1 Mathematical Articles

- Contain ordinary text and mathematical content.
- Need realization on paper and on computer networks.

Comment

What are our requirements for mathematical articles?
An example will indicate the scope of the task.

2 A Segment of an Article

The gamma function may be defined in a suitable right-half plane by an integral, which essentially amounts to Fourier transform relative to the multiplicative group of positive real numbers of the reciprocal exponential function, and then extended to a meromorphic function in the complex plane with entire reciprocal. The following formula represents a variant of its Weierstrass product expansion:

\[
\int_0^\infty t^x e^{-t} \frac{dt}{t} = \frac{1}{x} \prod_{k=1}^{\infty} \left(1 + \frac{1}{k}\right)^x \prod_{k=1}^{\infty} \left(1 + \frac{x}{k}\right).
\]

The product manifests simple poles at zero and each of the negative integers.

Comment

What are our goals for article preparation?

3 Goals

From a Single Source:

1. A typeset version of high quality for paper preprints.
2. An online version for a web preprint.
3. A typeset version of high quality meeting the requirements of a journal chosen after the
time of writing.

4 Goals for Online Versions
1. catalogable.
2. universally accessible.
3. re-scalable and re-sizable.
4. searchable for mathematical content.
5. “clippable” for mathematical content.

Comment
How may an article containing a segment like that of the previous example be prepared?

5 Mathematical Mainstream 1980–2000

Donald Knuth’s $\TeX$

$\LaTeX$

Plain $\TeX$

Other $\TeX$ Variants:

AMSTeX, ConTeXt, Omega, Texinfo

Comment
$\LaTeX$ is the most widely used $\TeX$ variant.

Texinfo is the $\TeX$ variant most suited to multiple presentations from a single source, but it is
limited in regard to mathematics

Context by Hans Hagen is new and worth a look. It was the basis of an astounding graphics
presentation at TUG 2001.

Context and latex both provide ways to write formatters for XML documents.
6 How May One Provide Multiple Formattings?

Early Ideas:

1. Intuitive Authoring Systems: the WYSIWYG Idea
2. Write HTML and Translate From There.
3. Write \TeX{} and Translate From There.

Comment

WYSIWYG stands for “what you see is what you get”. But what one sees on a screen may not be the same as what one sees on paper. This issue is compounded for source documents that admit multiple formattings.

A standard answer to those who inquire about WYSIWYG is that WYSIWYG really means WYSIAYG: “what you see is all you get”.

7 WYSIWYG: The Good News

- Standard with Common “Word Processors”
- Available for \TeX{}.
- Available for SGML Systems.
- Easy for Easy Tasks.
- OK for Very Simple Documents.

8 WYSIWYG: The Bad News

- Hard for Hard Tasks.
- Slow for Those Who Write Frequently.
- Insufficiently Rich for Mathematics.
- Insufficiently Abstract.
- Inapplicable to Multiple Formattings.
- Some say WYSIAYG: “what you see is all you get”.

3
9 WYSIWYG vs. Structured Markup

- Format-specific hanging indentation commands.
- Use of abstract list structures.

10 Translating HTML

- Reliable
- But:
  1. No Math in HTML
  2. HTML Generally Less Rich Than \[ \text{\LaTeX} \]
  3. Special characters are translation headaches.
     Examples: # $ % & ~ ^ { } \{ \} < >

11 Translating \[ \text{\LaTeX} \]

- Almost Impossible
- Good Structure a Help
- May Require Human Intervention
- Need to Proof Read Twice
- \texttt{htlatex} in the \TeX\mbox{}Live 6 distribution is remarkably good.

12 XML

\textbf{eXtensible Markup Language}

- Data Under a Template for Translation.
- Enforced Separation of Content and Presentation.
- Universal Exchange.
- Originated by
  World Wide Web Consortium (W3C)
  Sun Microsystems
13 XML

- Many Templates
- Synonym for XML Template: Document Type
- Two worlds
  1. Classical Documents:
     Examples: HTML, Docbook, TEI, ...
  2. Electronic Data Interchange (EDI)
     Example: Graham William’s \texttt{\LaTeX} Catalogue found on CTAN
     \url{http://CTAN/help/Catalogue/catalogue.html}

14 The GELLMU Project

- Superseded my earlier ideas:
  - Strictly controlled dialect of \texttt{\LaTeX}.
  - Adapting Texinfo (already suitable for multiple formattings).
- Relation to the Goals:
  - No present full realization of online version goals.
  - Proof that full realization of all goals and more is possible.

15 GELLMU

\textbf{Generalized Extensible \LaTeX-Like MarkUp}

- A markup interface for writing (SGML or) XML.
- \LaTeX\-like notation more succinct than that of XML.
- Extensible using GELLMU’s \texttt{\newcommand} with arguments. (SGML has no analogue of macros with arguments.)

16 Difference Between \LaTeX Source and GELLMU Source

- Article prepared under a template for processing by \texttt{latex}, the Program.
- Article prepared under a template for processing by many programs.

It is a small step from \LaTeX source to GELLMU source.
17 The Idea of \LaTeX-Like MarkUp

- Text + Commands
- A *command* is a function that operates on text.
- A command may take a non-negative number of arguments.

