What does Mathematical Notation actually mean, and how can computers process it?

James Davenport
Hebron & Medlock Professor of Information Technology[1]

University of Bath (U.K.)

29 January 2014

[1] Thanks to many people: typesetters, editors, OpenMath and MathML colleagues, TEXnicians

Overview

Disclaimer: I have read very little Hungarian mathematics, and this is a brief introduction to a very large (and diverse) subject: however, I used to typeset mathematics at school, and have been in OpenMath for 20 years, and MathML for 15

1. Mathematical notation

   and some of its flaws

2. How it is currently displayed/ represented

   MathML (Presentation/Content); OpenMath

3. How it might be understood

The subjects do overlap

(The outsider's perception of) Mathematical Notation

Unambiguous, unchanging, precise, world-wide (or more so)

- "as clear as 2+2=4"
- Google the phrase "mathematically precise"
- Various science-fiction stories (e.g. Pythagoras' Theorem)
- And in real life — mathematicians *can* and *do* communicate via notation
- The computing discipline of "Formal Methods" tries to reduce computer programming to mathematics/logic

And indeed there's a lot of truth in this

Certainly not unchanging

$+$	is less than 500 years old [Sti44] (also $-$ and $\surd$)
$=$	is slightly younger [Rec57]
Recorde	wrote $2\overline{a+b}$: $2(a+b)$ is later
$(\ldots)$ won	because it is (much!) easier for manual typesetting
Calculus	had/has two conflicting notations $\dot{x}$ or $\frac{\mathrm{d}x}{\mathrm{d}t}$.
Relativity	introduced the summation convention: $\sum_{i=1}^{3} c_i x^i$ is just $c_i x^i$ (but $c_\mu x^\mu$ is short for $\sum_{\mu=0}^{3} c_\mu x^\mu$) [Ein16]

And practically every mathematician introduces some notation: natural selection (generally) applies

Not quite so international

Idea	Anglo-Saxon	French	German
half-open interval	$(0,1]$	$]0,1]$	varies
single-valued function	$\arctan$	Arctan	$\arctan$
multi-valued function	Arctan	$\arctan$	Arctan
$\{0,1,2,\ldots\}$	$\mathbb{N}$	$\mathbb{N}$	$\mathbb{N}\cup\{0\}$
$\{1,2,3,\ldots\}$	$\mathbb{N}\setminus\{0\}$	$\mathbb{N}\setminus\{0\}$	$\mathbb{N}$

Or universal: $\sqrt{-1}$ is *i* to most people, but *j* to Electrical Engineers, and the MatLab system allows both

And these problems occur at an early age [Lop08]

MATHEMATICAL NOTATION COMPARISONS BETWEEN U.S. AND LATIN AMERICAN COUNTRIES

OPERATION DESCRIPTION
DIVISION

Many students come into the U.S. schools using algorithms learned in their country of origin. For example, students in many Latin American countries are expected to do and exhibit more mental computation as the following algorithm illustrates. To assist educators in recognizing different procedural knowledge as valid, we explain how this algorithm works

Format 1	**Format 2**	
$3\sqrt{74}$	$74 \lfloor\underline{3}$	In this algorithm, students will divide 3 into 74 and may write it in one of two ways.
$\begin{array}{r} 2 \\ 3\overline{)74} \\ 1 \end{array}$	$\begin{array}{l\|l} 74 & \underline{3} \\ 1 & 2 \end{array}$	▪ Students typically begin to formulate and answer questions such as: How many times can 3 go into 7? Another way of asking is if we divide 70 into 3 sets, how many are in each set. ▪ Students write the 2 in the tens place, above the 7, on Format 1, but the 2 goes below the divisor when written in Format 2 style. Notice

<table>
<tr><td></td><td></td><td>subtract. The only part that is written on paper is the remainder, 1 ten. Notice its location on both formats.</td></tr>
<tr><td>$\begin{array}{r}2\\ 3\overline{)74}\\ 14\end{array}$</td><td>$\begin{array}{l|l}74 & 3\\ \cline{2-2} 14 & 2\end{array}$</td><td>

- The 4 is brought down and students consider 14 next.
- Notice where the 14 is written on both formats.

