Arguing on Issues with Mathematical Knowledge Items in a Semantic Wiki

Christoph Lange
School of Engineering and Science
Jacobs University Bremen
D-28759 Bremen, Germany
ch.lange@jacobs-university.de

Tuukka Hastrup and **Stéphane Corlosquet**
Digital Enterprise Research Institute
National University of Ireland
Galway, Ireland
{tuukka.hastrup,stephane.corlosquet}@deri.org

Abstract

In informal collections of collaboratively created knowledge like wikis, there is no well-defined way of reporting issues with knowledge items. When something is wrong or needs improvement, users are hardly supported to communicate this in a focused way: discussing about an issue, coming up with ideas on how to solve it, agreeing on the best idea, and finally putting this idea into practice in a retraceable way. This workflow is sometimes standardised in terms of best practices and social conventions, but not supported by the system in terms of knowledge management.

We present an approach to improving this in a semantic wiki, where not only the articles contain structured knowledge, but also the discussions *about* this knowledge are structured using an argumentation ontology. We show how, by domain-specific extensions of this ontology, the wiki system can not only support a focused discussion about issues but also assist with putting solutions approved by the community into practice in many common cases. This is demonstrated on the prototype of the SWiM system in a use case from mathematical knowledge management.

1 Introduction

This article deals with *issues* with knowledge items in collaborative knowledge management systems – how to report them, how to argue about them, and how to resolve them[1]. Examples of issues could be that a knowledge item is hard to understand or consists of wrong facts, that redundancies with other knowledge items are identified, or that a subpart of one knowledge item is considered to deal with a topic of special interest, deserving to be promoted to a knowledge item of its own. Users of the system report issues and argue about them, propose solutions that are again subject to discussion, until finally a solution is approved and implemented. Such a discourse can be lengthy and hard to keep focused, as issues can be "wicked problems", exposing traits like not allowing for a "definininitive formulation", having solutions that are "not true-or-false but good-or-bad", and the nonexistence of an "immediate and [. . .] ultimate test of a solution" [Rittel and Webber, 1973]. A solution is usually materialised in an improved version of the affected knowledge item or a new knowledge item. Later, other users, who want to understand why some knowledge item is modelled in a particular way, can trace back the discourse that led to its creation or modification. Thus, the discussions about issues with knowledge items become part of the collective experience of the community.

We investigate a model for structured argumentation about issues with knowledge items and present a system that can assist users with the implementation of solutions in common cases. We try to keep our model general but particularly investigate it in the setting of a wiki for mathematical knowledge. In a wiki, one page usually holds knowledge about one distinct topic, or about a set of closely related topics [Ebersbach *et al.*, 2008]. Some typical knowledge items in mathematics are definitions of symbols or concepts, theorems, and proofs. Issues with them can be that a knowledge item is wrong, incomprehensible, presented in an uncommon style, or redundant (cf. section 4.2).

We continue with a motivation why issue handling is insufficient in existing wikis, an introduction to semantic wikis and argumentation ontologies, which we consider helpful in that regard, a theoretical outline of our approach, and a description of a prototypical implementation. Then, we walk through a typical use case to demonstrate the approach in our implementation.

2 Motivation: Wikis and Wikipedia

In unstructured collaborative knowledge bases, such as conventional wikis, there is no well-defined way of reporting issues with knowledge items. Many systems allow for *tagging* knowledge items (e. g. as "needs improvement"), or for *commenting* on them in more detail. In wikis, users can insert a warning message directly into a page affected by an issue, and probably add a detailed explanation or justification to the discussion page that is usually associated with every content page. Technical support is mostly limited to the possibility of creating building blocks for such warning messages, which can then be transcluded into pages, and to a button that allows for adding a new section to a discussion page which otherwise does not enforce any structure – see, e. g., MediaWiki [MediaWiki, 2008]. The discourse itself proceeds without technical support. The community is left alone with devising reasonable issue warning messages and establishing a workflow of reporting, discussing, and solving issues and documenting the solutions. This is mostly done by jointly agreeing on best practices in conflict resolution and authoring, and making them official policies for the community [Kittur *et al.*, 2007].

As a concrete example, consider a Wikipedia article that violates the Wikipedia principle of a neutral point of

[1] We will henceforth refer to a piece of knowledge about any distinct subject of interest as a "knowledge item".

view [Wikipedia, 2008d][2]. Some author who is concerned about this can tag the article by inserting the building block "Neutralität" (neutrality [Wikipedia, 2008b]). It is then recommended to justify why the neutrality of the article is debated by adding a respective section to the discussion page of the article. Within that section, the general conventions for discussions pages apply [Wikipedia, 2008c]: The author has to make clear what section of the article his discussion post applies to, he has to verbalise his report in a comprehensible way, and finally has to append his signature (a link to his user profile with a timestamp). An author who wants to discuss an existing issue has to look up the corresponding section on the discussion page and then indent his reply by one more level than the post he is replying to. Solutions to issues would be proposed in natural language only, and if users come to vote on proposals, they would do it in an ad hoc manner, e. g. using list items prefixed with "yes" or "no". A solution for restoring the neutrality of a controversial article could be citing reliable arguments in favour of the view that has been less represented so far. Eventually, one author who is trusted by the community would judge whether there is a consensus about a particular solution, or simply count the votes, and then implement the solution approved by the community, again without any assistance from the system. A justification for the resulting revision of a page can be given by a descriptive editing comment that links to the section of the discussion page where the respective issue was discussed [Wikipedia, 2008a]. However, authors do not always do this, which sometimes makes it hard to retrace decisions.

