MathML and Ecma Math (OMML)

MurrayS3 (https://social.msdn.microsoft.com/profile/MurrayS3) October 6, 2006

Before comparing these two math XMLs, I'm excited to point you at Jennifer Michelstein's nifty post on Equations in Word 2007 (http://blogs.msdn.com/microsoft_office_word/archive/2006/10/04/Equations-in-Word-2007.aspx). It comes complete with instructive videos demonstrating how to enter mathematical expressions using the Equation Ribbon as well as using the linear format (http://www.unicode.org/notes/tn28/UTN28-PlainTextMath-v2.pdf) discussed in my earlier postings. Jennifer plans to add other posts on this subject to the Microsoft Office Word (http://blogs.msdn.com/microsoft_office_word) blog. It's worth pointing out that Outlook 2007 also has Word's math facility since Outlook 2007 uses Word 2007 for editing all email body text.

I'm currently busily connecting RichEdit's math facility to external clients, using three file formats: RTF, MathML, OMML (Office MathML). These are all high-level formats compared to the linear format we use for equation input. Naturally RichEdit has to handle RTF math, since RTF has been RichEdit's main file format since RichEdit 1.0 shipped back in 1993 for the Windows 3.1 mail client. RTF math is essentially just OMML with RTF braces {} instead of XML tags. With it, it's easy to copy technical text to and from Word 2007. RichEdit is used by the Microsoft Math graphing calculator, which ships with the Encarta Student Edition, but the calculator communicates with RichEdit using RichEdit's text object model (TOM) instead of a file format. RichEdit also handles text in OneNote, but the math facility isn't turned on there yet.

Meanwhile it would be great to connect RichEdit with the incredible math engines out there like OpenMaple, Mathematica, etc., and clearly MathML is the lingua franca for that. So I've been developing MathML and OMML writers for RichEdit and Misha (Mikhail Baranovsky) is developing the readers (they're harder!) In the process, we've been comparing these two math XMLs.

Naturally there's been a lot of discussion (http://www.robweir.com/blog/2006/08/math-you-cant-use.html) as to why we even have OMML (http://www.ecma-international.org/news/TC45_current_work/Office%20Open%20XML%20Part%204%20-%20Markup%20Language%20Reference.docx), since MathML is really good. Brian Jones has addressed that issue in some detail in his Open XML Formats (http://blogs.msdn.com/brian_jones/) blog. The main problem is that Word needs to allow users to embed arbitrary span-level material (basically anything you can put into a Word paragraph) in math zones and MathML is geared toward allowing only math in math zones. A subsidiary consideration is the desire to have an XML that corresponds closely to the internal format, aiding performance and offering readily achievable robustness. Since both MathML and OMML are XMLs, XSLTs can (and have) been created to convert one into the other. So it seems you can have your cake and eat it too. Thank you XML!

The spirits of MathML and OMML are somewhat different. By MathML, I'm referring to the MathML presentation tag set (http://www.w3.org/TR/2003/REC-MathML2-20031021/chapter3.html), rather than the content tag set (http://www.w3.org/TR/2003/REC-MathML2-20031021/chapter4.html), since our main emphasis is on presentation. A quick summary of the difference in spirits is

1. MathML built-up objects may be described by an infix notation, while OMML's are described by a prefix notation

2. MathML built-up object arguments are defined positionally, while OMML's are tagged explicitly

With MathML when you find an <mrow>, you look at the next tag(s) to see what's inside. If you find an <mo> entry, you have an operator, which you look up in your operator table to figure what kind of a possibly built-up object is involved. It could be an open fence (parentheses, brackets, braces, etc.), an n-ary operator, a function-apply (for trigonometric and other math functions), or one of many operators that don't get built up. For fences MathML also has the <mfenced> tag, which is essentially the same as OMML's delimiter <d> tag.

Each way of representing fences has its advantages, the infix approach allowing embellished fences (such as underlined fences) and the <mfenced> approach allowing a sequence of separated arguments. OMML's <d> can't represent embellished fences, but fortunately for OMML, they aren't common. At least, we couldn't find any in my shelves of physics and math books, but you never know for sure about these mathematicians ☺. Built-up expressions like subscripts and superscripts are represented by prefix notations in both MathML and OMML.

