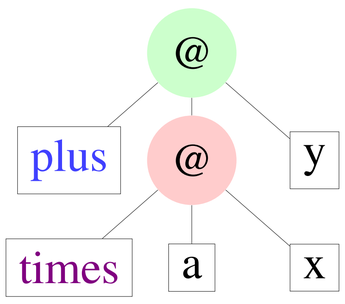
Presenting Mathematical Content with Flexible Elisions

Michael Kohlhase <m.kohlhase>, **Christoph Lange** <ch.lange>, **Florian Rabe** <f.rabe>
Knowledge Adaptation and Reasoning for Content, Computer Science
Jacobs University, Bremen, Germany <...@jacobs-university.de>

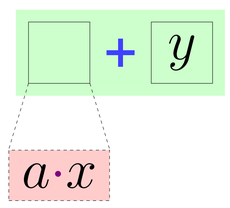
Mathematicians frequently elide brackets or symbols in formulae to concentrate on essential facts and to avoid distracting experienced mathematicians with notation that can easily be deduced from context. We propose a content dictionary format for content markup languages like Content MathML or OpenMath that supports flexary mixfix operators and is capable of handling flexible elisions.

Presentation as Composition and Elision

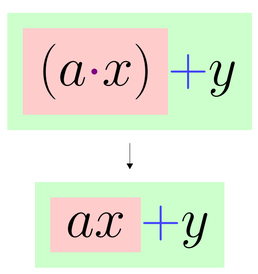
0. Content representations (variables, symbols, applications, binders)



1. 2D *composition* of presentations (formula tree → layout tree)



2. *elision* of parts that can be deduced from the context



Characteristics of Mathematical Symbols

fixity: pre-/post-/in-/mixfix
brackets: left and right, mostly round
associativity: fully/left/right

↯ These are not real *brackets*:

$$]a;b],\ \{x \in \mathbb{N} | x > 5\},\ \binom{n}{k}, \ldots$$

Notation Definitions: symbol- vs. template-based

```
<presentation      (OMDoc 1.2, '06)
 for="power" theory="arith1"
 role="application" class="1">
  <style format="TeX">
    <recurse select="*[2]"/>
    <text>^{</text>
    <recurse select="*[3]/>
    <text>}</text>
  </style>
<presentation>
```

```
<Notation>                (Naylor & Watt '01)
  <version style="1">
    <tex>\arg{a1}{n}^{\arg{a2}{m}}</tex>
  </version>
  <semantic-template>
    <OMA><OMS cd="power" name="arith1"/>
      <OMV name="n" id="a1/>
      <OMV name="n" id="a2/>
    </OMA>
  </semantic-template>
</Notation>
```

- easier than XSLT
- roughly equivalent

A Mixfix Presentation Model for Flexary Terms

$$f = w_0 \boxed{p_1|_{\pi_1}} w_1 \boxed{p_2|_{\pi_2}} \ldots \boxed{p_n|_{\pi_n}} w_n : p \qquad \text{(Paulson '05, Isabelle)}$$

w_i: strings in the output language, p: output precedence, boxes: argument specifications (rendered recursively), p_i: input precedences. Argument is bracketed when its operator binds weaker than the enclosing operator.

$$\boxed{p|_1} \backslash \text{vdash\_} \left\{ \boxed{p|_2} \right\} \boxed{p|_3} : \boxed{p|_4} : p$$

Typing judgment ($\Gamma \vdash_\Sigma t : \alpha$) in LaTeX

OpenMath and Content MathML require flexible arities!

$$\boxed{\times : p - 1|_1^{n-1}} \to \boxed{p|_n} : p$$

The space of n-ary functions

Flexible Elisions

Elision desired for redundant brackets, default or inferable argument values.

$$ax + y = (a \cdot x) + y \qquad \log x = \log_{10} x \qquad [\![t]\!] = [\![t]\!]^{\phi}_{\mathcal{M}}, \qquad \ldots$$

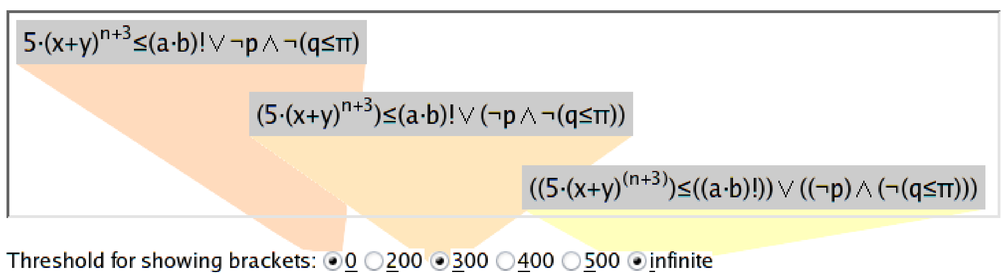
Flexible elisions serve experts and beginners.

