

# `cmath.sty`: An Infrastructure for building Inline Content Math in $\text{\S T E X}^*$

Michael Kohlhase & Deyan Ginev  
Jacobs University, Bremen  
<http://kwarc.info/kohlhase>

October 22, 2015

## Abstract

The `cmath` package is a central part of the  $\text{\S T E X}$  collection, a version of  $\text{T E X}/\text{\La T E X}$  that allows to markup  $\text{T E X}/\text{\La T E X}$  documents semantically without leaving the document format, essentially turning  $\text{T E X}/\text{\La T E X}$  into a document format for mathematical knowledge management (MKM).

This package supplies an infrastructure that allows to build content math expressions (strict content MathML or OpenMath objects) in the text. This is needed whenever the head symbols of expressions are variables and can thus not be treated via the `\symdef` mechanism in  $\text{\S T E X}$ .

---

\*Version v0.1 (last revised 2012/09/23)

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>The User Interface</b>	<b>3</b>
2.1	Variable Names . . . . .	3
2.2	Applications . . . . .	4
2.3	Binders . . . . .	4
<b>3</b>	<b>Limitations</b>	<b>4</b>
<b>4</b>	<b>The Implementation</b>	<b>5</b>
4.1	Package Options . . . . .	5
4.2	Variable Names . . . . .	6
4.3	Applications . . . . .	7
4.4	Binders . . . . .	9
4.5	Finale . . . . .	9

# 1 Introduction

TeX allows to build content math expressions via the `\symdef` mechanism [KGA15] if their heads are constants. For instance, if we have defined `\symdef{lt}[2]{#1<#2}` in the module `relation1`, then an invocation of `\lt3a` will be transformed to

```
<OMA>
  <OMS cd="relation1" name="lt"/>
  <OMI>3</OMI>
  <OMV name="a"/>
</OMA>
```

If the head of the expression (i.e. the function symbol in this case) is a variable, then we cannot resort to a `\symdef`, since that would define the functional equivalent of a logical constant. Sometimes, L<sup>A</sup>T<sub>E</sub>XML can figure out that when we write  $f(a,b)$  that  $f$  is a function (especially, if we declare them to be via the `functions=` key in the dominating statement environment [Koh15]). But sometimes, we want to be explicit, especially for  $n$ -ary functions and in the presence of elided elements in argument sequences. A related problem is markup for complex variable names, such as  $x_{\text{left}}$  or  $ST^*$ .

The `cmath` package supplies the L<sup>A</sup>T<sub>E</sub>X bindings that allow us to achieve this.

## 2 The User Interface

### 2.1 Variable Names

In mathematics we often use complex variable names like  $x'$ ,  $g_n$ ,  $f^1$ ,  $\tilde{\phi}_i^j$  or even  $foo$ ; for presentation-oriented L<sup>A</sup>T<sub>E</sub>X, this is not a problem, but if we want to generate content markup, we must show explicitly that those are complex identifiers (otherwise the variable name  $foo$  might be mistaken for the product  $f \cdot o \cdot o$ ). In careful mathematical typesetting,  $\sin$  is distinguished from  $\sin$ , but we cannot rely on this effect for variable names.

`\vname` `\vname` identifies a token sequence as a name, and allows the user to provide an ASCII (XML-compatible) identifier for it. The optional argument is the identifier, and the second one the LaTeX representation. The identifier can also be used with `\vnref` for referencing. So, if we have used `\vname[xi]{x_i}`, then we can later use `\vnref{xi}` as a short name for `\vname{x_i}`. Note that in output formats that are capable of generating structure sharing, `\vnref{xi}` would be represented as a cross-reference.<sup>1</sup>

Since indexed variable names make a significant special case of complex identifiers, we provides the macros `\livar` that allows to mark up variables with lower indices. If `\livar` is given an optional first argument, this is taken as a name. Thus `\livar[foo]{x}1` is “short” for `\vname[foo]{x_1}`. The macros `\livar`,

<sup>1</sup>EDNOTE: DG: Do we know whether using the same name in two `vname` invocations, would refer to two instances of the same variable? Presumably so, since the names are the same? We should make this explicit in the text. A different variable would e.g. have a name “xi2”, but the same body

<code>\nappa{f}{a_1,a_2,a_3}</code>	$f(a_1, a_2, a_3)$
<code>\nappe{f}{a_1}{a_n}</code>	$f(a_1, \dots, a_n)$
<code>\symdef{eph}[1]{e_{\#1}}^{\varphi(\#1)}</code> <code>\nappf{g}\eph14</code>	$g(e_1^{\varphi(1)}, \dots, e_4^{\varphi(4)})$
<code>\nappli{f}a1n</code>	$f(a_1, \dots, a_n)$
<code>\nappui{f}a1n</code>	$f(a^1, \dots, a^n)$

Figure 1: Application Macros

`\ulivar` serve the analogous purpose for variables with upper indices, and `\ulivar` for upper and lower indices. Finally, `\primvar` and `\pprimvar` do the same for variables with primes and double primes (triple primes are bad style).

