Francis Wright, May 2016

In this document I will briefly consider some very simple mathematical notation that is ambiguous and assess the extent to which LaTeX and MathML markup disambiguate it. I will use simplified markup for this illustration, which I have based on the markup generated by Maple. I am relying on MathJax to render LaTeX and presentation MathML markup in this document. Firefox renders presentation MathML natively, but not content MathML, which is not primarily intended to be rendered. Other browsers currently will not render MathML natively as far as I am aware.

I will use the delimiters "$…$" to indicate LaTeX mathematical input. The delimiters "$…$" are required to indicate MathML input, since MathML is an XML application (i.e. it is like HTML).

Products

The mathematical notation xy would normally be assumed to mean $x \times y$, i.e. the variable x multiplied by the variable y, and I will assume this is the intended meaning. But it could also mean a single variable with a two-letter name. How do we know which was intended? The normal LaTeX markup for this product would be "$ xy $", which renders as xy. In Maple, a space is required to indicate implied multiplication, i.e. "x y", but LaTeX ignores spaces in math mode, so the markup "$ xy $" is equivalent to "$ x y $" and both produce the output xy. This markup is non-semantic.

The presentation MathML markup for the product is

```
<math><mrow><mi>x</mi><mo>&InvisibleTimes;</mo><mi>y</mi></mrow></math>
```

which renders as $x\,y$, and the content MathML markup is

```
<math><apply><times/><ci>x</ci><c>y</ci></apply></math>
```

Both clearly use the word "times" in some form and so are semantic. The content MathML is a little more succinct and easily readable.

Functions

The expression $g(x + y)$ would normally be assumed to mean the function g applied to the expression $x + y$ as its argument, and I will assume this is the intended meaning. But it could also mean the variable g multiplied by the expression $x + y$. The normal LaTeX markup is "$ g(x+y) $", which renders as $g(x + y)$. This markup is non-semantic.

The presentation MathML markup for the function application is

```
<math><mrow><mi>g</mi><mo>&ApplyFunction;</mo><mfenced><mrow><mi>x</mi>
<mo>+</mo><mi>y</mi></mrow></mfenced></mrow></math>
```

which renders as $g\,(x + y)$, and the content MathML markup is

```
<math><apply><ci>g</ci><apply><plus/><ci>x</ci><ci>y</ci></apply></apply>
</math>
```

Both clearly use the word "apply" in some form and so are semantic. The content MathML is a little more succinct and easily readable.

Binomial coefficients and two-dimensional column vectors

The notation $\begin{pmatrix} a \\ b \end{pmatrix}$ is conventionally used to represent both a binomial coefficient and a two-dimensional column vector. (A higher-dimensional column vector is usually represented more distinctly using squarer brackets.)

Binomial coefficients

The best way to mark up a binomial coefficient using LaTeX is as "$ a \choose b $", which renders as $\binom{a}{b}$. This markup is perfectly semantic and corresponds to the way a mathematician would normally read a binomial coefficient (due to the way binomial coefficients arise in elementary probability theory).

The presentation MathML markup for the binomial coefficient is

```
<math><mfenced><mtable><mtr><mi>a</mi></mtr><mtr><mi>b</mi></mtr></mtable>
</mfenced></math>
```

which renders as $\left(\begin{array}{c} a \\ b \end{array} \right)$, and the content MathML markup is

```
<math><apply><csymbol>binomial</csymbol><ci>a</ci><ci>b</ci></apply></math>
```

The presentation MathML markup just refers to tabular notation and is non-semantic whereas the content MathML markup refers to the symbol "binomial" and is semantic.

Two-dimensional column vectors

The obvious LaTeX markup for a two-dimensional column vector would be "$ \left(\begin{array}{c} a \\ b \end{array} \right) $", which renders as $\left(\begin{array}{c} a \\ b \end{array} \right)$. This markup is just a special case of a rectangular array and is not very semantic, but it is distinct from the binomial coefficient markup.

The presentation MathML markup for a two-dimensional column vector is

```
<math><mfenced><mtable><mtr><mtd><mi>a</mi></mtd></mtr><mtr><mtd><mi>b</mi>
</mtd></mtr></mtable></mfenced></math>
```

which renders as $\left(\begin{array}{c} a \\ b \end{array} \right)$, and the content MathML markup is

```
<math><vector><ci>a</ci><ci>b</ci></vector></math>
```

Again, the presentation MathML markup just refers to tabular notation and is non-semantic whereas the content MathML markup uses the word "vector" and is semantic.

Summary

This very brief markup comparison indicates that LaTeX and presentation MathML markup are generally not semantic and hence not fully accessible, whereas content MathML markup is always fully semantic (which, of course, is its purpose) and hence fully accessible. However, content MathML is not normally rendered. There is nothing to stop content MathML being rendered, but it is not primarily intended to be rendered.

The best solution is to use a combination of presentation and content MathML markup and there are standard ways to do this. By default, Maple generates "Parallel Mode MathML", in which the content representation is included within the presentation representation as an XML annotation. This form of mathematical markup should be optimally accessible to both sighted and visually impaired readers using assistive technology. However, MathML can be very verbose compared with LaTeX. LaTeX plus content MathML would provide more succinct markup but the syntactic inconsistency of this combination is unappealing. But, in practice, mathematicians are unlikely to stop using LaTeX as their preferred mathematical markup.