Note that for procedures with a more serious effect, above all the deletion of an article, it is more highly regimented who may implement a solution. Only users with administrative permissions, which are awarded by public vote, may technically do so. However, in the remainder of this article we will not assume any such technical restrictions but assume well-behaved and cooperative users. Encouraging or enforcing orderly behaviour is an interesting research question in itself but not considered here.

In large knowledge collections like Wikipedia, these procedures work sufficiently thanks to the large user base; indeed, the quality of articles has been found to strongly correlate with the number of authors [Brändle, 2005]. We are aiming at knowledge collections that probably have a small user base but are operated by a system capable of certain knowledge management tasks: a system that has a basic understanding of what types of issues are reported with what types of knowledge items, what solutions are proposed, and whether people agree or disagree with these proposals.

3 Foundations

3.1 Semantic Wikis

In a semantic wiki, the knowledge is more structured than in a non-semantic one. Other than employing a knowledge representation based on semantic web technologies such as RDF [W3C, 2004] or ontologies (see e. g. [Fensel, 2003]), semantic wikis are quite diverse: In some systems, shallowly annotated text prevails, whereas in others, formal knowledge prevails and unstructured text only appears in comments or labels that describe formal concepts [Oren *et al.*, 2006; Buffa *et al.*, 2008]. Even others mix annotated text and highly formalised problem-solving knowledge [Baumeister and Puppe, 2008]. The most common approach is, however, to represent knowledge about one subject of interest – a "knowledge item", in the terminology of this article, – by one wiki page and to annotate pages and links between pages with types defined in an ontology. In this kind of semantic wikis it is advisable to keep pages small and refactor them if they tend to describe more than one knowledge item. The graph of typed pages and links is commonly represented in RDF [W3C, 2004]. Existing ontologies, such as Friend of a Friend [FOAF, 2008], are either preloaded into the wiki or imported later, or a new, custom ontology is built collaboratively during the annotation of the wiki pages. Semantic MediaWiki, for example, prefers the latter approach, where an ontology is implicitly extended whenever a page or a link is assigned a new type [Völkel *et al.*, 2006].

3.2 Semantic Discussion Threads with SIOC

To the best of our knowledge, there is currently only one semantic wiki where discussions are semantically structured, too. In IkeWiki [Schaffert, 2006], the relationship between the knowledge item represented by a page and the discussion about it is represented in the RDF graph, and the associated discussion page itself is not an opaque block of text, but a self-contained discussion forum. It consists of threads and posts, with RDF links to the user profiles of their authors. This is achieved using the SIOC ontology (Semantically Interlinked Online Communities [Breslin *et al.*, 2006]), which models generic aspects of various forms of discussion forums. SIOC itself does not yet allow for typing discussion posts as issues, proposed solutions, agreements, etc., but is meant to be extended by dedicated ontologies that assign more specific types to discussion posts [Bojārs and Breslin, 2007, sect. 4.2]. Indeed, making every post a distinct RDF resource and preserving the threaded structure of a discussion serves as a basis for adding an argumentative layer, as we will show in the following.

3.3 Argumentation Ontologies

An early approach at formalising argumentation about issues was IBIS (Issue-Based Information System [Kunz and Rittel, 1970]), which particularly aimed at wicked problems [Rittel and Webber, 1973][3] The DILIGENT argumentation ontology was conceived in the context of the ontology engineering methodology of the same name as an extension of IBIS that makes arguments more focused, thus making design decisions more traceable and allowing for inconsistent argumentations to be detected [Tempich *et al.*, 2005; 2007]. We found the DILIGENT vocabulary suitable for our setting of structured knowledge engineering. But note that our understanding of an *issue* differs from the DILIGENT methodology: There, issues are issues of conceptualising new aspects of a domain, and ideas refer to the formalisation of these conceptualisations, according to the well-known definition of an ontology [Gruber, 1993], whereas in our case issues are issues with more or less formalised knowledge items that already exist.

[2]As the audience of this article is mainly German and the different language editions of Wikipedia have developed slightly different conventions, we give references to the German Wikipedia. Pointers to related information in other Wikipedias can be found on the respective pages by following the links to other languages ("Andere Sprachen").

[3]Note that not all characteristics of wicked problems apply in our case, as they were originally investigated in governmental planning [Rittel and Webber, 1973], a domain that lends itself less well to formal modelling than knowledge engineering.

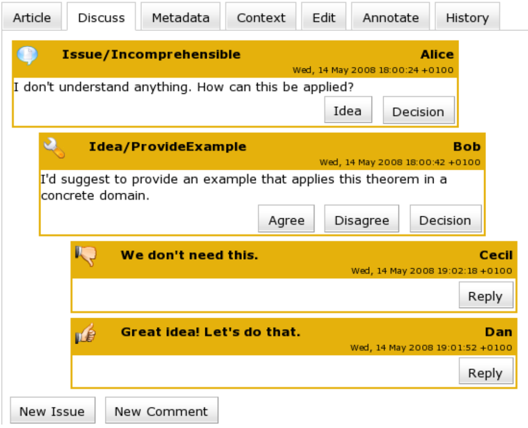

Figure 1: In the middle of a discourse

A discourse in terms of the DILIGENT argumentation ontology is structured as follows: When an issue has been raised, collaborators can express their agreement or disagreement with it, i. e. whether they consider this issue important, justified, and legitimate. An issue can be resolved by implementing a proposed and – again by posting agreements – approved idea in the space of knowledge items (called "ontology entities" in DILIGENT) and concluding the discussion thread with an explanation of the decision taken. This decision will link to the issue that has been solved and to the idea i that was realised. If that idea was to create or modify a knowledge item k, a link "i resolves into k" will be created. Besides merely agreeing or disagreeing with an issue or idea, collaborators can also *argue* about it, i. e. justify it by examples or evaluations, or challenge it by alternative proposals or counter-examples, and others can again agree or disagree with these arguments.

