Documents with flexible Notation Contexts as Interfaces to Mathematical Knowledge

Michael Kohlhase, Christine Müller, Normen Müller
Computer Science, Jacobs University Bremen, Germany
m.kohlhase/c.mueller/n.mueller@jacobs-university.de

Abstract

In this paper we explore the use of documents as interfaces to mathematical knowledge. We propose an augmented document model that facilitates the explication of a document's notation context, i.e. the selection of appropriate presentations for all symbols in the document. By separating the notation context from the structure and content of a document, we aim at identifying the author's selection of notation, i.e. his notation practices. This facilitates both, the identification of communities of practice (CoP) that share specific notation preferences and, conversely, the adaptation of documents to the notation preferences of identified CoPs, e.g. groups of readers and co-authors. Furthermore, explicating a document's notation context allows for referencing and, particularly, reusing contexts, which reduces the author's workload during the authoring process.

1 Introduction

In this paper we will re-evaluate the concept of a document in the age of digital media. Since the development of writing, documents have become one of the primary means of passing on knowledge. If we take the view that knowledge is a state of the (human) mind where concepts and experiences are highly interlinked, and interpret "documents" broadly to include spoken language or action sequences played out[1], then we can see that documents are almost the only means to communicate knowledge: quite literally, *documents are interfaces to knowledge*. Here, we will look at *mathematical documents* as interfaces to *mathematical knowledge*, since they have special complexities that reveal the underlying nature of documents and knowledge.

Before Gutenberg, the notion of a document was relatively clear: a document is a physical object created by applying ink to paper or parchment for the purposes of transporting information in the form of images or written language. With the invention of letterpress printing, the production of verbatim copies became so simple, that we have to distinguish two aspects of documents: the *information object* (IO, i.e. the common information conveyed by all copies), and the *physical object* (PO, e.g. the book), which are distinct: for instance, if you destroy one of the physical books (e.g. by burning it), then you do not necessarily destroy the information object which can live on in other copies. In the age of electronic typesetting, the picture becomes more complex still, since the *digital representation* (DR) of a document appears; it is situated somewhere between the IO and the PO, and is distinct from both. Moreover, the perceived one-to-one correspondence between the IO and the PO becomes tenuous, since the DR is structurally different from the PO (e.g. the screen presentation for printout). Consider for instance document parts like the table of contents or index in LaTeX, which are spread over the DR and are assembled during the formatting process that creates the PO. We will neglect the difference between the DVI file and the printout for the sake of this argument, but make use of the fact that we can also use the term PO for "*presentation object*" which includes the former without excluding the latter. Generalizing the IO-PO correspondence also allows to re-interpret the IO as the knowledge conveyed by the document and try to approximate it with a *content representation* (CR), that tries to capture the structure of the knowledge rather than the structure of the PO. We see first steps in this direction already in LaTeX and SGML, which strive to separate *function* from *form* in documents and employ styling mechanisms to compute the document form (i.e. the PO) from the CR. This trend finds its (current) conclusion in markup languages like OpenMath [BCC+04] and MathML [ABC+03] for mathematical formulae and OMDoc [Koh06] for mathematical documents which offer means to "identify the structure, the meaning of text fragments, and their relations to other knowledge" [Koh06].

In the context of such languages and dynamic presentation media (e.g. on the screen), we can now dream about "*active documents*", where we can interact with a document directly, e.g. instantiating a formula with concrete values or graphing a function to explore it or "*living/evolving documents*" which monitor the change of knowledge about a topic. All of these developments can strengthen the role of documents as interfaces to mathematical knowledge by empowering the reader in her interaction with the content, but considerably weaken our understanding of what a "document" might be.

In this paper, we will drill in on the intuition that a "document" should be an object, whose extent and

[1] After all, these can be captured by the digital media.

context are fixed by a "publication" process, and which contains enough information that a PO can be generated for it (e.g. for printing into a "pre-Gutenberg" document). We will use the process of presenting mathematical formulae as a razor to distinguish conceptual aspects documents and propose a novel document architecture, which we will develop concretely in the context of our OMDOC format.

