium⁵.

This document is an introductory user guide. The relation of regular GELLMU to the GELLMU Project overall is explained in the Manual⁶, and advanced user information on regular GELLMU is found there as well. This document covers fundamentals not mentioned in the manual. It is available in PDF⁷, XHTML+MATHML⁸, and classical HTML⁹ versions. Online copies of this Introductory Guide and of the Manual may be found at CTAN¹⁰ and the GELLMU web site¹¹.

To use this document as an example for learning refer to its source markup¹² and to use the project manual as an example refer to the manual's source markup.

2 First Steps

2.1 Acquiring and Installing

GELLMU is implemented as cross-platform free software licensed under the GNU General

¹This should not be construed to mean, however, that standard L^AT_EX processing could not incorporate this type of source markup at some time in the future.

²URI: <http://www.w3.org/Math/>

³The author prefers the freely available cross-platform browser *Firefox* (<http://www.mozilla.com/>) from the Mozilla Project.

⁴URI: <http://www.w3.org/WAI/>

⁵URI: <http://www.w3.org/>

⁶URI: `glman.html`

⁷URI: `userdoc.pdf`

⁸URI: `userdoc.xhtml`

⁹URI: `userdoc.html`

¹⁰URI: <http://www.tex.ac.uk/tex-archive/help/Catalogue/entries/gellmu.html>

¹¹URI: <http://www.albany.edu/~hammond/gellmu/>

¹²URI: `userdoc.glm`

Public License¹³. Its package is available from the Comprehensive TeX Archive Network¹⁴ (CTAN) and from the GELLMU web site.

The package requires a number of other cross-platform free software packages including GNU Emacs, *Perl*, and some standard items of SGML/XML software, chiefly *Open SP*. Locations may be found in the Manual, key "**requiredsoftware**".

Linux: The required packages are generally part of a full GNU/*Linux* distribution. The package should be installed in `/usr/local/gellmu` and symlinks to driver scripts should be made from a suitable place in one's command path.

MacOS-X and other Unix variants: The only difference from GNU/*Linux* is that some of the supporting packages may need to be acquired and installed.

MS Windows: The best strategy is to acquire and install a **full** *Cygwin*¹⁵ distribution. Then proceed as with *Linux*.

Additional information may be found in the Manual¹⁶.

2.2 Writing an Article

The minimal form of a regular GELLMU article is this:

```
\documenttype{article}
\title{Some title}

\begin{document}

First Paragraph.

Second Paragraph.

. . .

\end{document}
```

An article must begin with `\documenttype{article}`. The initial part of an article up to the required `\begin{document}` is the article's *preamble*, and the rest of the article is the article's *body*.

Almost everything in the *body* of an article must be part of a paragraph. Most typically a paragraph is begun with a blank line. Paragraphs themselves may be parts of sectional units.

Beware special characters. For example `'%` converts the rest of its line to comment status. To use `'%` in text write `"\%`". The special characters are with small exceptions the same as those in \LaTeX . Note, however, that in dealing with multiple downstream formats any character that is not alpha-numeric becomes potentially special. Therefore, there is a named form for every non-alphanumeric ASCII character. For example, `"\tild;"` is the name of ASCII `'~'`, which

¹³URI: <http://www.gnu.org/copyleft/>

¹⁴URI: <http://www.ctan.org>

¹⁵URI: <http://www.cygwin.com/>

¹⁶URI: [glman.html](#)

is sometimes handy when forming a URL.¹⁷

The names of special characters may be found by browsing the SGML document type definition in the file `gellmu.dtd` at the top of the distribution tree.¹⁸

3 Basic Markup Within Sentences

Although there is no markup for a sentence as a unit, there is markup for the end of a sentence or a “full stop”. The normal way to indicate the end of a sentence is to use an ASCII period ‘.’ followed by either a newline or two or more blank spaces. An end of sentence is marked up as the SGML empty element *eos*. This is automatically generated under the condition indicated above. For most purposes it will not matter if an ASCII period is followed by only a single blank space, but in that case no *eos* will be generated.