Our approach is to enhance SIOC discussions in a semantic wiki based on IkeWiki with the DILIGENT argumentation ontology and domain-specific extensions to it, as we will show in the following.

4 Approach

4.1 Domain-Specific Ontology

We assume that every knowledge item has at least one principal type and that these principal types are disjoint with each other. Concrete types heavily depend on the application domain; in this article, we focus on mathematics and consider mathematical statements as atomic knowledge items, a statement being a definition, theorem, proof, etc. We take these types from the OMDoc ontology [Lange, 2008b; 2008c] that we have modelled for representing mathematical knowledge originating from documents in the OMDoc semantic markup language [Kohlhase, 2006] on the semantic web. This ontology abstracts from the syntax of the OMDoc language and aims at formally modelling aspects of its semantics that cannot be expressed in an XML schema (e. g. that the target of a link from a proof to the resource that is proven must be a theorem, and that certain links between knowledge items constitute dependencies) but on the other hand are too informal for being meaningful to, e. g., an automated theorem prover. Additional types of knowledge items, e. g. its status in terms of project management (e. g. "draft", "under review", "published"), will not be taken into account for reporting and resolving issues.

We assume that there can be a threaded, SIOC-structured discussion about every knowledge item and that the discussion posts can be typed as instances of classes from the DILIGENT argumentation ontology. In our current model, we consider the classes *Issue*, *Idea*, *Agree*, *Disagree*, and *Decision* (fig. 2). To establish a bridge between the domain of knowledge and the argumentation about it, we created an ontology of domain-specific subclasses of DILIGENT's *Issue* and *Idea* class (fig. 2). A particular type of issue is considered applicable to certain types of knowledge items; we model this in the ontology as well. For example, an issue with a mathematical proof can be that it is wrong, whereas the notation of a symbol cannot be wrong but inappropriate, misleading, or hard to read or write. Furthermore, we assume that to a pair of a knowledge item type and an issue type, certain types of ideas can be applied. For example, if a proof is wrong, it could be deleted and replaced by a correct proof, or it could be kept as an instructive bad example. Obviously, we do not expect to cover *all* possible cases with a finite set of predefined issue and idea types, but the most common ones.

4.2 A Survey on Issues in Mathematics

To get an understanding of common issues and solutions in mathematical knowledge management, we are currently conducting a survey among domain experts[4]. We collect information about the previous experience of the participants with mathematical knowledge bases, the support for tracking and solving issues in the tools they have used, types of knowledge items they have dealt with, types of issues they have encountered, how these issues were solved, and reasons why issues remained unsolved. So far, 48 people have participated, 24 having answered all questions. A majority is experienced in contributing to libraries of software tools like automated theorem provers, but many participants have also contributed to wiki-like knowledge bases. The most commonly experienced granularity of knowledge items is either a course unit, a mathematical theory (i. e. a few related definitions and axioms), or a mathematical statement. Only in a few cases there was support for automated issue tracking and solving. The prevalent type of knowledge item that the participants have ever found affected by issues was a definition – theorems, proofs, examples, notation definitions for rendering symbols, and theories also being quite common. The most common issue was that a knowledge item was simply wrong, followed by being incomprehensible, its truth being uncertain, being underspecified, or redundant. Issues were mostly solved by directly improving the affected knowledge item (as opposed, e. g., to creating another one), by splitting it into more than one, or by deleting it altogether. Still, some participants have experienced issues being unresolved and mostly attributed this to an insufficient tool support for restructuring knowledge items.

4.3 Automated Assistance

Whenever there is a discourse about a knowledge item, the system will check whether there is an issue that is both unresolved (meaning that no decision on it has been posted yet) and not challenged as invalid by the existence of a majority of disagreement replies to it. If ideas have been posted on how to resolve this issue, the most popular in terms of the ratio of agreements to disagreements will be selected. Formally, any issue s satisfying $\mathrm{D}(s) = \emptyset \wedge \left(\mathrm{P}^-(s) \neq \emptyset \Rightarrow \left|\mathrm{P}^+(s)\right| > \left|\mathrm{P}^-(s)\right|\right)$ is considered legitimate, and the idea i in $\underset{i \in \mathrm{Id}(s),\, \mathrm{P}^+(i) \neq \emptyset}{\arg\max} \frac{\left|\mathrm{P}^+(i)\right|}{\left|\mathrm{P}^-(i)\right|+1}$