</td></tr>
<tr><td>$\begin{array}{r}24\\ 3\overline{)74}\\ 14\end{array}$</td><td>$\begin{array}{l|l}74 & 3\\ \cline{2-2} 14 & 24\end{array}$</td><td>

- Students now find that 3 will go into 14 three (3) times. They write 4 in the quotient’s place.

</td></tr>
<tr><td>$\begin{array}{r}24\\ 3\overline{)74}\\ 14\\ 2\end{array}$</td><td>$\begin{array}{l|l}74 & 3\\ \cline{2-2} 14 & 24\\ 2 & \end{array}$</td><td>

- Students again mentally subtract 12 from 14 and write only the remainder: 2.

</td></tr>
</table>

in fact there are many variations of long division

The MathML community know of 10, such as `stackedleftlinetop`: see `http://www.w3.org/Math/draft-spec/mathml.html#chapter3_presm.mlongdiv.ex`
Note the utility of being able to re-use one example with different presentations.

And it's certainly subject area specific

For example $(2,4)$ might be

- Set Theory: The ordered pair "first 2, then 4"
- (Geometry): The point $x = 2, y = 4$
- (Vectors): The 2-vector of 2 and 4
- Calculus: Open interval from 2 to 4
- Group Theory: The transposition that swaps 2 and 4
- Number Theory: The greatest common divisor of 2 and 4

In general, these are **spoken** differently: the written text "we draw a line from (2,4) to (3,5)" is spoken "we draw a line from the point (2,4) to the point (3,5)' . This makes "text to speech" very difficult for (advanced) mathematics: consider "Since $H_i \leq G$ for $i \leq n$"

Our Notation isn't perfect I (Landau Notation)

Orders of growth (The "Landau Notation" [Bac94])

- $\surd$ $O(f(n))$ for $\{g(n) | \exists N, A : \forall n > N \; |g(n)| < Af(n)\}$
- $\surd$ And similarly Ω, Θ etc.
- But we write "$n = O(n^2)$" when we should write "$n \in O(n^2)$"

  Generally spoken "n is big-O of n squared", not equals

This isn't the traditional use of "$=$", for example "$n = O(n^2)$" but *not* "$O(n^2) = n$"

Causes grief every time I have to explain this (I lecture the first-year Maths course that introduces this), and many books don't give the simple definition $\Theta(f(n)) = O(f(n)) \cap \Omega(f(n))$

[Lev07] is the only text I know to be "correct"

Our Notation isn't perfect II: Iterated functions

- $\sqrt{}$ $\sin(x^2)$: square x, then apply sin
- $\sqrt{}$ $(\sin x)^2$: apply sin to x, then square the result
- $\sqrt{}$ $\sin(\sin(x))$: apply sin to x, then apply sin again

$\sin^2 x$ is generally used to mean $(\sin x)^2$:
"[This] is by far the most objectionable of any" [Bab30]

If anything, it should mean $\sin(\sin(x))$:

since this is the sense in which we write $\sin^{-1}(x)$ — apply the inverse operation of sin, not $1/\sin(x)$

An example of mathematical notation?

$$\pi = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{292 + \ddots}}}}$$

which is nearly always written as

$$\pi = 3 + \frac{1}{7+}\,\frac{1}{15+}\,\frac{1}{1+}\,\frac{1}{292+}\cdots$$

Much easier for (manual) typesetting, and uses less space

So how might a computer display mathematical notation?

Historically	Some kind of image: GIF/JPEG
Typesetting	Many attempts, then $\TeX$ [Knu84]
Principle	boxes with width, height and depth
depth	is vital: recall continued fraction
Since 1998	(at least in theory) MathML (Presentation) [Con99]
But	back then browsers didn't have depth — still a significant problem, and Chrome, for example, sometimes does and sometimes doesn't support MathML
And	the range of fonts is often inadequate, or nonstandard
MathJax	is a very pragmatic solution [Mat11]

Linebreaking: a major challenge

How should a mathematical expression be broken across across multiple lines?