18 Examples of \LaTeX-Like Commands

Example of a command taking no argument:
\[ \texttt{\LaTeX} \]

Example of a command taking one argument:
\[ \texttt{this is \texttt{\emph{emphasized}} text} \]
\[ \texttt{this is \texttt{emphasized} text} \]

Example of a command (for math) taking two arguments:
\[ \frac{az + b}{cz + d} \]

19 GELLMU Modes

1. Basic
2. Advanced
   (a) Regular
   (b) Other (less fully developed)

20 Regular GELLMU: System Stages

1. \LaTeX-Like Source.
2. Syntactic Translation to SGML.
3. Translation of SGML to Enriched XML.
4. Various Formattings of Enriched XML.
21 Customizing

- Each stage presents opportunities for customizing.
- Each output format is the result of successive transforms.
- Additional intermediate transforms can be provided.
- These slides were prepared using a special formatting of regular GELLMU’s article.
- The source markup\(^1\) for these slides is as readable as ordinary $\LaTeX$ source.

22 The Syntactic Translator

source markup $\rightarrow$ XML or SGML

\begin{align*}
\texttt{\textbackslash foo\{ ... \}} & \rightarrow \texttt{<\textit{foo> ... </foo>}} \\
\texttt{\textbackslash foo;} & \rightarrow \texttt{<\textit{foo/>}} \\
\texttt{\textbackslash foo} & \rightarrow \texttt{<\textit{foo>}} \\
\texttt{\textbackslash foo:} & \rightarrow \texttt{</\textit{foo>}} \\
\texttt{\textbackslash foo[a=\"x\" ...]} & \rightarrow \texttt{<\textit{foo a=\"x\" ...>}}
\end{align*}

23 Syntactic Differences from $\LaTeX$

- Command names (element names) may contain numbers.
- Example: $\frac{2}{3}$ is a command name.
- Arguments must be delimited with braces or brackets.
- No white space between command name and first argument delimiter.
- No white space between delimiters of successive arguments.
- Bracketed arguments may not be optional.

24 Use of GELLMU in Basic Mode for XHTML

Write:

\begin{verbatim}
the WWW \a[href="http://www.w3.org/"
]{Consortium} site
\end{verbatim}

for generating the XML:

\footnote{URI: correct.glm}
the WWW <a href="http://www.w3.org/">
>Consortium</a> site

to produce:

the WWW Consortium\textsuperscript{2} site

25 \texttt{\newcommand with XHTML}

Definitions

\begin{verbatim}
\newcommand{\emph}[1]{\emph{#1}}
\newcommand{\w3ref}[2]{\%a[\href="http://www.w3.org/#1"]{#2}}
\end{verbatim}

Invocations

Using GELLMU’s \texttt{\emph{newcommand}}

one can reduce the markup required

for an anchor to \texttt{\w3ref{W3C}’s}

\texttt{\w3ref{Math/}{MathML} site}.

**Rendering:** Using GELLMU’s newcommand one can reduce the markup required for an anchor to W3C’s MathML site.

26 Why is GELLMU’s article “Didactic”? 

- Intended as a first XML document type for $\LaTeX$ authors
- Sits in the middle between
  1. What $\LaTeX$ authors are accustomed to.
  2. What high end XML people think is needed.
- Room to adjust and expand.

27 The Gamma Function: Its Weierstrass Product

\[
\int_0^\infty t^xe^{-t}\frac{dt}{t} = \frac{1}{x} \prod_{k=1}^\infty \frac{(1 + \frac{1}{x})^x}{(1 + \frac{k}{x})}
\]

\textsuperscript{2}URI: http://www.w3.org/
28 Markup for the Gamma Identity

Regular GELLMU source for the identity:

\[
\int_0^\infty t^x e^{-t} \frac{d t}{t} = \frac{1}{x} \prod_{k=1}^{\infty} \frac{(1 + \frac{1}{k})^x}{1 + \frac{x}{k}} \prod
\]

29 Gamma: Derived XML Markup

<displaymath>
<int>
<msub>0</msub><msup><infty/></msup>
$t^x e^{-t} \frac{d t}{t}$
</int>
<equals/>
<frac><numr>d t</numr><denm>t</denm></frac>
</int>
<equals/>
<frac><numr>1</numr><denm>x</denm></frac>
<prod>
<msub>k=1</msub><msup><infty/></msup>
<frac><numr><bal>1 + \frac{1}{k}\right)^x\numr><denm><bal>1 + \frac{x}{k}\right>\denm></frac>
</prod>
</displaymath>
30 Gamma: in MathML

(not by automatic translation)
31 Viewing MathML

Viewing support for MathML in web pages is not yet widely available. The above item can be rendered by:

- W3C's Amaya: wprod.html or wprod.xml.
- Mozilla's MathML development track: wprod.xml (only).
- With special plugin for MSIE: wprod.html (only).

32 Generating MathML from article

- Ad hoc wprod.html was made from GELLMU source: wprod.glm.
- The short article form (slide 28) of GELLMU source above could be given automatic translation to MathML.
- An automatic translation should go through content MathML and from there to presentation MathML.
- An automatic translation would not be under the umbrella of general XML processing.

33 Reliable Generation of MathML

Reliable translation will require:

A substantial non-XML, but XML-aware, parsing of all math zones in a GELLMU source document.

Occasional math parsing hints from authors in their markup.

Desirable, sometimes required:

1. Source markup labeling of math symbols.
2. Source markup typing of math symbols.

34 MathML Generation Issues

- Will authors cooperate?
- Will standard web user agents cooperate?

35 Two Final Notes

- For more information: http://www.albany.edu/~hammond/gellmu/
- GELLMU source for these slides is on the web: http://math.albany.edu:8000/~hammond/Presen/Correct/correct.glm