[4]See http://tinyurl.com/5qdetd

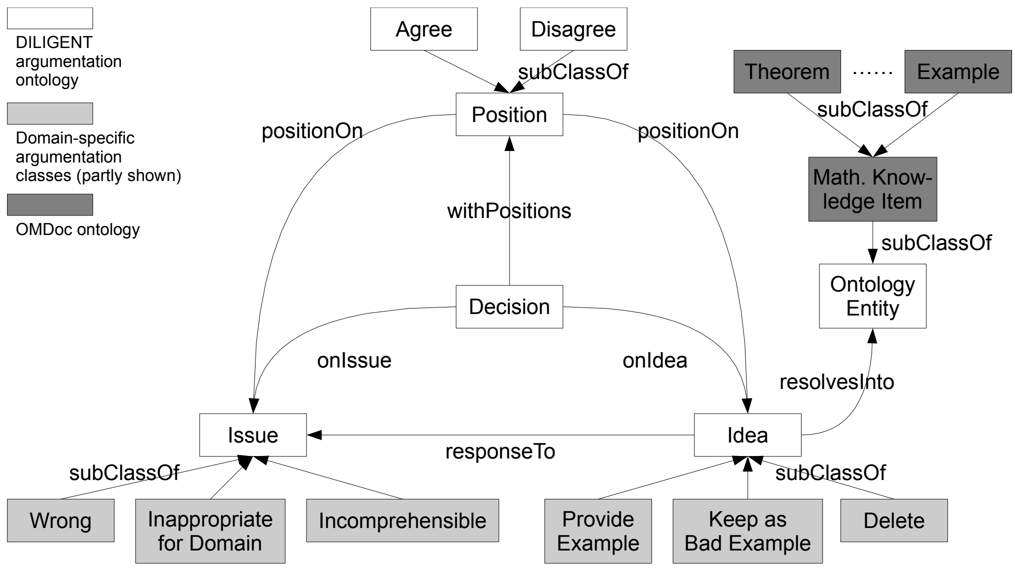

Figure 2: The DILIGENT/OMDoc argumentation ontology

wins, where Id, D, P^+, and P^- denote sets of ideas, decisions, agreements and disagreements with an issue or idea, respectively[5]. The system will provide assistance to any volunteering author to implement the winning solution in the space of knowledge items, e. g. by automatically creating a template for a new knowledge item that the author can then complete. If an author follows the steps proposed by the system, the system will conclude the respective discourse by posting an automatically generated decision. Still, we leave the freedom to the community to implement solutions manually, when users feel that the automatic support is not adequate to the wickedness of the current problem. In this case, the author to resolve an issue has to document this decision manually. Any thread that has been concluded by a decision will no longer be considered by the system.

5 Implementation

We have implemented a proof of concept in the mathematical semantic wiki SWiM. SWiM extends IkeWiki (cf. section 3.1) with mathematical semantic markup. It supports editing, viewing, and importing/exporting mathematical knowledge that is represented in the OpenMath [Buswell *et al.*, 2004] and OMDoc markup languages [Kohlhase, 2006] and has ontologies corresponding to these languages built in. One SWiM page usually holds one mathematical statement. From a saved or imported page, SWiM extracts an RDF outline in terms of these ontologies and uses this RDF graph to support semantic navigation through mathematical knowledge items and to offer an infrastructure for more sophisticated semantic services [Lange, 2008c].

We have extended the user interface for discussions by the possibility to make not just untyped comments but to post issues or ideas of specific types and to state one's agreement or disagreement (fig. 1). The OWL-DL implementation of the DILIGENT argumentation ontology, in the current prototype without the classes for fine-grained argumentation about issues and ideas, but enhanced by mathematics-specific issue and idea types (cf. fig. 2), is preloaded into SWiM; discussion posts are represented in RDF as instances of these types. Thus, the structure of the argumentation forms an overlay network on top of the raw structure of the threads represented in SIOC. On the top level of a discussion page, the user is invited to post issues of types that are applicable to the type t_k of the knowledge item to be discussed; a reply to an issue t_{is} can have one of the idea types that are applicable to (t_k, t_{is}). Thanks to IkeWiki's built-in ontology editor, privileged members of the community can even dynamically and interactively adapt the argumentation ontology to the community's needs. The formulæ for determining unsolved legitimate issues and "winning" ideas are implemented as sequences of SPARQL [Prud'hommeaux and Seaborne, 2008] queries to the RDF graph of the discussion about the knowledge item currently viewed. The assistance with implementing a solution is hard-coded into SWiM for now. Some functions, such as deleting an article, had been available before, whereas the creation of a related knowledge item (e. g. an example) has been implemented as a first proof of concept for additional assistance.

To demonstrate the system, we consider the situation that there is an incomprehensible theorem. The user Alice wants to report that issue. She opens the discussion page for the theorem and posts a new issue. As a type of the issue, she can select any type that is applicable to theorems. Then, Bob replies by clicking the "Idea" reply button in the issue post and selecting an idea type; any type of idea that is applicable to incomprehensible theorems can be selected. Cecil disagrees with the idea, Dan agrees (fig. 1). Now assume that Eric replies to the idea with another agreement and that, after that, Frank visits the theorem: By then SWiM will have identified the idea to provide an example for the theorem as the best one to resolve the issue and display a message that proposes this, offering a link to start a semi-automatic assistant (fig. 3). If Frank decides to provide the example and clicks on the link, a new example page, pointing to the original theorem, would be created, and he can fill out the template (fig. 4).

6 Evaluation and Discussion

As this work is still in progress, we focus our evaluation on discussing the extent of argumentation and assistance that our ontology and the implementation support. In section 8, we will provide an outlook to domain-specific case studies

[5] It is not yet clear what idea should be preferred when there is more than one such i.

