Annotating Rhetorical and Argumentative Structures in Mathematical Knowledge

Summary of my work at DERI (Apr–Oct 2008)

EECS Seminar

Christoph Lange

Jacobs University, Bremen, Germany

KWARC – Knowledge Adaptation and Reasoning for Content

October 14, 2008

My Home: Mathematical Knowledge Management

- Ph. D. student with Prof. Michael Kohlhase
- Our group does "Mathematical Knowledge Management"
  - dealing with mathematical knowledge
  - formality ranges from human-friendly to computer-verifiable

My Project

- Collaboration on semiformal knowledge
- Using semantic web technologies (a semantic wiki, in particular)

What I Wanted to Learn About the Semantic Web

- engineering ontologies for scientific documents
- user interfaces for annotating and browsing
- relation of social interaction to knowledge

Where?

At DERI, they do this (and more)

DERI (Digital Enterprise Research Institute)

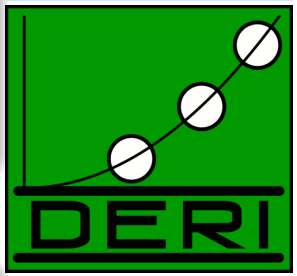
Largest semantic web research institute worldwide (130 members)

Applied Research

- eLearning
- semantic reality (sensor networks, ubiquitous computing)
- web services
- industrial applications
- *Semantic Information Systems and Language Engineering*
- *Social Software*

Foundational Research

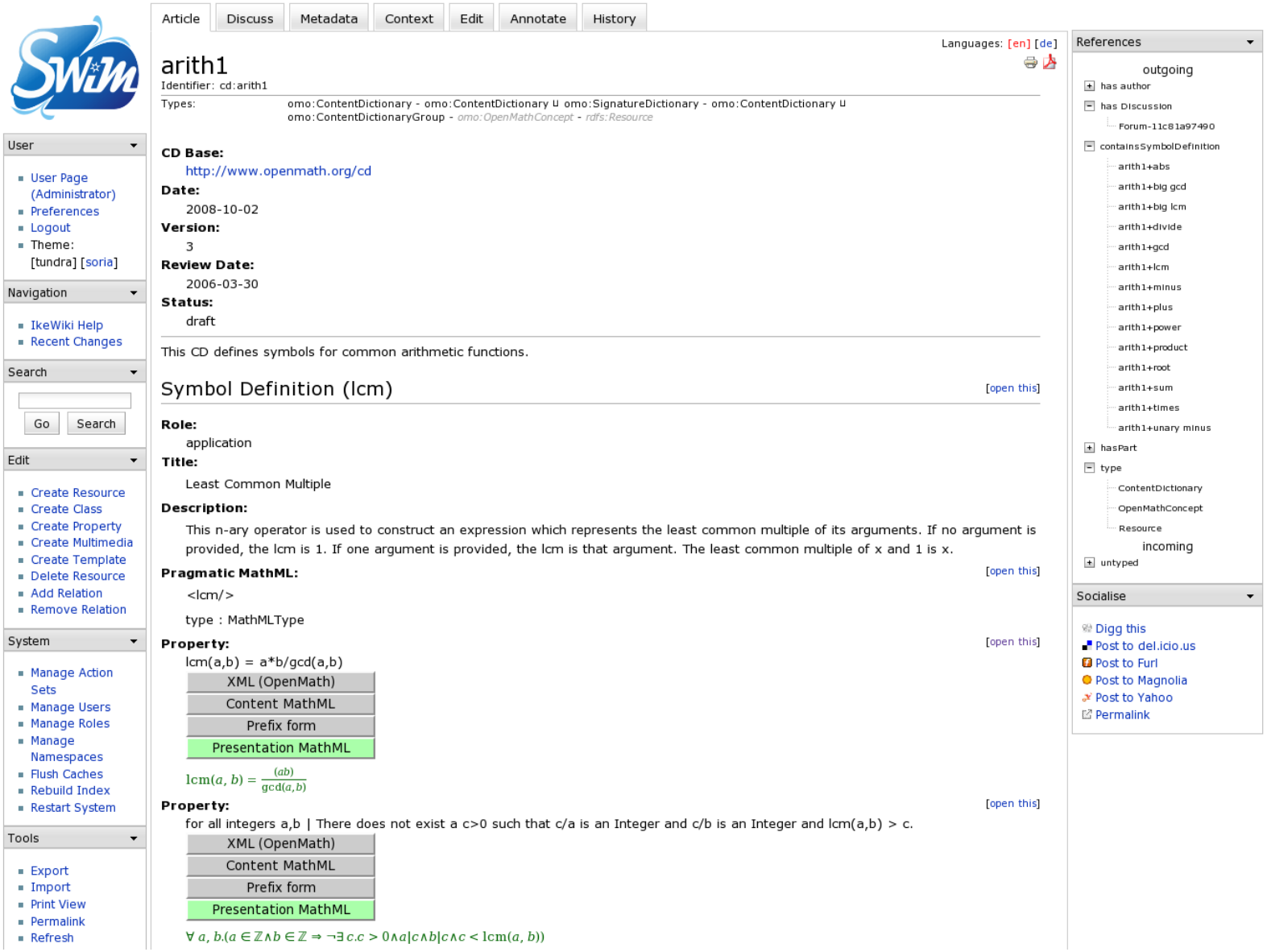
- data intensive infrastructures
- information mining and retrieval
- reasoning and querying



