modules.sty: Semantic Macros and Module Scoping in STEX*

Michael Kohlhase & Deyan Ginev & Rares Ambrus Jacobs University, Bremen http://kwarc.info/kohlhase

April 20, 2014

Abstract

The modules package is a central part of the STEX collection, a version of T_EX/I_FT_EX that allows to markup T_EX/I_FT_EX documents semantically without leaving the document format, essentially turning T_EX/I_FT_EX into a document format for mathematical knowledge management (MKM).

This package supplies a definition mechanism for semantic macros and a non-standard scoping construct for them, which is oriented at the semantic dependency relation rather than the document structure. This structure can be used by MKM systems for added-value services, either directly from the STFX sources, or after translation.

^{*}Version v1.3 (last revised 2013/12/12)

Contents

1	Introduction										
2	The	he User Interface									
	2.1	Package Options	3								
	2.2	Semantic Macros	4								
	2.3	Testing Semantic Macros	6								
	2.4	Symbol and Concept Names	6								
	2.5	Modules and Inheritance	7								
	2.6	Dealing with multiple Files	8								
	2.7		10								
	2.8	Including Externally Defined Semantic Macros	11								
	2.9	Views	11								
	2.10		12								
3	Lim	imitations & Extensions 13									
	3.1	Perl Utility sms	13								
	3.2	Module Signatures loaded even if Modules are	13								
	3.3	Only one level of path simplification	13								
	3.4	Qualified Imports	13								
	3.5	Error Messages	14								
	3.6	Crossreferencing	14								
	3.7		14								
4	The	Implementation	15								
	4.1	Package Options	15								
	4.2	Modules and Inheritance	15								
	4.3	Semantic Macros	18								
	4.4	Defining Math Operators	21								
	4.5	Testing Semantic Macros	22								
	4.6		22								
	4.7	Dealing with Multiple Files	23								
	4.8		24								
	4.9		27								
	4.10	v v	27								
			28								
			29								
		1	29								

1 Introduction

Following general practice in the TEX/IATEX community, we use the term "semantic macro" for a macro whose expansion stands for a mathematical object, and whose name (the command sequence) is inspired by the name of the mathematical object. This can range from simple definitions like \def\Reals{\mathbb{R}} for individual mathematical objects to more complex (functional) ones object constructors like \def\SmoothFunctionsOn#1{\mathcal{C}^\infty(#1,#1)}. Semantic macros are traditionally used to make TEX/IATEX code more portable. However, the TEX/IATEX scoping model (macro definitions are scoped either in the local group or until the rest of the document), does not mirror mathematical practice, where notations are scoped by mathematical environments like statements, theories, or such. For an in-depth discussion of semantic macros and scoping we refer the reader [Koh08].

The modules package provides a LATEX-based markup infrastructure for defining module-scoped semantic macros and LATEXML bindings [Mil] to create OM-DOC [Koh06] from STEX documents. In the STEX world semantic macros have a special status, since they allow the transformation of TEX/LATEX formulae into a content-oriented markup format like OPENMATH [Bus+04] and (strict) content MATHML [Aus+10]; see Figure 1 for an example, where the semantic macros above have been defined by the \symdef macros (see Section 2.2) in the scope of a \begin{module} [id=calculus] (see Section 2.5).

IATEX	\SmoothFunctionsOn\Reals
PDF/DVI	$\mathcal{C}^{\infty}(\mathbb{R},\mathbb{R})$
OPENMATH	% <oma> % <oms cd="calculus" name="SmoothFunctionsOn"></oms> % <oms cd="calculus" name="Reals"></oms> % </oma>
MATHML	% <apply> % <csymbol cd="calculus">SmoothFunctionsOn</csymbol> % <csymbol cd="calculus">Reals</csymbol> % </apply>

Example 1: OpenMath and MathML generated from Semantic Macros

2 The User Interface

The main contributions of the modules package are the module environment, which allows for lexical scoping of semantic macros with inheritance and the \symdef macro for declaration of semantic macros that underly the module scoping.

2.1 Package Options

showmods

EdN:1

The modules package takes two options: If we set showmods¹, then the views (see

 $^{^{1}}$ EdNote: This mechanism does not work yet, since we cannot disable it when importing modules and that leads to unwanted boxes. What we need to do instead is to tweak the sms utility to use an

qualifiedimports

Section 2.9) are shown. If we set the qualified imports option, then qualified imports are enabled. Qualified imports give more flexibility in module inheritance, but consume more internal memory. As qualified imports are not fully implemented at the moment, they are turned off by default see Limitation 3.4. The option noauxreq prohibits the registration of \@requiremodules commands in the aux file. They are necessary for preloading the module signatures so that entries in the table of contents can have semantic macros; but as they sometimes

showmeta

noauxrea

If the showmeta is set, then the metadata keys are shown (see [Koh13a] for details and customization options).

2.2 Semantic Macros

\symdef

The is the main constructor for semantic macros in STEX. A call to the \symdef macro has the general form

```
\symdef[\langle keys \rangle] \{\langle cseq \rangle\} [\langle args \rangle] \{\langle definiens \rangle\}
```

cause trouble the option allows to turn off preloading.

where $\langle cseq \rangle$ is a control sequence (the name of the semantic macro) $\langle args \rangle$ is a number between 0 and 9 for the number of arguments $\langle definiens \rangle$ is the token sequence used in macro expansion for $\langle cseq \rangle$. Finally $\langle keys \rangle$ is a keyword list that further specifies the semantic status of the defined macro.

The two semantic macros in Figure 1 would have been declared by invocations of the \symdef macro of the form:

Note that both semantic macros correspond to OPENMATH or MATHML "symbols", i.e. named representations of mathematical concepts (the real numbers and the constructor for the space of smooth functions over a set); we call these names the **symbol name** of a semantic macro. Normally, the symbol name of a semantic macro declared by a **\symdef** directive is just $\langle cseq \rangle$. The key-value pair $name=\langle symname \rangle$ can be used to override this behavior and specify a differing name. There are two main use cases for this.

name

The first one is shown in Example 3, where we define semantic macros for the "exclusive or" operator. Note that we define two semantic macros: \xorOp and \xor for the applied form and the operator. As both relate to the same mathematical concept, their symbol names should be the same, so we specify name=xor on the definition of \xorOp.

local

A key local can be added to $\langle keys \rangle$ to specify that the symbol is local to the module and is invisible outside. Note that even though \symdef has no advantage over \def for defining local semantic macros, it is still considered good style to use \symdef and \abbrdef, if only to make switching between local and exported semantic macros easier.

\abbrdef

The \abbrdef macro is a variant of \symdef that is only different in semantics,

internal version that never shows anything during sms reading.

not in presentation. An abbreviative macro is like a semantic macro, and underlies the same scoping and inheritance rules, but it is just an abbreviation that is meant to be expanded, it does not stand for an atomic mathematical object.

We will use a simple module for natural number arithmetics as a running example. It defines exponentiation and summation as new concepts while drawing on the basic operations like + and - from LaTeX. In our example, we will define a semantic macro for summation \Sumfromto, which will allow us to express an expression like $\sum i = 1^n x^i$ as \Sumfromto{i}1n{2i-1} (see Example 2 for an example). In this example we have also made use of a local semantic symbol for n, which is treated as an arbitrary (but fixed) symbol.

Example 2: Semantic Markup in a module Context

\symvariant

The \symvariant macro can be used to define presentation variants for semantic macros previously defined via the \symdef directive. In an invocation

```
\symdef [\langle keys \rangle] \{\langle cseq \rangle\} [\langle args \rangle] \{\langle pres \rangle\} \\ symvariant \{\langle cseq \rangle\} [\langle args \rangle] \{\langle var \rangle\} \{\langle varpres \rangle\} \\
```

the first line defines the semantic macro $\langle cseq \rangle$ that when applied to $\langle args \rangle$ arguments is presented as $\langle pres \rangle$. The second line allows the semantic macro to be called with an optional argument $\langle var \rangle$: $\langle cseq \rangle$ [var] (applied to $\langle args \rangle$ arguments) is then presented as $\langle varpres \rangle$. We can define a variant presentation for $\langle var \rangle$; see Figure 3 for an example.

```
\label{eq:lid=xbool} $$ \operatorname{Index}_{xor0p}_{\operatorname{lude}} = \operatorname{Index}_{xor0p}_{\operatorname{lude}} $$ \operatorname{Index}_{xor0p}_{\operatorname{lude}} = \operatorname{Index}_{\operatorname{lude}} $$ \operatorname{Index}_{\operatorname{lude}} = \operatorname{Index}_{\operatorname{lude}} $$ \operatorname{Index}_{\operatorname{lude}}
```

Example 3: Presentation Variants of a Semantic Macro

\resymdef

Version 1.0 of the modules package had the \resymdef macro that allowed

to locally redefine the presentation of a macro. But this did not interact well with the beamer package and was less useful than the \symvariant functionality. Therefore it is deprecated now and leads to an according error message.

2.3 Testing Semantic Macros

\symtest

One of the problems in managing large module graphs with many semantic macros, so the module package gives an infrastructure for unit testing. The first macro is \symtest, which allows the author of a semantic macro to generate test output (if the symtest option is set) see figure 4 for a "tested semantic macro definition". Note that the language in this purely generated, so that it can be adapted (tbd).

Example 4: A Semantic Macro Definition with Test

Just as the \symdef declarations define semantic macros for mathematical sym-

\abbrtest

The \abbrtest macro gives the analogous functionality for \abbrdef.

