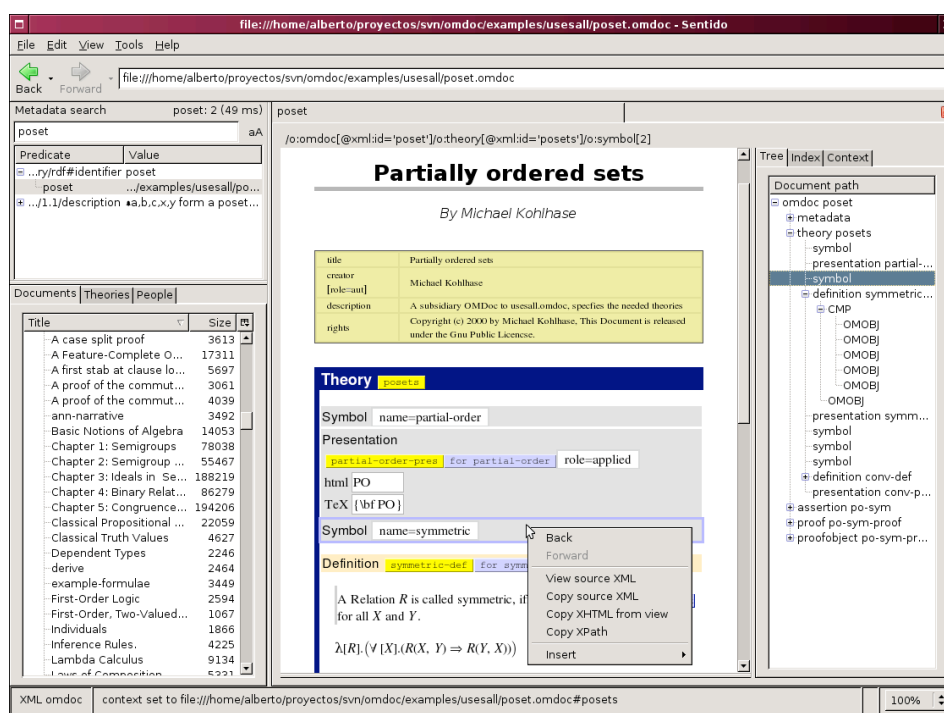


## 26.3 Sentido: an Integrated Environment for OMDoc

Project Home	<a href="http://www.matracas.org/sentido/index.en.html">http://www.matracas.org/sentido/index.en.html</a>
Authors	Alberto González Palomo Toledo, Spain <sup>4</sup>

SENTIDO is an integrated environment for browsing, searching, and editing collections of OMDOC documents. It is implemented as an extension for the MOZILLA/FIREFOX browsers to avoid the biggest problems found when using QMATH: the need to compile the program for installing, the batch mode of interaction that made small corrections consume much of the author's time, and the lack of any support for document navigation and search.



**Fig. 26.3.** SENTIDO after indexing the OMDoc repository in the library (left) and loading a document from it (center and right).

Figure 26.3 shows a typical session initiated by searching in the document library (described below in more detail) and opening one of the results. The context menu displays the options for browsing back and forward, viewing the XML (OMDOC) source of the selected element, copying it to the system clipboard, copying its MATHML rendering or an XPATH expression that identifies it, and inserting new elements.

<sup>4</sup> The author is currently employed part-time in the ACTIVE MATH project, developed by Saarland University and the DFKI, but this work was done on his own, without their supervision or support.

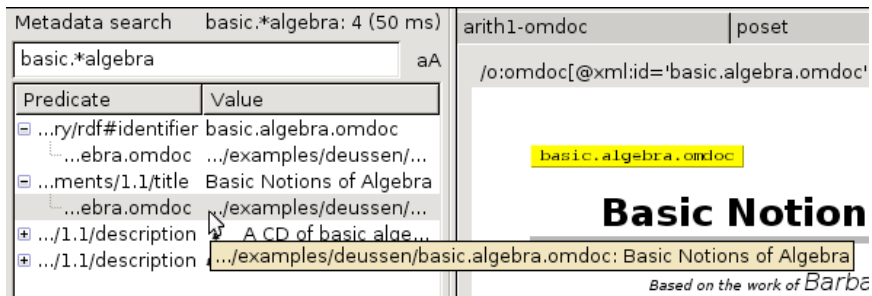
### 26.3.1 The User Interface

The window is made to resemble the web browser, and consists of two main panes: the smaller one on the left contains the interface for the “document library”, and the right one the “document view” and associated information like the document tree, element identifiers index, and context at the current cursor position.