Presenting Mathematical Formulae In [NW01] Naylor and Watt present an approach based on meta style sheets that was deployed in the *Notation Selection Tool* [Not07]: The approach utilizes a MATHML-based markup of arbitrary notations in terms of their content and presentation and, based on the manual selection of users, generates user-specific XSLT style sheets [Kay06] for the adaptation of documents. In [MLUM05] Manzoor et al. emphasize the need for maintaining uniform and appropriate notations in collaborative environments, in which various authors contribute mathematical material. They address the problem by providing authors with respective tools for editing notations as well as by developing a framework for a consistent presentation of symbols.

In this paper we propose the markup of the *notation context* of a document, i.e. the explication of the selected presentation for symbols by the author. By separating the notation context from the content and structure of the document, we present an alternative approach that reduces the workload of authors and supports the comprehensibility of documents by readers: For example, one could imagine that documents could be easily switched to the notation context preferred by a community of practice (COP) [Wen05], e.g. groups of readers and co-authors. Alternatively, authors could be enabled to reuse the notation contexts of other documents or the notation preference of a specific user or COP.

The Document/ Knowledge Model in OMDOC**:** The semantic markup language OMDOC1.2. contains two OMDOC types[2] in order to mark up the knowledge contained in a mathematical document and its structure: Content OMDOCs are "knowledge-centered documents that contain the knowledge conveyed in a document" [Koh06]. In contrast, narrative OMDOCs are used to "reference the knowledge[-centered documents] and add the theoretical and didactic structure of a document" [Koh06]. The combination of the narrative structure and the (mathematical) content of a document as the formal representation of a document model, has been defined by Normen Müller as NARCONS, i.e. two-dimensional graphs consisting of a **nar**rative layer and a **con**tent layer [Mül06]. Several terms have been created to denote knowledge in the content layer of a NARCON: Michael Kohlhase [Koh06] uses the term *knowledge item*, while Normen Müller proposes the term *information units* or INFOMS, i. e. "tangible/visual text fragments potentially adequate for reuse, constituting the content of documents" [Mül06]. We use the term INFOM to refer to *fragments of content* that are represented in *Content* OMDOC. INFOMS are stored in an OMDOC repository, which we call *content commons* according to the terminology of the educational knowledge repository CONNEXIONS [CNX07]. In this paper we revise our previous definition of documents and present an augmentation of NARCONS (cf. section 2).

Notation Context and Notation Practice: The *notation context* of a document shall be defined as a mapping of symbols to presentations. The selection of respective presentation for symbols highly depends on the author's *notation practice*, i.e. his individual way of selecting notations, which he acquires throughout his life and which is influenced by a number of factors.

Various researchers have addressed the challenge of identifying and describing mathematical practice. In [SW06] Watt and Smirnova introduce possible *reasons* for multiple notations of the same mathematical concept, namely *area of application*, *national conventions*, *level of sophistication*, the *mathematical context*, and the *historical period*: For example, to introduce an imaginary unit, a mathematician uses the symbol i. In contrast, an electrical engineer uses j to avoid confusion with the symbol I for electric current. i and j are two alternative presentation for the symbol "imaginary unit". Hence, the *notation context* of a mathematical document will most likely refer to the notation i, while the notation selected for a document by an electrical engineer refers to j. Based on the *national convention* we distinguish notations that are commonly used in different language. For example, a German researcher presents the symbol "binomial coefficient" with the notation $\binom{n}{k}$. In contrast, a Russian researcher uses the alternative presentation C_k^n, while a French researcher will most likely use C_n^k.

In contrast to the five *"reasons"* in [SW06], the OMDOC-based eLearning environment ACTIVEMATH [MLUM05] distinguishes four *"contexts" categories* that influence the adaptation of notations, namely *language*, *different patterns of the argument*, the *author's style*, and *notations of the same collection*. For example, depending on his individudal style, an author will use a/b or $a : b$ (cf. [MLUM05]). In addition, the ACTIVEMATH group specified a prioritization of notation collections, which we referred to as *notation contexts*: *system defaults*, *author*, *book*, *group*, and *individual* style (highest priority). Furthermore, in ACTIVEMATH authors can annotate different layout preferences for notations, such as *color*, *font*, or *border*, which shall also be considered parts of the notation practice.