- *visibility levels* (for brackets: *precedence difference*) per *elision group*, user can choose visibility threshold per group
- static output format (e. g. dead tree): choice at generation
- dynamic output format: elision annotations; interactive choice

Flexible Bracket Elision Demo

OMDoc OpenMathDocuments

- Powered by OMDoc
- Tested with Firefox 2.0 and Opera 9.0
- Authors: Michael Kohlhase, Christoph Lange, Florian Rabe



Threshold for showing brackets: ◉0 ○200 ◉300 ○400 ○500 ◉infinite

Operator	Mixfix declaration
x^y	[199][∞]:200
!	[300]!:300
·	[400]·[400]:400
+	[500]+[500]:500

Operator	Mixfix declaration
¬	¬[600]:600
≤	[700]≤[700]:700
∧	[1000]∧[1000]:1000
∨	[1200]∨[1200]:1200

The Flexary Mixfix Model in a CD Format

Made OMDoc more powerful, deprecated embedded XSLT. Syntactic sugar for fixities and format-independent settings.
How is notation definition for symbol determined? – Look up (1) exact match, (2) "default" presentation. Choice among ≥ 1 presentations is non-trivial (→ Kohlhase/Müller/Müller '07).

Example: the typing jugdment $\Gamma \vdash_\Sigma t : T$ in LaTeX

```
<symbol name="typing-judgment" role="application"/>
<presentation for="#typing-judgment"
 role="application" format="latex">
  <map begin="1"/>
  <text>\vdash_{</text><map begin="2"/><text>}</text>
  <map begin="3"/>
  <text>:</text>
  <map begin="4"/>
</presentation>
```

Input:

```
<OMA>
  <OMS name="typing-judgment" cd="typ"/>
  <OMS name="emptyset" cd="sets"/>
  <OMV name="Σ"/>
  <OMS name="true" cd="boolean"/>
  <OMS name="Boolean" cd="boolean"/>
</OMA>
```

Output:

```
\emptyset\vdash_{Σ}
  \mathit{true}:\mathit{Boolean}
```

Rendered: $\emptyset \vdash_\Sigma \mathit{true} : \mathit{Boolean}$

Flexary notation and multiple output formats:

```
<symbol name="times" role="application"/>
<presentation for="#times" role="constant" format="ascii">
  <text>*</text>
</presentation>
<presentation for="#times" role="constant" format="latex">
  <text>\ast</text>
</presentation>
<presentation for="#times" role="application"
 precedence="400" format="ascii latex">
  <text egroup="lbrack">(</text>
  <map begin="1" end="-1">
    <separator><map begin="0"/></separator>
    <recurse precedence="400"/>
  </map>
  <text egroup="rbrack">)</text>
</presentation>
```

Input:

```
<apply><power/>
  <apply><times/>
    <ci>x</ci><ci>y</ci>
  </apply>
  <cn>2</cn>
</apply>
```

Output:

LaTeX: $(a * b)^2$

ASCII: `(a*b)^2`

Bracket elision in Presentation MathML:

```
<presentation for="#plus" precedence="500">...</presentation>
<presentation for="#times" precedence="400">
  <element name="mo" egroup="lbrack"><text>(</text></element>
  ...
</presentation>
```

Input:

```
<OMA>
  <OMS name="plus" cd="arith1"/>
  <OMA>
    <OMS name="times" cd="arith1"/>
    <OMV name="a"/>
    <OMV name="x"/>
  </OMA>
  <OMV name="y"/>
</OMA>
```

Output:

```
<mrow>
  <mrow>
    <mo style="display:none"
     omdoc:elevel="100">(</mo>
    <mi>a</mi><mo>·</mo><mi>x</mi>
    <mo style="display:none"
     omdoc:elevel="100">)</mo>
  </mrow>
  <mo>+</mo><mi>y</mi>
</mrow>
```

Our approach is knowledge-based, extensible, adaptive, mathematical, and efficient. We propose a manageable declarative notation for content dictionaries that allows for generating interactive human-oriented presentations, adaptable to the level of experience of the reader. We propose this notation for the MathML 3 recommendation.