2.4 Symbol and Concept Names

bols, the modules package provides an infrastructure for mathematical concepts that are expressed in mathematical vernacular. The key observation here is that concept names like "finite symplectic group" follow the same scoping rules as mathematical symbols, i.e. they are module-scoped. The \termdef macro is an analogue to \symdef that supports this: use \termdef[$\langle keys \rangle$] { $\langle cseq \rangle$ } { $\langle cseq \rangle$ } to declare the macro $\langle cseq \rangle$ that expands to $\langle concept \rangle$. See Figure 5 for an example, where we use the \capitalize macro to adapt $\langle concept \rangle$ to the sentence beginning.². The main use of the \termdef-defined concepts lies in automatic cross-referencing facilities via the \termdef and \symmet macros provided by the statements package [Koh13d]. Together with the hyperref package [RO], this provide cross-referencing to the definitions of the symbols and concepts. As discussed in section 3.6, the \symdef and \termdef declarations must be on top-level

in a module, so the infrastructure provided in the modules package alone cannot be used to locate the definitions, so we use the infrastructure for mathematical

\termdef

\capitalize

EdN:2

\termref \symref

statements for that.

\termdef[name=xor]{xdisjunction}{exclusive disjunction}
\capitalize\xdisjunction is commutative: \$\xor{p}q=\xor{q}p\$

Example 5: Extending Example 3 with Term References

 $^{^2\}mathrm{EdNote}$: continue, describe $\langle keys \rangle$, they will have to to with plurals,...once implemented

2.5 Modules and Inheritance

module

Themodule environment takes an optional KeyVal argument. Currently, only the id key is supported for specifying the identifier of a module (also called the module name). A module introduced by \begin{module} [id=foo] restricts the scope the semantic macros defined by the \symdef form to the end of this module given by the corresponding \end{module}, and to any other module environments that import them by a \importmodule{foo} directive. If the module foo contains \importmodule directives of its own, these are also exported to the importing module.

\importmodule

Thus the \importmodule declarations induce the semantic inheritance relation. Figure 7 shows a module that imports the semantic macros from three others. In the simplest form, $\ightharpoonup \text{importmodule}\{\langle mod \rangle\}$ will activate the semantic macros and concepts declared by \symdef and \termdef in module $\langle mod \rangle$ in the current module¹. To understand the mechanics of this, we need to understand a bit of the internals. The module environment sets up an internal macro pool, to which all the macros defined by the \symdef and \termdef declarations are added; \importmodule only activates this macro pool. Therefore \importmodule $\{\langle mod \rangle\}$ can only work, if the TEX parser — which linearly goes through the STEX sources — already came across the module $\langle mod \rangle$. In many situations, this is not obtainable; e.g. for "semantic forward references", where symbols or concepts are previewed or motivated to knowledgeable readers before they are formally introduced or for modularizations of documents into multiple files. To enable situations like these, the module package uses auxiliary files called STEX module signatures. For any file, $\langle file \rangle$.tex, we generate a corresponding STFX module signature $\langle file \rangle$. sms with the sms utility (see also Limitation 3.1), which contains (copies of) all \begin/\end{module}, \importmodule, \symdef, and \termdef invocations in $\langle file \rangle$.tex. The value of an STFX module signature is that it can be loaded instead its corresponding STFX document, if we are only interested in the semantic macros. So \importmodule[load= $\langle filepath \rangle$] { $\langle mod \rangle$ } will load the STEX module signature (filepath).sms (if it exists and has not been loaded before) and activate the semantic macros from module $\langle mod \rangle$ (which was supposedly defined in $\langle filepath \rangle$.tex). Note that since $\langle filepath \rangle$.sms contains all \importmodule statements that $\langle filepath \rangle$.tex does, an \importmodule recursively loads all necessary files to supply the semantic macros inherited by the current module.³

importmodulevia

EdN:3

The \importmodule macro has a variant \importmodulevia that allows the specification of a theory morphism to be applied. $\mbox{\sc importmodulevia} \{\langle thyid \rangle\} \{\langle assignments \rangle\}$ specifies the "source theory" via its identifier $\langle thyid \rangle$ and the morphism by $\langle assignments \rangle$. There are three kinds:

\vassign

symbol assignments via $\sign{\langle sym \rangle} {\langle exp \rangle}$, which defines the symbol $\langle sym \rangle$ introduced in the current theory by an expression $\langle exp \rangle$ in the source theory.

\tassign

term assignments via \tassign[$\langle source-cd \rangle$] { $\langle tname \rangle$ }, which

¹Actually, in the current T_FX group, therefore \importmodule should be placed directly after the \begin{module}.

 $^{^3{}m EdNote}$: MK: document the other keys of module

assigns to the term with name $\langle tname \rangle$ in the current theory a term with name $\langle source\text{-}tname \rangle$ in the theory $\langle source\text{-}cd \rangle$ whose default value is the source theory.

\ttassign

term text assignments via $\t sign {\langle tname \rangle} {\langle text \rangle}$, which defines a term with name $\langle tname \rangle$ in the current theory via a definitional text.

```
\begin{module}[id=ring]
\begin{importmodulevia}{monoid}
 \vassign{rbase}\magbase
  \vassign{rtimesOp}\magmaop
  \vassign{rone}\monunit
\end{importmodulevia}
\symdef{rbase}{G}
\symdef[name=rtimes]{rtimesOp}{\cdot}
\symdef{rtimes}[2]{\infix\rtimesOp{#1}{#2}}
\symdef{rone}{1}
\begin{importmodulevia}{cgroup}
 \vassign{rplus}\magmaop
 \vassign{rzero}\monunit
 \vassign{rinv0p}\cginv0p
\end{importmodulevia}
\symdef[name=rplus]{rplusOp}{+}
\symdef[name=rminus]{rminusOp}{-}
\symdef{rminus}[1]{\infix\rminusOp{#1}{#2}}
\end{module}
```

Example 6: A Module for Rings with inheritance from monoids and commutative groups

\metalanguage

The \metalanguage macro is a variant of importmodule that imports the meta language, i.e. the language in which the meaning of the new symbols is expressed. For mathematics this is often first-order logic with some set theory; see [RabKoh:WSMSML10] for discussion.

2.6 Dealing with multiple Files

The infrastructure presented above works well if we are dealing with small files or small collections of modules. In reality, collections of modules tend to grow, get reused, etc, making it much more difficult to keep everything in one file. This general trend towards increasing entropy is aggravated by the fact that modules are very self-contained objects that are ideal for re-used. Therefore in the absence of a content management system for LATEX document (fragments), module collections tend to develop towards the "one module one file" rule, which leads to situations with lots and lots of little files.

Moreover, most mathematical documents are not self-contained, i.e. they do not build up the theory from scratch, but pre-suppose the knowledge (and notation) from other documents. In this case we want to make use of the semantic macros from these prerequisite documents without including their text into the current document. One way to do this would be to have LATEX read the prerequisite documents without producing output. For efficiency reasons, STEX chooses a different route. It comes with a utility sms (see Section 3.1) that exports the modules and macros defined inside them from a particular document and stores them inside .sms files. This way we can avoid overloading LaTeX with useless information, while retaining the important information which can then be imported in a more efficient way.

\importmodule

For such situations, the \importmodule macro can be given an optional first argument that is a path to a file that contains a path to the module file, whose module definition (the .sms file) is read. Note that the \importmodule macro can be used to make module files truly self-contained. To arrive at a file-based content management system, it is good practice to reuse the module identifiers as module names and to prefix module files with corresponding \importmodule statements that pre-load the corresponding module files.

```
\begin{module}[id=foo]
\importmodule[load=../other/bar]{bar}
\importmodule[load=../mycolleaguesmodules]{baz}
\importmodule[load=../other/bar]{foobar}
...
\end{module}
```

Example 7: Self-contained Modules via importmodule

In Example 7, we have shown the typical setup of a module file. The \importmodule macro takes great care that files are only read once, as STEX allows multiple inheritance and this setup would lead to an exponential (in the module inheritance depth) number of file loads.

Sometimes we want to import an existing OMDoc theory² $\widehat{\mathcal{T}}$ into (the OMDoc document $\widehat{\mathcal{D}}$ generated from) a STEX document \mathcal{D} . Naturally, we have to provide an STEX stub module \mathcal{T} that provides \symdef declarations for all symbols we use in \mathcal{D} . In this situation, we use\importOMDocmodule[$\langle spath \rangle$] { $\langle OURI \rangle$ } { $\langle name \rangle$ }, where $\langle spath \rangle$ is the file system path to \mathcal{T} (as in \importmodule, this argument must not contain the file extension), $\langle OURI \rangle$ is the URI to the OMDoc module (this time with extension), and $\langle name \rangle$ is the name of the theory $\widehat{\mathcal{T}}$ and the module in \mathcal{T} (they have to be identical for this to work). Note that since the $\langle spath \rangle$ argument is optional, we can make "local imports", where the stub \mathcal{T} is in \mathcal{D} and only contains the \symdefs needed there.

Note that the recursive (depth-first) nature of the file loads induced by this setup is very natural, but can lead to problems with the depth of the file stack in

\importOMDocmodule

²OMDoc theories are the counterpart of ST_FX modules.