## 2.2 Applications

To construct a content math application of the form  $f(a_1, \dots, a_n)$  with concrete arguments  $a_i$  (i.e. without elisions), then we can use the `\nappa` macro. If we have elisions in the arguments, then we have to interpret the arguments as a sequence of argument constructors applied to the respective positional indexes. We can mark up this situation with the `\nappf` macro: `\nappf{<fun>}{<const>}{<first>}{<last>}` where `<const>` is a macro for the constructor is presented as `<fun>(<const><first>, \dots, <const><last>)`; see Figure 1 for a concrete example, and Figure 1.<sup>2</sup>

For a simple elision in the arguments, we can use `\nappe{<fun>}{<first>}{<last>}` will be formatted as `<fun>(<first>, \dots, <last>)`. Note that this is quite un-semantic (we have to guess the sequence), so the use of `\nappe` is discouraged.

A solution to this situation is if we can think of the arguments as a finite sequence  $a =: (a_i)_{1 \leq i \leq h}$ , then we can use `\nappli{<fun>}{<seq>}{<start>}{<end>}`, where `<seq>` is the sequence, and the remaining arguments are the start and end index. The works like `\nappli`, but uses upper indices in the presentation.

## 2.3 Binders

## 3 Limitations

In this section we document known limitations. If you want to help alleviate them, please feel free to contact the package author. Some of them are currently discussed in the  $\text{\LaTeX}$  GitHub repository [sTeX].

1. none reported yet

<sup>2</sup>EDNOTE: MK@MK: we need a meta-cd `cmath` with the respective notation definition here. It is very frustrating that we cannot even really write down the axiomatization of flexary constants in `OpenMath`.

```
\symdef{eph}[1]{e_{#1}^{\phi(#1)}}
\nappf{g}\eph14
```

currently generates

```
<OMA>
  <OMS cd="cmath" name="apply-from-to"/>
  <OMV name="g"/>
  <OMBIND>
    <OMS cd="fns1" name="lambda"/>
    <OMBVAR><OMV name="x"/></OMBVAR>
    <OMA><OMS cd="???" name="eph"/><OMV name="x"/></OMA>
  </OMBIND>
  <OMI>1</OMI>
  <OMI>4</OMI>
</OMA>
```

**Example 1:** Application Macros

## 4 The Implementation

The `cmath` package generates to files: the  $\text{\LaTeX}$  package (all the code between `<*package>` and `</package>`) and the  $\text{LaTeXML}$  bindings (between `<*ltxml>` and `</ltxml>`). We keep the corresponding code fragments together, since the documentation applies to both of them and to prevent them from getting out of sync.

For  $\text{LaTeXML}$ , we initialize the package inclusions.

```
1 <*ltxml>
2 # -*- CPERL -*-
3 package LaTeXML::Package::Pool;
4 use strict;
5 use LaTeXML::Package;
6 </ltxml>
```

### 4.1 Package Options

We declare some switches which will modify the behavior according to the package options. Generally, an option `xxx` will just set the appropriate switches to true (otherwise they stay false).<sup>3</sup>

```
7 <*package>
8 \ProcessOptions
9 </package>
10 <*ltxml>
11 ProcessOptions();
12 </ltxml>
```

---

<sup>3</sup>EdNOTE: we have no options at the moment

The next measure is to ensure that some  $\text{\LaTeX}$  packages are loaded. For  $\text{\LaTeX}$ ML, we also initialize the package inclusions, there we do not need `ntheorem`, since the XML does not do the presentation.

```

13 <*package>
14 \RequirePackage{presentation}
15 </package>
16 <*ltxml>
17 RequirePackage('presentation');
18 </ltxml>

