A Controlled Natural Language Interface for Semantic Media Wiki Using the Rabbit Language

Jie Bao[1], Paul R. Smart[2] and Dave Braines[3]

[1] Department of Computer Science, Rensselaer Polytechnic Institute,
Troy, NY 12180, USA.
baojie@cs.rpi.edu
[2] School of Electronics and Computer Science, University of Southampton,
Southampton, SO17 1BJ, United Kingdom.
ps02v@ecs.soton.ac.uk
[3] Emerging Technology Services, IBM United Kingdom Ltd, Hursley Park,
Winchester, Hampshire, SO21 2JN, United Kingdom.
dave_braines @uk.ibm.com

1 Introduction

Semantic wikis extend the idea of collaborative content editing (made popular by systems such as Wikipedia) to the realm of semantically-enriched representations and formal knowledge models. While a conventional wiki includes structured text and untyped hyperlinks, a semantic wiki is based on the representation of metadata elements. Semantic MediaWiki (SMW) [1] is probably the most popular and mature semantic wiki. It relies on the same wiki engine as Wikipedia and uses elements from the Resource Description Framework (RDF) and Web Ontology Language (OWL) namespaces.

Despite their potential value as collaborative knowledge editing systems, semantic wikis present a number of usability challenges. In particular, how can we enable users to create and edit structured knowledge content (in the form of RDF models and OWL ontologies) without reneging on the kind of simple user interaction mechanisms that makes conventional wiki systems, such as Wikipedia, so popular? One answer to this question is to capitalize on the availability of controlled natural languages (CNLs) that provide some support for ontology model development. CNLs such as Rabbit [2], Sydney OWL Syntax (SOS) [3] and Attempto Controlled English (ACE) [4], all support the creation of semantically-enriched knowledge models, while preserving the production and comprehension benefits of natural languages. As such, CNL interfaces for semantic wiki systems may provide a potent mechanism for encouraging the large-scale participation of user communities in the creation of semantically-enriched online content. In this vein, Kuhn [5] presents a prototype application, called AceWiki, that co-opts the principles of a semantic wiki with the usability features of the ACE CNL.

In this paper, we present our efforts to develop a CNL interface system for semantic wikis using the Rabbit CNL. Our approach is similar to Kuhn [5], although in order to support the expressivity required for OWL ontology development we were obliged to extend the representational substructure of the SMW system. Our ultimate aim is to provide a system that is capable of supporting multiple extant CNLs, e.g.

ACE, Rabbit, SOS, etc. Thus, although our current system only supports the Rabbit CNL, we have developed the system architecture and technology components with a view to accommodating other CNLs using a template-based mechanism. Ultimately, we hope to support end-users in defining their own CNL formalisms for structured knowledge entry in a collaborative knowledge editing environment.

2 System Architecture

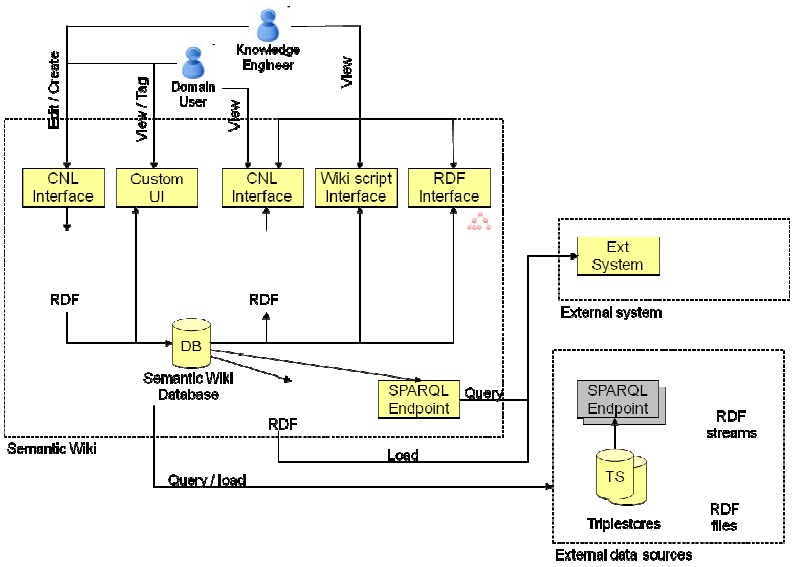

Fig. 1. Architectural overview of a generic CNL interface system for SMW

Figure 1 illustrates the architecture of our system. The system consists of the following components:

- **CNL Interface:** This is an editing interface that allows users to create and edit knowledge statements using the target CNL (Rabbit in the case of our initial prototype).
- **Custom UI:** This is a Controlled Natural Language Generation (CNLG) module that can generate CNL descriptions of the knowledge statements associated with a wiki page.
- **Wiki Script Interface:** This interface supports the generation of semantic wiki scripts from the knowledge statements entered by an end-user.
- **RDF Interface:** This is an interface that supports an RDF-based view of a wiki page. It is intended for those users who are familiar with RDF models.
- **Database (DB):** This is the wiki database that stores both semantic and non-semantic (the free text contents of wiki pages) data.
- **Import/Export Modules:** A number of import/export modules handle the communication with external tools and knowledge technology components.

