Proceedings Paper:
Denaux, R, Dimitrova, V and Cohn, AG (2013) Interacting with ontologies and linked data through controlled natural languages and dialogues. In: Do-Form: Enabling Domain Experts to Use Formalised Reasoning - AISB Convention 2013. AISB Convention 2013, 03-05 Apr 2014, University of Exeter 2013. Society for the study of artificial intelligence , 18 - 20. ISBN 9781908187321

Interacting with Ontologies and Linked Data through Controlled Natural Languages and Dialogues

Ronald Denaux, Vania Dimitrova and **Anthony G. Cohn** [1]

Abstract. This paper describes a suite of tools developed at the University of Leeds which aim to make it easier for domain experts to be involved in the creation and use of ontologies. The paper summarises the main features of the tools and gives a short summary of our evaluations and experiences using the tools with domain experts.

1 Expressing Knowledge through a Controlled Natural Language

At the core of our suite of tools for supporting domain experts when using ontologies is the use of a Controlled Natural Language (CNL) as the main way to express knowledge in a way that:

- is easy to understand and write by domain experts and
- can be automatically translated into a logical form, in particular an ontology language

We have adopted Rabbit, a CNL designed by the Ordnance Survey – the mapping agency of Great Britain – to enable domain experts (e.g. cartographers) to contribute to the development of ontologies for describing the Ordnance Survey's topographic data [8]. Although inspired by the topographical domain, Rabbit has been designed to be domain independent. The language is designed to be directly mappable to OWL, the web ontology language.

1.1 Tool Support

As part of our adoption of the Rabbit CNL, we have collaborated with the Ordnance Survey to build a **parser** for the language that translates textual inputs into OWL constructs. This parser uses Natural Language Processing (NLP) techniques to improve the recognition of Rabbit sentences. These techniques are used, for example, to recognise that concepts tend to be expressed as nouns (or noun phrases); similarly, relationships tend to be expressed as verbs. Another example of the use of NLP techniques is in recognising that the same concept may be expressed as either a singular noun or a plural noun.

An important aspect of the Rabbit parser is that, when the input text is not syntactically correct, it is often able to generate **error messages that are easy to understand** and help the domain expert to correct the input (see for example, Figure 1). We refer to [3] for more information about the Rabbit parser.

Finally, other Rabbit tool support we provide is **automatic translation of valid Rabbit sentences into OWL** constructs and **generation of Rabbit sentences** to show the contents of existing ontolo-
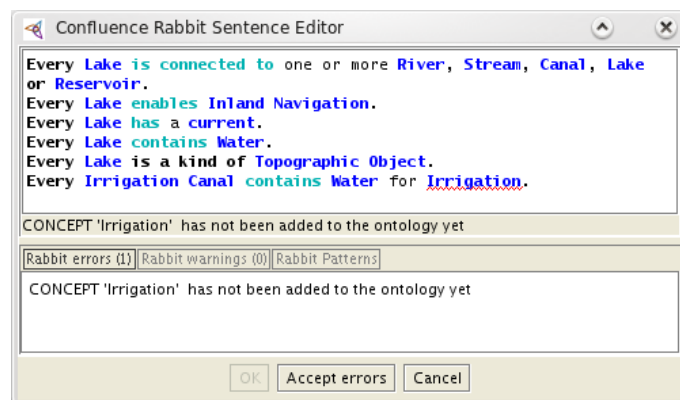

Figure 1. Rabbit Editor component showing various Rabbit sentences, syntax highlighting and error feedback.

gies[2]. Since we have developed the tool support in house, we are able to adapt and extend the tools if necessary.

1.2 Conclusion

The Ordnance Survey has studied how easy the Rabbit language is to understand [8] and to write [7] with positive results. In the next sections, we will discuss our experiences using the language to support ontology authoring and use of linked data.

Although we are using Rabbit as the CNL in our tools, there are other CNLs which provide comparable tool support and expressivity. We refer to [11] for a comparison to the main alternative CNLs for OWL.