Author TeX, and LaTeX, provide no support for breaking displayed equations, and not much for "in-line" equations

when I reformat a document, re-breaking equations is a significant part of the effort

System the author of a web page has no control over the screen-size of the browser, so the browser *has* to break the expression

The author can give hints, and the MathML standard provides suggestions, but this is an unsolved problem (and an important one for e-books!)

MathML (Presentation)

This specifies the 'presentation' elements of MathML, which can be used to describe the layout structure of mathematical notation. $f(x)$, `f(x)` in TEX, would (best) be represented in MathML as

```
<mrow>
  <mi> f </mi>
  <mo> &ApplyFunction; </mo>
  <mrow>
    <mo> ( </mo>
    <mi> x </mi>
    <mo> ) </mo>
  </mrow>
</mrow>
```

Note that it is clear precisely what the argument of f is: this matters for line breaking and speech rendering — "f of x", as well as meaning

But it **is** presentation

and, I would argue, largely written presentation, though MathML→speech is definitely better than predecessors, and good for “K-12” (school) mathematics

```
<mrow>
  <mo> ( </mo>
  <mn> 2 </mn>
  <mo> , </mo>
  <mn> 4 </mn>
  <mo> ) </mo>
</mrow>
```

(spoken “open bracket, two, comma, four, close bracket”)
is just as ambiguous as $(2, 4)$ (indeed, it’s really the same thing) To ask what the mathematics “means”, we need MathML (Content)

MathML (Content)

“an explicit encoding of the underlying mathematical meaning of an expression, rather than any particular rendering for the expression” [Con14]

Consider $(F+G)x$: this could be

multiplication or function application

```
<apply><times/>
  <apply><plus/>
    <ci>F</ci>
    <ci>G</ci>
  </apply>
  <ci>x</ci>
</apply>
```

```
<apply>
  <apply><plus/>
    <ci>F</ci>
    <ci>G</ci>
  </apply>
  <ci>x</ci>
</apply>
```

No need for brackets, as `<apply>` groups, and the meaning is explicit: in the first we have application of `<times/>` while in the second we are applying $F+G$

OpenMath: 1994–

This grew out of the computer algebra community: exchanging mathematics between different algebra systems
Extensibility was key: very few basic concepts

Basic objects `OMI` integers, `OMF` (IEEE) floating point numbers, `OMSTR` (Unicode) strings, `OMB` byte arrays, `OMV` (mathematical) variables, `OMS` OpenMath symbols

`OMA` (the concept of) function application

`OMATTR` attributes of an object

`OMBIND` binding variables (λ, $\sum_i$ etc.)

`OMERR` error objects

All else is built from these: even addition is just a symbol

OpenMath symbols

A symbol (or several) is defined in a *Content Dictionary* (CD), which lists the symbols and, formally or informally, their meaning

- `<OMS name="plus" cd="arith1"/>` the “addition” operator
- `<OMS name="times" cd="arith1"/>` the “multiplication” operator
- `<OMS name="times" cd="arith2"/>` non-commutative multiplication
- `<OMS name="log" cd="transc1"/>` the complex logarithm, with an informal specification of the branch cut (following [AS64])
- `<OMS name="arctan" cd="transc1"/>` the inverse tangent, with a **formal** relationship with log.

Anyone can wrte a Content Dictionary: `private`, `experimental` and can become `official`

MathML (Content) evolution

MathML was the first XML application

1.0: 1998 "K–12" (Kindergarten to High School) Mathematics: 90 elements

2.0: 2000 rather more calculus: 127 elements

2.0 2nd ed: 2003 ability to extend via OpenMath

3.0: 2010 Full interoperability with OpenMath

3.0 2nd ed: 2014 (some bug fixes)

so now `<times/>` is just a shorthand for
`<OMS name="times" cd="arith1"/>`
OpenMath workshop at CICM 2014
(`http://cicm-conference.org/2014/cicm.php`) will consider closer integration

How might a computer understand written mathematics?

The technical term is **parsing** and there are papers, books and numerous tools (`flex`, `bison` etc.) to do this, for over fifty years

But two-dimensional parsing? Little literature and no tools

It's not even clear what the specification would be

A few packages, both for reverse-engineering PDF [BSS12, Suz11] and for handwritten mathematics [HW13]

Generally a mass of heuristics, often with machine-learning

Even the one-dimensional parsing is hard:

What does juxtaposition mean?

