

Travelling Domain Experiment: Preliminary Results for OilEd

Sean Bechhofer
Information Management Group
University of Manchester
Kilburn Building
Oxford Road
Manchester M13 9PL

Introduction

OilEd [Bechhofer01b] is a tool that allows the construction of DAML+OIL ontologies. It is targeted primarily at ontology construction (rather than knowledge base population), and thus concentrates on the language constructors related to classes and the relationships between them. A key aspect of **OilEd** is its use of the FaCT reasoner which allows the user to produce classification hierarchies and check classes for inconsistency.

As **OilEd** is not intended for knowledge base construction, this experiment concentrated primarily on trying to represent information about the classes that arise in the example scenario, and the relationships and constraints that appear between them. In addition, as data type support is limited, little was done to model attributes with concrete ranges. As a result, the model produced only really covers a small subset of the application domain – additional functionality would certainly be needed if this were to be used within an application. However, the results demonstrate how we can use DAML+OIL, in particular axioms and expressive power such as negation, for the representation of some quite sophisticated constraints.

OilEd is not an application for query or ontology delivery, so it is difficult to assess in a satisfactory way the model produced in terms of query or use in an application. However, as discussed below, we can see how definitions in the resulting model might be used in applications.

Building the Model

Building the Model consisted of 3 basic stages:

Knowledge Acquisition: identifying the basic classes and properties;

Definition: identifying the relationships between the basic classes;

Constraints: identifying the constraints that limit the ways in which descriptions can be formed.

The process was iterative. Once the basic definitions were in place, these were refined and further elaborated. The reasoner was used regularly throughout the process to assist in organisation of the model and check for consistency.

Vehicles and Accommodation

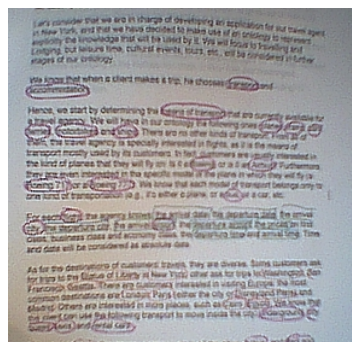


Figure 1 KA

The first step was basic knowledge acquisition. The NL description was examined in order to determine what the major classes occurring in the model were likely to be. This basically involved identifying things like the various kinds of Vehicle (Bus, Car etc) and the kinds of Accommodation (Hotel, B&B). Figure 1 shows a marked up section of the scenario. The concepts identified were added to the model as basic primitive classes.

In addition to the concepts, a number of basic relationships and attributes were identified. In the scenario, many of these are what DAML+OIL calls datatype properties, for example arrival or departure dates and times, number of rooms, price etc. The support for data types in **OilEd** is minimal. Relationships can be introduced and asserted to be data type properties and a range can be given, but

no complex data types such as ranges are supported.

Once the basic classes had been added, abstract classes were introduced into the model, for example in order to represent the different kinds of vehicle that can be used. These were introduced as primitives, e.g. **Air Vehicle**, **Land Vehicle** and **Water Vehicle**. One possible approach would be to model these as defined classes – for example a **Land Vehicle** is a vehicle that travels across land. A decision was taken that this was not necessary (at this point). If it became useful to explicitly represent and model these properties, these definitions could subsequently be refactored “in place”. An axiom was added stating that the **Vehicle** class was covered (disjointly) by the three vehicle types (see the discussion below about round tripping). This precludes us from including, for example, amphibious vehicles and asserts that there are no other kinds of **Vehicles** than the three introduced. Again, for this particular scenario, this was deemed acceptable.

Places

The notion of a **Geographical Location** encompasses **Countries**, **Cities** and **Attractions**. A rather general notion of containment is provided – **Geographical Locations** can be in other **Geographical Locations**. This relationship is transitive, thus an **Attraction** which is in a **City** which is in a **Country** is also in that **Country**.

Hotels and **B&Bs** are kinds of **Places to Stay**. These have a number of datatype properties (relating to **address**, **telephone number** etc). As discussed above, this particular model makes no attempt to deal with these relationship other than by simply introducing them.

Particular places (**USA**, **Manchester**, **EuroDisney**) have been introduced as individuals in the model. Although **OilEd** is not intended for large scale knowledge base construction, individuals can be introduced for use in “one-of” expressions. These can then be used in definitions such as a **Trip from Europe to America** (as discussed below).

Trips and Journeys

Once the basic artefacts of the model were in place (**Vehicles** and **Places**), concepts representing travel were introduced. A **Trip** is the basic atomic unit. Every **Trip** must have a departure and arrival point. **Trips** can use at most one vehicle. **Trips** can be combined into **Journeys**. Ideally a constraint should be in place that the trips that make up a journey all fit together in a sensible fashion, but this is not currently represented.

Subclasses of trips can now be defined: for example a **Flight** is a **Trip** that uses an **Air Vehicle**. A **Taxi Ride** is a trip using a **Taxi**. Axioms are used to constrain the start and end points of **Trips**. For example, any **Flight** must begin and end at an **Airport**. Similarly, a **Sailing** (a **Trip** using a **Water Vehicle**) must begin and end at a **Sea Port**.

Further compositional concepts can now be defined. An **Overland Journey** is one which consists solely of trips using **Land Vehicles**. A **Long Journey** is one which involves at least three trips. The expressiveness of DAML+OIL coupled together with the editing facilities of **OilEd** allow us to make these definitions without having to introduce intermediate descriptions. For example, the definitions of **Overland Journey** as shown in Figure 2 uses a universal restriction whose filler is an existential restriction.