\requiremodules

the TeX formatter (it is usually set to something like 15³). Therefore, it may be necessary to circumvent the recursive load pattern providing (logically spurious) \importmodule commands. Consider for instance module bar in Example 7, say that bar already has load depth 15, then we cannot naively import it in this way. If module bar depended say on a module base on the critical load path, then we could add a statement \requiremodules{../base} in the second line. This would load the modules from ../base.sms in advance (uncritical, since it has load depth 10) without activating them, so that it would not have to be re-loaded in the critical path of the module foo. Solving the load depth problem.

\sinput

In all of the above, we do not want to load an sms file, if the corresponding file has already been loaded, since the semantic macros are already in memory. Therefore the modules package supplies a semantic variant of the \input macro, which records in an internal register that the modules in the file have already been loaded. Thus if we consistently use \sinput instead of \input or \include for files that contain modules⁴, we can prevent double loading of files and therefore gain efficiency. The \sinputref macro behaves just like \sinput in the LATEXML workflow, but in the LATEXML conversion process creates a reference to the transformed version of the input file instead. \inputref and \sincluderef also only create references in the LATEXML workflow but do not register loaded files or avoid duplicate loading.

\sinputref

\inputref \includeref

Finally, the separation of documents into multiple modules often profits from a symbolic management of file paths. To simplify this, the modules package supplies the \defpath macro: $\defpath[\langle baseURI \rangle] \{\langle cname \rangle\} \{\langle path \rangle\}$ defines a command, so that $\langle csname \rangle \{\langle name \rangle\}$ expands to $\langle path \rangle / \langle name \rangle$. So we could have used

 \defpath

\defpath{OPaths}{../other}
\importmodule[load=\OPahts{bar}]{bar}

instead of the second line in Example 7. The variant \OPaths has the big advantage that we can get around the fact that T_EX/E^TEX does not set the current directory in \input, so that we can use systematically deployed \defpath-defined path macros to make modules relocatable by defining the path macros locally. The optional parameter $\langle baseURI \rangle$ is for the LATEXML transformation, which (if $\langle baseURI \rangle$ is specified) resolves $\langle path \rangle$ to an absolute URI according to [BFM05, section 5.2].

2.7 Using Semantic Macros in Narrative Structures

The \importmodule macro establishes the inheritance relation, a transitive relation among modules that governs visibility of semantic macros. In particular, it can only be used in modules (and has to be used at the top-level, otherwise it

³If you have sufficient rights to change your TEX installation, you can also increase the variable max_in_open in the relevant texmf.cnf file. Setting it to 50 usually suffices

⁴files without modules should be treated by the regular IATEX input mechanism, since they do not need to be registered.

\usemodule

is hindered by LATEX groups). In many cases, we only want to use the semantic macros in an environment (and not re-export them). Indeed, this is the normal situation for most parts of mathematical documents. For that STEX provides the \usepace \usepace \usepace \usepace \usepace \usepace \usepace \usepace \usepace \underset \underset

```
\begin{omtext}
  \usemodule[../algebra/rings.tex]{ring}
  We $R$ be a ring $(\rbase,\rplus,\rzero,\rminusOp,\rtimes,\rone)$, ...
\end{omtext}
```

Example 8: Using Semantic Macros in Narrative Structures

\adoptmodule

EdN:4

EdN:5

Still another import-like relation is the adoption of a module (see [Koh13c] for details). We use the \adoptmodule macro for that.⁴

2.8 Including Externally Defined Semantic Macros

In some cases, we use an existing LaTeX macro package for typesetting objects that have a conventionalized mathematical meaning. In this case, the macros are "semantic" even though they have not been defined by a \symdef. This is no problem, if we are only interested in the LaTeX workflow. But if we want to e.g. transform them to OMDoc via LaTeXML, the LaTeXML bindings will need to contain references to an OMDoc theory that semantically corresponds to the LaTeX package. In particular, this theory will have to be imported in the generated OMDoc file to make it OMDoc-valid.

\requirepackage

To deal with this situation, the modules package provides the \requirepackage macro. It takes two arguments: a package name, and a URI of the corresponding OMDOC theory. In the LATEX workflow this macro behaves like a \usepackage on the first argument, except that it can — and should — be used outside the LATEX preamble. In the LATEXML workflow, this loads the LATEXML bindings of the package specified in the first argument and generates an appropriate imports element using the URI in the second argument.

2.9 Views

A view is a mapping between modules, such that all model assumptions (axioms) of the source module are satisfied in the target module. ⁵

 $^{^4\}mathrm{EdNote}$: elaborate this, so that the documentation becomes (more) self-contained.

 $^{^5\}mathrm{EdNote}\colon$ Document and make Examples

2.10 Support for MathHub

Much of the STEX content is hosted on MathHub (http://MathHub.info), a portal and archive for flexiformal mathematics. MathHub offers GIT repositories (public and private escrow) for mathematical documentation projects, online and offline authoring and document development infrastructure, and a rich, interactive reading interface. The modules package supports repository-sensitive operations on MathHub.

Note that MathHub has two-level repository names of the form $\langle group \rangle / \langle repo \rangle$, where $\langle group \rangle$ is a MathHub-unique repository group and $\langle repo \rangle$ a repository name that is $\langle group \rangle$ -unique. The file and directory structure of a repository is arbitrary – except that it starts with the directory source because they are Math Archives in the sense of [Hor+11]. But this structure can be hidden from the STEX author with MathHub-enabled versions of the modules macros.

\importmhmodule

The importmhmodule macro is a variant of \importmodule with repository support. Instead of writing

```
\defpath{MathHub}{/user/foo/lmh/MathHub}
\importmodule[load=\MathHub{fooMH/bar/source/baz/foobar}]{foobar}
```

we can simply write (assuming that \MathHub is defined as above)

\importmhmodule[fooMH/bar]{baz/foobar}{foobar}

Note that the **\importmhmodule** form is more semantic, which allows more advanced document management features in MathHub.

If baz/foobar is the "current module", i.e. if we are on the MathHub path ...MathHub/fooMH/bar..., then stating the repository in the first optional argument is redundant, so we can just use

\importmhmodule{baz/foobar}{foobar}

\mhcurrentrepos

Of course, neither LATEX nor LATEXMLknow about the repositories when they are called from a file system, so we can use the \mhcurrentrepos macro to tell them. But this is only needed to initialize the infrastructure in the driver file. In particular, we do not need to set it in in each module, since the \importmhmodule macro sets the current repository automatically.

\usemhmodule \adoptmhmodule

The $\$ and $\$ adoptmhmodule macros are the analogs to $\$ and $\$ adoptmodule.

Caveat if you want to use the MathHub support macros (let's call them mhvariants), then every time a module is imported or a document fragment is included from another repos, the mh-variant \importmhmodule must be used, so that the "current repository" is set accordingly. To be exact, we only need to use mhvariants, if the imported module or included document fragment use mh-variants.

\mhinputref

For this, the modules package supplies the mh-variants \mhinputref and

\mhinput

\mhinput of the \inputref macro introduced above and normal LATEX \input macro.

3 Limitations & Extensions

In this section we will discuss limitations and possible extensions of the modules package. Any contributions and extension ideas are welcome; please discuss ideas, requests, fixes, etc on the STEX TRAC [sTeX].

3.1 Perl Utility sms

Currently we have to use an external perl utility sms to extract STEX module signatures from STEX files. This considerably adds to the complexity of the STEX installation and workflow. If we can solve security setting problems that allows us to write to STEX module signatures outside the current directory, writing them from STEX may be an avenue of future development see [sTeX, issue #1522] for a discussion.

3.2 Module Signatures loaded even if Modules are

Currently, the module signature $\langle filepath \rangle$. sms is loaded even if $\langle filepath \rangle$. tex has already been loaded. The Problem is that the \input \inputref macros do not register the files. I guess, for \input we may not want that, but for \inputref we should; and maybe we should also have a registering variant \rinput that does, then we can leave the choice to the user.

3.3 Only one level of path simplification

It seems that module simplification only covers one level, in particular foo/bar/../.. is only simplified to foo/..., which is not enough in practice. As a consequence, module signatures are loaded unnecessary in the presence of relative paths.

3.4 Qualified Imports

In an earlier version of the modules package we used the usesqualified for importing macros with a disambiguating prefix (this is used whenever we have conflicting names for macros inherited from different modules). This is not accessible from the current interface. We need something like a \importqualified macro for this; see [sTeX, issue #1505]. Until this is implemented the infrastructure is turned off by default, but we have already introduced the qualifiedimports option for the future.

qualifiedimports

3.5 Error Messages

The error messages generated by the modules package are still quite bad. For instance if thy A does note exists we get the cryptic error message

```
! Undefined control sequence.
\module@defs@thyA ...hy
\expandafter \mod@newcomma...
1.490 ...ortmodule{thyA}
```

This should definitely be improved.

3.6 Crossreferencing

Note that the macros defined by \symdef are still subject to the normal TEX scoping rules. Thus they have to be at the top level of a module to be visible throughout the module as intended. As a consequence, the location of the \symdef elements cannot be used as targets for crossreferencing, which is currently supplied by the statement package [Koh13d]. A way around this limitation would be to import the current module from the STEX module signature (see Section 2.5) via the \importmodule declaration.