Communities of Practice: In [KK06] Andrea Kohlhase and Michael Kohlhase propose the application of the economic theory of *communities of practice (*COP*)* [Wen05] to the area of mathematics. According to their discussions, *mathematical practice* is *inscribed* into documents, e. g. by selecting specific notations or referencing other mathematical publications.

[2]The OMDOC group does do not claim to have invented this concept, it is part of the XML folklore and can already be found e.g. in [VD04]. But the OMDOC format probably implements this idea in the cleanest way.

Analyzing a *collection of documents* will potentially lead to *clusters of shared practices*, i.e. *communities of practice*, that define the mathematical practice of researchers. For example, the participants of the MathUI Workshop establish a community of notation practice by adapting to the notation of fellow researchers and, vice versa, by influencing the community with their individual habits. The longer a researcher takes part in the annual MathUI Workshop, the more likely he will adapt to this set of shared notations, hence moving from the outer border of the MathUI CoP to the inner parts. However, [KK06] further emphasize the problem of *inscribing practices*, e. g. by *enforcing* the use of specific notations in an editor, and call for the implementation of workflows that preserve the *fluid movement* in and between CoPs.

This paper aims at resuming the discussion on mathematical CoPs and takes one step towards a CoP *model* that preserves the *dynamics* and *flexibility* of *mathematical communities* in terms of their *notation practice* (cf. section 2).

2 Modeling the Notation Context in OMDoc

We have extended our NarCon model by a third dimension — the presentation layer and developed a concrete syntax for inclusion in OMDoc2.0: In OMDoc1.2, the root element `omdoc` does double duty: encapsulating content- and narrative- and mixed documents without an elaborated mechanism that allows to distinguish between the cases. Moreover, the document structure is marked up by the `omgroup` element that has the same content model as the `omdoc` element. In the proposed document representation, we bid farewell to the illusion that we can represent narrative and content structure in one document, and make the three dimensions of a document explicit. Documents will be structured by `omdoc` elements which have three optional children: a `narrative` element specifying the narrative structure of the document, a `content` caching its Infoms, and a `pcontext`[3] for explicating the notation context of the document. We propose that the `type` attribute on the `omdoc` element be generated from the *document ontology* [Mül07] of OMDoc.

The following RelaxNG [vdV04] grammar summarizes the proposed document structure.

id . att	= attribute xml:id {xsd:ID}
omdoc	= element omdoc {id. att ?, attribute type{xsd:NCName}, pcontext ?, narrative ?, content ?}
pcontext	= element pcontext {id. att ?, attribute base list {xsd:anyURI∗},ref∗}
narrative	= element narrative {id. att ?, narrative . class ∗}
content	= element content {content. class ∗}
narrative . class	= metadata|omdoc|ref| tableofcontents |authorindex |. . .
content . class	= omtext| definition |theory |. . .

The `narrative` element specifies the structure of a document via `omdoc` elements (e. g. for chapters or sections) and via `ref` elements that reference (i.e. include) document fragments from the *document commons*[4] or Infoms from the sibling `content` element. These can be freely inter mixed with special narrative elements for tables of content, indices, etc.

The `content` element is used to *cache* content of the document that is referenced in the `narrative` elements of the document. For example, an `omtext` element that already exists in the *content commons*, can be referenced in the `narrative` element and may optionally be copied to the `content` element. Alternatively, the author can write a new `omtext` element himself, which is then added to the `content` element and eventually added to the *content commons*[5].