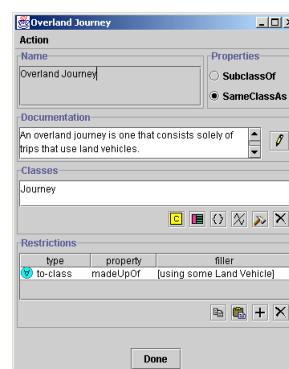


Figure 2 Overland Journey

We can now begin to describe additional constraints on the way in which trips can be put together. For example, it is impossible to travel from **North America** to **Europe** by train, car, bike or motorbike. This is represented by an axiom which states that any **Trip** from a place in **North America** to a place in **Europe** cannot use a **Land Vehicle**.

Discussion

Although this was a small scale experiment, and concentrated on a subset of the described domain, there were some interesting aspects. The model took around a day to produce, from examining the NL description to final production of the DAML+OIL model. The classified ontology was also exported as HTML and vanilla RDFS (including the inferred superclasses found by the classifier), and a plot of the hierarchy was produced using output from **OilEd** and **dotty**¹.

¹ <http://www.research.att.com/sw/tools/graphviz/>

Some Examples & Queries

As discussed above, **OilEd** is not intended for knowledge base construction, and does not directly support queries. However, we can use **OilEd** to define concept expressions which represent queries which might be posed by an application. These can then be classified against the ontology to check their consistency and their relationship with other descriptions – this use of the classifier was the key aspect explored in the experiment.

For example, a **Non Flyer** is defined as somebody who will only take **Journeys** that are made up of **Trips** that do not use **Air Vehicles**. A **Non Airbus Traveller** is someone who will not take a trip that involves something made by **Airbus**. An interesting side effect of this definition is the discovery that a **Non Flyer** is also a **Non Airbus Traveller** (as an axiom states that **Airbus** only make aircraft).

Such definitions will be useful for representing the preferences of travellers in an application. In the example model, **XX Non Flyer** is provided as a specialisation of **Non Flyer** that takes a trip that includes a **Flight**. This is spotted as an inconsistency by the classifier.

In a similar vein, the concept **XX Europe to America** is defined as a **Trip** from **Europe** to **America** using a **Bus**. This is again spotted as an inconsistency.

Another simple example of an inconsistency is the concept **XX Journey**, which is defined as being an **Overland Journey** that includes a **Flight**.

Round Tripping

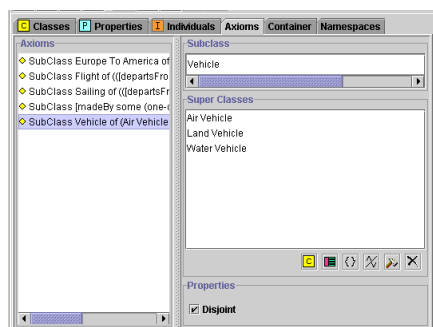


Figure 3 Original Axiom

ontology. When the model is saved and then re-read by the tool, the information is presented as an explicit superclass (of **Vehicle**) and a disjointness axiom. Of course, the semantics of this ontology are identical to those of the original, but the information is being presented in a different way. This is, admittedly, partly a tools issue – the tool could perhaps do more to try and preserve the information about the so-called *semiotic* information [Euzenet00] in the ontology, through extensions to the RDF used (although this would also require mechanisms for grouping and referring to collections of statements in the RDF model).

As discussed in [Bechhofer01a], there can be problems when “round-tripping” from **OilEd** in to DAML+OIL and back again. This was illustrated here by the axioms relating to **Vehicles**. In the original model, a covering axiom was added stating that **Land**, **Water** and **Air Vehicles** formed a disjoint covering of **Vehicle**. This is represented in DAML+OIL using a subClass axiom with a disjointUnionOf Class description. However, there is no information in the DAML+OIL that tells us how this was actually presented in the original

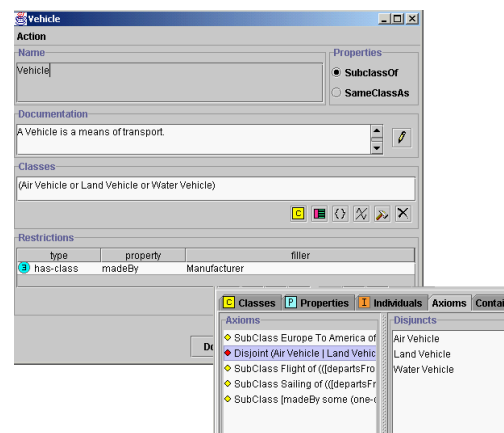


Figure 4 After Round Tripping

References

- [Bechhofer01a] Sean Bechhofer, Carole Goble, Ian Horrocks. *DAML+OIL is not enough*. SWWS-1, Semantic Web working symposium, Stanford (CA), July 29th-August 1st, 2001
- [Bechhofer01b] Sean Bechhofer, Ian Horrocks, Carole Goble, Robert Stevens. *OilEd: a Reason-able Ontology Editor for the Semantic Web*. Proceedings of KI2001, Joint German/Austrian conference on Artificial Intelligence, September 19-21, Vienna. Springer-Verlag LNAI Vol. 2174, pp 396–408. 2001.
- [Euzenet00] Jerome Euzenet. Towards formal knowledge intelligibility at the semiotic level. In *ECAI 2000 Workshop Applied Semiotics: Control Problems, Berlin (DE)*, pages 59–61, 2000.