SmILE

kwarc JACOBS UNIVERSITY

Before: SWiM, a Semantic Wiki for Mathematics

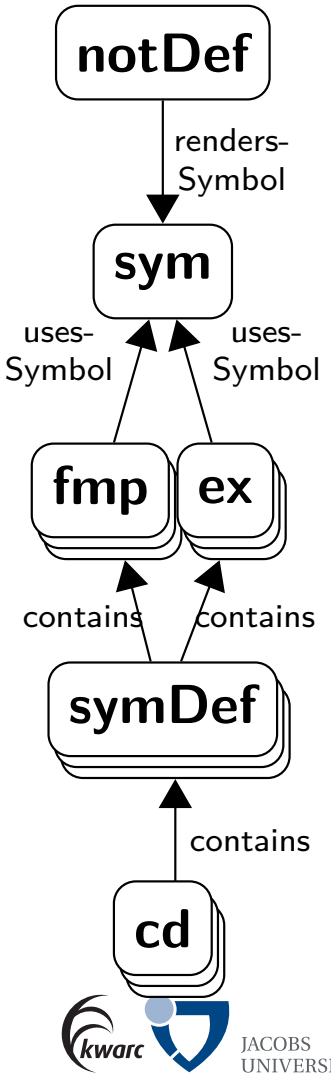


Ontologies for Mathematical Documents (1)

Previous State

I had a basic ontology that modelled structures of mathematical knowledge; mainly *statements* (definition, theorem, proof, examples)

- used in SWiM for navigation, queries, internal bookkeeping



Ontologies for Mathematical Documents (1)

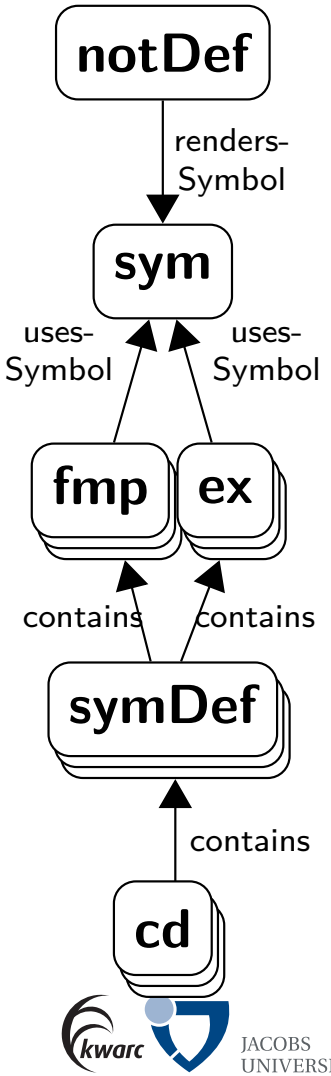
Previous State

I had a basic ontology that modelled structures of mathematical knowledge; mainly *statements* (definition, theorem, proof, examples)

- used in SWiM for navigation, queries, internal bookkeeping

Next Challenge

- Semi-formal knowledge often comes in *documents* that also contain *text*
- There is a document structure (chapter, section, cross-reference), and a rhetorical structure, both of which can be independent from the mathematical structure.



Ontologies for Mathematical Documents (2)

Getting the Model Right

$$\text{document ont.} \leftrightarrow \text{annotation ont.} \leftrightarrow \begin{cases} \text{rhetorical ont.} \\ \text{mathematical ont.} \end{cases}$$

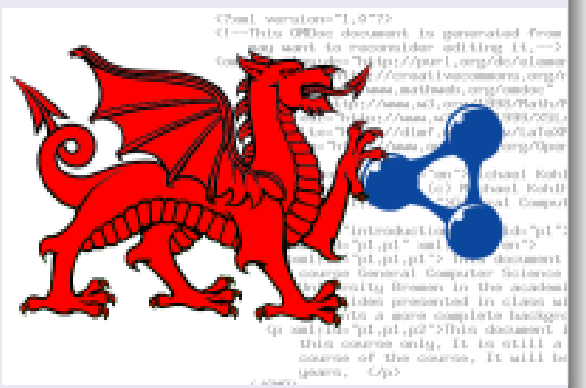
(following the SALT approach)

