

ROO: Involving Domain Experts in Authoring OWL Ontologies

Ronald Denaux
Vania Dimitrova
Anthony G. Cohn
School of Computing, University of Leeds
Woodhouse Lane, Leeds, LS2 9JT, UK
{rdenaux, vania, agc@comp.leeds.ac.uk}

Catherine Dolbear
Glen Hart
Ordnance Survey Research
Romsey Rd, Southampton, SO16 4GU, UK
{Catherine.Dolbear,
Glen.Hart}@ordnancesurvey.co.uk

ABSTRACT

This demonstration presents ROO, a tool that facilitates domain experts' definition of ontologies in OWL by allowing them to author the ontology in a controlled natural language called Rabbit. ROO *guides* users through the ontology construction process by following a methodology geared towards domain experts' involvement in ontology authoring, and exploiting intelligent user interfaces techniques. An experimental study with ROO was conducted to examine the usability and usefulness of the tool, and the quality of the resultant ontologies. The findings of the study will be presented in a full paper at the ISWC08 research track [2].

Categories and Subject Descriptors

I.2.4 [Knowledge Representation Formalisms and Methods]:
Ontology Authoring

General Terms

Design, Languages

Keywords

Ontology Authoring, Controlled Natural Language Interfaces,
Evaluation of Ontology Building Tools, Geographical Ontologies

1. INTRODUCTION

The construction of ontologies requires a large amount of time and effort, which hinders the ability and willingness of organisations to apply Semantic Web technologies to large-scale data integration and sharing initiatives[1]. Most existing ontology construction tools are designed to be used by specialists with appropriate knowledge engineering skills but who may lack the necessary domain expertise to create the relevant ontologies. Domain experts have to provide relevant knowledge sources, or have to collaborate in knowledge elicitation techniques to discover information, while knowledge engineers encode the ontology using available tools. This approach can hinder the ontology construction process and may have a negative impact on the quality of the resultant ontology (e.g. poor documentation, inconsistency of terminology used, incorrect or incomplete knowledge constructs). Hence, tools that facilitate the ontology construction process and are tailored to the needs of domain experts can speed up the authoring process and improve the quality of the resultant ontologies.

Recently, controlled language (CL) interfaces have been provided for entering knowledge constructs in an intuitive way; see the recent review in [4]. The major deficiency of these tools is that they focus solely on the CL aspect while ignoring the whole ontology construction process. We will present a fairly stable version of a tool called ROO (Rabbit to OWL Ontology authoring) that not only provides a CL interface but offers systematic guidance throughout the whole ontology construction process. The development of ROO is driven by the needs of a large public organisation – the national mapping agency of Great Britain, Ordnance Survey. The Ordnance Survey is developing a modular topographic domain ontology to facilitate the description and reuse of its topographic data by third parties [3]. At the heart of ontology development at Ordnance Survey is the *active involvement of domain experts* (e.g. geographers, ecologists and emergency planners) [7].

2. THE ROO TOOL

ROO is a Protégé plug-in that guides domain experts with little or no knowledge engineering experience to build conceptual ontologies. The distinctive characteristics of ROO are:

- catering for the needs of domain experts without knowledge engineering skills but also enabling users with experience in ontology development to inspect and edit the resultant ontology;
- exploiting techniques from intelligent user interfaces to assist the ontology construction process by following an ontology authoring methodology (the current implementation follows the methodology used at Ordnance Survey for developing several large ontologies with the active involvement of domain experts);
- providing an intuitive interface to enter knowledge constructs in a controlled language (CL), Rabbit¹, designed to meet the needs of domain experts and compliant with OWL 1.1[5].

ROO is developed within the Confluence project and is distributed as open source².

¹ The language is named after Rabbit in Winnie the Pooh, who is actually cleverer than Owl.

² The tool can be downloaded from <http://sourceforge.net/projects/confluence>, a demo is available at <http://www.comp.leeds.ac.uk/confluence/>

ROO builds on the strengths and minimises the usability threats of existing CL tools for ontology construction. Figure 1 shows the main architectural elements of the currently implemented ROO tool and their interactions.

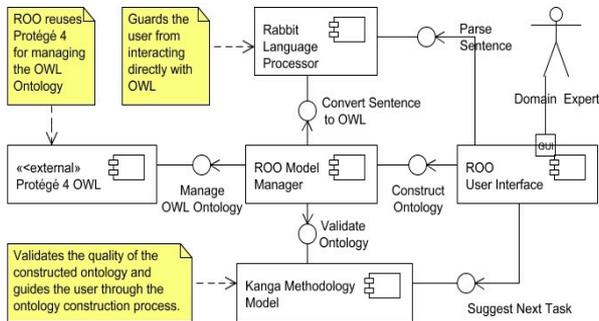


Figure 1. UML 2.0 component diagram illustrating architectural elements, interfaces and inter-element connections in ROO.