The foregoing paragraph provides several instances of markup within sentences other than *eos*. There are (1) a quoted phrase, (2) two abbreviations, (3) a quoted character, and (4) two instances of emphasized text. The source for that paragraph is:

```
Although there is no markup for a sentence as a unit, there is markup
for the end of a sentence or a ‘full stop’. The normal way to
indicate the end of a sentence is to use an \abbr{ASCII} period
\quochar{.} followed by either a newline or two or more blank spaces.
An end of sentence is marked up as the \sgml empty element \emph{eos}.
This is automatically generated under the condition indicated above.
For most purposes it will not matter if an \ascii period is followed
by only a single blank space, but in that case no \emph{eos} will be
generated.
```

The “full stop” is a marked up quoted phrase. The named form of a quoted phrase is *quophrase*. The markup `\quophrase{full stop}` is the equivalent expanded form. An abbreviation is indicated with the command *abbr*. There is, of course, no requirement that an abbreviation be marked up; it is merely good practice.

The name of markup for emphasized text is *emph*. A common alternate form of emphasis is called *bold*. One way the two differ is that *emph* has order two effect, i.e., *emph* within *emph* removes emphasis while *bold* is not permitted within itself. Note that ***bold*** may be emphasized.

In the cited marked up text above the instances `\sgml` and `\ascii` are user-defined *newcommands* occurring in the preamble (or at least before the invocations):

```
\newcommand{\ascii}{\abbr{ASCII}}
\newcommand{\sgml}{\abbr{SGML}}
```

The command *newcommand* is a meta-command in GELLMU source. All *newcommand* instances are fully expanded by the syntactic translator before it writes any output.

The markup `\quochar{x}` produces a quoted character: in this case ‘x’. There is no compelling

¹⁷In GELLMU source, as in L^AT_EX source “~” is markup for “non-breaking space”. Note further that ASCII tilde “\tild;” should not be confused with either the accent command “\tilde” or the math command “\sim” (rendered as ~) that is typically used for mathematical operators.

¹⁸The SGML document type definition is more elaborately commented than its XML mirror found in the file `xml/xgellmu.dtd`.

reason why *quophrase* could not have been used instead of *quochar*, but it is customary in technical documentation to use a different presentation style for quoted characters than for quoted phrases. Beyond that for non-traditional forms of processing it can be very useful to provide for semantic distinctions such as the distinction between a quoted phrase in normal discourse and a quoted string in technical discourse. More particularly, it is significant that in some languages for computer programming there is an important distinction between a character and a string of unit length. Some of the markup commands useful with technical documentation are

command	example in source	example rendered
<i>quochar</i>	use <code>\quochar{\%}</code> to comment	use ‘%’ to comment
<i>quostr</i>	<code>\quostr{unsigned long num;}</code>	unsigned long num;
<i>qquostr</i>	type <code>\qquostr{done}</code> to exit	type "done" to exit
<i>path</i>	<code>\path{C:\bsl;TEX\bsl;TEXMF}</code>	C:\TEX\TEXMF
<i>urlanch</i>	see <code>\urlanch{userdoc.glm}</code>	see userdoc.glm

With regular L^AT_EX standard practice is to use *texttt* for the examples just cited involving *quochar*, *quostr*, *path*, and, apart from web-related considerations, *urlanch*. The name *url* has been used variously in regular L^AT_EX; therefore, *url* has not been introduced as a name in regular GELLMU so that an author may safely use it, if desired, as a macro name.

4 Sectioning

Unless one wants fancy non-default behavior, one may provide sections, subsections, and sub-subsections as in L^AT_EX. Thus, for example,

```

\documenttype{article}
\title{Some title}

\begin{document}

\section{Title of first section}
First Paragraph.

Second Paragraph.

. . .

\section{Title of second section}
First Paragraph.

Second Paragraph.

. . .

\end{document}

```

With this style of markup one does not indicate the end of a section although there is a way, described in the Manual, key "**secusage**", to gain finer control over sectional units.

5 Lists

As with L^AT_EX GELLMU *article* provides lists with the names *description*, *enumerate*, and *itemize*. For example, the markup

```
\begin{enumerate}
\item bird
\item cat
\item dog
\end{enumerate}
```

yields:

1. bird
2. cat
3. dog

while the markup

```
\begin{description}
\item[Parrot] a type of bird
\item[Persian] a type of house cat
\item[Pointer] a type of dog
\end{description}
```

yields:

Parrot a type of bird
Persian a type of house cat
Pointer a type of dog

A list is normally part of a paragraph, but a list may occur outside of a paragraph in a sectional unit or even at the top level of an article's body.

An *item* in a list may contain paragraphs but, as the foregoing examples show, may contain loose text.

While L^AT_EX provides an excellent construction facility for customized lists through its *list* environment, which may be used in conjunction with its *newenvironment* command, there is no facility for the provision of on-the-fly markup in SGML. Therefore, the *article* document type provides several other kinds of list.