SALT (Semantically Annotated LaTeX)

semantic authoring framework for creating scientific publications

Implementation

- Expansion of the ontology
- Rules for extracting these concepts from OMDoc documents to RDF
- $\rightsquigarrow$ Krextor XML$\rightarrow$RDF extraction library



User Interfaces for Annotating and Browsing

Improved Annotation Support

More and easier annotation support in the editor

- toolbars for easy selection of types of mathematical knowledge
- from phrase to theory level
- deleting annotations
- auto-completion of link targets (prepared)

User Interfaces for Annotating and Browsing

Improved Annotation Support

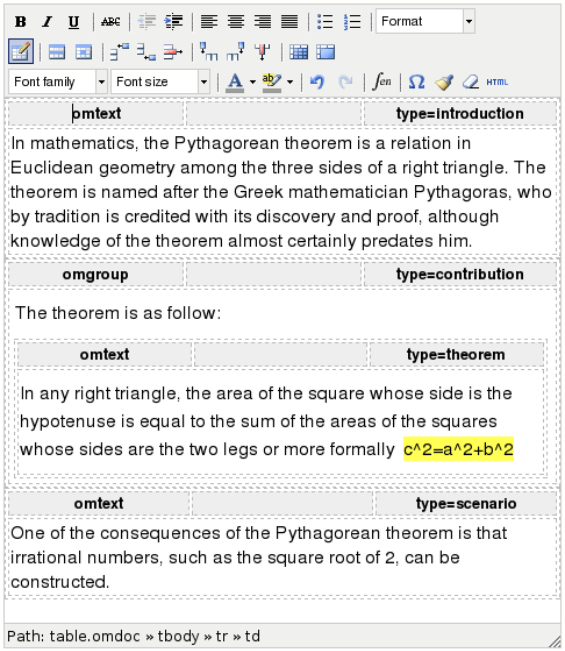
More and easier annotation support in the editor

- toolbars for easy selection of types of mathematical knowledge
- from phrase to theory level
- deleting annotations
- auto-completion of link targets (prepared)

Rhetorical Annotation and Visualisation

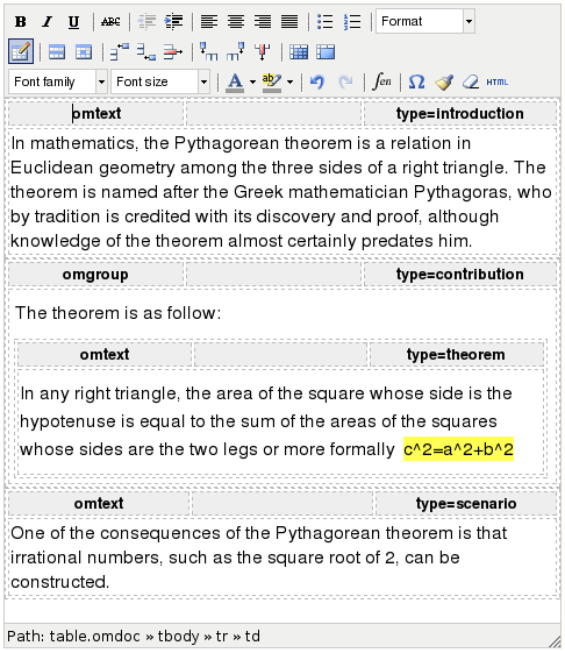
- improved and extended syntax for annotating SALT-/RST-like rhetorical structures in OMDoc
  using the SALT ontology within the host language OMDoc, not LaTeX
- ideas for an editing interface
- visualisation of rhetorical relations and blocks implemented
  $\longrightarrow$ *active documents*