The `pcontext` element is a collection of one or more references to `presentation` elements in the sibling `content` element or *content commons*. The optional `base` attribute contains a whitespace-separated list of URIs pointing to `pcontext` of other NarCons, hence it can be used to inherit presentation contexts. The effective presentation context is computed by cascading (left-to-right) the effective presentation contexts referenced and the `presentation` elements referenced locally. For an example for the prioritization of presentation contexts and elements see the next page.

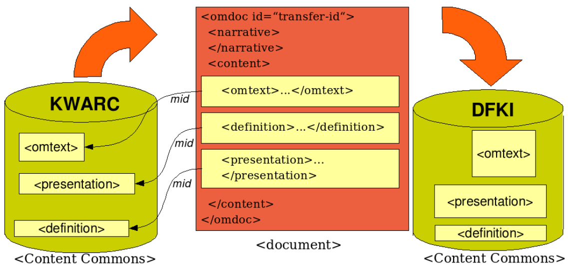

Figure 1: Documents as Interfaces

Documents as Interfaces to Mathematical Knowledge: Figure 1 visualizes the idea of documents as interfaces to mathematical knowledge. The Infoms in the *content commons* of the *KWARC group* are embedded into an OMDoc document, in particular, they are cached in the `content` element of an `omdoc` element, which in this case does not necessarily include a `narrative` element. To make the references to the oringinal Infoms, which are now (re-)used in the document, *RESTful* [RST07], unique, and aware of the used Infom state, we re-introduce the OMDoc1.0 `mid` attribute. Each Infom within the *con-*

[3]The element is named presentation context since it will also be used to specify further presentation-specific contexts besides the notation contexts, e. g. layout preferences such as color or fonts. For the further course of the paper, both terms, notation and presentation contexts, are used synonymously.

[4]A *document commons* is a repository of documents, e. g. in OMDoc format.

[5]For now we assume that the author manually references the existing elements from the *content commons*, i.e. by indicating the respective URI. Also, when adding new elements to the content commons (that potentially already exist in the repository), these are currently not matched with the exiting ones but simply stored as new entries. In the near future, the consistent referencing of existing elements as well as the alignment of new/ modified elements among existing ones will be managed automatically.

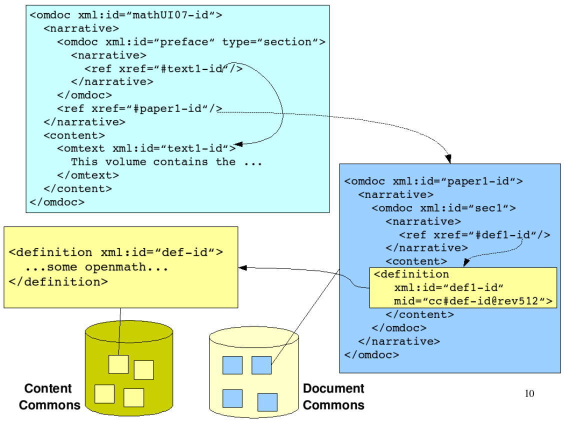

Figure 2: Documents and Content Commons

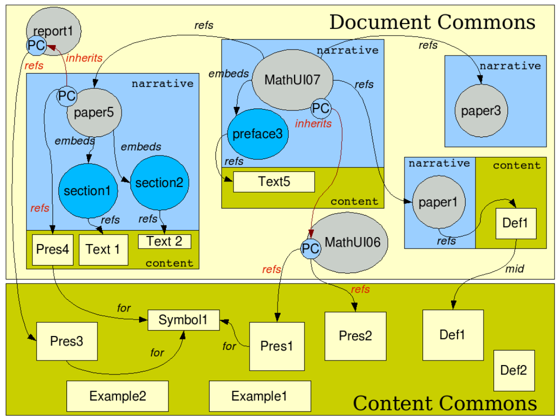

Figure 3: Narrative Structure of the MKM-07 proceedings

tent commons is keyed by a unique `URI@REV`[6] and can be requested by simple, stateless HTTP request (`PUT`, `POST`, `GET` or `DELETE`). By this means, a document may be used to transfer mathematical knowledge between *content commons*, e.g. between the *KWARC* and the *DFKI* group.

