Evaluation experiment of ontology tools' interoperability with the WebODE ontology engineering workbench

Óscar Corcho, Asunción Gómez-Pérez, Danilo José Guerrero-Rodríguez, David Pérez-Rey, Alberto Ruiz-Cristina, Teresa Sastre-Toral, M. Carmen Suárez-Figueroa[(*)]

Laboratorio de Inteligencia Artificial
Facultad de Informática
Universidad Politécnica de Madrid
Campus de Montegancedo sn.
Boadilla del Monte, 28660. Madrid, Spain
(*)Contact author: mcsuarez@delicias.dia.fi.upm.es

Abstract. This paper presents the results of the interoperability experiment proposed in EON2003, using the following ontology tools: Protégé-2000 and WebODE. We will show which knowledge is preserved and which knowledge is lost in the import/export processes between tools when using RDF(S) as an intermediate language.

1 Introduction

Protégé-2000 1.8[1] [6] and WebODE 2.0[2] [4, 1] are ontology platforms which are able to import and export ontologies in different languages (RDF(S), DAML+OIL, etc.). These ontology platforms and their RDF(S) import and export services have been used in our interoperability experiment.

This document analyzes how ontologies are exchanged (exported and imported) between the previous ontology tools using RDF(S) [2, 5]. We have studied which type of knowledge is preserved and which knowledge is lost during ontology export and import in such tools. In our experiment we have reused the travel ontology built in WebODE for the EON2002 workshop [3].

2 Interoperability experiment with WebODE and Protégé-2000

In order to analyze the interoperability between WebODE and Protégé-2000, we have carried out the following process:

1. Reuse the travel ontology built in WebODE for the EON2002 Workshop [3], and export such ontology to RDF(S) using the WebODE RDF(S) export service.

---

[1] http://protege.stanford.edu/

[2] http://webode.dia.fi.upm.es/

2. Import this RDF(S) ontology in Protégé-2000.
3. Export the ontology from Protégé-2000 to RDF(S).
4. Import the Protégé-2000 RDF(S) ontology in WebODE, and analyze the differences between the original ontology (reused ontology) and the ontology that results from this circular import/export process.

Figure 1 shows the circular import/export process that we have carried out in the first part of our interoperability experiment.
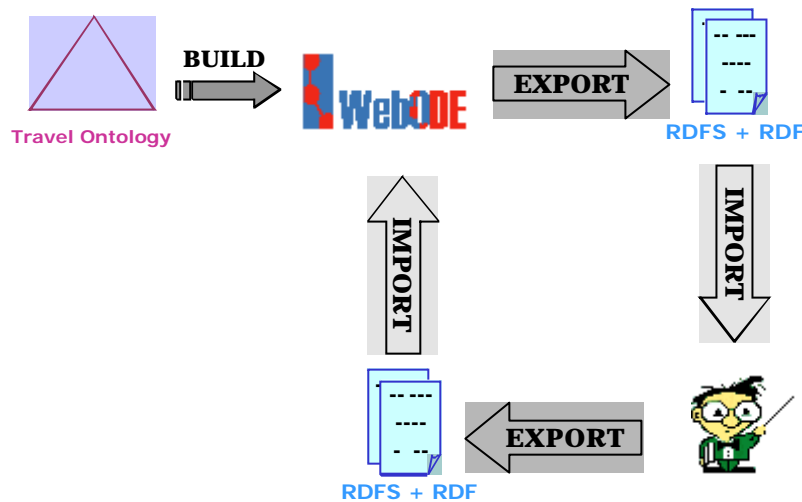

Figure 1. Circular import/export process using WebODE and Protégé-2000.

2.1 Step 1. Export to RDF(S) using WebODE

The WebODE ontology in the travel domain described in [3] and shown in figure 2 have been first exported automatically to RDF(S).
We have studied the generated RDF(S) files, and we can mention the following features:
- WebODE generates a ZIP file that contains:
  - One file for the conceptualization of the ontology (*travel_fromWebODE.rdfs* which contains the classes and properties of the ontology).
  - One file for each instance set that the user has decided to export (which contain the instances of that instance set). In our case, we have exported one of the instance sets, the one corresponding to the travel agency in New York (*travelAgencyNY_fromWebODE.rdf*).
- As a difference with the RDF(S) export function of other tools, such as Protégé-2000, WebODE does not export all the knowledge of the ontology as it is defined in the original ontology, but only those pieces of knowledge that can be directly represented with the standard knowledge model of RDF Schema. Consequently, axioms defined in the original ontology are not exported, disjoint and exhaustive