Annotation

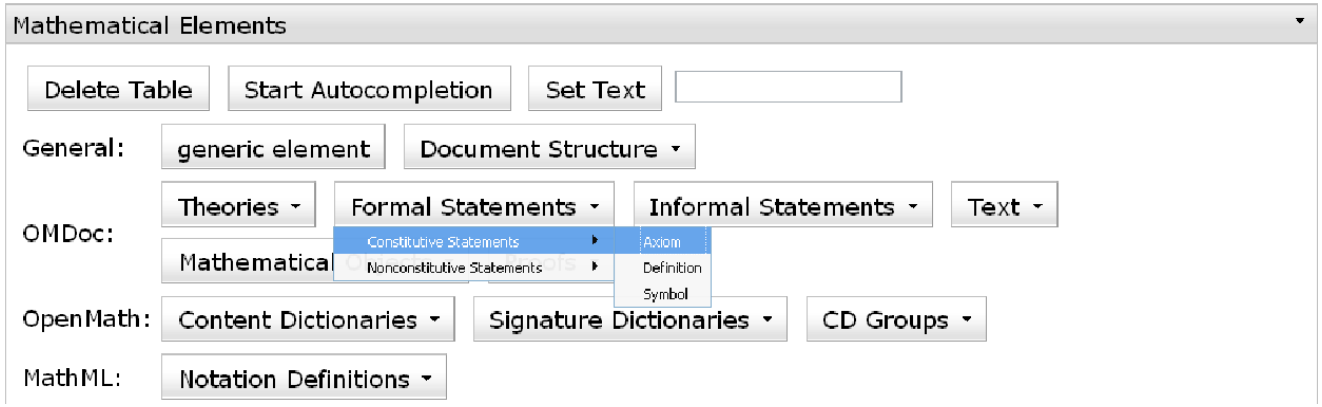


Sections in the editor

Annotation

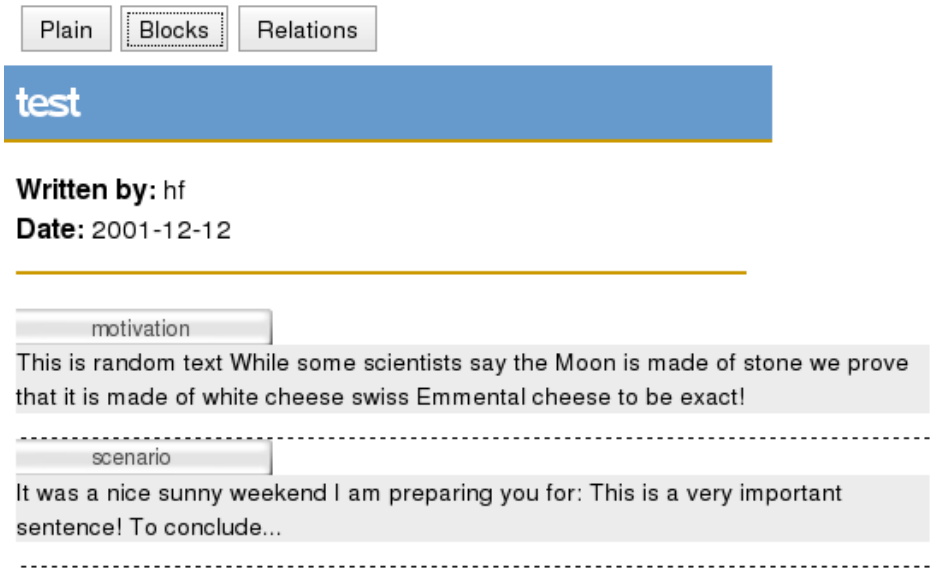


Sections in the editor
(Implementation by Gordan Ristovski)



