Mapping Attempto Controlled English to OWL DL

Kaarel Kaljurand
Institute for Computational Linguistics
University of Zurich
kalju@ifi.unizh.ch

Norbert E. Fuchs
Institute for Computational Linguistics
University of Zurich
fuchs@ifi.unizh.ch

ABSTRACT

We describe ongoing work on the mapping between Attempto Controlled English (ACE) and OWL DL. ACE is a well-studied controlled language, with a parser that converts ACE texts into Discourse Representation Structure (DRS). We describe a relatively direct mapping of a subset of the DRS language to OWL DL. This mapping renders ACE an interesting companion to existing OWL front-ends.

1. INTRODUCTION

Existing OWL tools (e.g. Protégé, SWOOP, SemanticWorks) are user-friendly graphical point-and-click editors but for complex class descriptions they require the user to possess a large knowledge of Description Logics. E.g. [3] list the problems that users encounter when working with OWL and express the need for a "pedantic but explicit" paraphrase language.

We envision a text based system that allows the users to express the ontologies in the most natural way — in natural language. Such a system would be easy to use since it does not presuppose a knowledge of mathematical concepts such as disjointness or transitivity. The system would be tightly integrated with an OWL reasoner, but the output of the reasoner (if expressed in OWL as a modification of the ontology) would again be verbalized in natural language, so that all user interaction takes place in natural language.

As a basis of the natural language, we have chosen Attempto Controlled English (ACE), a subset of English that can be converted through its DRS representation into first-order logic representation and automatically reasoned about (see [1] for more information). The current version of ACE offers language constructs like countable and mass nouns, collective and distributive plurals, generalized quantifiers, indefinite pronouns, noun phrase/verb phrase/sentence negation, and anaphoric references to noun phrases through proper names, definite noun phrases, pronouns, and variables. The intention behind ACE is to minimize the number of syntax and interpretation rules needed to predict the resulting DRS, or for the end-user, the reasoning results. The small number of ACE function words have a clear and predictable meaning and the remaining content words are classified only as verbs, nouns, adjectives and adverbs. Still, ACE has a relatively complex syntax compared to the OWL representation e.g. in the OWL Abstract Syntax specification, but as ACE is based on English, its grammar rules are intuitive i.e. already known to English speakers.

Some existing results show the potential and the need for a natural language based interface to OWL. [2] paraphrase OWL class hierarchies but their target is not a controlled language and cannot be edited and parsed back into a standard OWL representation. [4] propose writing ontologies in a controlled language, but do not provide a natural syntax for writing TBoxes. In the following, we describe a mapping from a subset of ACE (OWL ACE) to OWL DL (in RDF/XML notation) and conclude with an overview of the remaining work.[1]

2. FROM ACE TO OWL

Figure 1 shows an ACE text and its corresponding DRS that makes use of a small number of predicates, most importantly `object` derived from nouns and `predicate` derived from verbs. The predicates share information by means of discourse referents (denoted by capital letters) and are further grouped by embedded DRS boxes, that represent implication (derived from *if... then...* or *every*), negation (derived from various forms of English negation), and disjunction (derived from *or*). Conjunction — derived from relative clauses, explicit *and*, or the sentence end symbol — is represented by the co-occurrence in the same DRS-box.

The mapping to OWL does not modify the existing DRS construction algorithm but only the interpretation of the DRS. It considers everything in the toplevel DRS to denote individuals or relations between them. Individuals are introduced by nouns, so that propernames map to individuals with type *owl:Thing* and common nouns to an anonymous individual with the type derived from the corresponding noun (e.g. class *Man*). Properties are derived from

Demos and Posters of the *3rd European Semantic Web Conference (ESWC 2006)*, Budva, Montenegro, 11th - 14th June, 2006

[1]A demo of this mapping is available among the Attempto tools at http://www.ifi.unizh.ch/attempto/tools

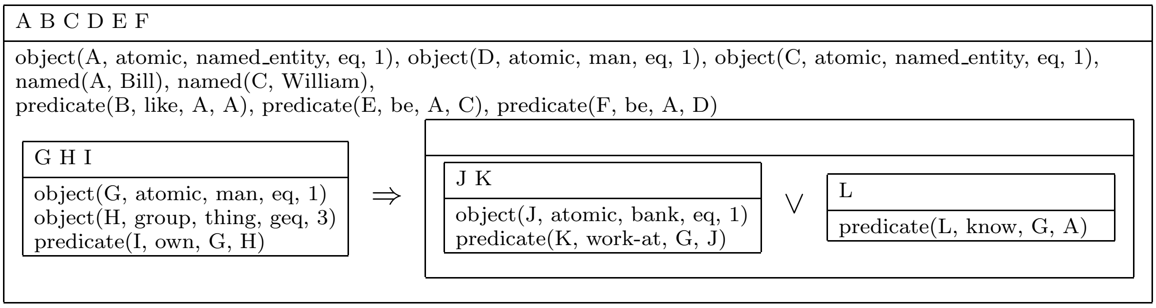

Figure 1: DRS corresponding to the ACE text "Bill who is a man likes himself. Bill is William. Every man who owns at least 3 things works-at a bank or knows Bill." Note that the DRS has been simplified for layout purposes. Also, the example is somewhat artificial to demonstrate concisely the features of OWL as expressed in ACE.