Number formation 23 $(2 \cdot 10 + 3)$

Word formation sin

function application $\sin x$ (`<sin/>⁡x`)

Multiplication xy (`x⁢y`)

Concatenation M_{ij} (`i⁣j`)

Addition $4\frac{1}{2}$ (`4⁤`...)

(for technical reasons, this isn't `4&InvisiblePlus;`)

Juxtaposition "explained" [Dav14, Table 1]

left weight	right weight	meaning	example
normal	normal	lexical	sin
normal	italic	application	$\sin x$
italic	italic	multiplication	xy (but M_{ij})
italic	normal	multiplication	$a \sin x$
digit	digit	lexical	42 (but M_{42})
digit	italic	multiplication	$2x$
digit	normal	multiplication	$2 \sin x$
normal	digit	application	$\sin 2$
		(but note the precedence in	$2 \sin 3x$)
italic	digit	error	$x2$
		(but reconsider)	x^2 or x_2?
digit	fraction	addition	$4\frac{1}{2}$
italic	greek	application^{-1}	$a\phi$
		(as in group theory)	i.e. $\phi(a)$
italic	(	unclear	$f(y+z)$ or $x(y+z)$

Consequences

- Compare "$\sin x$" (`$\sin x$`) with "${\rm sin}x$" (`${\rm sin}x$`)
- The (trained!) eye is very sensitive to these differences of spacing
- Note also that the font drives the **meaning** of juxtaposition
- Hence the requirement to digitise mathematics more carefully than normal text (at least 400dpi, preferably 600dpi, whereas normal text is fine at 300dpi)
- "All variables are equal" (α-conversion) isn't true in practice: $f(y+z)$ versus $x(y+z)$, however, there's no theory here (except in relativistic summation)

We've come a long way from just images, but there's still a long way to go: in particular *searching* for formulae is still an unsolved problem (MathSearch workshops/challenges)

Bibliography I

M. Abramowitz and I. Stegun.
Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing.
US Government Printing Office, 1964.

C. Babbage.
On notations.
Edinburgh Encyclopaedia, 15:394–399, 1830.

P. Bachmann.
Die analytische Zahlentheorie.
Teubner, 1894.

Bibliography II

J. Baker, A. Sexton, and V. Sorge.
Maxtract: Converting PDF to LaTeX, MathML and Text.
In J. Jeuring *et al.*, editor, *Proceedings CICM 2012*, pages 421–425, 2012.

World-Wide Web Consortium.
Mathematical Markup Language (MathML[tm]) 1.01 Specification: W3C Recommendation, revision of 7 July 1999.
`http://www.w3.org/TR/REC-MathML/`, 1999.

World-Wide Web Consortium.
Mathematical Markup Language (MathML) Version 3.0: editors' second edition.
`http://www.w3.org/Math/draft-spec/`, 2014.

Bibliography III

J.H. Davenport.
Nauseating Notation.
`http://staff.bath.ac.uk/masjhd/Drafts/Notation.pdf`, 2014.

A. Einstein.
Die Grundlage der allgemeinen Relativitaetstheorie (The Foundation of the General Theory of Relativity).
Annalen der Physik Fourth Ser., 49:284–339, 1916.

R. Hu and S. Watt.
Determining Points on Handwritten Mathematical Symbols.
In J. Carette *et al.*, editor, *Proceedings CICM 2013*, pages 168–183, 2013.

Bibliography IV

D.E. Knuth.
The $\TeX$book.
Computers and Typesetting Vol. A, 1984.

A. Levitin.
Introduction to the design and analysis of algorithms.
Pearson Addison-Wesley, 2007.

N.R. Lopez.
Todos: Mathematics for All.
Harris County Department of Education, 2008.

MathJax Consortium.
MathJax: Beautiful math in all browsers.
http://www.mathjax.org/, 2011.

Bibliography V

R. Recorde.
The Whetstone of Witte.
London, 1557.

Stifelius.
Arithmetica Integra.
Nurimberg, 1544.

M. Suzuki.
Infty (2011).
`http://www.inftyproject.org`, 2011.