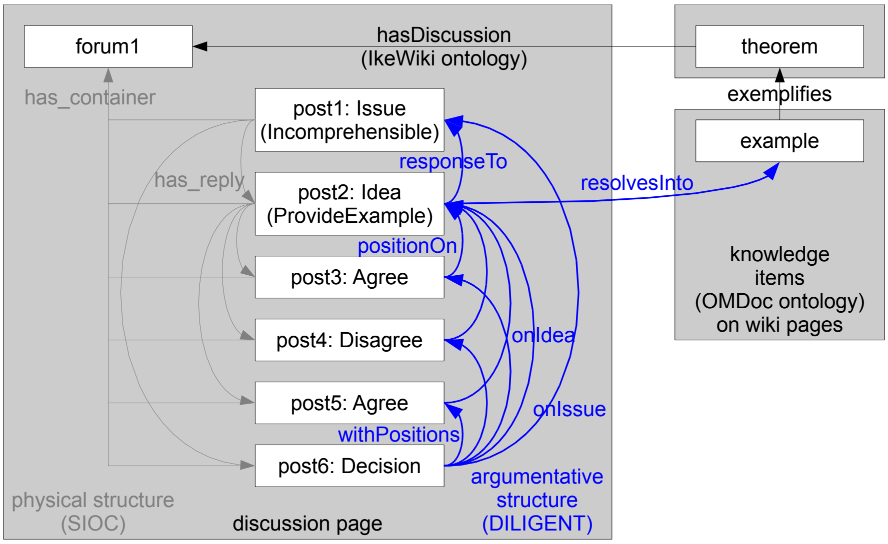

Figure 5: RDF graph of a sample use case

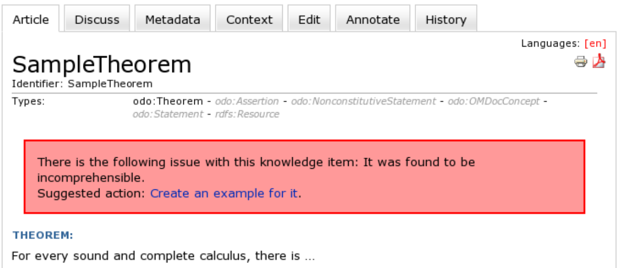

Figure 3: Warning about an issue and the offer to solve it

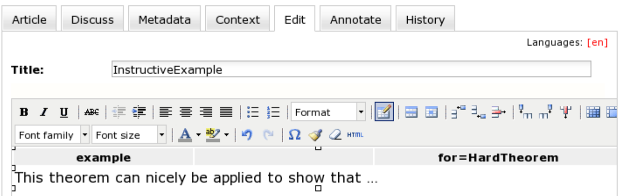

Figure 4: Editing the newly created example

that we are planning to conduct ourselves. Further feedback is anticipated from users of IkeWiki, where we have recently integrated the basic argumentative discussion functionality without domain-specific extensions.

For the sake of simplicity, arguments about issues and ideas have been left out in our first prototype in favour of simple agreement or disagreement, but we are planning to complete the coverage of the DILIGENT argumentation ontology. Following the findings of the DILIGENT authors, we do not, however, consider the currently restricted set of available domain-specific issue and idea types an obstacle; we think it just has to be refined to cover the most common situations in mathematical knowledge management. Concerning their restricted set of argument types (e. g. "challenge" and "justification"), the DILIGENT authors have found out in their case studies that this made discussions more effective and focused [Tempich *et al.*, 2005]; we assume that the same will turn out for our issue and idea types. Note that our current formula for selecting the best idea is quite simplistic. Once arguments will have been introduced, it will no longer suffice and should be replaced by a more sophisticated weighting function like the one proposed in [Gordon and Karacapilidis, 1997].

As issues in our model refer to knowledge items, and the assistance the system can offer depends on the type of knowledge item, the knowledge must already be structured to some extent. In two situations, this is likely not to be the case: When knowledge about some *new* topic has not even been conceptualised, or when it has been conceptualised and put on a wiki page, but not yet formalised (here: annotated with a type). To improve on this, we consider providing a global discussion space, where conceptualisation issues can be raised, as it was originally intended with DILIGENT (cf. section 3.3), and to introduce a generic issue type "needs formalisation", which can be filed with any knowledge item that does not yet have a type from the domain-specific ontology. *Assistance* in the latter case would likely require techniques like natural language processing (NLP), which we have not yet considered: Once a knowledge item or discussion post has been given a type, we base all further decisions about assistance exclusively on this type. Obviously, this is only meaningful if the informal text of a knowledge item does not contradict its formal type. Currently, we hold the users responsible for that, but NLP would help here, too.

Finally, we object that, while the *model* that we adopted allows for arguing on wicked problems, those issues for which our system is actually able to offer semi-automatic assistance are not really wicked ones by definition. For a solution to be supported by the system, users are, first of all, required to clearly pinpoint the issue using a type from the domain-specific extension of the argumentation ontology – contradicting the traits of a wicked problem to have "no definitive formulation" and to be "essentially unique" in the sense that no pattern of solution could be applied to a whole class of problems [Rittel and Webber, 1973]. For every type of issue, we support a set of predefined solution patterns, but a wicked problem does "not have an enumerable (or an exhaustively describable) set of potential solutions" [Rittel and Webber, 1973]. In supporting solutions we currently focus on one issue with one knowledge item and neglect the previous history or related issues and knowledge items; however, "every wicked problem can be considered to be a symptom of another problem" [Rittel and Webber, 1973]. Then, our system currently assumes that an issue has definitely been solved once a "decision" reply exists – contradicting the definition that "wicked problems have no stopping rule" [Rittel and Webber, 1973]. However, a user who considers an existing solution inadequate is always free to file a new issue on the affected knowledge item.

7 Related work

Several projects are related to ours, in different aspects: the Foucault@Wiki study [Pentzold and Seidenglanz, 2006] about Wikipedia discourses, Lekapidia [Tempich *et al.*, 2007] and the Cicero system [Dellschaft *et al.*, 2008] about DILIGENT argumentations in a wiki, Collaborative Protégé [Tudorache and Noy, 2007] for collaborative ontology engineering, Drupal as a content management system with a plugin for issue tracking [Drupal,], and finally panta rhei [panta rhei, 2008] as a different approach towards a discussion platform for structured mathematical knowledge.

Foucault@Wiki was a study that analysed Wikipedia discourses: how edit summaries and discussion posts related to changes made to Wikipedia articles, and what types of changes occurred [Pentzold and Seidenglanz, 2006]. From this, a model for Wikipedia argumentations – in fact an informal site-specific argumentation ontology – was derived. This analysis had to be made without machine support, as Wikipedia articles are largely unstructured, and discussions and edit summaries are given in natural language, and the space of possible arguments is unrestricted, as opposed to the finite set of types in a formal argumentation ontology. Note that the goal was not to design software support for discourses or for improving knowledge items, i. e. Wikipedia articles. We consider that hard to achieve, as Wikipedia articles frequently contain multiple sections that do not focus on one knowledge item only, which makes it hard to express what part of an article a discussion post refers to.