2 Ontology Authoring Methodology Support for Domain Experts

After developing Rabbit, the Ordnance Survey found that domain experts were able to create ontologies using Rabbit, but these ontologies contained modelling errors due to the lack of knowledge about ontology modelling processes by domain experts [8]. The Ordnance Survey adapted existing ontology engineering methodologies to give the domain experts a central role in the ontology authoring process, and we developed tool support to guide domain experts through this methodology. This resulted in an ontology editor called ROO (Rabbit to OWL Ontology authoring).

---

[1] University of Leeds, England, email: r.denaux@leeds.ac.uk

[2] This only works if the ontologies provide labels for their classes and properties.

ROO provides an alternative user interface for Protege, one of the leading ontology editors. The main advantage of ROO is that it is designed for domain experts:

- knowledge is entered and presented using the Rabbit CNL
- the interface has been simplified to

  1. guide the user through the ontology authoring process: the main interface consists of three tabs, one for defining the scope and purpose of the ontology, a second tab for defining knowledge sources and a third tab to define the ontology concepts and relationships.

  2. avoid OWL specific terminology: for example by using the term 'concept' instead of the OWL specific 'class'. As another example, OWL provides options to use annotations with language tags and various XML types; ROO uses defaults for various annotations to simplify the input by domain experts.

Another feature of ROO for guiding domain experts through the ontology authoring methodology is the **Guide Dog**: when the domain expert is unsure about what to do next, the Guide Dog analyses the current ontology and suggests a next task to perform. Example tasks are: to define the purpose of the ontology, to add a natural language definition for a concept or to add a Rabbit sentence to define a concept.

We performed a comparative evaluation study comparing ROO to a similar CNL-based ontology authoring tool. In the study, Geography and Environmental Studies students and researches at the University of Leeds were given a task to create an ontology about Hydrology and Water Polution. The result were encouraging, showing that domain experts with no previous exposure to ROO were able to create an initial ontology. The results validated out approach for providing tool support for editing CNL sentences and following an ontology authoring ontology [6, 4].

ROO is open source[3] and has more than a 1000 downloads. The ontology editor has successfully been used by several staff members and students at the University of Leeds to build ontologies. For example as part of an EU-project[4], ROO was used by a staff member of the Leeds University Business School to create an ontologies about Activity Theory and Interpersonal Communication.

3 Feedback about Logical Aspects of Ontology Authoring

While Rabbit enables domain experts to directly contribute their knowledge to an ontology and ROO guides them through the process of building ontologies, the problem of modelling errors in the resulting ontologies remained. In practice, this means that the contributions made by domain experts have to be evaluated and corrected by knowledge engineers who have training in the formal logics of ontology languages. However, finding and evaluating such ontology bugs can be difficult, especially if the knowledge engineer is not an expert of the domain at hand. To alleviate this problem, we added tool support for providing semantic feedback: **interactive feedback about the logical consequences of integrating a new fact into an existing ontology**.

In order to provide semantic feedback, we have defined a framework that extends the syntactic analysis performed by the Rabbit

---

parser with an integration analysis based on various existing ontology reasoning services. When the domain expert enters a valid Rabbit sentence, the resulting new ontological fact is classified into one of the following categories:

- **the fact is already in the ontology**: this can be because the syntax of the sentence differs from the syntax used in the ontology;
- the fact is implied by the ontology: ontologies consist of a number of stated facts, but those facts can combine and imply new facts that have not been explicitly stated. Knowing about such inferences can be useful to domain experts to detect unwanted inferences but also to avoid redundancy in the ontology;
- **the fact is inconsistent with the ontology**: in this case the new fact contradicts other facts in the ontology. Domain experts typically are not aware of such logical contradiction, thus getting feedback helps them to become aware of existing parts of the ontology and to avoid introducing bugs into the ontology.
- **the fact is novel, but no 'relevant' implications can be found**: this case helps the domain expert to realise that the ontology needs to be more interconnected in order to make more inferences possible. The domain expert may also be expecting the system to make inferences, getting this type of feedback helps them to realise the limitations of the ontology language and its reasoning capabilities.
- **the fact is novel and has 'relevant' implications**: in this case, the system can provide a list of new implications that follow from the new fact. This helps the domain expert to become aware of existing facts in the ontology and to get a feeling for the reasoning capabilities of the ontology languages. Furthermore, it helps to avoid unwanted inferences.