transitive verbs. A special meaning is assigned to the copula 'be' which introduces an identity between individuals. An embedded implication-box introduces a *subClassOf* relation between classes: the head of the implication maps to the subclass description, the body to its superclass description. Transitive verbs introduce a property restriction with *someValuesFrom* a class denoted by the object of the verb, and the copula introduces a class restriction. Negation and disjunction boxes in the implication-box introduce *complementOf* and *unionOf*, respectively. Any embedding of them is allowed. The plural form of the word 'thing' which can be modified by a number allows to define cardinality restrictions. Thus the DRS of figure 1 has the following meaning (in Description Logics notation):

$$\text{bill} \in \top,\ \text{m1} \in \text{Man},\ \text{william} \in \top,$$
$$\text{bill} = \text{m1},\ \text{bill} = \text{william},$$
$$\text{likes(bill, bill)}$$

$$\text{Man} \sqcap \text{owns} \geq 3 \sqsubseteq$$
$$\exists\ \text{worksAt Bank} \sqcup \exists\ \text{knows}\ \{\text{bill}\}$$

ACE can also describe OWL properties (super property, inverse property and transitivity) but this sounds quite "mathematical", e.g. transitivity is expressed as "If a thing A is taller than a thing B and B is taller than a thing C then A is taller than C.". On the other hand there does not seem to be a better way in natural languages.

Note that the mapping does not target all the syntactic variety defined in the OWL specification, e.g. elements like *disjointWith* or *equivalentProperty* cannot be directly expressed in ACE, but their semantically equivalent constructs can be generated.

Given that ACE is easy to learn and use, can we say the same about OWL ACE? With regards to full ACE, OWL ACE introduces a number of restrictions: there is no support for ditransitive and intransitive verbs, prepositional phrases, adverbs, intransitive adjectives and most forms of plurals. Furthermore, there are restrictions to the DRS structure which are more difficult to explain to the average user, e.g. disjunction is not allowed to occur at the toplevel DRS ("John sees Mary or John sees Bill."). A further restriction could require the predicates in the implication-box to share one common discourse referent as the subject argument, and not to share the object arguments. This would allow us to exclude sentences like "If a man sees a mouse then a woman does not see the mouse." which does not seem to map nicely to an ontology language but instead to a rule language. Then again, this restriction is too strong as it would exclude property expressions ("Everybody who loves somebody likes him/her.") and a way to express *allValuesFrom* ("Everything that a herbivore eats is a plant.").

3. FUTURE WORK

The current mapping lacks support for datatype properties and enumerations (*oneOf*). Furthermore, there is only a limited support for *someValuesFrom* and *allValuesFrom*, meaning that not all the possible configurations of these constructs can be generated with ACE. We will add support of those constructs along with support of URIs for naming classes, properties and individuals.

We will also implement the mapping from OWL to ACE which must handle all OWL constructs, some of which the ACE-to-OWL mapping does not produce. The mapping from OWL to ACE must also deal with the naming conventions of OWL constructs.

4. REFERENCES

[1] Attempto project. Attempto website, 2006. `http://www.ifi.unizh.ch/attempto`.

[2] Daniel Hewlett, Aditya Kalyanpur, Vladamir Kovlovski, and Chris Halaschek-Wiener. Effective Natural Language Paraphrasing of Ontologies on the Semantic Web. In *End User Semantic Web Interaction Workshop (ISWC 2005)*, 2005.

[3] Alan L. Rector, Nick Drummond, Matthew Horridge, Jeremy Rogers, Holger Knublauch, Robert Stevens, Hai Wang, and Chris Wroe. OWL Pizzas: Practical Experience of Teaching OWL-DL: Common Errors & Common Patterns. In *EKAW 2004*, 2004.

[4] Rolf Schwitter. Controlled Natural Language as Interface Language to the Semantic Web. In *2nd Indian International Conference on Artificial Intelligence (IICAI-05)*, Pune, India, December 20–22nd 2005.