Lekapidia was a case study in which the authors of the DILIGENT argumentation ontology preloaded a semantic wiki (coefficientMakna) with the argumentation ontology. Then, they replayed the collaborative engineering of a simple dessert recipe ontology, which had earlier been developed using the DILIGENT methodology, in their wiki, and found out that the wiki "significantly reduces the effort to capture the arguments in a structured way" [Tempich *et al.*, 2007]. Special attention was paid to detecting inconsistent argumentations (e. g. when one user first votes in favour of one argument a but then introduces a new one that contradicts a [Tempich *et al.*, 2005]) and fostering consensus, both of which we have not yet investigated in detail.

A collection of semantically structured mathematical knowledge can be considered an ontology, particularly if it contains formal definitions of mathematical concepts. The argumentation ontology in SWiM is particularly enhanced for the domain of mathematics, which allows for creating more specific issues and ideas in a machine-understandable way and for implementing semi-automatic assistance in implementing certain of these specific ideas into the system.

Cicero is an extension for Semantic MediaWiki that allows for a DILIGENT-like argumentation about issues in projects (not issues with knowledge items) [Dellschaft *et al.*, 2008]. Cicero offers versatile options for voting and deciding. Before a first deadline set by a project administrator, users can propose solutions for an issue and argue about them; then, voting on the issues – single vote or multiple choice – takes place until a second deadline is reached. The ontology is currently only available in the wiki, not implemented externally.

Collaborative Protégé is a collaborative ontology editor featuring a "change and annotation ontology" [Tudorache and Noy, 2007]. This is not an argumentation ontology in the strict sense, but also allows for documenting design decisions. Two ways of voting on change proposals are available: a "5-star" rating and a "yes/no" vote.

Drupal is a modular framework and content management system [Drupal,]. Thanks to its great variety of modules, Drupal can be adapted flexibly to different requirements. Knowledge mainly consists of nodes, comments, and taxonomy vocabularies, each of which can easily be enhanced. The primary type of content is a node. Out of the box, Drupal can let users comment on the content of the node, or reply to other comments resulting in threaded conversations – which is unsupported by most wikis. Drupal also offers many hooks to let developers jump in during the page generation, validation or submission process and that way react precisely according to the input of the user. The project issue tracking module is already using these features, along with some categorisation[6].

panta rhei is an interactive and collaborative reader for mathematical documents, currently being evaluated in the educational context of a computer science lecture [panta rhei, 2008; Müller, 2008]. With SWiM, it shares the same domain of application – in fact, either system can import OMDoc – and in the possibility to create typed discussion posts about knowledge items; differences lie in the use of knowledge items – arranged for reading in panta rhei, whereas the main concern of SWiM is to edit them – and the different knowledge representation: SWiM relies on an underlying RDF model and an ontology for browsing, searching, and editing workflows, whereas panta rhei uses hand-made SQL database queries.

A panta rhei page usually contains one exercise or a sequence of one to a few lecture slides, mostly containing some related mathematical statements, such as the definitions of a few symbols with an example where they can occur, or a theorem followed by its proof. However, content authors can assign identifiers to any subitem of a page to make it annotatable. SWiM is less flexible here: Only knowledge items that have their own wiki page are annotatable. While imported documents are automatically split into pages of statement size, authors would have to do this manually to achieve a finer granularity.

In panta rhei, threaded discussion items can be posted on any annotatable knowledge item; in addition, there is a global forum. Each post has to be typed as, e. g., "advice", "answer", "comment", "example", or "question"[7]. The set of possible types of a post is not restricted by the type of the knowledge item nor the post it replies to. Compared to SWiM's use of an argumentation ontology, which encourages a targeted discussion towards solutions, this potentially makes discussion threads less focused. panta rhei currently uses the types of posts for statistical purposes and for search. Statistics are currently not computed in SWiM; semantic search is possible and powerful thanks to the support for inline SPARQL queries, but not yet friendly to users who do not know SPARQL. Another annotation-related feature of panta rhei without a SWiM counterpart is the ability to *rate* knowledge items on a scale from 0 to 10 w. r. t. several measures like their difficulty or helpfulness. While the argumentations in SWiM implicitly rate knowledge items, there

[6] `http://drupal.org/project/project_issue`

[7] This is a custom vocabulary made for panta rhei, not an argumentation ontology.

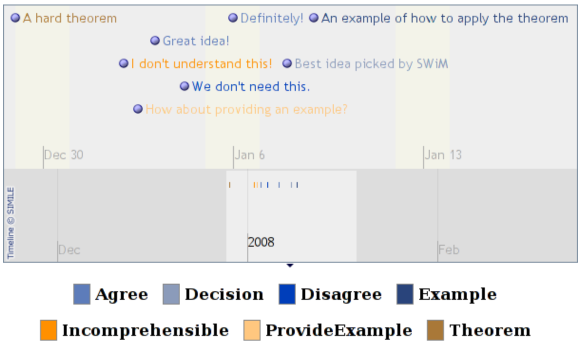

Figure 6: The Exhibit timeline view of fig. 5, starting at a hard theorem and ending at a helpful example.

is no agile, one-click user interface for this.

8 Conclusion and Outlook

We have justified the need for structured and annotated discussions in a collaborative structured knowledge base. With a domain-specific argumentation ontology, we have covered common types of issues that can occur with mathematical knowledge items, common types of ideas on how to solve these issues, and a semi-automatic assistance in implementing a solution approved by knowledge engineers. The discourse that led to the implementation of a solution can be traced back transparently from the affected knowledge items, which supports the experience management of the community. We have implemented and demonstrated a first proof of concept in the OMDoc-based semantic wiki SWiM but consider the methodology easily transferable to other collaborative knowledge management systems, as it is largely based on an abstract ontology.