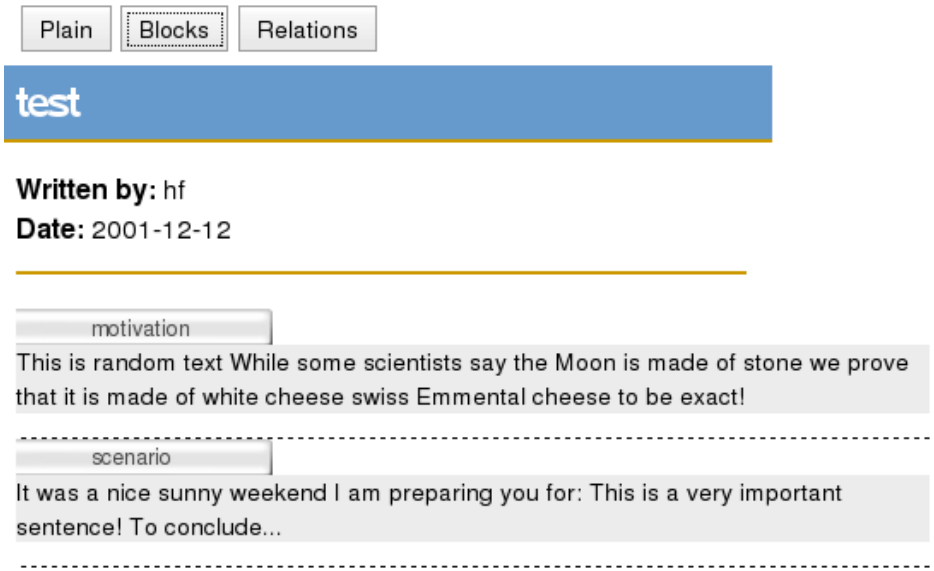
The toolbar

Visualisation of Rhetorical Structures



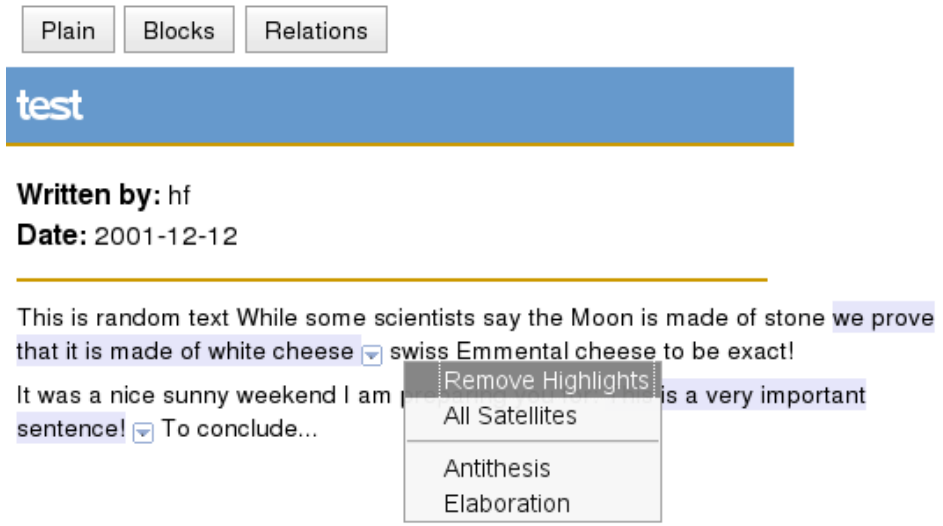
Rhetorical Blocks (SALT)

Visualisation of Rhetorical Structures



Rhetorical Blocks (SALT)

Rhetorical Relations
(SALT, implementing RST)



(Implementation by Jana Giceva)

Argumentation about Mathematical Knowledge

Idea

Need for structured wiki discussions, well-defined workflow for solving problems with knowledge in a wiki

My Case

- a wiki page is an item of mathematical knowledge, e. g. a theorem
- issues discussed will be quite specific: e. g. "This theorem is hard to understand" (or wrong, or inadequately presented, . . .)

Argumentation about Mathematical Knowledge

Idea

Need for structured wiki discussions, well-defined workflow for solving problems with knowledge in a wiki

My Case

- a wiki page is an item of mathematical knowledge, e. g. a theorem
- issues discussed will be quite specific: e. g. "This theorem is hard to understand" (or wrong, or inadequately presented, . . .)

Related Topic

There is also argumentation *within* artifacts of scientific knowledge, but so far I focused more on argumentation *about* them.

Resource | Edit | **Discussion** | History

Resource | Edit | **Discussion** | History

Issue	Alice	2008–05–30	[Idea][Argument][Agree][Disagree][Decision]
It's hard to find out how to improve content (= resources) in wikis ☹			

Resource | Edit | **Discussion** | History

Issue Alice 2008–05–30 [Idea][Argument][Agree][Disagree][Decision]

It's hard to find out how to improve content (= resources) in wikis ☹

Agree Bob 2008–05–31

Indeed, besides automated approaches it's hard to get focused feedback from users.

Resource | Edit | **Discussion** | History

Issue Alice 2008–05–30 [Idea][Argument][Agree][Disagree][Decision]

It's hard to find out how to improve content (= resources) in wikis ☹

Agree Bob 2008–05–31

Indeed, besides automated approaches it's hard to get focused feedback from users.

Idea Claire 2008–06–01 [Argument][Agree][Disagree][Decision]

So let's make wiki discussions semantic!

Resource | Edit | **Discussion** | History

Issue Alice 2008–05–30 [Idea][Argument][Agree][Disagree][Decision]

It's hard to find out how to improve content (= resources) in wikis ☹

Agree Bob 2008–05–31

Indeed, besides automated approaches it's hard to get focused feedback from users.

Idea Claire 2008–06–01 [Argument][Agree][Disagree][Decision]

So let's make wiki discussions semantic!

Argument Dave 2008–06–02 [Agree][Disagree]

We could take types from an argumentation ontology for the posts.

Resource | Edit | **Discussion** | History

Issue Alice 2008–05–30 [Idea][Argument][Agree][Disagree][Decision]

It's hard to find out how to improve content (= resources) in wikis ☹

Agree Bob 2008–05–31

Indeed, besides automated approaches it's hard to get focused feedback from users.

Idea Claire 2008–06–01 [Argument][Agree][Disagree][Decision]

So let's make wiki discussions semantic!

Argument Dave 2008–06–02 [Agree][Disagree]

We could take types from an argumentation ontology for the posts.

Argument Eric 2008–06–03 [Agree][Disagree]

And every discourse should be connected to resources corresponding to the wiki page, and there should be domain-specific *Idea* and *Issue* subclasses.

Resource | Edit | **Discussion** | History

Issue Alice 2008–05–30 [Idea][Argument][Agree][Disagree][Decision]

It's hard to find out how to improve content (= resources) in wikis ☹

Agree Bob 2008–05–31

Indeed, besides automated approaches it's hard to get focused feedback from users.