An example input Rabbit sentence with the corresponding feedback is shown below:

Rabbit Input: *Every Teaching Hospital is a kind of Hospital.*
Axiom category [Novel Axiom with new Relevant Implications]
ROO Feedback: This assertion is **novel**: it has not been added to the ontology yet. This input **implies** 6 new relevant facts. Have a look at the list of implications to make sure you agree with the implications. If you do not agree, it may be that you are using the wrong terminology. **Check the new implications**: • *Every Teaching Hospital has footprint a Footprint.* • *Organisation and Teaching Hospital are mutually exclusive.* • *Training Centre and Teaching Hospital are mutually exclusive.* • *Every Teaching Hospital is a kind of Topographic Object.* • *Every Teaching Hospital is a kind of Place.* • *Teaching Hospital and University (Institution) are mutually exclusive.*

A full description of the integration analysis performed as well as of the semantic feedback provided for each of the categories is provided in [5]. We have integrated the semantic feedback in ROO and we performed an evaluation to find out what novice users without prior knowledge in ontologies found about the feedback. We compared their impressions of the feedback with experienced knowledge engineers and we found that both novice users as well as experts find the interactive semantic feedback informative, timely and useful [5]. We are not aware of another semantic system that provides this type of information in a manner that is understandable to domain experts.

A current limitation of this tool support is that the feedback helps domain experts to become aware of possible problems, but it does not provide sufficient information to resolve the problems as this requires further analysis of the problem. Also, the evaluation study was performed in a controlled environment; we look forward to evalute the semantic feedback in a production-like context.

4 CNL-Based Dialogue Interface to Ontologies and Linked Data

The final tool in our suite to support domain experts is an **ontology and CNL-based dialogue manager**. The need for such a dialogue manager came from the realisation that CNL-based ontology authoring is similar to a dialogue where the domain expert is telling facts about the domain to the ontology editor, which responds with syntactic or semantic feedback as described above. While our experiences with such an approach for ontology authoring are promising, one limitation is that – depending on the situation – the feedback provided can be repetitive, overwhelming, redundant or insufficient. The dialogue manager enables us to define different feedback strategies for different situations in order to provide more appropriate feedback.

The dialogue manager is based on years of research in the field of dialogue systems [10] and can be seen as an extension to the syntactic analysis provided by the Rabbit parser and the semantic analysis described in Section 3. The dialogue manager adds a level of input analysis in which **each input by the domain expert is assigned an intention**: in ROO until now, the intention is assumed to be to inform the system about facts about the domain, but other intentions may be to query the system about existing facts in the ontology or to tell the system that the domain expert disagrees with an existing fact in the ontology. The dialogue manager also **keeps track of the dialogue history**, which enables the definition of alternative interaction strategies based on the previous inputs. For example, if the domain expert has informed the system about a new fact that is novel without 'relevant' implications, the dialogue manager can now elicit new facts that can make new inferences possible, as shown in turn 10 in the example dialogue below:[5]:

---

1 ROO: Tell me a new fact to add to the ontology.
2 ROO: You can do one of the following:

- Enter a new fact or
- Enter *Skip* if you do not want to tell a new fact or
- Enter *Help* to receive a hint.

3 DE: Bus Line 95 has end point Bodington Hall.
4 ROO: This fact is indeed new.
5 ROO: Normally, entering a new fact enables me to make further inferences but I cannot do this in this case.
6 ROO: Can you tell me something that I should be able to infer?
7 ROO: You can do one of the following:

- Enter a new fact or
- Enter *Skip* if you prefer not to answer this question or
- Enter *Help* to receive a hint.

8 DE: Bus Line 95 stops near Bodington Hall.
9 ROO: Actually, "Bus Line 95 stops near Bodington Hall" is a new fact.
10 ROO: Tell me more general facts about `Bus Route` or `Hall of Residence` or `has end point` or `stops near`.

At the moment, the dialogue manager has been implemented and can be used to simulate the ontology authoring interaction in ROO, as well as to perform simple elicitations as shown above. In the near future we are planning to perform evaluations to determine the benefit of such a dialogue interface for ontology authoring.