```

## 4.2 Variable Names

`\vname` a name macro; the first optional argument is an identifier  $\langle id \rangle$ , this is standard for  $\text{\LaTeX}$ , but for  $\text{\LaTeX}$ ML, we want to generate attributes `xml:id="cvar.<id>"` and `name="<id>"`. However, if no id was given in we default them to `xml:id="cvar.<count>"` and `name="name.cvar.<count>"`.

```

19 <*package>
20 \newcommand\vname[2] [] {#2%
21 \def\@opt{#1}%
22 \ifx\@opt\empty\else\expandafter\gdef\csname MOD@name@#1\endcsname{#2}\fi}
23 </package>
24 <*ltxml>
25 # return: unique ID for variable
26 sub cvar_id {
27   my ($id) = @_ ;
28   $id = ToString($id);
29   if (!$id) {
30     $id=LookupValue('cvar_id') || 0;
31     AssignValue('cvar_id', $id + 1, 'global'); }
32   "cvar.$id"; }#$
33 DefConstructor(' \vname[] {}',
34   "<ltx:XMWrap role='ID' xml:id='&cvar_id(#1)'>#2</ltx:XMWrap>",
35   requireMath=>1);
36 </ltxml>

```

`\vnref`

```

37 <*package>
38 \def\vnref#1{\csname MOD@name@#1\endcsname}
39 </package>
40 <*ltxml>
41 # \vnref{<reference>}
42 DefMacro(' \vnref{}', '\@XMRef{cvar.#1}');
43
44 </ltxml>

```

4

EdN:4

---

<sup>4</sup>EdNOTE: the following macros are just ideas, they need to be implemented and documented

`\uivar` constructors for variables.

```

45 <*package>
46 \newcommand\primvar[2] [] {\vname[#1]{#2^\prime}}
47 \newcommand\pprimvar[2] [] {\vname[#1]{#2^\prime\prime}}
48 \newcommand\uivar[3] [] {\vname[#1]{#{#2}_{#3}}}
49 \newcommand\livar[3] [] {\vname[#1]{#{#2}_{#3}}}
50 \newcommand\ulivar[4] [] {\vname[#1]{#{#2}_{#3}_{#4}}}
51 </package>
52 <*ltxml>
53 # variants for declaring variables
54 DefMacro('uivar[]{}{}', '\vname[#1]{#{#2}_{#3}}');
55 DefMacro('livar[]{}{}', '\vname[#1]{#{#2}_{#3}}');
56 DefMacro('ulivar[]{}{}', '\vname[#1]{#{#2}_{#3}_{#4}}');
57 DefMacro('primvar[]{}{}', '\vname[#1]{#2^\prime}');
58 DefMacro('pprimvar[]{}{}', '\vname[#1]{#2^\prime\prime}');
59 </ltxml>

```

### 4.3 Applications

EdN:5

```

5 \napp*
60 <*package>
61 \newcommand\nappa[3] [] {\prefix[#1]{#2}{#3}}
62 \newcommand\nappe[4] [] {\nappa[#1]{#2}{#3,\ldots,#4}}
63 \newcommand\nappf[5] [] {\nappe[#1]{#2}{#3{#4}}{#3{#5}}}
64 \newcommand\nappli[5] [] {\nappe[#1]{#2}{#3_{#4}}{#3_{#5}}}
65 \newcommand\nappui[5] [] {\nappe[#1]{#2}{#3^{#4}}{#3^{#5}}}
66 </package>
67 <*ltxml>
68 # \nappa{<function>}{<(var)(, \1)*>}
69 # @#1(#2)
70 DefConstructor('nappa[]{}{}',
71 "<ltx:XMAApp>"
72 ".<ltx:XMTok name=' #2' />"
73 ".<ltx:XMAArg>#3</ltx:XMAArg>"
74 ".</ltx:XMAApp>");
75
76 # \@napp@seq{<function>}{start <const>}{end <const>}
77 # @#1(@sequence(#2,sequencefromto,#3))
78 DefConstructor('@napp@seq[]{}{}{}',
79 "<ltx:XMAApp>"
80 ".<ltx:XMTok name=' #2' />"
81 ".<ltx:XMAArg>"
82 ".<ltx:XMAApp>"
83 ".<ltx:XMTok meaning='sequence' />"
84 ".<ltx:XMAArg>#3</ltx:XMAArg>"
85 ".<ltx:XMAArg><ltx:XMTok meaning='sequencefromto' /></ltx:XMAArg>"
86 ".<ltx:XMAArg>#4</ltx:XMAArg>"