Dependencies between Document and Content Commons: Let us now consider the dependencies between a `definition` element in the *content commons*, a paper stored in the *document commons*, and a currently edited workshop proceedings (cf. Figure 2). The `narrative` element of the proceedings references the OMDOC representation of *paper1*. The `narrative` element further includes an `omdoc` child-element that references an `omtext` element that is cached in the `content` element of the proceedings and has not yet been stored in the *content commons*. In contrast, the `narrative` element of the paper references a `definition`, which is a copy of an already existing and reused element from the *content commons*. The original `definition` is identified by the `mid` attribute of its clone.

Narrative Structure and Content Management: Figure 3 displays the (cascading) narrative structure of the *MathUI07 proceedings*. The nodes of the *narrative graph* references `omdoc` elements in the *document commons*, e. g. papers, or embed `omdoc` child-elements specifying further narrative structures such as a sections or paragraphs of a paper. The content of the `omdoc` elements is stored in the sibling `content` elements and/ or the *content commons*. For example, the content of the *MathUI07* and *paper5* is only stored in their local `content`, while the content of *paper1* is stored in its local `content` referencing the original INFOM in the *content commons*. This also applies to `presentation` elements: The `pcontext` of the *MathUI06 proceedings* references `presentation` elements in the *content commons* which are not cached in the local `content` element, while the referenced `presentation` element of *paper5* is only stored in its local `content`. In consequence, the content of documents may include copies of the *content commons*, but may also include new content that is not part of the *content commons*. In order to transfer a document independently from the *content commons*, its `content` element needs to cache all content of the document.

Cascading Presentation Contexts and Prioritization: The cascading structure of the OMDOC representation of documents allows for the global and local specification of presentation contexts: Each nodes of the narrative graph can be associated with a specific `pcontext`, whereas the local `pcontext` overwrite more general `pcontexts`. For example in the proceedings figure above, the `pcontext` of the `omdoc` element representing the *MathUI07 proceedings* is inherited from the `pcontext` of the *MathUI06*. The `pcontext` of *paper5* is implicitly inherited from the *MathUI06 proceedings* since the document is part of the narrative structure of the *MathUI07 proceedings*. Furthermore, it is extended by inheriting the `pcontext` of *report1* as well as a local reference to a `presentation` element. The prioritization among `pcontexts` works as follows: The local `pcontext` of *paper5* overwrites the inherited `pcontext` of *paper5*, which further overwrites the implicitly inherited `pcontext` of *MathUI07 proceedings*: In particular, the prioritization of the presentation for *symbol1* are: *Pres1*, which is overwritten by *Pres3*, which is further overwritten by the locally referenced *Pres4* with the highest priority.

The prioritization on `omdoc` element level provides authors with the following options: They can either manually specify all the `presentations`, which consequently leads to rather *static document*. Alternatively, they can choose to omit most of the presentation specification, which leads to *flexible documents* that can be embedded in the `pcontext` of other documents. For example, while *paper5* overwrites the inherited `pcontext` of the proceedings, *paper1* adopts the referenced `presentation` elements.

The following listings provides the OMDOC representation of the *MathUI07 proceedings*. In addi-

[6] Subversion [SVN07] calls `REV` the peg revision to identify a unique line of history.

tion to the *preface* and the referenced papers, the narrative of the *proceedings* includes two special elements, i.e. a *table of contents* and an *author index*:

Listing 1: Representation of the MathUI proceedings

```
<omdoc xml:id="mathUI07" type="proceedings">
  <pcontext xml:id="pc−mathUI07" base="#mathUI06"/>
  <narrative>
    <metadata><dc:title>MathUI07</dc:title></metadata>
    <omdoc xml:id="preface" type="preface">
      <narrative>
        <metadata><dc:title>Preface</dc:title></metadata>
        <ref xref="#text5"/>
      </narrative>
    </omdoc>
    <tableofcontents/>
    <ref xref="#paper1"/>
    <ref xref="#paper3"/>
    <ref xref="#paper5"/>
    <authorindex/>
  </narrative>
  <content>
    <omtext xml:id="text5">
      <CMP>This volume contains the proceedings of the ... </CMP>
    </omtext>
  </content>
</omdoc>
```