The following table summarizes the built-up objects in the Office math model along with the OMML and target MathML tags

Built-up Office Math Object	OMML tag	MathMl
Accent	acc	mover/munder
Bar	bar	mover/munder
Box	box	menclose (approx)
BoxedFormula	borderBox	menclose
Delimiters	d	mfenced
EquationArray	eqArr	mtable (with alignment groups)
Fraction	f	mfrac
FunctionApply	func	&FunctionApply; (binary operator)
LeftSubSup	sPre	mmultiscripts (special case of)
LowerLimit	limLow	munder
Matrix	m	mtable
Nary	nary	mrow followed by n-ary mo
Phantom	phant	mphantom and/or mpadded
Radical	rad	msqrt/mroot
GroupChar	groupChr	mover/munder
Subscript	sSub	msub
SubSup	sSubSup	msubsup
Superscript	sSup	msup
UpperLimit	limUpp	mover

OMML tags are always written with a math namespace prefix like "m:" and I recommend this convention for MathML as well. The reason is that these XMLs are useful in many contexts, not just in HTML(5) and using namespace prefixes allows the XML parser to delegate to the appropriate tag-set owner.

Comparing the two ways of representing the built-up fraction, we see how OMML has explicit argument tags, whereas MathML determines arguments by position. The built up version of the fraction a/b in OMML is represented by (aside from possible properties)

```
<m:oMath xmlns:m="http://schemas.openxmlformats.org/officeDocument/2006/math
(http://schemas.openxmlformats.org/officeDocument/2006/math)">
  <m:f>
    <m:num>
      <m:r>
        <m:t>a</m:t>
      </m:r>
    </m:num>
    <m:den>
      <m:r>
        <m:t>b</m:t>
      </m:r>
    </m:den>
  </m:f>
</m:oMath>
```

where we see how the numerator and denominator are tagged explicitly. In MathML, these arguments are given by the next entity and the one after that, respectively:

```
<m:mfrac>
  <m:mi>a</m:mi>
  <m:mi>b</m:mi>
</m:mfrac>
```

This comparison reveals that OMML can be more verbose than MathML. A less verbose comparison results for the fraction (a+b)/c, since in OMML it's

```
<m:f>
  <m:num>
    <m:r><m:t>a+b</m:t></m:r>
  </m:num>
  <m:den>
    <m:r><m:t>c</m:t></m:r>
  </m:den>
</m:f>
```

whereas in MathML, it's

```
<m:mfrac>
  <m:mrow>
    <m:mi>a</m:mi>
    <m:mo>+</m:mo>
    <m:mi>b</m:mi>
  </m:mrow>
  <m:mi>c</m:mi>
</m:mfrac>
```

Here the <m:mrow> is needed for the numerator to make it the first entity following the <m:mfrac>. For both a/b and (a+b)/c, the linear format sure has the shortest representation!

Another difference between MathML and OMML is in the positioning of the radical (root) degree and prescript arguments relative to their respective bases. In OMML these arguments are positioned so that the left and right arrow keys traverse the objects unidirectionally. Specifically for the radical, the degree argument precedes the radicand, while for MathML it follows the radicand. By having it precede, a right arrow key at the start of the radical moves into the degree and then into the radicand, exactly the way one would expect geometrically. Similarly OMML's prescripts precede the base, whereas in MathML's multiscripts object they follow the base.

Tags MathML (https://blogs.msdn.microsoft.com/murrays/tag/mathml/) OMML (https://blogs.msdn.microsoft.com/murrays/tag/omml/) RichEdit (https://blogs.msdn.microsoft.com/murrays/tag/richedit/) Word 2007 (https://blogs.msdn.microsoft.com/murrays/tag/word-2007/)

Comments (13)

Name *

Email *

Website

Post Comment

Brian Jones: Open XML Formats (http://blogs.msdn.com/brian_jones/archive/2006/10/12/comparison-of-openxml-math-and-mathml.aspx)
October 12, 2006 at 9:58 pm (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-43)
Murray Sargent who was the architect of the new math functionality in Office 2007 has a blog post where

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=43#respond)

Brian Jones: Open XML Formats (http://blogs.msdn.com/brian_jones/archive/2006/10/18/they-re-bringing-out-the-big-guns.aspx)
October 18, 2006 at 9:56 pm (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-73)
Anyone else been following the latest blog posts from IBM and Sun discussing the Office Open XML formats

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=73#respond)

Nick
December 19, 2006 at 8:19 pm (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-263)
Hi,look:

Want to know about eye surgery, check this site: lasik (http://lasik-laser-surgery.blogspot.com">lasik) surgery

End ^) Bay

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=263#respond)

Bill
December 25, 2006 at 8:57 pm (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-273)
Hi,look:

Information tenuate dospan tenuate (http://tenuatedospanpill.blogspot.com">tenuate) dospan Jast for you!

End ^) Bay

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=273#respond)

hiranthap@hotmail.com
June 4, 2007 at 4:52 am (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-803)
Is there any implementation on converting OMML to MathML?

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=803#respond)

Pablo Fernicola
October 9, 2007 at 2:55 pm (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-1283)
Hirantha, from an earlier post by Murray:

Office 2007 ships XSLTs to convert OMML to MathML (omml2mml.xsl) and MathML to OMML (mml2omml.xsl). These XSLTs are used, for example, by Word for MathML clipboard support. They are stored in the subdirectory C:Program FilesMicrosoft OfficeOffice12.