The lists called *menu* and *Menu* consist of *items*. These lists have no item labels. They differ only in intention for spacing in presentation formats.

It often makes sense for *menu* to be used inside an *item* of another list. Moreover, a single item *menu* will render in either the L^AT_EX or HTML formattings as “hanging indentation”.

A *defnlist* is intended to be more like the list called *dl* in HTML than like *description*. It must consist of one or more *term*, *desc* pairs with *term* mandatory and *desc* optional.

There are also other lists that will not be mentioned here except to point out that when the name *verbatim* is enabled as a metacommand, rather than as a command, by calling the syntactic translator with a non-default function¹⁹ such as `verblis`t or `latex-faq`, the verbatim material

¹⁹The most convenient way to call the syntactic translator with a non-default function is to make the first line

is organized as a *verblis*t.

For more information about lists browse the document type definitions `gellmu.dtd` and `xml/xgellmu.dtd` being mindful that default processing in the didactic production system sometimes does a small amount of list manipulation between the SGML and XML versions of an *article*.

6 *Tabular* Environment Emulation

Most users having a L^AT_EX background will find that *tabular* with column arguments limited to ‘l’, ‘r’, ‘c’, and ‘p’ is reasonably functional. Improvement made during 2006 were made possible by the increased availability to users of web browsers supporting version 2 of the cascading style sheets (CSS) specification.

For example, the following markup underlies a small *lcc* table in the section on mathematics.

```
\display{\begin{tabular}{lcc}
\bold{Name} & \bold{Opener}          & & \bold{Closer}  \\
math        & \quostr{\bsl; }                   & & \quostr{\bsl;)} \\
tmath      & \quostr{\$}                       & & \quostr{\$}    \\
displaymath & \quostr{\bsl\lsb;}                & & \quostr{\bsl\rsb;} \\
\end{tabular}}
```

(In this markup *bsl* is a name for the backslash character and *lsb*, *rsb* the names of the square brackets.)

There is presently no concept of “float”.

More information is available in the Manual, key “`tabular`”.

7 Labels, Cross References, and Anchors

The basic idea is very simple. Use

```
\label{mylabelkey}
```

to **mark** a location that is associated with the named label key, and use

```
\ref{mylabelkey}
```

or one of several other methods to reference that location. The command *ref* may be regarded as the default method. In an online version, however, it has **not** been designed to create a selectable anchor but rather simply to indicate section number.

For example, in this document there is a label with the key “`math`” at the beginning of the section (8) on mathematics. The previous sentence contains a reference with the markup `\ref{math}` to that section. This markup results in the number of that section appearing in the first sentence. For clarity the markup for the entire first sentence of this paragraph is:

of the source file point to the name of the special function. For example, the function “`gellmu-verblis`” will be called if the first line of the source file is “`%!verblis`”.

For example, in this document there is a label with the key `\qqostr{math}` at the beginning of the section (`\ref{math}`) on mathematics.

If, on the other hand, one wants an online version to contain a selectable anchor to section 8 to the section on math, then in the preamble of the document one might place the newcommand definition

```
\newcommand{\iref}[2]{\anch[iref="#1"]{#2}}
```

and then, as in the foregoing portion of this sentence use the markup

```
\iref{math}{selectable anchor to section \ref{math}}.
```

Note that this markup uses the label key "math" twice: once as the first argument to the newcommand called *iref* and once as the sole argument of the command *ref* that itself appears in the second argument of *iref*.

On the question of whether to use *ref* or one of several other commands to make a reference, there are several considerations:

1. How should the reference be indicated in an online format that will be viewed in an ebook or a web browser?
2. How should the reference be indicated in a typeset format where there is neither a cursor nor the possibility of automatic scrolling?

More details on this subject may be found in the Manual, key "labelref".

8 Mathematics

GELLMU *article* provides 5 containers for mathematical content: *math*, *tmath*, *displaymath*, *equation*, *eqnarray*. All may be referenced by name. However, L^AT_EX-like markup is supported:

Name	Opener	Closer
<i>math</i>	<code>\(</code>	<code>\)</code>
<i>tmath</i>	<code>\$</code>	<code>\$</code>
<i>displaymath</i>	<code>\[</code>	<code>\]</code>

The containers *math* and *tmath* are essentially equivalent non-displayed containers. The current L^AT_EX formatter provides a tiny amount of extra horizontal space around *math* but not around *tmath*.