4.1 Beyond Ontology Authoring

The dialogue manager makes it possible to establish dialogues for other purposes than ontology authoring. For example, in the context of the ImReal project[6], we are using the dialogue manager to assess the cultural exposure of learners based on countries they have visited.

---

[5] The example is in the domain of Points of Interest in Leeds
[6] `http://imreal-project.eu/`

We are reusing a service that determines a list of visited countries of a user based on their Flickr and Twitter profiles[9].

Instead of letting the domain expert build an ontology, in this case we have extracted various ontologies from DBPedia [1] containing general facts about countries (currency, languages, income inequality and human development) as well as cultural facts, in particular gestures that occur in specific countries. We use these ontologies to generate a quiz that is presented to the user in the form of a dialogue. Based on the user's anwers, the dialogue attempts to determine how much cultural exposure the learner had to the visited countries. We encourage the reader to see [2] for more details on this dialogue, the data sources used and, in particular, a video showing an example dialogue session.

5 Conclusion

We have presented a suite of tools for enabling the interaction of domain experts with ontologies and linked data. Our line of research has focused on CNL-based interaction between domain experts and ontologies as well as interaction to facilitate ontology authoring. However, the various techniques described can be adapted to other tasks where domain experts need to interact with logic systems as shown with in our latest work on a dialogue for assessing the cultural exposure of learners, where we have successfully reused linked data (i.e. DBPedia) to drive the dialogue.

REFERENCES

[1] C Bizer, J Lehmann, G Kobilarov, S Auer, C Becker, R Cyganiak, and S Hellmann, 'DBpedia-A crystallization point for the Web of Data', *Web Semantics: Science, Services and Agents on the World Wide Web*, **7**(3), 154–165, (2009).

[2] Ronald Denaux. Modelling User Cultural Exposure through Dialogues, 2012. `http://imash.leeds.ac.uk/services/CultaDis/` accessed on 10/12/2012.

[3] Ronald Denaux, Vania Dimitrova, Anthony Cohn, Catherine Dolbear, and Glen Hart, 'Rabbit to OWL: Ontology Authoring with a CNL-Based Tool', in *Controlled Natural Language*, ed., Norbert Fuchs, volume 5972 of *Lecture Notes in Computer Science*, 246–264, Springer Berlin / Heidelberg, (2010).

[4] Ronald Denaux, Catherine Dolbear, Glen Hart, Vania Dimitrova, and Anthony G. Cohn, 'Supporting domain experts to construct conceptual ontologies: A holistic approach', *Web Semantics: Science, Services and Agents on the World Wide Web*, **9**(2), 113–127, (July 2011).

[5] Ronald Denaux, Dhaval Thakker, Vania Dimitrova, and Anthony G. Cohn, 'Interactive Semantic Feedback for Intuitive Ontology Authoring', in *7th International Conference on Formal Ontology in Information Systems*, Graz, (2012).

[6] Vania Dimitrova, Ronald Denaux, Glen Hart, Catherine Dolbear, Ian Holt, and Anthony G. Cohn, 'Involving Domain Experts in Authoring OWL Ontologies', in *Proceedings of the 7th International Semantic Web Conference, ISWC*, pp. 1–16, (2008).

[7] Paula Engelbrecht, Glen Hart, and Catherine Dolbear, 'Talking rabbit: a user evaluation of sentence production', in *Proceedings of the 2009 conference on Controlled natural language*, pp. 56–64. Springer, (June 2010).

[8] G. Hart, M. Johnson, and C. Dolbear, 'Rabbit: Developing a control natural language for authoring ontologies', in *Proceedings of the 5th European Semantic Web Conference*, pp. 348–360. Springer, (2008).

[9] C Hauff and G J Houben, 'Geo-Location estimation of flickr images: social web based enrichment', *Advances in Information Retrieval*, 85–96, (2012).

[10] Staffan Larsson and David R Traum, 'Information state and dialogue management in the TRINDI dialogue move engine toolkit', *Natural Language Engineering*, **6**(3–4), 323–340, (2000).

[11] R. Schwitter, K. Kaljurand, A. Cregan, C. Dolbear, and G. Hart, 'A Comparison of three Controlled Natural Languages for OWL 1.1', in *4th OWL Experiences and Directions Workshop (OWLED 2008 DC)*, Washington DC, (2008).