The OMDOC specification below represents *paper5* of the *MathUI07*: It inherits its presentation context from the proceedings pcontext, but extends/ overwrite it by the inherited pcontext of the referenced report and a local reference to a presentation element.

```
<omdoc xml:id="paper5" type="paper">
  <pcontext xml:id="paper5−pc" base="#report−pc">
    <ref xref="#pres4−id"/>
  </pcontext>
  <narrative>
    <metadata>
      < dc:title >Documents with flexible Notation ... </ dc:title >
    </metadata>
    <omdoc xml:id="section1" type="section">
      <narrative>
        <metadata><dc:title>Introduction</dc:title></metadata>
        <ref xref="#text1"/>
      </narrative>
      <content>
        <omtext xml:id="text1">
          <CMP>In this paper we explore...</CMP>
        </omtext>
      </content>
    </omdoc>
    <omdoc xml:id="section2" type="section">
      <narrative>
        <metadata><dc:title>Outlook</dc:title></metadata>
        <ref xref="#text2"/>
      </narrative>
      <content>
        <omtext xml:id="text2">
          <CMP>To conclude...</CMP>
        </omtext>
      </content>
    </omdoc>
  </narrative>
  <content>
   <presentation xml:id="pres4−id">...</presentation>
  </content>
</omdoc>
```

Presentation Contexts within OMDOC elements: The association of presentation contexts to narrative nodes is one way to specify the presentation context. However, in some cases this is not sufficient, requiring for a more local specification: In OMDOC this is solved by extending the phrase and ref element with an optional attribute pc that references the respective presentation element. Please note that the presentation contexts given by pc attributes overwrites the specification in the pcontext elements: For example in listing 2, the pc attributes, which references the presentation element for the French C_k^n and Russian C_n^k notation, overwrite the more global specification in the pcontext, i.e. the German presentation $\binom{n}{k}$.

Listing 2: A Multi-context Presentation

```
<omdoc xml:id="section3" type="section">
  <narrative>
    <omdoc type="para">
      <pcontext xml:id="pc1"><ref xref="#bk−D"/><pcontext>
      <narrative><ref xref="#text1"/></narrative>
      <content>
        <omtext xml:id="text1">
          <CMP>We will discuss the binomial coefficient which
               can be represented in three alternative ways: i.e.
          <OMOBJ xml:id="foo">
            <OMA>
              <OMS cd="bk" name="bk"/>
              <OMV name="n"/></OMV name="k"/>
            </OMA>
          </OMOBJ> (Germany),
          <ref xref="#foo" pc="#bk−F"/> (France), and
          <ref xref="#foo" pc="#bk−R"/> (Russia) </CMP>
        </omtext>
      </content>
    </omdoc>
  </narrative>
</omdoc>
```

After transforming the OMDOC representation into PDF, the original document would appear as shown below:

> We will discuss the binomial coefficient which can be represented in three alternative ways: i.e. $\binom{n}{k}$ (Germany), C_k^n (France), and C_n^k (Russia).

Figure 4: The presentation of listing 2

3 Conclusion and Future Work

In this paper we have initiated the redefinition of *documents* towards a more *dynamic* and *living* view. So far, the term mostly referred to static articles and publications, neglecting the new opportunities offered by modern technologies. In the first step, we have explicated the narrative and content layer and extended the *document model* by a third dimension, i.e. the presentation layer. We further specified how the notation contexts of a document can be modeled on different cascading layers, i.e. on the omdoc element level by using pcontext elements as well as within omdoc elements by specifying pc attributes. However, several issues remain, which we will address in future work:

Personalization of Documents: The specification of a document's notation context can be used to automatically adapt the presentation to a specific (personalized) context, e. g. of a group or user. However, this requires to *prioritize the context specifications*, which is not trivial: We already mentioned, that within the cascading structure of documents a local pcontext overwrite a more global pcontext; and the pc attribute overwrites the pcontext of the omdoc element. However, the prioritization of a document's notation contexts among various contexts, such as system default, the author's and group's notation context as well as the individual (reader's) contexts has not been sufficiently solved yet.