As our system is built on RDF, semantic web technologies can be used for integration with other systems. For example, the Exhibit framework [Huynh *et al.*, 2007] can consume RDF data and provide us with a timeline visualisation of the argumentation process (fig. 6). If we follow the linked data guidelines for semantic web publishing, the knowledge items and discussions in our system will be visible to linked data crawlers such as Sindice [Tummarello *et al.*, 2007] and browsers such as Tabulator [Berners-Lee *et al.*, 2006]. If we can crosslink with other sites, discussions can refer to knowledge items across system boundaries.

Two practical settings, where we will evaluate our system, are the Flyspeck and OpenMath projects. Flyspeck is a large-scale proof formalisation effort concerned with developing a machine-verifiable representation of a proof of the Kepler sphere packing conjecture [Lange *et al.*, 2008a]. We have gathered first requirements for supporting this project with a wiki and consider support for discussing formalisation issues and for refactoring formal structures highly important. In OpenMath, SWiM will be used for editing content dictionaries (collections of definitions of mathematical symbols). Conceptualisation and formalisation of symbols and their notations need to be supported [Lange, 2008a].

A key question that we anticipate these case studies will answer is to what extent the automated identification of "winning" solutions and the support offered for implementing them will be found satisfactory by knowledge engineers – whether the 80/20 rule will apply (i. e. 80% of the everyday issues will be solvable with semi-automatic assistance implemented for 20% of all possible solutions), or whether a large percentage of issues turns out to be too wicked. In the latter case we still hope that the argumentation ontology will support a more focused and productive discussion about wicked problems than in a system with unstructured discussions and thus facilitate finding and implementing solutions even without further automatic assistance.

Acknowledgments We would like to thank John Breslin for advice with SIOC and argumentation, Markus Luczak-Rösch for guidance with using the DILIGENT argumentation ontology, Christine Müller for reviewing and helping to compare our system to panta rhei, Max Völkel for ideas about argumentation in a wiki, and Uldis Bojārs for suggestions how to extend our approach to the web.This work was supported by JEM-Thematic-Network ECP-038208.

References

[Baumeister and Puppe, 2008] Joachim Baumeister and Frank Puppe. Web-based knowledge engineering using knowledge wikis. In *AAAI 2008 Spring Symposium "Symbiotic Relationships between Semantic Web and Knowledge Engineering"*, pages 1–13, 2008.

[Bechhofer *et al.*, 2008] Sean Bechhofer, Manfred Hauswirth, Jörg Hoffmann, and Manolis Koubarakis, editors. *ESWC*, volume 5021 of *LNCS*. Springer, 2008.

[Berners-Lee *et al.*, 2006] Tim Berners-Lee, Yuhsin Chen, Lydia Chilton, Dan Connolly, Ruth Dhanaraj, James Hollenbach, Adam Lerer, and David Sheets. Tabulator: Exploring and Analyzing linked data on the Semantic Web. In *3rd Semantic Web User Interaction Workshop (SWUI)*, 2006.

[Bojārs and Breslin, 2007] Uldis Bojārs and John G. Breslin. SIOC core ontology specification. W3C Member Submission, World Wide Web Consortium (W3C), 2007. `http://www.w3.org/Submission/2007/SUBM-sioc-spec-20070612/`.

[Brändle, 2005] Andreas Brändle. Zu wenige Köche verderben den Brei. Eine Inhaltsanalyse der Wikipedia aus Perspektive der journalistischen Qualität, des Netzeffekts und der Ökonomie der Aufmerksamkeit. Master's thesis, Universität Zürich, 2005.

[Breslin *et al.*, 2006] John G. Breslin, Stefan Decker, Andreas Harth, and Uldis Bojārs. SIOC: an approach to connect web-based communities. *International Journal of Web Based Communities*, 2(2):133–142, 2006.

[Buffa *et al.*, 2008] Michel Buffa, Fabien Gandon, Guillaume Ereteo, Peter Sander, and Catherine Faron. SweetWiki: A semantic wiki. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2008.

[Buswell *et al.*, 2004] 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.

[Dellschaft *et al.*, 2008] Klaas Dellschaft, Hendrik Engelbrecht, José Monte Barreto, Sascha Rutenbeck, and Steffen Staat. Cicero: Tracking design rationale in collaborative ontology engineering. In Bechhofer et al. [2008], pages 782–786.

[Drupal,] Drupal.org – community plumbing. `http://drupal.org`.

[Ebersbach *et al.*, 2008] Anja Ebersbach, Markus Glaser, and Richard Heigl. *Wiki: Web Collaboration*. Springer, 2008.

[Fensel, 2003] Dieter Fensel. *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce*. Springer, 2003.

[FOAF, 2008] Friend of a Friend (FOAF) project. http://www.foaf-project.org/, seen June 2008.

[Gordon and Karacapilidis, 1997] Thomas F. Gordon and Nikos Karacapilidis. The zeno argumentation framework. In *Sixth International Conference on Artificial Intelligence and Law*, pages 10–18. ACM Press, 1997.

[Gruber, 1993] Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, June 1993.

[Huynh *et al.*, 2007] David Huynh, David Karger, and Rob Miller. Exhibit: Lightweight structured data publishing. In Williamson et al. [2007].

[Kittur *et al.*, 2007] Aniket Kittur, Bongwon Suh, Bryan A. Pendleton, and Ed H. Chi. He says, she says: conflict and coordination in wikipedia. In Mary Beth Rosson and David J. Gilmore, editors, *CHI*, pages 453–462. ACM, 2007.

