

MathML Intent Attribute

Murray Sargent

30-Nov-23

Mathematical notation is incredibly successful partly because it expresses relationships in concise ways. Nevertheless, it can be ambiguous and its representation in Presentation MathML can mirror the ambiguities. Accordingly, MathML 4.0 adds the intent attribute which can reduce or resolve the ambiguities particularly for math speech. Some intent attribute values can be derived directly from the input notations of LaTeX and [UnicodeMath](#). Others can be inferred from mathematical notation. Still others must be declared explicitly by the content author or by a math-knowledgeable copy editor. Implied and inferred attributes can be overridden by the author by specifying intents explicitly. This document describes how intent attribute values are supported by [UnicodeMathML](#), a web UnicodeMath-to-MathML implementation.

Roundtripping UnicodeMath semantics

The major purpose of the intent attribute is to enable Presentation MathML (henceforth, MathML) for creating unambiguous math speech. But a nice side benefit is that it enables round tripping some UnicodeMath math semantics through MathML that would otherwise be lost or hard to extract. For example, UnicodeMath has the concept of a math function, such as a trigonometric function. When these functions are converted to [OfficeMath](#), the result is a math-function object. That object is persisted in the OMML format with the <func> tag. Well-formed MathML represents a math function by an <mrow> containing the math function name, the function-apply operator (U+2061), and the math-function argument list. While a MathML reader can parse such an <mrow> to figure out that the <mrow> represents a math function, it's easier if the <mrow> has the attribute intent="function".

Similarly, an <mrow> that contains an n-ary (large-op) object, such as an integral or summation, consists of <msub>, <msup>, <msubsup>, <munder>, <mover>, or <munderover> with the n-ary symbol as the base, followed by the n-aryand (integrand, summand, ...). That too can be parsed by a MathML reader, but it's easier if the <mrow> has the attribute intent="integral", "summation", or in general, "n-ary". In OMML, n-ary objects are represented by the <nary> element.

Another example is roundtripping the slanted, open-face symbols $\mathbb{D}d\mathfrak{eij}$ (U+2145..U+2149). The symbol \mathfrak{d} represents the differential d. Unlike the OfficeMath display engine, web math display engines display these characters as in the Unicode Standard: slanted and open face. But human readers want to see math italic in US publications, upright in European publications, and as is in US Patent applications. By including an intent attribute with the symbol, the symbol can be roundtripped in a MathML <mi> entity containing the corresponding symbol with the desired display format.

Intent attributes implied by UnicodeMath

The attributes described in the previous section are examples of intent attributes that are implied by UnicodeMath. Such attributes are felicitous since the author gets them automatically from an intent-aware UnicodeMath input method. The constructs in this category include equation arrays ($\backslash eqarray$), matrices ($\backslash matrix$, $\backslash pmatrix$, ...), determinants ($\backslash vmatrix$), cases ($\backslash cases$), absolute value ($\backslash abs$), binomial coefficients ($\backslash choose$), differential d (\mathfrak{d} U+2146) and related symbols ($\mathbb{D}d\mathfrak{eij}$ U+2145..U+2149), fences

((), [], {}, ...), math functions (trigonometric functions, exp, ...), and n-ary (large-op) expressions (integrals, summations, products, ...).

Intent attributes inferred from UnicodeMath

Authors also get inferred intent attributes automatically. Such attributes include those for open and/or closed intervals ([a,b) (a,b] [a,b] [a,b[]a,b]]a,b[) and Leipzig derivatives such as dy/dx .

Intent attributes requiring author mark up

In general authors can specify intents and argument references via the `\intent` and `\arg` control words with the syntax

`\intent("..."<expression>)`

`\arg [A-Za-z]+ <expression>`

To resolve ambiguous `\arg` mark up, e.g., for the base of a superscript object vs the whole superscript object, parentheses can delimit the `\arg` argument, that is,

`\arg([A-Za-z]+ <expressions>)`

Examples

In the following UnicodeMath examples, the symbols (a), (a), (c), and (i) are input by the control words `\abs`, `\arg`, `\choose`, and `\intent`, respectively. Ideally in English, the first example, $\partial f(x,y)/\partial x$ should be spoken as “partial derivative of f of x, y with respect to x”, that is, not spoken with a comma-separated list. The list notation is global in character but should be spoken in a localized way.

UnicodeMath	Display math	MathML
$\partial f(x,y)/\partial x$	$\frac{\partial f(x,y)}{\partial x}$	<pre> <math display="block"> <mrow> <mfrac intent="partial- derivative(\$f,1,x)"> <mrow> <mi>\partial</mi> <mrow arg="f"> <mi>f</mi> <mo>&#x2061;</mo> <mrow intent="fenced"> <mo>(</mo> <mrow> <mi>x</mi> <mo>,</mo> <mi>y</mi></mrow> <mo>)</mo> </mrow></mrow></mrow> <mrow> <mi>\partial</mi> <mi>x</mi></mrow></mfrac> </mrow></math> </pre>

$d^2f(x)/dx^2$	$\frac{d^2f(x)}{dx^2}$	$\frac{d^2f(x)}{dx^2}$
$\partial^2\psi(x,t)/\partial x\partial t$	$\frac{\partial^2\psi(x,t)}{\partial x\partial t}$	$\frac{\partial^2\psi(x,t)}{\partial x\partial t}$

$\partial^2/\partial x^2 f(x,y)$	$\frac{\partial^2}{\partial x^2} f(x,y)$	$\frac{\partial^2 f(x,y)}{\partial x^2}$
$d(\tan x)/dx$	$\frac{d(\tan x)}{dx}$	$\frac{d(\tan x)}{dx}$
$(-\infty, 3]$	$(-\infty, 3]$	$(-\infty, 3]$

		x^3
b^n	x^n	b^n
A^T	A^T	A^T
$ax^2 + bx + c = 0$	$ax^2 + bx + c = 0$	$ax^2 + bx + c = 0$
$\binom{n}{k}$	$\binom{n}{k}$	$\binom{n}{k}$
$\binom{n}{k^2}$	$\binom{n}{k^2}$	$\binom{n}{k^2}$

		$\frac{k^2}{2}$
$(a)x$	$ x $	$ x $
$(a)((a)x - (a)y)$	$ x - y $	$ x - y $