3.7 No Forward Imports

STEX allows imports in the same file via $\mbox{importmodule}(\mbox{mod})$, but due to the single-pass linear processing model of TeX, \mbox{mod} must be the name of a module declared before the current point. So we cannot have forward imports as in

```
\begin{module}[id=foo]
  \importmodule{mod}
    ...
\end{module}
    ...
\begin{module}[id=mod]
    ...
\end{module}
```

a workaround, we can extract the module $\langle mod \rangle$ into a file mod.tex and replace it with $sinput\{mod\}$, as in

```
\begin{module}[id=foo]
  \importmodule[load=mod]{mod}
    ...
\end{module}
    ...
\sinput{mod}
```

then the \importmodule command can read mod.sms (created via the sms utility) without having to wait for the module $\langle mod \rangle$ to be defined.

4 The Implementation

The modules package generates two files: the LATEX package (all the code between <code>*package</code>) and <code>\/package</code>) and the LATEXML bindings (between <code>*ltxml</code>) and <code>\/ltxml</code>). We keep the corresponding code fragments together, since the documentation applies to both of them and to prevent them from getting out of sync.

4.1 Package Options

We declare some switches which will modify the behavior according to the package options. Generally, an option xxx will just set the appropriate switches to true (otherwise they stay false).

- 1 (*package)
- 2 \DeclareOption{showmeta}{\PassOptionsToPackage{\CurrentOption}{metakeys}}
- 3 \newif\ifmod@show\mod@showfalse
- 4 \DeclareOption{showmods}{\mod@showtrue}
- 5 \newif\ifaux@req\aux@reqtrue
- 6 \DeclareOption{noauxreq}{\aux@reqfalse}
- 7 \newif\ifmod@qualified\mod@qualifiedfalse
- 8 \DeclareOption{qualifiedimports}{\mod@qualifiedtrue}

Finally, we need to declare the end of the option declaration section to LATEX.

9 \ProcessOptions

LATEXML does not support module options yet, so we do not have to do anything here for the LATEXML bindings. We only set up the PERL packages (and tell emacs about the appropriate mode for convenience

The next measure is to ensure that the sref and xcomment packages are loaded (in the right version). For LATEXML, we also initialize the package inclusions.

- 10 \RequirePackage{sref}
- 11 \RequirePackage{xspace}
- 12 \RequirePackage{mdframed}

4.2 Modules and Inheritance

We define the keys for the module environment and the actions that are undertaken, when the keys are encountered.

module:cd This KeyVal key is only needed for LATEXML at the moment; use this to specify a content dictionary name that is different from the module name.

- 13 \addmetakey{module}{cd}% no longer used
- 14 \addmetakey{module}{load}% ignored
- 15 \addmetakey*{module}{title}
- 16 \addmetakey*{module}{creators}
- 17 \addmetakey*{module}{contributors}

module:id For a module with $[id=\langle name \rangle]$, we have a macro \module@defs@ $\langle name \rangle$ that acts as a repository for semantic macros of the current module. I will be called by \importmodule to activate them. We will add the internal forms of the semantic

macros whenever \symdef is invoked. To do this, we will need an unexpended form \this@module that expands to \module@defs@ $\langle name \rangle$; we define it first and then initialize \module@defs@ $\langle name \rangle$ as empty. Then we do the same for qualified imports as well (if the qualifiedimports option was specified). Furthermore, we save the module name in \mod@id and the module path in \ $\langle name \rangle$ @cd@file@base which we add to \module@defs@ $\langle name \rangle$, so that we can use it in the importing module

```
18 \define@key{module}{id}{%
```

- 19 % \@ifundefined{module@defs@#1}{}{\PackageError{modules}%
- 20 % {The module name #1 has already been used in this scope; please rename it!}}%
- 21 \edef\this@module{\expandafter\noexpand\csname module@defs@#1\endcsname}%
- 22 \global\@namedef{module@defs@#1}{}%
- 23 \ifmod@qualified
- 24 \edef\this@qualified@module{\expandafter\noexpand\csname module@defs@qualified@#1\endcsname}%
- 25 \global\@namedef{module@defs@qualified@#1}{}%
- 26 \fi
- 27 \def\mod@id{#1}%
- 28 \expandafter\edef\csname #1@cd@file@base\endcsname{\mod@path}%
- 29 \expandafter\g@addto@macro\csname module@defs@#1\expandafter\endcsname\expandafter%
- 30 {\expandafter\def\csname #1@cd@file@base\expandafter\endcsname\expandafter{\mod@path}}}

module@heading

Then we make a convenience macro for the module heading. This can be customized.

- 31 \newcounter{module} [section]
- 32 \newcommand\module@heading{\stepcounter{module}%
- 33 \ifmod@show%
- $34 \neq \{\text{Module} \rightarrow \text{Module} \}$
- 35 \sref@label@id{Module \thesection.\themodule [\mod@id]}%
- 36 \ifx\module@title\@empty :\quad\else\quad(\module@title)\hfill\\\fi%
- 37 \fi}%mod@show

module Finally, we define the begin module command for the module environment. All the work has already been done in the keyval bindings, so this is very simple.

- 39 \ifmod@show\surroundwithmdframed{module}\fi

@module A variant of the **module** environment that does not create printed representations (in particular no frames)

 $40 \newenvironment{@module}[1][]{\netasetkeys{module}{#1}}{}$

\activate@defs

To activate the \symdefs from a given module $\langle mod \rangle$, we call the macro \module@defs@ $\langle mod \rangle$. But to make sure that every module is activated only once, we only activate if the macro \module@defs@ $\langle mod \rangle$ is undefined, and define it directly afterwards to prohibit further activations.

- 41 \def\activate@defs#1{%
- $42 \end{fined} $$ 42 \end{fined} $$ activated $$ {\csname module@defs@#1\endcsname} $$ $$ $$ $$$
- 43 \Onamedef{moduleO#1Oactivated}{true}}

\export@defs

EdN:6

\export@defs{ $\langle mod \rangle$ } exports all the \symdefs from module $\langle mod \rangle$ to the current module (if it has the name $\langle currmod \rangle$), by adding a call to \module@defs@ $\langle mod \rangle$ to the registry \module@defs@ $\langle currmod \rangle$.

- 44 \def\export@defs#1{\@ifundefined{mod@id}{}%
- 45 {\expandafter\expandafter\expandafter\g@addto@macro\expandafter%
- 46 \this@module\expandafter{\activate@defs{#1}}}}

\importmodule

The $\mbox{importmodule}[\langle file\rangle] {\langle mod\rangle}$ macro is an interface macro that loads $\langle file\rangle$ and activates and re-exports the $\mbox{symdefs}$ from module $\langle mod\rangle$. It also remembers the file name in $\mbox{mod@path}$. As we will need to keep a record of the currently imported modules (top-level only), we divide the functionality into a user-visible macro that records modules in the $\mbox{imported@modules}$ register and an internal one ($\mbox{@importmodule}$) that does the actual work.

- 47 \gdef\imported@modules{}
- 48 \srefaddidkey{importmodule}
- 49 \addmetakey{importmodule}{load}
- 50 \addmetakey[sms]{importmodule}{ext}
- 51 \newif\if@export
- 52 \newcommand\importmodule[2][]{\metasetkeys{importmodule}{#1}%
- 53 \ifx\imported@modules\@empty\edef\imported@modules{#2}%
- $54 \else\edgnodules \fi%$
- 55 \@exporttrue\@importmodule[\importmodule@load]{#2}{\importmodule@ext}\ignorespaces}

\@importmodule

 $\ensuremath{\mbox{\sc Nod}} = [\langle file \rangle] = [\langle mod \rangle]$

- 56 \newcommand\@importmodule[3][]{%
- $57 {\edf(@load{#1}\ifx(@load(@empty)else{\mod@showfalse\requiremodules(@load{#3})fi})} \\$
- 58 \activate@defs{#2}\if@export\export@defs{#2}\fi}

\usemodule

\usemodule acts like \importmodule for the LATEX side, except that the sms utility does not transfer it to the module signatures and it does not export the symdefs.

59 \newcommand\usemodule[2][]{\@exportfalse\importmodule[#1]{#2}}

\importmodules

This variant just imports all the modules in a comma-separated list (usually \imported@modules)

60 \newcommand\importmodules[1]{\@for\@I:=#1\do{\importmodule\@I}}%

\importmodulevia

The importmodulevia environment just calls \importmodule, but to get around the group, we first define a local macro \@doit, which does that and can be

 $^{^6\}mathrm{Ed}\mathrm{NoTE}\colon$ MK: I have the feeling that we may be exporting modules multiple times here, is that a problem?

called with an \aftergroup to escape the environment grouping introduced by importmodulevia.

- 61 \newenvironment{importmodulevia}[2][]{\gdef\@doit{\importmodule[#1]{#2}}%
- 62 \ifmod@show\par\noindent importing module #2 via \@@doit\fi}
- 63 {\aftergroup\@@doit\ifmod@show end import\fi}

vassign

64 \newcommand\vassign[3][]{\ifmod@show\ensuremath{#2\mapsto #3}, \fi}

tassign

65 \newcommand\tassign[3][]{\ifmod@show #2\ensuremath{\mapsto} #3, \fi}

ttassign

66 \newcommand\ttassign[3][]{\ifmod@show #2\ensuremath{\mapsto} ''#3'', \fi}

\importOMDocmodule

EdN:7

for the LATEXML side we can just re-use \importmodule, for the LATEXML side we have a full URI anyways. So things are easy.⁷

67 \newcommand\importOMDocmodule[3][]{\importmodule[#1]{#3}}

\metalanguage

\metalanguage behaves exactly like \importmodule for formatting. For LA-TEXML, we only add the type attribute.

