OpenMath in a (Semantic) Web

James H. Davenport
Department of Computer Science
University of Bath
Bath BA2 7AY England
J.H.Davenport@bath.ac.uk
Thanks to JEM for inviting me

February 1, 2008

Some history (1)

OpenMath started life as a protocol for communication between computer algebra systems. It has since expanded in scope to include communication between:

- Computer algebra systems;
- Reasoning systems (theorem provers);
- Pedagogical tools;
- Mathematical databases;
- Layout systems such as LaTeX.

Lack of human involvement is the key.

Some history (2)

The world-wide web developed as a means of communicating between human beings, even if many of the 'pages' were generated by computers. This has led to strategies such as

- "if it's ambiguous ask the user"
- "if it's ambiguous pick the most likely"
- "ambiguity? what's that??".

The belief that there's a browser driven by a human being is the key.

Some history (3)

Much hype about a 'semantic web' — what might it mean for mathematics?

- The semantics *must* not need a human being to interpret.
- The presentation *must* be capable of adapting to the user's culture (anglosaxon/french, normal/EE etc.).
- There *should* not be a split between 'meant for humans' and 'meant for machines'.
- ...

Semantics

We tend to think, and non-mathematicians are very prone to think, that the semantics of mathematics are well-defined and universal, and that mathematical notation represents this universally.

- $[0, 1)$ or $[0, 1[$?
- Arctan or arctan?
- $\mathbf{N}$ or $\mathbf{N} \cup \{0\}$?
- i or j?
- $n = O(n^2)$ so $O(n^2) = n$?

$[0, 1)$ or $[0, 1[$?

`<OMS name="interval_co" cd="interval1"/>` represents the content. The rendering is not specified at all.

Arctan or arctan?

- `<OMS name="arctan" cd="transc1"/>` is a (specific) single-valued function.

- `<OMS name="arctan" cd="transc3"/>` is a multi-valued function.

Which is represented by arctan and which by Arctan is up to the renderer.

$\mathbf{N}$ or $\mathbf{N} \cup \{0\}$?

Here OpenMath has to make a decision, and in fact `<OMS name="N" cd="setname1"/>` *is* $\{0, 1, 2, \ldots\}$. However, there *should* be a representation of $\{1, 2, \ldots\}$.

i or j?

- Electrical engineers will need to customise their renderers!

- The plus side is that they will therefore have access to mathematics not especially directed at them!!

$$n = O(n^2) \text{ so } O(n^2) = n?$$

Well, every mathematician 'talks the talk' that this isn't really "=", but "$\in$", but only one 'walks the walk'. The solution is

```
<OMS name="LandauIn" cd="asymp1"/>,
```

whose semantics are those of $\in$, but whose rendering is that of $=$.

Surface or Deep Structure?

- I carry the book.
- I carried the book.
- The book was carried by me.

Surface or Deep Structure?

- I carry the book.
- [Ego] fero librum.
- I carried the book.
- [Ego] tuli librum.
- The book was carried by me.
- Liber latus meo est.

Surface or Deep Semantics?

- $a_1 \cup a_2 \cup a_3$
- $\bigcup\{a_1, a_2, a_3\}$
- $\bigcup_{i=1}^{3} a_i$

Surface or Deep Semantics? The OpenMath View

- $a_1 \cup a_2 \cup a_3$
- `<OMS name="union" cd="set1"/>`
- $\bigcup\{a_1, a_2, a_3\}$
- `<OMS name="big_union" cd="set3"/>`

or `<OMS name="apply_to_list" cd="fns2"/>`

- $\bigcup_{i=1}^{3} a_i$
- `big_union` on `make_list`

Surface or Deep Semantics?
The MathML View

- $a_1 \cup a_2 \cup a_3$
- `<apply> <union/> <i>`a_1`</i>...</apply>`
- $\bigcup\{a_1, a_2, a_3\}$
- `<apply> <union/> <bvar>i</bvar> <domain ...> <set> <i>`a_1`</i>...</set>`
- $\bigcup_{i=1}^{3} a_i$
- `<apply> <union/> <bvar>i</bvar> <lowlimit>...`

Surface or Deep Semantics?
$\bigcup$'s comparison

MathML seems to have one symbol, which represents (slightly clumsily in the second case) $\bigcup$'s deep semantics. OpenMath seems to have different symbols for the different surface semantics.

Note that $\cup\{\{a\},\{b\}\} = \{\{a\},\{b\}\}$, while $\bigcup\{\{a\},\{b\}\} = \{a,b\}$, so some clumsiness is inevitable in the second case.

Surface or Deep Semantics?

- $a_1 + a_2 + a_3$
- $\Sigma\{a_1, a_2, a_3\}$ (??)
- $\Sigma_{i=1}^{3} a_i$

Surface or Deep Semantics? The OpenMath View

- $a_1 + a_2 + a_3$
- `<OMS name="plus" cd="arith1"/>`
- $\Sigma\{a_1, a_2, a_3\}$ (??)
- `<OMS name="apply_to_list" cd="fns2"/>`
- $\Sigma_{i=1}^{3} a_i$
- `<OMS name="sum" cd="arith1"/>`

Surface or Deep Semantics?
The MathML View

- $a_1 + a_2 + a_3$
- `<apply> <plus/> <i>`a_1`</i>...</apply>`
- $\Sigma\{a_1, a_2, a_3\}$ (??)
- `<apply> <sum/> <bvar>i</bvar> <domain ...> <set> <i>`a_1`</i>...</set>`
- $\Sigma_{i=1}^{3} a_i$
- `<apply> <sum/> <bvar>i</bvar> <lowlimit>...`

Surface or Deep Semantics?
Σ's comparison

This is clearly mathematics' form of an irregular verb. Mathematics' answer to a deponent verb is presumably $\prod$, the 'big' version of `&InvisibleTimes;`.

Ordinary mathematical notation finds the second case "unusual". In this case, MathML looks much more like OpenMath — surface semantics.

Conclusion

- 'Deep semantics' is very appealing, but

  $$\cup\{\{a\},\{b\}\} = \{\{a\},\{b\}\}\bigcup\{\{a\},\{b\}\} = \{a,b\}$$

  shows that it is ambiguous, even if *humans* rarely care.

- Disentangling the various uses of MathML's `<union/>` is problematic: some systems might be able to handle straight-forward cases, but not those with bound variables involved.

- Mathematics itself has irregular verbs, even if most constructs are ‘regular’.

- One question is whether irregularity can be solved purely at the notational level, e.g.

```
<OMS name="apply\_to\_list" cd="fns2"/>
  <OMS name="plus" cd="arith1"/>
```

  is recognised as Σ. If so, how do we tell the difference between conditional convergence and absolute convergence?

- If you don't think this matters, does the difference between Riemann integrals and Cauchy Principal Value integrals matter?

- Can you explain the difference?

- Is it possible to formalise 'deep semantics', as MathML sometimes tries to do, or should we stick with OpenMath's 'surface semantics' ?