Just as with ordinary L^AT_EX there can be difficulty locating a missing ‘\$’ if one uses these for *tmath*. Unbalanced ‘\$’ characters usually lead the syntactic translator to complain (correctly, though not with as much user-friendliness as one might wish) about not being able to balance a brace. In this case if *tmath* is seen in the displayed surrounding text, a missing ‘\$’ is likely.

MathML rendering support (with MathML's presentation markup) in HTML versions, i.e., the modern XML form of HTML extended with MATHML, is new in the summer of 2004.

A program to generate an output stream providing content-level MathML markup, i.e., semantic markup transparent for computer algebra systems, has not yet been written.

Those who use mathematics may find the following illustrations instructive.

The markup

```
Just as with \tex; or \latex; it is easy to provide a mathematical
formula, such as $y = x^3/(x^2 + 1)$\aoc either inline or displayed.
```

produces:

Just as with T_EX or L^AT_EX it is easy to provide a mathematical formula, such as $y = x^3/(x^2 + 1)$, either inline or displayed.

This markup

```
Here the same formula is displayed.
\[ y = \frac{x^3}{x^2 + 1} \eos\]
```

yields:

Here the same formula is displayed.

$$y = \frac{x^3}{x^2 + 1} .$$

The *eos* corresponds to an empty SGML element for “end of sentence” (or “full stop”). Normally one simply uses ‘.’, but inside a displayed math zone there is semantic-level risk of confusion with various possible uses of ‘.’ as a mathematical symbol.

Markup:

```
The square root of a compound fraction:
\[ \sqrt{\frac{\frac{a}{b}}{\frac{c}{d}}} \eos \]
```

Rendered:

The square root of a compound fraction:

$$\sqrt{\frac{\frac{a}{b}}{\frac{c}{d}}} .$$

In the current XHTML+MATHML formatting by default radical signs in source markup are converted by default to fractional exponents because of an intellectual property issue with fonts available for use by the open source *Mozilla* browser and its spinoffs including *Firefox*, *Galeon*, and *Nautilus*.²⁰

Using the markup for numbered equations

The quadratic equation

²⁰Similar intellectual property issues pertain more widely to fonts. In particular, for this reason by default in the XHTML+MATHML output character coloring is used to cover possible font failures for the blackboard bold, calligraphic, and fraktur math fonts; the colors corresponding colors are, respectively, red, blue, and green.

determination of what is a single symbol. By long convention in \TeX the string “abcd” in math is understood as the product (in whatever sense of “product”) of four symbols. If a string with more than one character — for example, “Hom” — is to be a single symbol, it may be marked up with *mbox* or, in GELLMU source, be assigned a symbol name having custom print representation using *mathsym* in the preamble. In GELLMU source *mathbf* should only contain a single symbol.

The markup

```
\displaystyle{\nabla\times\vec{\mbox{\bold{B}}}}$
  \ with ‘mbox’ and ‘bold’ versus
  \ $\nabla\times\vec{\mathbf{B}}$ \ with ‘mathbf’,.}
```

produces

$\nabla \times \vec{\mathbf{B}}$ with “mbox” and “bold” versus $\nabla \times \vec{B}$ with “mathbf”.

Diagrams are done using *array*. They must, therefore, be rectangular. The markup

```
\newcommand{\ra}{\rightarrow}
\newcommand{\da}{\downarrow}
\[ \begin{array}{ccccccc}
0 & \ra & A' & \ra & A & \ra & A'' & \ra & 0 \\
~ & & & \da & & \da & & \da & \\
0 & \ra & B' & \ra & B & \ra & B'' & \ra & 0 \\
\end{array} \]
```

yields

$$\begin{array}{ccccccc}
0 & \rightarrow & A' & \rightarrow & A & \rightarrow & A'' & \rightarrow & 0 \\
& & \downarrow & & \downarrow & & \downarrow & & \\
0 & \rightarrow & B' & \rightarrow & B & \rightarrow & B'' & \rightarrow & 0
\end{array}$$

Comments about *array*, *eqnarray*, *tabular*

1. In this markup note the appearance of the character ‘~’ at the beginning of the second row of the array. The modular design of the didactic production system requires that the syntactic translator may only deal with markup syntax. The ‘~’, which is markup in GELLMU source as in \LaTeX for “non-breaking space”, in the first cell of the second row serves to keep the first cell from being empty. There is a technical error in the SGML output of the syntactic translator when the first cell in an *array* or *tabular* row is empty unless there is an explicit opentag for the first cell. The technical error will trigger a validation error though it should not interfere with the generation of a correct parsed output.²¹
2. Also in the foregoing markup note that the \LaTeX -like use of the character ‘&’ as the introducer of a new cell must in GELLMU source be followed by whitespace. Otherwise there is risk of confusion with the alternative markup use of ‘&’ in GELLMU source (though not in \LaTeX source) to initiate an SGML entity reference.