68 \let\metalanguage=\importmodule

\adoptmodule

\adoptmodule macro behaves exactly like \importmodule for formatting. For ${\rm LATEXML.}$

69 \let\adoptmodule=\importmodule

Semantic Macros 4.3

\mod@newcommand We first hack the LATEX kernel macros to obtain a version of the \newcommand macro that does not check for definedness. This is just a copy of the code from latex.ltx where I have removed the \@ifdefinable check.⁵

- 70 \let\mod@newcommand=\newcommand
- 71 % \def\mod@newcommand{\@star@or@long\mod@new@command}
- 72 % \def\mod@new@command#1{\@testopt{\@mod@newcommand#1}0}
- 73 % \def\@mod@newcommand#1[#2]{\kernel@ifnextchar [{\mod@xargdef#1[#2]}{\mod@argdef#1[#2]}}
- 74 % \long\def\mod@argdef#1[#2]#3{\@yargdef#1\@ne{#2}{#3}}
- 75 % \long\def\mod@xargdef#1[#2][#3]#4{\expandafter\def\expandafter#1\expandafter{%
- 76 % \expandafter\@protected@testopt\expandafter #1\csname\string#1\endcsname{#3}}%
- 77 % \expandafter\@yargdef\csname\string#1\endcsname\tw@{#2}{#4}}

Now we define the optional KeyVal arguments for the \symdef form and the actions that are taken when they are encountered.

⁷EDNOTE: MK@DG: this macro is seldom used, maybe I should just switch arguments.

⁵Someone must have done this before, I would be very happy to hear about a package that provides this.

```
conceptdef
```

```
78 \srefaddidkey{conceptdef}
79 \addmetakey*{conceptdef}{title}
80 \addmetakey*{conceptdef}{subject}
81 \addmetakey*{conceptdef}{display}
82 \def\conceptdef@type{Symbol}
83 \newcommand\conceptdef[2][]{\metasetkeys{conceptdef}{#1}%
84 \ifx\conceptdef@display\st@flow\else{\stDMemph{\conceptdef@type} #2:}\fi%
85 \ifx\conceptdef@title\@empty~\else~(\stDMemph{\conceptdef@title})\par\fi}
```

EdN:8

symdef:keys

The optional argument local specifies the scope of the function to be defined. If local is not present as an optional argument then \symdef assumes the scope of the function is global and it will include it in the pool of macros of the current module. Otherwise, if local is present then the function will be defined only locally and it will not be added to the current module (i.e. we cannot inherit a local function). Note, the optional key local does not need a value: we write \symdef[local]{somefunction}[0]{some expansion}. The other keys are not used in the LATEX part.

```
86 \newif\if@symdeflocal
87 \srefaddidkey{symdef}
88 \define@key{symdef}{local}[true]{\@symdeflocaltrue}
89 \define@key{symdef}{assocarg}{}
90 \define@key{symdef}{bvars}{}
91 \define@key{symdef}{bargs}{}
92 \addmetakey{symdef}{fname}
93 \addmetakey*{symdef}{title}
94 \addmetakey*{symdef}{description}
95 \addmetakey*{symdef}{subject}
96 \addmetakey*{symdef}{display}
```

EdN:9

\symdef The the \symdef, and \@symdef macros just handle optional arguments.

```
97 \def\symdef{\@ifnextchar[{\@symdef}{\@symdef[]}}
```

 $98 \end{figure} $98 \$

next we locally abbreviate \mod@newcommand to simplify argument passing.

99 \def\@mod@nc#1{\mod@newcommand{#1}[1]}

now comes the real meat: the \@@symdef macro does two things, it adds the macro definition to the macro definition pool of the current module and also provides it.

100 \def\@@symdef[#1]#2[#3]#4{%

We use a switch to keep track of the local optional argument. We initialize the switch to false and set all the keys that have been provided as arguments: name, local.

101 \@symdeflocalfalse\metasetkeys{symdef}{#1}%

 $^{^8{\}rm EdNote}$: MK@DG: maybe we need to add DefKeyVals here? $^9{\rm EdNote}$: MK@MK: we need to document the binder keys above.

First, using $\mbox{modQnewcommand}$ we initialize the intermediate macro $\mbox{moduleQ}(sym)$ QpresQ, the one that can be extended with \symvariant

102 \expandafter\mod@newcommand\csname modules@#2@pres@\endcsname[#3]{#4}%

and then we define the actual semantic macro. Note that this can take an optional argument, for which we provide with $\c sym \c sym \$

- 103 \expandafter\def\csname #2\endcsname%
- $104 {\tt (Csname modules@#2\endcsname} {\tt (Csname modules@#2\endcsname[])}{\tt (Csname modules@#2\endcsname[])}{\tt (Csname modules@#2\endcsname[])}{\tt (Csname modules@#2\endcsname)}{\tt (Csname modules@#2$
- 105 \expandafter\def\csname modules@#2\endcsname[##1]%
- 106 {\csname modules@#2@pres@##1\endcsname}%

Finally, we prepare the internal macro to be used in the \symmetric call.

- $107 \verb|\expandafter@mod@nc\csname| mod@symref@#2\expandafter\endcsname\\| expandafter%| mod@symref@#2\expandafter\csname\\| expandafter\csname\\| expan$
- 108 {\expandafter\mod@termref\expandafter{\mod@id}{#2}{#1}}%

We check if the switch for the local scope is set: if it is we are done, since this function has a local scope. Similarly, if we are not inside a module, which we could export from.

- 109 \if@symdeflocal\else%
- 110 \@ifundefined{mod@id}{}{%

Otherwise, we add three functions to the module's pool of defined macros using $\gomegaddto@macro$. We first add the definition of the intermediate function $\mbox{modules}@\langle sym\rangle@pres@$.

- 111 \expandafter\g@addto@macro\this@module%
- 112 {\expandafter\mod@newcommand\csname modules@#2@pres@\endcsname[#3]{#4}}%

Then we add add the definition of $\langle sym \rangle$ in terms of the function $\langle \mathfrak{C}(sym) \rangle$ to handle the optional argument.

- 113 \expandafter\g@addto@macro\this@module%
- 114 {\expandafter\def\csname#2\endcsname%
- 115 {\@ifnextchar[{\csname modules@#2\endcsname}}\%csname modules@#2\endcsname[]}}}%

We add add the definition of $\langle Q(sym) \rangle$, which calls the intermediate function.

- 116 \expandafter\g@addto@macro\this@module%
- 117 {\expandafter\def\csname modules@#2\endcsname[##1]%
- 118 {\csname modules@#2@pres@##1\endcsname}}%

We also add $\mbox{mod@symref@}(sym)$ macro to the macro pool so that the \symref macro can pick it up.

- 119 \expandafter\g@addto@macro\csname module@defs@\mod@id\expandafter\endcsname\expandafter%
- 120 {\expandafter\@mod@nc\csname mod@symref@#2\expandafter\endcsname\expandafter%
- 121 {\expandafter\mod@termref\expandafter{\mod@id}{#2}{##1}}}%

Finally, using \g@addto@macro we add the two functions to the qualified version of the module if the qualifiedimports option was set.

- 122 \ifmod@qualified%
- $123 \verb|\expandafter\g@addto@macro\this@qualified@module\%|$
- 124 {\expandafter\mod@newcommand\csname modules@#2@pres@qualified\endcsname[#3]{#4}}%
- 125 \expandafter\g@addto@macro\this@qualified@module%

```
126 {\expandafter\def\csname#2atqualified\endcsname{\csname modules@#2@pres@qualified\endcsname}}%
                         127 \fi}% mod@qualified
                         128 \fi% symdeflocal
                           So now we only need to show the data in the symdef, if the options allow.
                         130 \ifx\symdef@display\st@flow\else{\noindent\stDMemph{\symdef@type} #2:}\fi%
                         131 \ifx\symdef@title\@empty~\else~(\stDMemph{\symdef@title})\par\fi
                         132 \fi}% mod@show
                         133 \def\symdef@type{Symbol}
\symvariant \symvariant{\langle sym \rangle}[\langle args \rangle] {\langle var \rangle}{\langle cseq \rangle} just extends the internal macro
                           \mbox{modules@}(sym)\mbox{Qpres@} defined by \symdef{}(sym)]{}(args)]{}...} with a variant
                           \mbox{modules@}(sym)@pres@(var) which expands to (cseq). Recall that this is called
                           by the macro \langle sym \rangle [\langle var \rangle] induced by the \symdef.<sup>10</sup>
                         134 \def\symvariant#1{\@ifnextchar[{\@symvariant{#1}}{\@symvariant{#1}[0]}}
                         135 \def\@symvariant#1[#2]#3#4{%
                         136 \expandafter\mod@newcommand\csname modules@#1@pres@#3\endcsname[#2]{#4}%
                           and if we are in a named module, then we need to export the function
                           \mbox{modules@}(sym)\mbox{@pres@}(opt) just as we have done that in \symdef.
                         137 \@ifundefined{mod@id}{}{%
                         138 \expandafter\g@addto@macro\this@module%
                         \resymdef This is now deprecated.
                         140 \def\resymdef{\@ifnextchar[{\@resymdef}{\@resymdef[]}}
                         141 \end{figuresymdef} \begin{tabular}{l} 141 \end{figuresymdef} \begin{tabular}{l} 142 \end{figuresymdef} \begi
                         142 \def\@@resymdef[#1]#2[#3]#4{\PackageError{modules}
                                   {The \protect\resymdef macro is deprecated,\MessageBreak
                                        use the \protect\symvariant instead!}}
      \abbrdef The \abbrdef macro is a variant of \symdef that does the same on the IATEX
                         145 \let\abbrdef\symdef
                                        Defining Math Operators
 \DefMathOp
                         146 \define@key{DefMathOp}{name}{\def\defmathop@name{#1}}
                         147 \newcommand\DefMathOp[2][]{%
                         148 \setkeys{DefMathOp}{#1}%
                         149 \symdef [#1] {\defmathop@name} {#2}}
                              ^{10}\mathrm{EdNote}: MK@DG: this needs to be implemented in LaTeXML
                              <sup>11</sup>EdNote: MK@MK,DG: DefMathOp needs to be documented above, what can we do with it?
```