For example, applying the prioritization in [MLUM05] takes away the intention of the author: If we allow a user to overwrite the presentation context in listing 2 by his individual style, i.e. the German presentation of the binomial coefficient, the adaptation would end in nonsense:

> We will discuss the binomial coefficient which can be represented in three alternative ways: i.e. $\binom{n}{k}$ (Germany), $\binom{n}{k}$ (France), and $\binom{n}{k}$ (Russia).

Figure 5: A nonsensical presentation of listing 2

We are currently working on supporting authors to preserve specific notations, i.e. to indicate whether a presentation can be adapted or not. One possible way is to add priorities to the presentation contexts as e.g. in CSS or XSLT.

Extensional vs. Intensional Model: In the first step we use an *extensional model* to specify the notation context, i.e. the author has to specify all presentations for symbols in his document by either inheriting existing specifications or referencing `presentation` elements. In the next step, we will also evaluate and potentially model an *intensional approach*, i.e. the author will be enabled to specify a context by e.g. indicating the language, the area of application, or the level of sophistication. His intensional specification will then trigger an automatic identification of `presentation` elements for symbols in his document. One possible way to support the intensional specification is adding `type` attributes to the `presentation` element (cf. [MLUM05]). For example, the type `de` specifies German presentations, `physics` specifies presentation for symbols in physics, or `elementary-level` indicates the level of sophistication. In order to allow for a more flexible specification in OMDOC2.0, the syntax of `presentation` elements has been respecified in [KLR07].

From Notation Contexts to CoP Models: We have extended the specification of notation context and practice by our predecessors, e. g. Watts/Smirnova [SW06] and ACTIVEMATH [MLUM05], towards a more general model: Instead of predefining system defaults, a book's or an author's context, we provide a model that allows for *reusability* of existing contexts, e. g. of a paper or book. In the next step, we aim at providing means to specify the notation context of an individual and communities of practice.

One potential mean to specify an author's or CoP notation context is using `omdoc` elements as a collection of references to their preferred or *most frequently used* presentations. The following listing provides an example for such an `omdoc` element. The `pcontext` of the document references the `presentation` elements used for symbols within the document written by the author or CoP. For example, the author "Christine Mueller" has used symbols of the *MathUI06*, a *report*, and a *specific* `presentation` element. The respective `presentation` elements are stored in the content commons and/ or the `content` element of the document.

```
<omdoc xml:id="cmueller−pcontext" type="notationcontext">
  <narrative>
  <metadata>
    < dc:title >The notation context of Christine Mueller</ dc:title >
  </metadata>
  <narrative>
  <pcontext>
    <ref xref="#mathUI06"/>
    <ref xref="#report"/>
    <ref xref="#pr4"/>
  <pcontext>
  <content>
    <presentation xml:id="pr4" for="#bk" role="applied">
     ... </presentation>
  </content>
</omdoc>
```

The identification of CoPs can be based on an analysis of a collection of notation contexts gained from a document collection. For now, the identification of a CoP's notation contexts is simply a prioritized merge of all `pcontext` elements and `pc` attributes of their documents. However, we took the first step towards a concrete application of the theory of *communities of practice* towards a scientific and technical field. We are currently working on an improved method for establishing CoPs, e.g. by means of an cluster analysis to identify the CoP's most frequently used presentations.

The identification of individual contexts is analogous: In the beginning, authors will manually specify their preferred presentation context. We will then implement routines that automatically detect an author's preferred notations. In order to store an author's notation contexts we propose the implementation of *ePortfolio*s[7]. We are currently developing an educational platform that manages ePortfolios of students, which will be use to evaluate the specification and deployment of notation contexts as well as to take further steps towards a scientific CoP model.