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=1283#respond)

Jesper Lund Stocholm (http://idippedut.dk)
January 31, 2008 at 6:20 am (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-1603)
Hiranta,

I just finished an article about OMML (and conversion to MathML) and I think it might answer some of your questions. You can see my article here: http://idippedut.dk/post/2008/01/Do-your-math—OOXML-and-OMML.aspx (http://idippedut.dk/post/2008/01/Do-your-math---OOXML-and-OMML.aspx)

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=1603#respond)

A 'mooh' Point (http://idippedut.dk/post/2008/01/Do-your-math---OOXML-and-OMML.aspx)
February 12, 2008 at 4:26 am (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-1623)
Do your math – OOXML and OMML (Updated 2008-02-12)

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=1623#respond)

Douglas Capozzalo (http://sites.google.com/site/douglascapozzalo/)
August 23, 2008 at 12:39 pm (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-2113)
Thanks for your post, Murray.

So when do you expect to be able to "paste" an equation made in Word into Excel to get results? Thanks.

Best regards,

Douglas

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=2113#respond)

MurrayS3 (https://social.msdn.microsoft.com/profile/MurrayS3)
May 1, 2009 at 2:58 pm (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-2853)
It would be great to be able to copy an equation in a Word document and paste it into Excel for computation. That's certainly on my wish list 🙂 Maybe it'll happen some day. I'm thinking that Content MathML might be better for Excel computation than Presentation MathML. But for many formulas there are heuristics that let you transform Presentation MathML into Content MathML reliably.

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=2853#respond)

Eduardo
January 11, 2011 at 9:10 am (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-5103)
Can mathml2omml.xsl be used in conjunction with a software other than Office?

WE do not know which are the licensing restrictions for this particular XSLT…

Best regards,

Eduardo

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=5103#respond)

Alex
October 22, 2014 at 8:13 pm (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-8363)
How to convert from MathML in omml? Can you describe?

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=8363#respond)

Nikhil
March 14, 2015 at 10:16 pm (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/#comment-8643)
Hi,

There seems to be very few people who have actually been able to convert Mathml to OMML to render properly inside a Word document. I'm building a Ruby on Rails app (on a Mac) that wants to print a question (which is stored in the db as mathml) to a Word DOC.

While I can copy and paste mathml from the clipboard to Word and it works fine, I am unable to find a solution programatically. My plan is to convert the MathML to OMML and replace it in the document.xml file. I am stuck doing this. For one, I cannot find a XSLT on the Mac Office which does the conversion (I have Mac Office 2011) of MML2OMML.xslt. Does anyone know where I can find such a file on a Mac?

Reply (https://blogs.msdn.microsoft.com/murrays/2006/10/06/mathml-and-ecma-math-omml/?replytocom=8643#respond)

Follow Us


(https://blogs.msdn.microsoft.com/murrays/feed/)

Popular Tags

RichEdit (https://blogs.msdn.microsoft.com/murrays/tag/richedit/)

Math in Office (https://blogs.msdn.microsoft.com/murrays/tag/math-in-office/)

MathML (https://blogs.msdn.microsoft.com/murrays/tag/mathml/)

Word 2007 (https://blogs.msdn.microsoft.com/murrays/tag/word-2007/)

OMML (https://blogs.msdn.microsoft.com/murrays/tag/omml/)

Archives

January 2017 (https://blogs.msdn.microsoft.com/murrays/2017/01/) (1)
December 2016 (https://blogs.msdn.microsoft.com/murrays/2016/12/) (1)
All of 2017 (https://blogs.msdn.microsoft.com/murrays/2017/) (1)
All of 2016 (https://blogs.msdn.microsoft.com/murrays/2016/) (12)
All of 2015 (https://blogs.msdn.microsoft.com/murrays/2015/) (12)
All of 2014 (https://blogs.msdn.microsoft.com/murrays/2014/) (12)
All of 2013 (https://blogs.msdn.microsoft.com/murrays/2013/) (12)
All of 2012 (https://blogs.msdn.microsoft.com/murrays/2012/) (12)
All of 2011 (https://blogs.msdn.microsoft.com/murrays/2011/) (12)
All of 2010 (https://blogs.msdn.microsoft.com/murrays/2010/) (16)
All of 2009 (https://blogs.msdn.microsoft.com/murrays/2009/) (15)
All of 2008 (https://blogs.msdn.microsoft.com/murrays/2008/) (17)
All of 2007 (https://blogs.msdn.microsoft.com/murrays/2007/) (16)
All of 2006 (https://blogs.msdn.microsoft.com/murrays/2006/) (10)