Idea Claire 2008–06–01 [Argument][Agree][Disagree][Decision]

So let's make wiki discussions semantic!

Argument Dave 2008–06–02 [Agree][Disagree]

We could take types from an argumentation ontology for the posts.

Argument Eric 2008–06–03 [Agree][Disagree]

And every discourse should be connected to resources corresponding to the wiki page, and there should be domain-specific *Idea* and *Issue* subclasses.

Agree Anonymous 2008–06–04

That's great, then the wiki could assist with the realisation of an approved idea. And old decisions would be documented.

Resource | Edit | **Discussion** | History

Issue Alice 2008–05–30 [Idea][Argument][Agree][Disagree][Decision]

It's hard to find out how to improve content (= resources) in wikis ☹

Agree Bob 2008–05–31

Indeed, besides automated approaches it's hard to get focused feedback from users.

Idea Claire 2008–06–01 [Argument][Agree][Disagree][Decision]

So let's make wiki discussions semantic!

Argument Dave 2008–06–02 [Agree][Disagree]

We could take types from an argumentation ontology for the posts.

Argument Eric 2008–06–03 [Agree][Disagree]

And every discourse should be connected to resources corresponding to the wiki page, and there should be domain-specific *Idea* and *Issue* subclasses.

Agree Anonymous 2008–06–04

That's great, then the wiki could assist with the realisation of an approved idea. And old decisions would be documented.

Decision Christoph 2008–06–05

So let's do it! (Available in SWiM)

Domain-Specific Argumentation

Assumptions

- Possible problems depend on the type of knowledge item
- Possible solutions depend on the type of knowledge item *and* the type of problem
- Standard problems have standard solutions, with which software can assist

Domain-Specific Argumentation

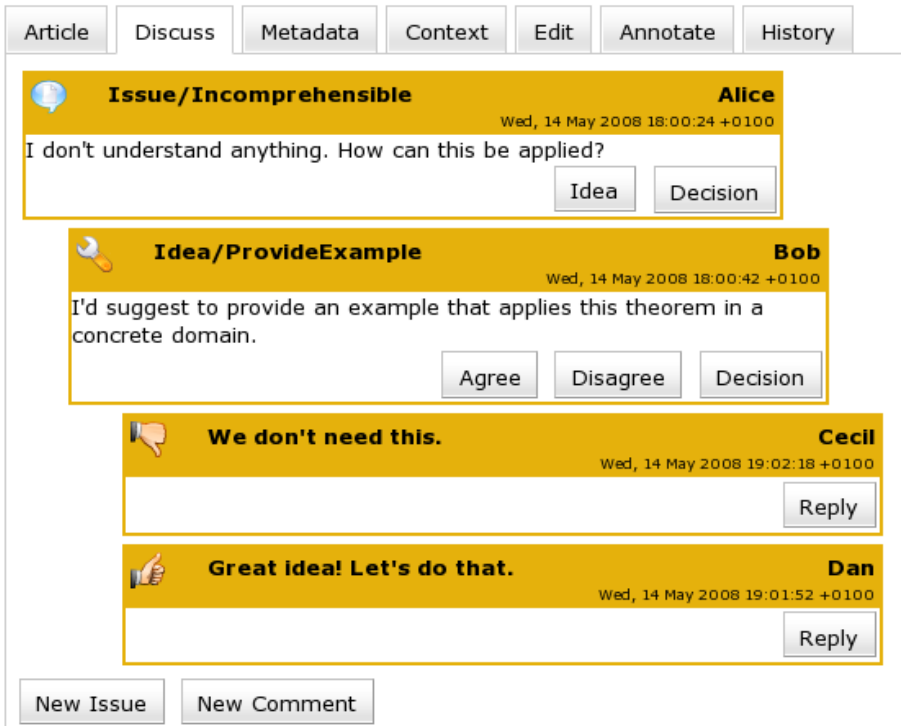
Assumptions

- Possible problems depend on the type of knowledge item
- Possible solutions depend on the type of knowledge item *and* the type of problem
- Standard problems have standard solutions, with which software can assist

Survey (`tinyurl.com/5qdetd`)