Acknowledgments We would like to thank the other members of our KWARC group for the fruitful discussions and continuous feedback on our work.

[7] ePortfolios are digital collections of private documents, such as individual notation contexts or CoP information.

References

[ABC$^+$03] Ron Ausbrooks, Stephen Buswell, David Carlisle, Stéphane Dalmas, Stan Devitt, Angel Diaz, Max Froumentin, Roger Hunter, Patrick Ion, Michael Kohlhase, Robert Miner, Nico Poppelier, Bruce Smith, Neil Soiffer, Robert Sutor, and Stephen Watt. Mathematical Markup Language (MathML) version 2.0 (second edition). W3C recommendation, World Wide Web Consortium, 2003. Available at `http://www.w3.org/TR/MathML2`.

[BCC$^+$04] Stephen Buswell, Olga Caprotti, David P. Carlisle, Michael C. Dewar, Marc Gaetano, and Michael Kohlhase. The Open Math standard, version 2.0. Technical report, The Open Math Society, 2004. `http://www.openmath.org/standard/om20`.

[CNX07] CONNEXIONS. Project homepage at `http://www.cnx.org`, seen February 2007.

[Kay06] Michael Kay. XSL Transformations (XSLT) Version 2.0. W3C Candidate Recommendation, World Wide Web Consortium (W3C), June 2006. Available at `http://www.w3.org/TR/2006/CR-xslt20-20060608/`.

[KK06] Andrea Kohlhase and Michael Kohlhase. Communities of Practice in MKM: An Extensional Model. In Jon Borwein and William M. Farmer, editors, *Mathematical Knowledge Management, MKM'06*, number 4108 in LNAI, pages 179–193. Springer Verlag, 2006.

[KLR07] Michael Kohlhase, Christoph Lange, and Florian Rabe. Presenting mathematical content with flexible elisions. submitted, 2007.

[Koh06] Michael Kohlhase. OMDOC – *An open markup format for mathematical documents [Version 1.2]*. Number 4180 in LNAI. Springer Verlag, 2006.

[MLUM05] Shahid Manzoor, Paul Libbrecht, Carsten Ullrich, and Erica Melis. Authoring Presentation for OPENMATH. In Michael Kohlhase, editor, *Mathematical Knowledge Management, MKM'05*, number 3863 in LNAI, pages 33–48. Springer Verlag, 2005.

[Mül06] Normen Müller. An Ontology-Driven Management of Change. In *Wissens- und Erfahrungsmanagement LWA (Lernen, Wissensentdeckung und Adaptivität) conference proceedings*, 2006.

[Mül07] Normen Müller. Towards an Ontology-Driven Management of Change. Exposé of PhD research proposal, June 2007.

[Not07] Notation Selection Tool, seen February 2007. available at `http://www.orcca.on.ca/MathML/NotationSelectionTool/`.

[NW01] Bill Naylor and Stephen M. Watt. Meta style sheets for the conversion of mathematical documents into multiple forms. In *Proceedings of the International Workshop on Mathematical Knowledge Management*, 2001.

[RST07] Representational state transfer. `http://en.wikipedia.org/wiki/Representational_State_Transfer`, seen May 2007.

[SVN07] Subversion, seen May 2007. available at `http://subversion.tigris.org/`.

[SW06] Elena Smirnova and Stephen M. Watt. Notation Selection in Mathematical Computing Environments. In *Proc. Transgressive Computing 2006: A conference in honor of Jean Della Dora (TC 2006)*, Granada, Spain, 2006.

[VD04] Katrien Verbert and Erik Duval. Towards a Global Component Architecture for Learning Objects: A Comparative Analysis of Learning Object Content Models. In *Proceedings of the EDMEDIA 2004 World Conference on Educational Multimedia, Hypermedia and Telecommunications*, pages 202–208, 2004.

[vdV04] Eric van der Vlist. *RELAXNG: A simple schema language for XML*. O'Reilly, 2nd edition, 2004.

[Wen05] Etienne Wenger. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, 2005.