ROO includes the following **usability features** to support ontology authoring by domain experts:

- providing suggestions by guessing intended constructs;
- showing the parsed structure to help the user recognise correct sentence patterns;
- providing a flexible way to parse English sentences using robust language technologies and automatically translate to OWL;
- using templates to facilitate the knowledge input process;
- immediate feedback and task-specific error messages.

ROO offers **task suggestions** based on monitoring the state of the ontology and the user's activities and suggests the most appropriate actions. For example, at the beginning of the ontology construction process, ROO will suggest that the scope and purpose of the ontology should be identified. In another case, when a user has already defined several concepts and relationships, ROO will suggest that the user should enter a natural language description for concepts missing such a description. The user interface is tailored accordingly, to simplify the execution of the suggested tasks.

The GUI in ROO extends existing Protégé GUI components. The user guidance is implemented as a rule system based on the JBoss Drools³ rule engine. The Rabbit language processor currently parses and converts nine basic Rabbit patterns into OWL. The parsing of the Rabbit sentences is based on an extension of the GATE⁴ suite for natural language parsing and information extraction.

3. EVALUATION AND FUTURE PLANS

ROO is being developed following an agile methodology that involves an iterative development of features. User evaluations are performed to examine the usability and the use of the Rabbit controlled language constructs[5] to define sample domain

ontologies. A recent evaluation study has been conducted at the University of Leeds with 16 volunteers from the departments of Geography (8 students) and Earth and Environment (8 students) [2] following ontology modeling tasks conducted by domain experts associated with Ordnance Survey. To examine the benefits of the support offered in ROO, the tool was compared to ACEView [6] which provides similar interaction means.

Further experimental studies are planned for the next iterations in the ROO development, when the full set of Rabbit sentences will be parsed and converted to OWL. These evaluation studies will involve domain experts at Ordnance Survey. We are further considering implementing additional support for collaborative ontology construction, ontology validation and inspection.

4. ACKNOWLEDGMENTS

This research is conducted under the Confluence project funded by Ordnance Survey, the mapping agency of the Great Britain, and is conducted by a team of researchers, domain experts and software developers at the University of Leeds and Ordnance Survey. Thanks go to previous members of the Confluence team – Katalin Kovacs, Sheng Zhou, Ilaria Corda, and Ian Holt, who helped clarifying the ROO functionality and conducting experimental studies. Thanks go to Paula Engelbrecht for sharing her material on the Environment Agency which helped creating realistic user scenarios for the design and evaluation of ROO.

5. REFERENCES

- [1] Alani, H., Dupplaw, D., Sheridan, J., O'Hara, K., Darlington, J., Shadbolt, N., Tullo, C. 2007. Unlocking the Potential of Public Sector Information with Semantic Web Technology. *In Proc. of ISWC 2007*.
- [2] Dimitrova, V., Denaux, R., Hart, G., Dolbear, C., Cohn, A.G. 2008. Involving Domain Experts in Authoring OWL Ontologies. *In Proceedings of ISWC2008*.
- [3] Dolbear, C., Hart, G., Goodwin, J., Zhou, S. and Kovacs, K. 2007. The Rabbit language: description, syntax and conversion to OWL. *Ordnance Survey Research Labs Technical Report IRI-0004*.
- [4] Funk, A., Davis, B., Tablan, V. and Bontcheva, K. 2007. Controlled Language IE Components version 2. *SEKT project deliverable D.2.2.2*.
- [5] Hart, G., Johnson, M. and Dolbear, C. 2008. Rabbit: Developing a Control Natural Language for Authoring Ontologies. *In Proceedings of ESWC2008*.
- [6] Kaljuran, K. Attempto Controlled English as a Semantic Web Language. PhD thesis, Faculty of Mathematics and Computer Science, University of Tartu, 2007.
- [7] Kovacs, K., Dolbear, C., Hart, G., Goodwin, J., Mizen, H.: A Methodology for Building Conceptual Domain Ontologies. *Ordnance Survey Research Labs Technical Report IRI-0002 2006*.

³ <http://labs.jboss.com/drools>

⁴ <http://gate.ac.uk>