²¹Beyond that in certain situations some web browsers appear not to deal well with empty table cells, and placing an HTML non-breaking space in such a cell is often a good thing for that reason. Of course, the special character ‘~’ for non-breaking space in GELLMU source gives rise in HTML output to the SGML non-breaking space character denoted with the entity reference “ ”.

There is multiscript support. The command *mscript* takes 5 arguments. The first is its base. The script order is counter-clockwise beginning with upper-left. For example,

$${}^{14}_2C_{5+} \quad {}^a_b\text{Hom}_c^d(X, Y)$$

is produced by

```
\[
\mscript{\mathbf{C}}{14}{2}{5+}{2}\quad
\mscript{\mbox{Hom}}{a}{b}{c}{d}(X, Y)
\]
```

AMSMath-like overset and underset: `\overset{A}{B}` then `\underset{A}{B}`

$$\overset{A}{B} \quad \underset{A}{B}$$

The command *overset* is the AMSMath replacement in L^AT_EX for L^AT_EX's native *stackrel*.

A map may be labeled using `\emph{overset}`:

```
\[ X \overset{f}{\longrightarrow} Y \]
```

or using `\emph{underset}`:

```
\[ X \underset{f}{\longrightarrow} Y \]
```

A map may be labeled using *overset*:

$$X \overset{f}{\rightarrow} Y$$

or using *underset*:

$$X \underset{f}{\rightarrow} Y$$

Binary or pseudo-binary operators:

vee: \vee setminus: \setminus ll: \ll gg: \gg approx: \approx asymp: \asymp
 rightarrow: \rightarrow leftarrow: \leftarrow implies: \Rightarrow revimplies: \Leftarrow iff: \Leftrightarrow

Other operators:

uparrow: \uparrow downarrow: \downarrow triangle: \triangle abuts: \Rightarrow

The command *vect* is for vector generation. Its usage is:

`\vect[open-close-separator spec]{coordinate} ...`

Thus, the markup `\vect{x}{y}{z}` produces (x, y, z) . On the other hand the markup `\vect[] {x}{y}{z}` produces xyz . This might be useful, for example, in providing reasonably semantic (but visibly traditional) markup for a tensor field. For example

$$T_{j_1 j_2 \dots j_q}^{i_1 i_2 \dots i_p}$$

When *vect*'s option is empty, vector opener, closer, and coordinate separators are all nil. *vect* might be used to provide the stylized, though semantically dubious, inner product markup `\vect[<>]{v}{w}` yielding

$$\langle v; w \rangle ,$$

or with the markup `\vect[\lsb;\rsb;]{a}{b}{c}` to customize notation for the homogeneous coordinate triple representing a point of \mathbf{P}^2 :

$$[a:b:c] .$$

Leray's spectral sequence:

If

$$X \longrightarrow B$$

with fibre F , then

$$E_2^{pq} = H^p(B, H^q(F)) \Rightarrow H^*(X) .$$

is produced by

If

`\[X \ \longrightarrow \ B \]`

with fibre F , then

`\[{\mscript{E}}{}{}{}{}{\vect}[]{}{} = H^p\mathaF(B, H^q\mathaF(F))`

`\abuts H^*\mathaF(X) \ \eos \]`

In this markup the command `\aF`, is not strictly necessary. It provides a deliberate way to indicate that, in the second instance above, the symbol X is argument-like in relation to the symbol H^* . In MathML rendering `aF` gives rise to presentation-level MathML's invisible operator *ApplyFunction*.

9 Beyond this Introductory Guide

More information about GELLMU is available in the Manual²². Some examples of topics to be found there include:

- Inclusion of graphic objects: key "graphics".
- Equations and equation arrays: key "eqn".
- Emulation of \LaTeX 's *newtheorem* environments: key "assert".
- Emulation of \LaTeX 's counters: key "counters".
- The flowchart for regular GELLMU processing: key "flow".

The author tries to keep fresh information on the GELLMU web site.

²²URI: glman.html