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

[Kunz and Rittel, 1970] Werner Kunz and Horst W. J. Rittel. Issues as elements of information systems. Working paper 131, Institute of Urban and Regional Development, University of California, Berkeley, July 1970.

[Lange *et al.*, 2008a] Christoph Lange, Sean McLaughlin, and Florian Rabe. Flyspeck in a semantic wiki – collaborating on a large scale formalization of the Kepler conjecture, 2008.

[Lange *et al.*, 2008b] Christoph Lange, Sebastian Schaffert, Hala Skaf-Molli, and Max Völkel, editors. *3rd Workshop on Semantic Wikis*, volume 360 of *CEUR Workshop Proc.*, 2008.

[Lange, 2008a] Christoph Lange. Mathematical Semantic Markup in a Wiki: The Roles of Symbols and Notations. In Lange et al. [2008b].

[Lange, 2008b] Christoph Lange. The OMDoc document ontology. http://kwarc.info/projects/docOnto/omdoc.html, seen August 2008.

[Lange, 2008c] Christoph Lange. SWiM – a semantic wiki for mathematical knowledge management. In Bechhofer et al. [2008], pages 832–837.

[MediaWiki, 2008] MediaWiki. http://www.mediawiki.org, seen June 2008.

[Müller, 2008] Christine Müller. The CS precourse project. http://cs-precourse.kwarc.info/, 2008.

[Oren *et al.*, 2006] Eyal Oren, Renaud Delbru, Knud Möller, Max Völkel, and Siegfried Handschuh. Annotation and navigation in semantic wikis. In Max Völkel, Sebastian Schaffert, and Stefan Decker, editors, *1st SemWiki Workshop*, volume 206 of *CEUR Workshop Proc.*, 2006.

[panta rhei, 2008] The panta rhei Project. http://kwarc.info/projects/panta-rhei/, 2008.

[Pentzold and Seidenglanz, 2006] Christian Pentzold and Sebastian Seidenglanz. Foucault@Wiki – first steps towards a conceptual framework for the analysis of wiki discourses. In Dirk Riehle and James Noble, editors, *WikiSym*, ACM, 2006.

[Prud'hommeaux and Seaborne, 2008] Eric Prud'hommeaux and Andy Seaborne. SPARQL query language for RDF. W3C Recommendation, 2008. http://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/.

[Rittel and Webber, 1973] Horst W. J. Rittel and Melvin M. Webber. Dilemmas in a general theory of planning. *Policy Sciences*, 4(2):155–169, June 1973.

[Schaffert, 2006] Sebastian Schaffert. IkeWiki: A semantic wiki for collaborative knowledge management. In *1st International Workshop on Semantic Technologies in Collaborative Applications (STICA)*, 2006.

[Tempich *et al.*, 2005] Christoph Tempich, H. Sofia Pinto, York Sure, and Steffen Staab. An argumentation ontology for DIstributed, Loosely-controlled and evolvInG Engineering processes of oNTologies. In Asunción Gómez-Pérez and Jérôme Euzenat, editors, *ESWC*, volume 3532 of *LNCS*, pages 241–256. Springer, 2005.

[Tempich *et al.*, 2007] Christoph Tempich, Elena Simperl, Markus Luczak, Rudi Studer, and H. Sofia Pinto. Argumentation-based ontology engineering. *IEEE Intelligent Systems*, 22(6):52–59, 2007.

[Tudorache and Noy, 2007] Tania Tudorache and Natasha Noy. Collaborative Protégé. In Williamson et al. [2007].

[Tummarello *et al.*, 2007] Giovanni Tummarello, Renaud Delbru, and Eyal Oren. Sindice.com: Weaving the open linked data. In Karl Aberer, Key-Sun Choi, Natasha Fridman Noy, Dean Allemang, Kyung-Il Lee, Lyndon J. B. Nixon, Jennifer Golbeck, Peter Mika, Diana Maynard, Riichiro Mizoguchi, Guus Schreiber, and Philippe Cudré-Mauroux, editors, *ISWC/ASWC*, volume 4825 of *LNCS*, pages 552–565. Springer, 2007.

[Völkel *et al.*, 2006] Max Völkel, Markus Krötzsch, Denny Vrandečić, Heiko Haller, and Rudi Studer. Semantic Wikipedia. In *15th WWW Conference*, 2006.

[W3C, 2004] Resource Description Framework (RDF). http://www.w3.org/RDF/, 2004.

[Wikipedia,] Wikipedia, the free encyclopedia. http://www.wikipedia.org.

[Wikipedia, 2008a] Hilfe: Zusammenfassung und Quelle, May 2008. http://de.wikipedia.org/w/index.php?title=Hilfe:Zusammenfassung_und_Que%lle&oldid=46731506.

[Wikipedia, 2008b] Wikipedia: Bewertungsbausteine. http://de.wikipedia.org/w/index.php?title=Wikipedia:Bewertungsbausteine%&oldid=44205795, March 2008.

[Wikipedia, 2008c] Wikipedia: Diskussionsseiten. http://de.wikipedia.org/w/index.php?title=Wikipedia:Diskussionsseiten&o%ldid=46659885, May 2008.

[Wikipedia, 2008d] Wikipedia: Neutraler Standpunkt. http://de.wikipedia.org/w/index.php?title=Wikipedia:Neutraler_Standpunk%t&oldid=46873335, June 2008.

[Williamson *et al.*, 2007] Carey L. Williamson, Mary Ellen Zurko, Peter F. Patel-Schneider, and Prashant J. Shenoy, editors. *16th WWW Conference*. ACM, 2007.