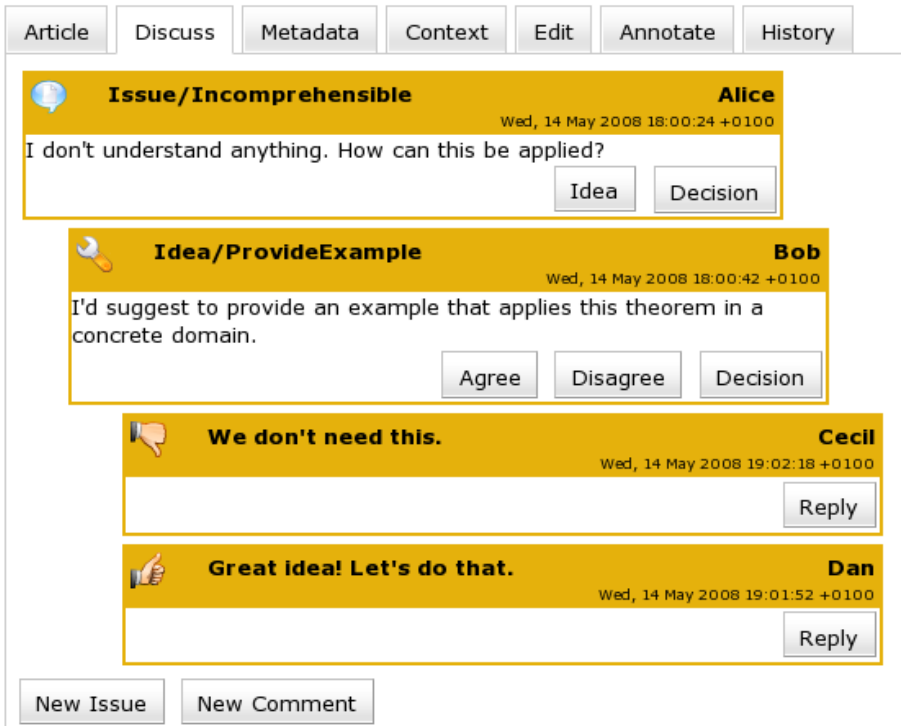
- Common issues: wrong, incomprehensible, uncommon style, underspecified, redundant, truth uncertain
- Common solutions: directly improve affected knowledge item, split it
- When issues remain unresolved, it's mostly due to insufficient restructuring support

Domain-Specific Argumentation (Example)

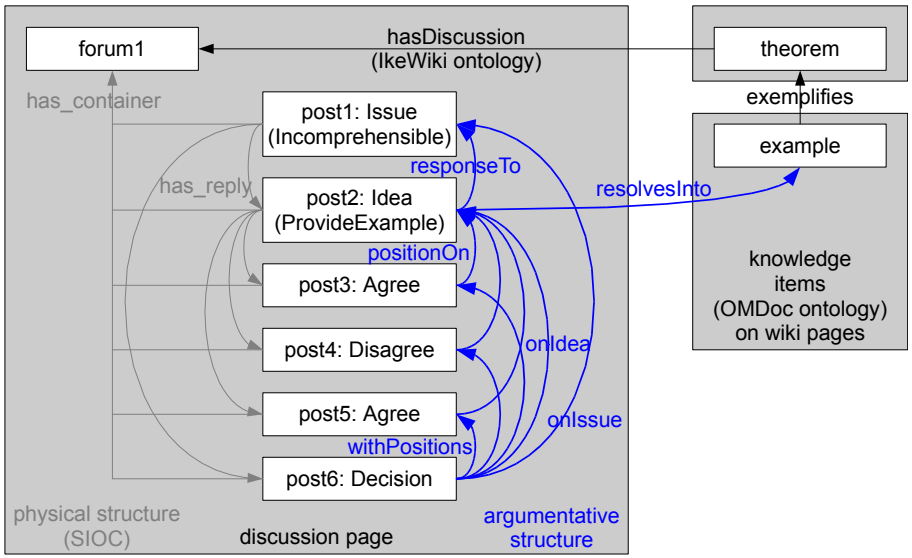


User Interface

Domain-Specific Argumentation (Example)



