MathML/XML series

# Glossary of some XML formats

Peter James Rowlett
School of Computing & Informatics
Nottingham Trent University
peter.rowlett@ntu.ac.uk

The world of XML is filled with acronyms and jargon. This article aims to give a brief summary of some XML formats that may be useful or interesting.

## XML

XML stands for eXtensible Markup Language [1]. Readers who know a little HTML may benefit from the statement 'XML is like HTML where you can make your own tags' (almost true; see below on XHTML). For those who don't, XML uses tags to enclose text giving it particular meaning. For instance, the following is some text marked as a "para":

```
<para>some text</para>
```

An XML format is a particular set of such tags. XML formats exist as international standards or can be created to suit a particular purpose. The meaning of the tags in a particular XML format is agreed on and programs using that XML format should behave consistently. For instance, "para" above might refer to a paragraph of text, a name of a parachuter or a paranormal phenomenon (or anything else, really), depending on the meaning given to it by the creator(s) of the particular XML format. In general, output is best in a recognised format so that it can be viewed by as many users as possible. Custom XML formats can be useful for internal storage and processing.

XML is simply a way of marking text and can be written in text editor.

## XSLT

eXtensible Stylesheet Language Transformations (XSLT) [2] are procedures, defined in XML, for converting one kind of XML into another. For example, an internal XML database might be written to hold records of some kind using tags that are appropriate. This data can be searched and sorted, since it is appropriately marked-up. For viewing on the Web, an XSLT can be written for conversion to XHTML.

## XHTML

eXtensible HyperText Markup Language (XHTML) is similar to HTML and used primarily for webpages [3]. XHTML structures documents into paragraphs, headings, lists, links, etc.

The main differences between XHTML and HTML are avoiding bad practice like missing out end tags and not overlapping tags (see Fig 1), and writing tags in lower

| (a) | (b) |
|---|---|
| Paragraph one.<br> Paragraph 2. | Paragraph one. <br> Paragraph 2. |
| Some text in bold<br>and italics too | Some text in bold<br>and italics too |

**Fig 1 Examples of (a) not XHTML; (b) Valid XHTML**

case. Complying with XHTML is largely about being strict and careful with your HTML. Easy to follow guides are available [4], as are validators [5].

### MathML

Mathematical Markup Language (MathML) was the subject of a previous article [6]. Basically, MathML is an internationally agreed standard for mathematical notation that has good potential and is seeing increasing use.

### OpenMath

OpenMath is a standard for storing mathematics and its semantic meaning [7]. There is a large overlap with MathML and the two standards work well in conjunction. In fact, OpenMath is seen as an "upgrade" from MathML, in that while MathML has pre-defined a limited set of mathematical symbols suitable for use up to early university level, OpenMath provides a mechanism for defining the semantics of symbols. MathML can be extended using OpenMath or Content MathML can be mapped to symbols in OpenMath [8].

### SVG

Traditionally, graphics are split into pixels and each pixel is given a value which describes its colour (rasterised). By contrast, vector graphics are represented using "geometrical primitives such as points, lines, curves, and polygons" [9]. Consequently, vector graphics can be resized or zoomed with no loss of quality, while rasterised graphics become pixelated. Vector graphics also tend to have smaller filesizes than rasterised graphics.

Scalable Vector Graphics (SVG) [10] is an XML format for describing two-dimensional graphics, both static and animated. Since SVG is an XML format, graphics can be produced in a textual format. For example, a circle is defined by a set of parameters: radius, the x and y coordinates of its centre, line colour and thickness and fill colour (see Fig 2).

SVG also defines standards for mobile devices and printing. SVG is interoperable with other XML formats such as MathML, meaning that SVG images can contain MathML code [11].

### X3D

eXtensible 3D (X3D) [12] is an XML language for describing 3-dimensional graphics. Similarly to SVG, it allows textual representation of 3D graphics.

### XSL-FO

eXtensible Stylesheet Language - Format Objects (XSL-FO) [2] is a print format defined in XML. Since conversion from other XML formats to XSL-FO is readily completed using XSLT, it will be advantageous to know that XSL-FO can be converted to PDF by the free program Format Object Processor (FOP) [13]. It is even possible to give FOP the XML and XSLT files and have it run the XSLT and output a PDF document. Obviously, PDF is an important document format for distribution of content.