EdN:10

EdN:11

²¹

4.5 Testing Semantic Macros

152 \newcommand\symtest[3][]{\metasetkeys{symtest}{#1}%

150 \addmetakey{symtest}{name}
151 \addmetakey{symtest}{variant}

\symtest

```
153 \par\noindent \textbf{Symbol}~%
                                                         154 \ifx\symtest@name\@empty\texttt{#2}\else\texttt{\symtest@name}\fi%
                                                         155 \ifx\symtest@variant\@empty\else\ (variant \texttt{\symtest@variant})\fi%
                                                         156 \ with semantic macro \texttt{\textbackslash #2}%
                                                         157 : used e.g. in \ensuremath{#3}}
             \abbrtest
                                                         158 \addmetakey{abbrtest}{name}
                                                         159 \newcommand\abbrtest[3][]{\metasetkeys{abbrtest}{#1}%
                                                          160 \par\noindent \textbf{Abbreviation}~%
                                                          161 \t \abbrest@name\empty\texttt{#2}\else\texttt{\abbrest@name}\fi% \end{minipage} \label{fi:mame} $$ 161 \t \abbrest@name\empty\texttt{#2}\else\texttt{\abbrest@name} $$ 161 \t \abbrest@name\empty\texttt{#2}\else\texttt{\abbrest@name} $$ 161 \t \abbrest@name\empty\texttt{$$ 161 \t \abbrest\texttt{$$ 161 \t \abbrest@name\empty\texttt{$$ 161 \t \abbrest\texttt{$$ 161 \t \abbrest\t \abbrest\texttt{$$ 161 \t \abbrest\texttt{$$ 161 \t \abbrest\texttt{$$ 161 \t \abbrest\texttt{$$ 161 \t \abbrest\t $$ 16
                                                         162 : used e.g. in \ensuremath{#3}}
                                                             4.6
                                                                                           Symbol and Concept Names
                                                            the \mod@path macro is used to remember the local path, so that the module
                                                             environment can set it for later cross-referencing of the modules. If \mod@path is
                                                             empty, then it signifies the local file.
                                                          163 \def\mod@path{}
                 \termdef
                                                         164 \def\mod@true{true}
                                                         165 \addmetakey[false]{termdef}{local}
                                                         166 \addmetakey{termdef}{name}
                                                         167 \newcommand\termdef[3][]{\metasetkeys{termdef}{#1}%}
                                                         168 \expandafter\mod@newcommand\csname#2\endcsname[0]{#3\xspace}%
                                                         169 \ifx\termdef@local\mod@true\else%
                                                         170 \end{fined} 
                                                         171 {\expandafter\mod@newcommand\csname#2\endcsname[0]{#3\xspace}}}%
                                                         172 \fi}
   \capitalize
                                                         173 \def\@capitalize#1{\uppercase{#1}}
                                                         174 \newcommand\capitalize[1]{\expandafter\@capitalize #1}
                                                           \mbox{\colored} \mbox{\color
\mod@termref
                                                             is defined. If it is, we make it the prefix of a URI reference in the local macro
                                                             \@uri, which we compose to the hyper-reference, otherwise we give a warning.
                                                          175 \def\mod@termref#1#2#3{\def\@test{#3}%
                                                         176 \@ifundefined{#1@cd@file@base}%
                                                         177 {\protect\G@refundefinedtrue%
                                                         178 \ClatexCwarning{\protect\termref with unidentified cd "#1": the cd key must reference an active
```

```
179 \def\@label{sref@#2 @target}}%
180 {\def\@label{sref@#2@#1@target}}%
181 \expandafter\ifx\csname #1@cd@file@base\endcsname\@empty% local reference
182 \sref@hlink@ifh{\@label}{\ifx\@test\@empty #2\else #3\fij\else%
183 \def\@uri{\csname #1@cd@file@base\endcsname.pdf\#\@label}%
184 \sref@href@ifh{\@uri}{\ifx\@test\@empty #2\else #3\fij\fij
```

4.7 Dealing with Multiple Files

Before we can come to the functionality we want to offer, we need some auxiliary functions that deal with path names.

4.7.1 Simplifying Path Names

The \mod@simplify macro is used for simplifying path names by removing $\langle xxx \rangle / \dots$ from a string. eg: $\langle aaa \rangle / \langle bbb \rangle / \dots / \langle ddd \rangle$ goes to $\langle aaa \rangle / \langle ddd \rangle$ unless $\langle bbb \rangle$ is ... This is used to normalize relative path names below.

\mod@simplify The macro \mod@simplify recursively runs over the path collecting the result in the internal \mod@savedprefix macro.

185 \def\mod@simplify#1{\expandafter\mod@simpl#1/\relax}

It is based on the \mod@simpl macro¹²

188 \ifx\@second\@empty\edef\mod@savedprefix{\mod@savedprefix#1}%

189 \else\mod@simplhelp#1/#2\relax\fi}

which in turn is based on a helper macro

```
190 \def\mod@updir{..}
```

EdN:12

 $191 \end{simple} $$191 \end{simple} $$191 \end{simple} $$1/42/$3\left(\frac{41}{def\end{#2}\end{#2}\end{#3}\% \right). $$$

192 %\message{mod@simplhelp: first=\@first, second=\@second, third=\@third, result=\mod@savedprefix

193 \ifx\@third\@empty% base case

194 \ifx\@second\mod@updir\else%

 $195 \ \texttt{\mod@savedprefix{\mod@savedprefix#1}\%} \\$

196 \else\edef\mod@savedprefix{\mod@savedprefix#1/#2}%

197 \fi% @second empty

198 \fi% @second = ..

199 \else% non-base case: @third non-empty

200 \ifx\@first\mod@updir%

 $201 \edgned@savedprefix{\mod@savedprefix#1/}\mod@simplhelp#2/#3\relax% of the content of the c$

202 \else% @first != ..

203 \ifx\@second\mod@updir\mod@simpl#3\relax%

 $204 \else\edef\mod@savedprefix{\mod@savedprefix#1/}\mod@simplhelp#2/#3\relax%.$

 $205 \fi$ % Ofirst

206 \fi% @first = ...

207 \fi}% non-base case

 $^{^{12}\}mathrm{EDNoTE}\colon$ what does the mod@blaaa do?; it seems to be a switch that checks if mod@savedprefix has been saved.

We directly test the simplification:

source	result	should be
aaa	aaa	aaa
//aaa	//aaa	//aaa
aaa/bbb	aaa/bbb	aaa/bbb
aaa/		
//aaa/bbb	//aaa/bbb	//aaa/bbb
/aaa//bbb	/bbb	/bbb
/aaa/bbb	/aaa/bbb	/aaa/bbb
aaa/bbb//ddd	aaa/ddd	aaa/ddd
aaa/bbb//	aaa/	

\defpath

 $208 \end{small} $$ 208 \rightarrow $$ \end{small} {\rm and } $$ 2\end{small} $$ 2\end{sm$

4.8 Loading Module Signatures

We will need a switch¹³

209 \newif\ifmodules

and a "registry" macro whose expansion represents the list of added macros (or files)

\mod@reg We initialize the \mod@reg macro with the empty string.

 $210 \qef\mod@reg{}\}$

\mod@update

EdN:13

This macro provides special append functionality. $\mbox{\mbox{modQupdate}} \{\langle string \rangle\}$ takes appends $\langle string \rangle$ to the expansion of the $\mbox{\mbox{modQreg}}$ macro in the following way: $\langle string \rangle \mbox{\mbox{\mbox{\mbox{modQreg}}}$.

 $211 \end{update} $$11 \end{u$

212 \else\xdef\mod@reg{#1@\mod@reg}\fi}

\mod@check

The \mod@check takes as input a file path (arg 3), and searches the registry. If the file path is not in the registry it means it means it has not been already added, so we make \ifmodules true, otherwise make \ifmodules false. The macro \mod@search will look at \ifmodules and update the registry for \modulestrue or do nothing for \modulesfalse.

213 \def\mod@check#1@#2///#3\relax{%

 $214 \end@one{#1}\def\mod@two{#2}\def\mod@three{#3}\%$

Define a few intermediate macros so that we can split the registry into separate file paths and compare to the new one

215 \expandafter%

216 \ifx\mod@three\mod@one\modulestrue%

217 \else%

 $218 \ \texttt{ifx} \ \texttt{detwo} \ \texttt{dempty} \ \texttt{modulesfalse} \ \texttt{else} \ \texttt{mod@check#2///#3} \ \texttt{ii} \ \texttt{mod@check#2///#3} \ \texttt{mod@check#$

219 \fi}

 $^{^{13}\}mathrm{EdNote}$: explain why?