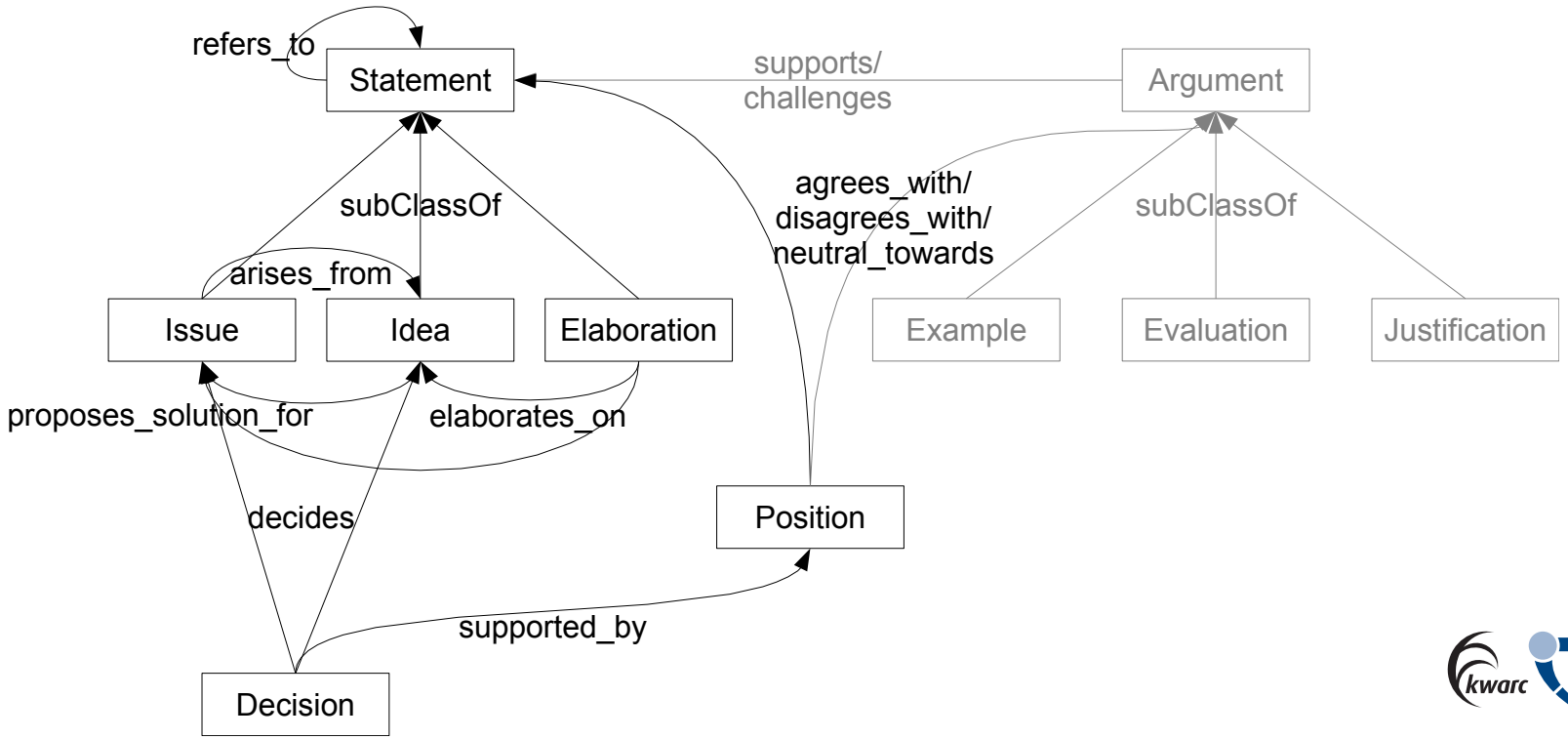
User Interface



RDF Graph

General Argumentation on Social Media Sites

- developing an argumentation module for SIOC (ontology for Semantically Interlinking Online Communities)
- joint work with Uldis Bojārs (SIOC) and Tudor Groza (SALT)
- use cases, model, guidelines for usage
- implementation and evaluation to be done



OpenMath Case Study

- lightweight mathematical ontology engineering (`http://wiki.openmath.org`)
  - no rhetorical structures, no documents
    but still a lot of structures to annotate!
  - definitions, formal properties, examples, notations
  - local argumentation
- small group of knowledge engineers (domain experts)
- specialised editors: structured definitions, formulas, metadata
- evaluation needed

Summary

What I hope(d) to learn – to use it for mathematical knowledge management:

- engineering ontologies for scientific documents ✓
- user interfaces for annotating and browsing ✓
- relation of social interaction to knowledge ✓

Summary

What I hope(d) to learn – to use it for mathematical knowledge management:

- engineering ontologies for scientific documents ✓
- user interfaces for annotating and browsing ✓
- relation of social interaction to knowledge ✓

What I hope to contribute to the semantic web:

- mathematics as a complex use case pointing out limits of the semantic web
- an ontology for a complex domain, with document structure, mathematical structure, and rhetorical structure
- domain-specific argumentation

Further Work

Active Documents

Interactive editing and previewing of notations

Further Work

Active Documents

Interactive editing and previewing of notations

Argumentation

Study relationship between argumentation within and about documents

Further Work

Active Documents

Interactive editing and previewing of notations

Argumentation

Study relationship between argumentation within and about documents

Ontologies

1. Scalable metadata syntax and semantics for OMDoc
   $\longrightarrow$ import metadata vocabularies as theories
2. Document these vocabularies in OMDoc
3. Model them in OMDoc
4. Export them back to the semantic web

Further Work

Active Documents

Interactive editing and previewing of notations

Argumentation

Study relationship between argumentation within and about documents

Ontologies

1. Scalable metadata syntax and semantics for OMDoc
   $\longrightarrow$ import metadata vocabularies as theories
2. Document these vocabularies in OMDoc
3. Model them in OMDoc
4. Export them back to the semantic web

Semantic Web Empowering MKM