Towards a Semantic Wiki for Science

Christoph Lange

Computer Science, Jacobs University Bremen, ch.lange@iu-bremen.de

Collaborative work environments (CWEs) for scientific knowledge have many applications in research and education. In recent years, successful platforms open for anyone appeared on the web, e. g. *Wikipedia* and *PlanetMath*, a wiki particularly tailored to mathematics, or *Connexions*, a CMS for general courseware[1]. Thanks to flexible content creation and linking, similar systems also support corporate knowledge management, but they lack services desirable for effective *scientific knowledge management*. For example, full text search is not suitable for mathematical or chemical formulae[2], and tagging pages does not help to find unproven theorems about triangles. Current semantic wikis [5] solve the latter problem by typing pages and links with terms from ontologies, but they do not support formula search, which would require *structural semantic markup* (SSM), a common approach in mathematical knowledge management.

Further semantic services that have been realised on the Semantic Web, but not yet in open CWEs, include dependency maintenance across changes and learning assistance by suggesting direct and indirect prerequisites to the scholar. How can the knowledge that *is available* in CWEs (e. g. the RDF graph behind a semantic wiki) be used for more than just displaying navigation links, some editing assistance, and semantic search? I will investigate whether a CWE can be turned into an integration platform for semantic services by first creating a uniform ontology abstraction layer *at its core*[3] and prototype such an application that supports SSM formats for various scientific domains based on the semantic *IkeWiki* [3], as wikis particularly support the stepwise formalisation workflow required for scientific SSM (cf. [3,1])[4].

SSM, already having many applications in mathematics (e. g. in the context of the OMDoc XML format [1]), is currently being extended towards other sciences. Research conducted in our group showed that a three-layered model of knowledge can be assumed in mathematics and physics, and probably in most other sciences: *Objects* (symbols, numbers, equations, molecules, etc.), *statements* (axioms, hypotheses, measurement results, examples, with relations like "proves", "defines", or "explains") and *theories* (collections of interrelated statements, defining the context for symbols) [1]. For Semantic Web software, these classes and relations need to be formalised in an *ontology*; I will base my system on the ontologies behind scientific markup languages, and, following the

[1] See http://www.{wikipedia,planetmath,cnx}.org.

[2] $c = \sqrt{a^2 + b^2}$ can mean the same as $x^2 + y^2 = z^2$.

[3] Ontology support is mostly *optional* in current systems.

[4] The related se(ma)^{2}wi [6] system is an experiment with a *Semantic MediaWiki* fed with mathematical knowledge formatted in OMDoc. The semantic structure of the formulae and the links between pages is lost during this conversion, though.

assumption that sciences have common traits like the notion of a "theory" or a "dependency" relation among theories, a generic *upper* ontology of these. To date, merely part of the ontologies behind SSM formats are given as human-readable specifications; I will formalise and generify them in OWL. In a scientific CWE, one page would usually contain one statement, one small theory, or a course module aggregating a few of them. A generic mapping mechanism between XML schemata and ontologies will be applied to *extract* knowledge that is relevant for semantic services from those XML pages to an RDF representation.

As SSM is inherently hard to edit manually, the interaction with the semantic services will be designed in a user-centered way, where the benefits of services like enhanced search and navigation are shared with the users in order to motivate them to contribute. One such service is an ontology-based auto-completion of link targets in the editor. Not all page names starting with the letters typed so far are suggested, but only those pages whose type matches the range of the relation the current link represents. Further planned services include a learning assistant that suggests to explore transitive dependencies, a dependency maintenance assistant, as well as connecting the system to external services already available, e. g. *MathWebSearch*[5]. A preliminary classification suggests that most of the cross-domain services can indeed be modeled on top of the abstraction layer provided by the above-mentioned upper ontology; a formal analysis of the demands of the services on knowledge representation will follow. A challenge is, however, making the different levels of reasoning required by the services (plain triple query for auto-completion vs. computing compositions of relations for dependency management) work smoothly in an inherently inconsistent collaborative setting.

An existing prototype of a wiki for OMDoc [2], featuring basic functionality like page editing, rendering as XHTML+MathML and typed navigation links from a user's perspective, and a basic OMDoc/XML to RDF mapping from a knowledge representation perspective, will be completely redesigned by introducing a generic ontology-based abstraction layer and integrating semantic services on top. It will be evaluated in a cross-domain case study with scientists and in an educational case study with students, leading to feedback for the ontology design. If the abstraction layer approach does facilitate the design and integration of semantic services that increase benefit and reduce users' investment, improving other CWEs, even in non-scientific domains, in a similar way will become possible.

1. M. Kohlhase. OMDoc – *An open markup format for mathematical documents [Version 1.2].* Number 4180 in LNAI. Springer, 2006.
2. C. Lange. SWiM – a semantic wiki for mathematical knowledge management. Technical report, Jacobs University Bremen, 2007.
3. S. Schaffert. Semantic social software – semantically enabled social software or socially enabled semantic web? In Sure and Schaffert [4].
4. Y. Sure and S. Schaffert, editors. *Semantics: From Visions to Applications*, 2006.
5. M. Völkel, S. Schaffert, and S. Decker, editors. *1st Workshop on Semantic Wikis*, volume 206 of *CEUR Workshop Proceedings*, Budva, Montenegro, June 2006.
6. C. Zinn. Bootstrapping a semantic wiki application for learning mathematics. In Sure and Schaffert [4].

[5] http://search.mathweb.org