```

---

<sup>5</sup>EDNOTE: document keyval args above and implement them in LaTeXML

```

87 . "</ltx:XMAApp>"
88 . "</ltx:XMAArg>"
89 . "</ltx:XMAApp>");
90
91 DefMacro('\' \nappe [] {} {} {}', '\@napp@seq[#1]{#2}{#3}{#4}');
92 DefMacro('\' \nappf [] {} {} {} {}', '\@napp@seq[#1]{#2}{#3}{#4}{#3}{#5}');
93 DefMacro('\' \nappli [] {} {} {} {}', '\@napp@seq[#1]{#2}{#3}_{#4}{#3}_{#5}');
94 DefMacro('\' \nappui [] {} {} {} {}', '\@napp@seq[#1]{#2}{#3}^{#4}{#3}^{#5}');
95 </ltxml>

```

EdN:6

\anapp\*

```

6
96 <*package>
97 \newcommand\anappa[3] [] {\assoc[#1]{#2}{#3}}
98 \newcommand\anappe[4] [] {\anappa[#1]{#2}{#3,\ldots,#4}}
99 \newcommand\anappf[5] [] {\anappe[#1]{#2}{#3}{#4}{#3}{#5}}
100 \newcommand\anappli[5] [] {\anappe[#1]{#2}{#3}_{#4}{#3}_{#5}}
101 \newcommand\anappui[5] [] {\anappe[#1]{#2}{#3}^{#4}{#3}^{#5}}
102 </package>
103 <*ltxml>
104 # \nappa{<function>}{<(const)(,\1)*>}
105 # @#1(#2)
106 DefConstructor('\' \anappa [] {} {}',
107   "<ltx:XMAApp>"
108   . "<ltx:XTok name=' #2' />"
109   . "<ltx:XMAArg>#3</ltx:XMAArg>"
110   . "</ltx:XMAApp>");
111 # \@napp@seq{<function>}{start <const>}{end <const>}
112 # @#1(@sequence(#2,sequencefromto,#3))
113 DefConstructor('\' \anapp@seq {} {} {}',
114   "<ltx:XMAApp>"
115   . "<ltx:XTok name=' #1' />"
116   . "<ltx:XMAArg>"
117   . "<ltx:XMAApp>"
118   . "<ltx:XTok meaning='sequence' />"
119   . "<ltx:XMAArg>#2</ltx:XMAArg>"
120   . "<ltx:XMAArg><ltx:XTok meaning='sequencefromto' /></ltx:XMAArg>"
121   . "<ltx:XMAArg>#3</ltx:XMAArg>"
122   . "</ltx:XMAApp>"
123   . "</ltx:XMAArg>"
124   . "</ltx:XMAApp>");
125 DefMacro('\' \anappe [] {} {} {}', '\@anapp@seq[#1]{#2}{#3}{#4}');
126 DefMacro('\' \anappf [] {} {} {} {}', '\@anapp@seq[#1]{#2}{#3}{#4}{#3}{#5}');
127 DefMacro('\' \anappli [] {} {} {} {}', '\@anapp@seq[#1]{#2}{#3}_{#4}{#3}_{#5}');
128 DefMacro('\' \anappui [] {} {} {} {}', '\@anapp@seq[#1]{#2}{#3}^{#4}{#3}^{#5}');
129 </ltxml>

```

---

<sup>6</sup>EDNOTE: document anapp\* and implement in LaTeXML (i.e. get the presentation information into the OM/MathML).



## 4.4 Binders

## 4.5 Finale

Finally, we need to terminate the file with a success mark for perl.

130 `<ltxml>1;`

# Index

Numbers written in *italic* refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; numbers in **roman** refer to the code lines where the entry is used.

LaTeXML, 3, 5, 6 XML, 3

## References

- [KGA15] Michael Kohlhase, Deyan Ginev, and Rares Ambrus. *modules.sty: Semantic Macros and Module Scoping in sTeX*. Tech. rep. Comprehensive T<sub>E</sub>X Archive Network (CTAN), 2015. URL: <http://www.ctan.org/get/macros/latex/contrib/stex/modules/modules.pdf>.
- [Koh15] Michael Kohlhase. *omtext: Semantic Markup for Mathematical Text Fragments in L<sup>A</sup>T<sub>E</sub>X*. Tech. rep. Comprehensive T<sub>E</sub>X Archive Network (CTAN), 2015. URL: <http://www.ctan.org/tex-archive/macros/latex/contrib/stex/omtext/omtext.pdf>.
- [sTeX] *KWARC/sTeX*. URL: <https://svn.kwarc.info/repos/stex> (visited on 05/15/2015).