We have implemented a prototype semantic wiki system based on the system architecture presented in Figure 1 (see http://tw.rpi.edu/dev/cnl/). The system currently supports the Rabbit CNL, but we are extending the system to support multiple CNLs, such as SOS and ACE-OWL. This extension is accomplished using a template-based mechanism which is introduced in the next section.

3 OWL Meta-Modeling on Semantic Wiki

In order to accommodate CNLs, such as Rabbit, within a semantic wiki system, we need to address a number of expressivity constraints associated with semantic wikis. SMW, for example, does not provide full support for OWL modeling formalisms, and this introduces a mismatch between the kind of knowledge statements that can be represented in Rabbit and the knowledge statements that can be created in SMW[1]. In order to address this limitation, we developed a meta-model extension to SMW, called SMW-mOWL. SMW-mOWL represents an OWL ontology using a set of semantic templates, each of which encodes ontology elements (i.e. classes, properties and individuals) in template instances (i.e. `Template:Class`, `Template:Property` and `Template:Individual`). Each axiom is also represented as an instance of a template. For example, `Template:Some` represents the `owl:somesValuesFrom` restriction. The motivation behind this meta-modeling approach is based on a number of design considerations. These include the following:

- **Correspondences to the OWL Abstract Syntax (OWL-AS):** SMW-mOWL is intended to have a direct correspondence to OWL-AS. Entities (classes, properties and individuals) and axioms in SMW-mOWL are represented as corresponding wiki templates. For example, the class:

  ```
  Class(Rabbit partial Animal
      restriction(eat someValuesFrom(FreshVegetable)))
  ```

  is represented in the semantic wiki as a page (`Category:Rabbit`) using the `Template:Class`, `Template:ClassRelation` and `Template:Some` templates. This results in the following meta-model representation of the class:

  ```
  {{Class|label=Rabbit|plural=Rabbits}}
  {{ClassRelation|type=subClassOf|class=Animal}}
  {{Some|on property=eat|on class=FreshVegetable}}
  ```

  This correspondence between OWL-AS and SMW-mOWL has a couple of advantages. Firstly, OWL-AS can be used as an intermediate syntax for knowledge exchange between SMW and other tools. Secondly, it provides an extensible framework for supporting multiple CNLs within the SMW environment.
- **UI Correspondences:** The design of SMW-mOWL is intended to have a direct correspondence to the SMW knowledge editing system. By utilizing what are called 'semantic forms' within SMW, each template can be edited using a form-

[1] It should be pointed that languages such as Rabbit and ACE are considerably more expressive than current versions of OWL.

based interface. Thus, having the OWL meta-model immediately provides us with an OWL ontology editor within the SMW environment.

- **Query Convenience:** The use of a template-based mechanism for SMW-mOWL allows us to store the knowledge model in the SMW database and to use SMW-QL (the query language for SMW) to retrieve specific information.

4 Conclusions and Future Work

This paper summarizes our initial efforts to develop a generic CNL interface for semantic wiki systems. Thus far we have developed an architecture to support the collaborative editing of community knowledge using semantic wikis. We have also developed a set of extensions to SMW in order to accommodate the expressivity features of many of CNLs. Finally, we have implemented a prototype of our system (see http://tw.rpi.edu/dev/cnl/) and are currently testing the system to ensure compliance with the Rabbit CNL specification. Future work will focus on the development of parsing and editing capabilities (i.e. parsing of predefined Rabbit models and developing a better user interface for the entry of Rabbit language constructs). In addition, we aim to extend the semantic template system, briefly described herein, in order to accommodate multiple CNLs (e.g. Rabbit, SOS, ACE, etc.) within a single semantic wiki system. This will, we suggest, encourage large-scale collaborative knowledge modeling and semantic annotation as part of the Semantic Web initiative.

References

1. Krötzsch, M., Vrandecic, D., Völkel, M., Haller, H., Studer, R.: Semantic Wikipedia. Journal of Web Semantics **5** (2007) 251-261
2. Hart, G., Johnson, M., Dolbear, C.: Rabbit: Developing a Control Natural Language for Authoring Ontologies. 5th European Semantic Web Conference (ESWC'08), Tenerife, Spain (2008)
3. Cregan, A., Schwitter, R., Meyer, T.: Sydney OWL Syntax—towards a Controlled Natural Language Syntax for OWL 1.1. OWL Experiences and Directions Workshop, Innsbruck, Austria (2007)
4. Kaljurand, K., Fuchs, N.E.: Bidirectional mapping between OWL DL and Attempto Controlled English. 4th Workshop on Principles and Practice of Semantic Web Reasoning, Budva, Montenegro (2006)
5. Kuhn, T.: AceWiki: A Natural and Expressive Semantic Wiki. Semantic Web User Interaction Workshop at CHI 2008, Florence, Italy (2008)