220 \def\mod@search#1{% We put the registry as the first argument for \mod@check and the other argument is the new file path. 221 \modulesfalse\expandafter\mod@check\mod@reg @///#1\relax% We run \mod@check with these arguments and the check \ifmodules for the result 222 \ifmodules\else\mod@update{#1}\fi} The macro operates almost as the mod@search function, but it does not update \mod@reguse the registry. Its purpose is to check whether some file is or not inside the registry but without updating it. Will be used before deciding on a new sms file 223 \def\mod@reguse#1{\modulesfalse\expandafter\mod@check\mod@reg @///#1\relax} \mod@prefix This is a local macro for storing the path prefix, we initialize it as the empty string. $224 \def\mod@prefix{}$ This macro updates the path prefix \mod@prefix with the last word in the path \mod@updatedpre given in its argument. 225 \def\mod@updatedpre#1{% 226 \edef\mod@prefix\mod@prefix\mod@pathprefix@check#1/\relax}} \mod@pathprefix@check \mod@pathprefix@check returns the last word in a string composed of words separated by slashes 227 \def\mod@pathprefix@check#1/#2\relax{% $228 \left(\frac{42}{\%} \right)$ no slash in string 229 \else\mod@ReturnAfterFi{#1/\mod@pathprefix@help#2\relax}% 230 \fi} It needs two helper macros: 231 \def\mod@pathprefix@help#1/#2\relax{% $232 \text{ ifx}\$ end of recursion 233 \else\mod@ReturnAfterFi{#1/\mod@pathprefix@help#2\relax}% 235 \long\def\mod@ReturnAfterFi#1\fi{\fi#1} \mod@pathpostfix@check \mod@pathpostfix@check takes a string composed of words separated by slashes and returns the part of the string until the last slash 236 \def\mod@pathpostfix@check#1/#2\relax{% slash $237 \left(\frac{2}{\infty}\right)$ 238 #1\else\mod@ReturnAfterFi{\mod@pathpostfix@help#2\relax}% 239 \fi} Helper function for the \pathpostfix@check macro defined above 240 \def\mod@pathpostfix@help#1/#2\relax{% 241 \ifx\\#2\\% 242 #1\else\mod@ReturnAfterFi{\mod@pathpostfix@help#2\relax}% 243 \fi}

\mod@search Macro for updating the registry after the execution of \mod@check

\mod@updatedpost

This macro updates \mod@savedprefix with leading path (all but the last word) in the path given in its argument.

244 \def\mod@updatedpost#1{%

245 \edef\mod@savedprefix\mod@savedprefix\mod@pathpostfix@check#1/\relax}}

4.8.1 Selective Inclusion

\requiremodules

this macro loads the modules in a file. It does quite a lot of work to make sure that they are only loaded once (actually this does not quite work so they are loaded only twice). \requiremodules $\{\langle file \rangle\}$ also deposits a call to \@requiremodules $\{\langle file \rangle\}$ to the aux file to make sure that the semantic macros are already loaded for the table of contents.

```
246 \newcommand\requiremodules[2]{%
```

- 247 {\mod@showfalse% save state and ensure silence while reading sms
- $248 \mbox{mod@updatedpre{#1}}\%$ add the new file to the already existing path
- 249 \let\mod@savedprefix\mod@prefix% add the path of the new file to the prefix
- 250 \mod@updatedpost{#1}%
- 251 \def\mod@blaaaa{}}% macro used in the simplify function (remove .. from the prefix)
- 252 \mod@simplify{\mod@savedprefix}% remove |xxx/..| from the path (in case it exists)
- 253 \mod@reguse{\mod@savedprefix}%
- 254 \ifmodules\else%
- 255 \let\newreg\mod@reg% use to compare, in case the .sms file was loaded before
- 256 \mod@search{\mod@savedprefix}% update registry
- 257 \ifx\newreg\mod@reg\else\input{\mod@savedprefix.#2}% load if not in registry
- 258 \ifaux@req\protected@write\@auxout{}{\string\@requiremodules{\mod@savedprefix}{#2}}\fi%
- 259 \fi% ifx
- $260 \fi}$ % ifmodules

\@requiremodules the internal macro that actually does the loading. We disable it at the end of the document, so that when the aux file is read again, nothing is loaded.

261 \newcommand\@requiremodules[2]{\if@tempswa\mod@showfalse\input{#1.#2}\fi}

262 {\renewcommand\@requiremodules[2]{}}

\sinput

- 263 \def\sinput#1{
- 264 {\mod@updatedpre{#1}% add the new file to the already existing path
- 265 \let\mod@savedprefix\mod@prefix% add the path to the new file to the prefix
- 266 \mod@updatedpost{#1}%
- 267 \def\mod@blaaaa{}\% macro used in the simplify function (remove .. from the prefix)
- 268 \mod@simplify{\mod@savedprefix}% remove |xxx/..| from the path (in case it exists)
- $269 \verb|\mod@reguse{\mod@savedprefix}||%$
- 270 \let\newreg\mod@reg% use to compare, in case the .sms file was loaded before
- 271 \mod@search{\mod@savedprefix}% update registry
- 272 \ifx\newreg\mod@reg%\message{This file has been previously introduced}
- 273 \else\input{\mod@savedprefix}%
- 274 \fi}}

EdN:14

```
275 \let\sinputref=\sinput
276 \let\inputref=\input
277 \let\includeref=\include
```

4.9 Including Externally Defined Semantic Macros

\requirepackage

278 \def\requirepackage#1#2{\makeatletter\input{#1.sty}\makeatother}

4.10 Views

We first prepare the ground by defining the keys for the view environment.

```
279 \srefaddidkey{view}
280 \addmetakey*{view}{title}
281 \addmetakey{view}{display}
282 \addmetakey{view}{from}
283 \addmetakey{view}{to}
284 \addmetakey{view}{creators}
```

\view@heading Then we make a convenience macro for the view heading. This can be customized.

```
285 \newcounter{view} [section]
286 \newcommand\view@heading [2] {\stepcounter{view}%
287 \noindent\ifx\view@display\st@flow\else
288 {\textbf{View} {\thesection.\theview} from \textsf{#1} to \textsf{#2}}%
289 \sref@label@id{View \thesection.\theview}%
290 \ifx\view@title\@empty\quad\else\quad(\view@title)\fi%
291 \par\noindent%
292 \fi\ignorespaces}%ifmod@show
```

view The view environment relies on the @view environment (used also in the STEX module signatures) for module bookkeeping and adds presentation (a heading and a box) if the showmods option is set.

```
293 \newenvironment{view}[3][]%
294 {\begin{@view}[#1] {#2}{#3}\view@heading{#2}{#3}}
295 {\end{@view}\ignorespaces}
296 \ifmod@show\surroundwithmdframed{view}\fi
```

Oview The Oview does the actual bookkeeping at the module level.

```
297 \newenvironment{@view}[3][]{% keys, from, to
298 metasetkeys{view}{#1}\sref@target%
299 \@importmodule[\view@from]{#2}\@importmodule[\view@to]{#3}{\view@ext}}
300 {}
```

 $^{^{14}\}mathrm{EDNote}$: the sinput macro is just faked, it should be more like requiremodules, except that the tex file is inputted; I wonder if this can be simplified.

```
viewsketch The viewsketch environment behaves like view, but only has text contents.
               301 \newenvironment{viewsketch}[3][]%
               302 {\begin{@view}[#1]{#2}{#3}\view@heading{#2}{#3}}
               303 {\end{@view}}
               304 \ifmod@show\surroundwithmdframed{viewsketch}\fi
                4.11
                       Support for MathHub
\importmhmodule
                The \importmhmodule saves the current value of \mh@currentrepos in a local
                macro \mh@@repos, resets \mh@currentrepos to the new value if one is given in
                the optional argument, and after importing resets \mh@currentrepos to the old
                value in \mh@@repos.
               305 \srefaddidkey{importmhmodule}
               306 \addmetakey{importmhmodule}{repos}
               307 \addmetakey[sms]{importmhmodule}{ext}
               308 \newcommand\importmhmodule[3][]{\metasetkeys{importmhmodule}{#1}%
               309 \edef\mh@currentrepos}%
               311 \importmodule[load=\MathHub{\mh@currentrepos/source/#2},
                                ext=\importmhmodule@ext,id=\importmhmodule@id]{#3}%
               313 \mhcurrentrepos\mh@@repos\ignorespaces}
                and now the analogs
  \usemhmodule
               314 \newcommand\usemhmodule[3][]{\metasetkeys{importmhmodule}{#1}%
               315 \edef\mh@@repos{\mh@currentrepos}%
               316 \ifx\importmhmodule@repos\@empty\else\mhcurrentrepos{\importmhmodule@repos}\fi%
               317 \usemodule [load=\MathHub{\mh@currentrepos/source/#2},
                             ext=\importmhmodule@ext,id=\importmhmodule@id]{#3}%
               319 \mhcurrentrepos\mh@@repos\ignorespaces}
\adoptmhmodule
               320 \newcommand\adoptmhmodule[3][]{\metasetkeys{importmhmodule}{#1}%
               321 \edef\mh@@repos{\mh@currentrepos}%
               322 \ifx\importmhmodule@repos\@empty\else\mhcurrentrepos{\importmhmodule@repos}\fi%
               323 \adoptmodule[load=\MathHub{\mh@currentrepos/source/#2},
                              ext=\importmhmodule@ext,id=\importmhmodule@id]{#3}%
               325 \mhcurrentrepos\mh@@repos\ignorespaces}
   \mhinputref
               326 \newcommand\mhinputref[2][]{\def\@repos{#1}%
               327 \edef\mh@currentrepos}%
               328 \ifx\@repos\@empty\else\mhcurrentrepos{#1}\fi%
               329 \inputref{\MathHub{\mh@currentrepos/source/#2}}%
               330 \mhcurrentrepos\mh@@repos\ignorespaces}
      \mhinput
```

331 \let\mhinput\mhinputref

\mhcurrentrepos \@mhcurrentrepos

\mhcurrentrepos is used to initialize the current repository. It calls the internal macro \@mhcurrentrepos and writes a call to the same to the aux file. So that the \importmodule calls there work correctly.