The document library is a knowledge base about documents, the theories defined in them, and people mentioned in their metadata as authors, editors, translators, etc. It is implemented as an RDF store with the documents organized in collections called “volumes” with references to documents, so that different volumes can have documents in common. The tabs labelled “Documents”, “Theories” and “People” display different views of the library content.

The bibliographic data for each document is stored using the Bibliographic Record Schema [Len04], which includes FOAF<sup>5</sup> entries for people.

The documents in the library are indexed by the search engine, which stores their metadata entries and theory identifiers in an abridged inverted index to speed up the searches to the point where “search as you type” becomes possible<sup>6</sup>. The search pattern accepts regular expression syntax, as shown in Figure 26.4.



**Fig. 26.4.** Metadata Search in SENTIDO: tooltips show the content of the cropped entries.

The document view is built using XHTML + MATHML that can be edited normally, with the changes being propagated to the internal OMDoc.

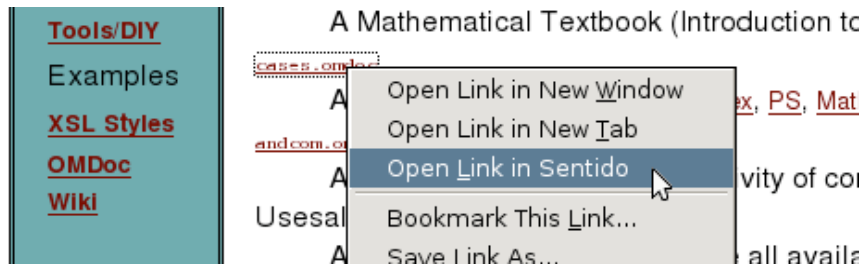
The view is built on demand (using XSLT) as the subparts of the document are unfolded in the document navigation tree found in the right part of the window. This has been found important in practice since many real

<sup>5</sup> “Friend of a friend”, described in their web page <http://www.foaf-project.org> as being “about creating a Web of machine-readable homepages describing people, the links between them and the things they create and do.”

<sup>6</sup> On the author’s 1 GHz laptop computer, the search times in a library of around two thousand documents are usually between 100 and 200 milliseconds.

uses of OMDoc involve documents that contain large lists of elements, like exercises, that are largely independent of each other and thus do not usually require being viewed at the same time, and the biggest delay in opening a medium to large sized document was by far the display of the XHTML view. Another motivation for this approach is to progress towards handling the source document more like a database, and customize its presentation for the task at hand.

SENTIDO adds some options to the context menu in the browser, to allow the user to open links to OMDoc files from web pages (see Figure 26.5).

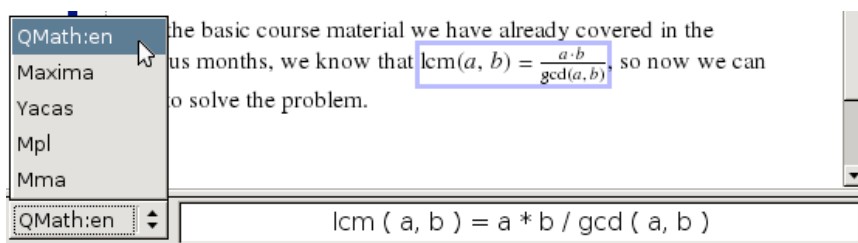


**Fig. 26.5.** MOZILLA's Context Menu after Installing SENTIDO.

### 26.3.2 Formula Editing

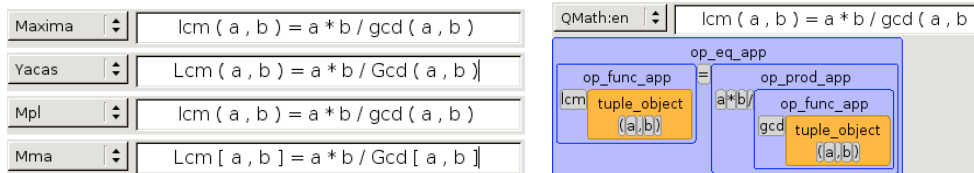
Mathematical expressions are entered using a selectable linear syntax, translated by a new version of the QMATH parser described in Section 26.2. This is a much more capable implementation based on finite-state cascades [Abn96].

There are five grammars included in the install package, that are used for translating back and forth between OPENMATH and the linear syntax of QMATH and the Computer Algebra Systems MAXIMA, YACAS, MAPLE™ and MATHEMATICA®. More syntaxes can be added by writing new grammars, with a format similar to QMATH “context” files.