### Other formats

This is of course only a tiny selection of XML formats, chosen since they are widely used and of possible interest to the MSOR user.

### References


1. XML: http://www.w3.org/XML/
2. XSL: http://www.w3.org/Style/XSL/
3. XHTML and HTML: http://www.w3.org/MarkUp/
4. *Differences Between XHTML And HTML* [online]. W3Schools, 2006. Available at: http://www.w3schools.com/xhtml/xhtml_html.asp [Accessed 30 June 2006].
5. *The W3C Markup Validation Service* [online]. W3C, 2006. Available at: http://validator.w3.org/ [Accessed 30 June 2006].


```
<circle r="50" cx="100" cy="100" stroke="black" stroke-width="1" fill="blue"/>
```

**Fig 2 Example of a circle defined in SVG**

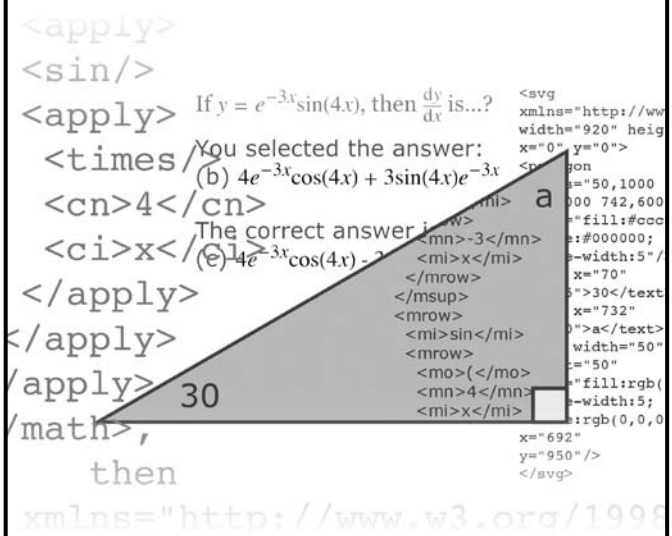
6. Rowlett, P.J. 'MathML / XML series: What is MathML?' *MSOR Connections 6(1).* Maths, Stats and OR Network, 2006.
7. *OpenMath* [online]. The OpenMath Society, 2004. Available at: http://www.openmath.org/ [Accessed 30 June 2006].
8. *OpenMath and MathML* [online]. The OpenMath Society, 2004. Available at: http://www.openmath.org/overview/om-mml.html [Accessed 30 June 2006].
9. *Vector graphics* [online]. Wikipedia, the free encyclopedia, 2006. Available at: http://en.wikipedia.org/wiki/Vector_graphics [Accessed 30 June 2006].
10. SVG: http://www.w3.org/Graphics/SVG/
11. *An XHTML + MathML + SVG Profile* [online]. W3C, 2002. Available at: http://www.w3.org/TR/XHTMLplusMathMLplusSVG/ [Accessed 30 June 2006].
12. *About X3D* [online]. Web3D Consortium, 2006. Available at: http://www.web3d.org/about/ [Accessed 30 June 2006].
13. *Apache FOP* [online]. The Apache Software Foundation, 2006. Available at: http://xmlgraphics.apache.org/fop/ [Accessed 30 June 2006].


The Higher Education Academy — Maths, Stats & OR Network

## MathML/XML Series details

The MathML/XML Series initiated by Peter Rowlett in November 2005 is supported and published by the Maths, Stats & OR Network. The MathML/XML Series intends to explore MathML and other XML formats of interest and how the MSOR Community is making use of these. Do you use MathML or another XML format in your teaching and learning? Will you share your experiences through MSOR Connections? Please contact the editors (**editors@mathstore.ac.uk**) if you would be willing to contribute an article to this series.

The Maths, Stats & OR Network maintains a list of MathML resources at:
**http://www.mathstore.ac.uk/mathml/**




# MathML/XML series