332 \newcommand\mhcurrentrepos[1]{\@mhcurrentrepos{#1}\protected@write\@auxout{}{\string\@mhcurrentrepos 333 \newcommand\@mhcurrentrepos[1]{\def\mh@currentrepos{#1}}

importmhmodulevia

```
334 \newenvironment{importmhmodulevia}[3][]{\gdef\@@doit{\importmhmodule[#1]{#2}{#3}}%
335 \ifmod@show\par\noindent importing module #2 via \@@doit\fi}
```

336 {\aftergroup\@@doit\ifmod@show end import\fi}

4.12 Deprecated Functionality

In this section we centralize old interfaces that are only partially supported any more.

module:uses

For each the module name xxx specified in the uses key, we activate their symdefs and we export the local symdefs. 15

```
337 \define@key{module}{uses}{%
```

 $338 \end{area} $$ \end{area}$

module:usesqualified

EdN:15

This option operates similarly to the module: uses option defined above. The only difference is that here we import modules with a prefix. This is useful when two modules provide a macro with the same name.

```
339 \define@key{module}{usesqualified}{%
```

340 \@for\module@tmp:=#1\do{\activate@defs{qualified@\module@tmp}\export@defs\module@tmp}}

\coolurion/off

```
341 \def\coolurion{\PackageWarning{modules}{coolurion is obsolete, please remove}}
342 \def\coolurioff{\PackageWarning{modules}{coolurioff is obsolete, please remove}}
```

4.13 **Experiments**

In this section we develop experimental functionality. Currently support for complex expressions, see https://svn.kwarc.info/repos/stex/doc/blue/comlex_ semmacros/note.pdf for details.

\csymdef For the LATEX we use \symdef and forget the last argument. The code here is just needed for parsing the (non-standard) argument structure.

```
343 \ef\csymdef{\csymdef}{\csymdef[]}}
```

 $344 \ef{\csymdef[#1]#2}(\csymdef[#1]{#2}){\csymdef[#1]{#2}[0]}$

 $345 \ensuremath{ \ensuremath{ \mbox{00csymdef} \mbox{#1] \mbox{#2} \mbox{#4}} }$

\notationdef For the LATEX side, we just make \notationdef invisible.

346 \def\notationdef [#1] #2#3{}

 $^{^{15}\}mathrm{EdNote}$: this issue is deprecated, it will be removed before 1.0.

The code for avoiding duplicate loading is very very complex and brittle (and does not quite work). Therefore I would like to replace it with something better. It has two parts:

- keeping a registry of file paths, and only loading when the file path has not been mentioned in that, and
- dealing with relative paths (for that we have to string together prefixes and pass them one)

For the first problem, there is a very nice and efficient solution using etoolbox which I document below. If I decide to do away with relative paths, this would be it.

\reqmodules

We keep a file path registry \@register and only load a module signature, if it is

```
347 \newcommand\reqmodules[2]{\ifinlist{#1}{\@register}{}{\listadd\@register{#1}\input{#1.#2}}}
```

for the relative paths, I have to find out the directory prefix and the file name. Here are two helper functions, which work well, but do not survive being called in an \edef, which is what we would need. First some preparation: we set up a path parser

```
348 \newcounter{@pl}
```

349 \DeclareListParser*{\forpathlist}{/}

\file@name

\file@name selects the filename of the file path: \file@name{/foo/bar/baz.tex} is baz.tex.

```
350 \def\file@name#1{\setcounter{@pl}{0}%
```

- 351 \forpathlist{\stepcounter{@pl}\listadd\@pathlist}{#1}
- $352 \end{area} $$1^{\phi}^{1}_{\pi^1}_{\add to counter_{0pl}_{-1}} $$$
- 353 \dolistloop{\@pathlist}}

\file@path \file@path selects the path of the file path \file@path{/foo/bar/baz.tex} is /foo/bar

```
354 \def\file@path#1{\setcounter{@pl}{0}%
```

- 355 \forpathlist{\stepcounter{@pl}\listadd\@pathlist}{#1}
- $357 \simeq {0pl}{1}{\#1}{\#1}}$
- 358 \dolistloop{\@pathlist}}
- 359 (/package)

what I would really like to do in this situation is

\NEWrequiremodules

but this does not work, since the \file@name and \file@path do not survive the \edef.

```
360 \def\@NEWcurrentprefix{}
```

- $361 \ensuremath{\tt MEWrequiremodules\#1{\tt def\tt @pref{\tt file@path{\#1}}}\%$
- 362 \ifx\@pref\@empty\else\xdef\@NEWcurrentprefix{\@NEWcurrentprefix/\@pref}\fi%
- 363 \edef\@input@me{\@NEWcurrentprefix/\file@name{#1}}
- 364 \message{requiring \@input@me}\reqmodule{\@input@me}}

Index

Numbers written in italic refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; numbers in roman refer to the code lines where the entry is used.

LATEXML,	3, 10)—	name			relation		
12,	15,	18	module,		7	inheritance	(se-	
						mantic),		7
MARKIMI	9	4	OMDOC, 3, OPENMATH,	9,	11			
МатнМL,	3,	4	OpenMath,	3,	4	semantic		
module						inheritance		
name,		7	Perl,		15	relation,		7

References

- [Aus+10] Ron Ausbrooks et al. Mathematical Markup Language (MathML) Version 3.0. Tech. rep. World Wide Web Consortium (W3C), 2010. URL: http://www.w3.org/TR/MathML3.
- [BFM05] Tim Berners-Lee, Roy T. Fielding, and Larry Masinter. Uniform Resource Identifier (URI): Generic Syntax. RFC 3986. Internet Engineering Task Force (IETF), 2005. URL: http://www.ietf.org/rfc/rfc3986.txt.
- [Bus+04] Stephen Buswell et al. The Open Math Standard, Version 2.0. Tech. rep. The OpenMath Society, 2004. URL: http://www.openmath.org/standard/om20.
- [Hor+11] Fulya Horozal et al. "Combining Source, Content, Presentation, Narration, and Relational Representation". In: *Intelligent Computer Mathematics*. Ed. by James Davenport et al. LNAI 6824. Springer Verlag, 2011, pp. 212-227. ISBN: 978-3-642-22672-4. URL: http://kwarc.info/frabe/Research/HIJKR_dimensions_11.pdf.
- [Koh06] Michael Kohlhase. OMDoc An open markup format for mathematical documents [Version 1.2]. LNAI 4180. Springer Verlag, Aug. 2006. URL: http://omdoc.org/pubs/omdoc1.2.pdf.
- [Koh08] Michael Kohlhase. "Using LATEX as a Semantic Markup Format". In: Mathematics in Computer Science 2.2 (2008), pp. 279-304. URL: https://svn.kwarc.info/repos/stex/doc/mcs08/stex.pdf.
- [Koh13a] Michael Kohlhase. metakeys.sty: A generic framework for extensible Metadata in LATEX. Tech. rep. Comprehensive TEX Archive Network (CTAN), 2013. URL: http://www.ctan.org/tex-archive/macros/latex/contrib/stex/metakeys/metakeys.pdf.
- [Koh13b] Michael Kohlhase. omdoc.sty/cls: Semantic Markup for Open Mathematical Documents in LaTeX. Tech. rep. Comprehensive TeX Archive Network (CTAN), 2013. URL: http://www.ctan.org/tex-archive/macros/latex/contrib/stex/omdoc/omdoc.pdf.
- [Koh13c] Michael Kohlhase. "Organizing Symbols between Slides, Lecture Notes, Encyclopedias, and Original Papers in Semantic Digital Libraries". KWARC Blue Note. 2013. URL: https://svn.omdoc.org/repos/ omdoc/doc/blue/slides-notes/note.pdf.
- [Koh13d] Michael Kohlhase. statements.sty: Structural Markup for Mathematical Statements. Tech. rep. Comprehensive TeX Archive Network (CTAN), 2013. URL: http://www.ctan.org/tex-archive/macros/latex/contrib/stex/statements/statements.pdf.
- [Mil] Bruce Miller. LaTeXML: A LATeXX to XML Converter. URL: http://dlmf.nist.gov/LaTeXML/ (visited on 03/12/2013).

- [RO] Sebastian Rahtz and Heiko Oberdiek. Hypertext marks in LATEX: a manual for hyperref. URL: http://tug.org/applications/hyperref/ftp/doc/manual.pdf (visited on 01/28/2010).
- [sTeX] Semantic Markup for LATEX. Project Homepage. URL: http://trac.kwarc.info/sTeX/ (visited on 02/22/2011).