**Fig. 26.6.** The formula editor under the document view, with the input syntax menu and the text field where the formula is typed, which updates continuously the internal OPENMATH representation and the MATHML view.

When the cursor enters a formula, the linear input field appears at the bottom of the document view, as seen in Figure 26.6. It contains a text field for editing, and a menu button for selecting the syntax, which can be done at any moment: the linear expression is regenerated instantaneously from its OPENMATH form, so it is possible to enter a formula using, for instance, MATHEMATICA<sup>®</sup> syntax, then select another syntax such as MAPLE<sup>™</sup>, and get the expression immediately translated, going through its OPENMATH representation (Figure 26.7).



**Fig. 26.7.** The formula is translated by SENTIDO each time the user selects another syntax (left, the vertical line is the blinking caret), and it is possible to view the parse tree (right), updated as the input is modified.

Insertion of formulae is achieved by typing the dollar symbol “\$”, which produces an empty formula readily editable so that the sequence of keystrokes is similar to typing T<sub>E</sub>X or QMATH text: one can type  $\$e^{(\pi*i)+1}=0\$$  and get  $e^{\pi i} + 1 = 0$  without having to look at the formula editor or use the mouse. The changes as the formula is being modified are stored, and the display updated from the OPENMATH form, at each point when there is a complete parse of the formula. This gives immediate feedback on how the program understands the input.

An important difference is that there is no need to care about “context files” any more. In QMATH, specifying a “context” had a double function: putting symbols in scope for disambiguation, and selecting a notation style for them. Those aspects are separated in SENTIDO: the in-scope symbols are automatically determined from the enclosing theory and those imported from it (recursively), and the notation is selectable by the user.

Note that the parser allows any characters supported by the browser rendering engine of MOZILLA/FIREFOX (a big subset of UNICODE), not just ASCII. For example, the number 3.14159265... can be entered either as  $\pi$  or with an ASCII form depending on the selected syntax: “pi” for QMATH, “%pi” for MAXIMA, or “Pi” for YACAS, MAPLE<sup>™</sup> and MATHEMATICA<sup>®</sup>.

### 26.3.3 Future Work and Availability

SENTIDO is a long term personal project that has been in development for several years (since 2004), entirely in the author’s spare time and using his

own computing resources, based on experiments<sup>7</sup> and notes collected during the development of QMATH. Therefore, we expect it to continue developing during the foreseeable future unless a better application appears that makes it redundant.

Its components are designed to be reusable, which is tested from time to time by producing spin-off applications that use subsets of its functionality in a self-contained way. One example is the small Computer Algebra System called ALGEBRA [Pala], that contains parts of SENTIDO such as the new parser combined with specific ones like the function plotter and the term rewriting engine.

Future developments will focus on what we consider the two main tasks for a development environment for semantic encoding of mathematical content:

- Ease the tedium of writing all the details needed for an unambiguous encoding of the content. This is where the flexible input parser comes into play: having a syntax redefinable at any point in the content simplifies the expression input, as the syntax can be adapted to the context in which an expression occurs.
- Provide some benefit once we have the semantic encoding which would not be present with an ambiguous encoding such as  $\text{T}_{\text{E}}\text{X}$ . Here we need to implement detailed checking and strong search capabilities. A next step would be to assist the writing process by inferring new content and informing the input interface about the context as mentioned above.

Some planned improvements in SENTIDO are:

- Make the browser open OMDOC documents linked from normal pages directly in SENTIDO, by implementing a stream handler for the `MIME` type `application/omdoc+xml`.
- Integrate ALGEBRA into SENTIDO, to add automated symbolic manipulation to the document editing process.
- Extend the checking being done on the theories: at the time of writing these lines, only the theory import relations are checked for loops and unknown theory references, which was already enough to locate several mistyped theory identifiers in the OMDOC repository.
- Implement useful features found in other projects such as THEOREMA [PB04]. This is strongly related to the two points above since THEOREMA implements many features needed for the task of content checking which are still missing in SENTIDO, and some of them are available in proof-of-concept form in ALGEBRA.

SENTIDO is Free Software distributed under the GNU General Public License (GPL [FSF91]).

---

<sup>7</sup> Some of those early experiments with MOZILLA inspired work done on adapting OPENOFFICE and  $\text{T}_{\text{E}}\text{X}_{\text{MACS}}$  for OMDOC in collaboration with George Goguadze [GP03]