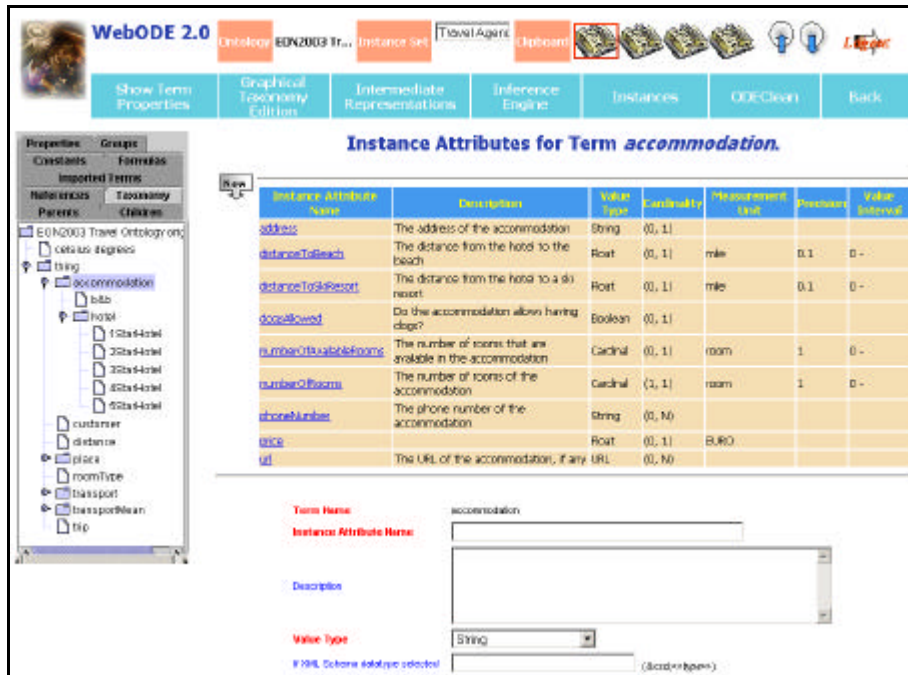decompositions and partitions are not exported as such but as subclass-of relationships, etc.

Figure 2. Edition of instance attributes of the concept *accommodation* with the WebODE ontology editor.

- In the RDF(S) export process, the user is requested the namespace of the ontology to be exported. We have used the namespace: *http://webode.dia.fi.upm.es/RDFS/EON2003_Travel_Ontology#*. The files exported contain the following predefined namespaces for the RDF and RDFS prefixes:
  - *rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#*
  - *rdfs: http://www.w3.org/2000/01/rdf-schema#*

We have found the following problems in the exported RDF and RDFS files:
- The concept *b&b* has a different identifier than the one used in WebODE, as follows:

      <rdf:Description rdf:about='#b&b'>

- The relation *usesTransportMean*, which is defined in WebODE between the following pairs of concepts: *(carRented, car)*, *(cityBus, bus)*, *(flight, airTransportMean)*, *(undergroundTransport, underground)*, *(transport, transportMean)*, is defined only once in the generated RDFS file. This is due to the fact that RDF does not allow homonymous property names. Besides, in RDFS this property does not have its domain nor its range defined.

- The same applies to class and instance attributes, which are necessarily attached to a concept in WebODE, so that we can have different attributes with the same name in different concepts. For instance, the class attribute *numberOfStars* is defined once in the RDFS file, while it is defined for five classes in WebODE (*1StarHotel*, *2StarHotel*, etc.). In this case, neither the domain nor the range are specified in the RDFS file.

- Finally, since the RDF(S) export function was developed when the treatment of datatypes was not clear in the RDFS specification, the current RDF(S) export function converts all the types of WebODE instance and class attributes to *rdfs:Literal*.

- WebODE constants are transformed into concepts in RDF(S). For instance, the constant *celsius degrees* is transformed into the concept *celsius_degrees*. Consequently, it loses its value.

2.2 Step 2. Import the RDF(S) files generated by WebODE into Protégé-2000

We have imported into Protégé-2000 the RDF and RDFS files generated in the previous stage of our experiment. During the import process, the following comments have been provided by Protégé-2000:

- Protégé-2000 has recognized four namespaces in the ontologies imported:
  - *rdf*, *rdfs*, and the base namespace of the ontology
  - One additional namespace that appears as the value of a property for a hotel: http://holidayinn.com.

```
<NS0:url rdf:resource='http://holidayinn.com/13492'/>
```

- Besides, the values of class attributes that were exported from WebODE to RDF(S) are not correctly imported (e.g., the number of stars of a hotel, the air company in charge of a flight, etc.). Protégé-2000 shows a warning that alerts the user that this "own slot" has not been defined in a metaclass, as shown in figure 3. Consequently, this information is lost.
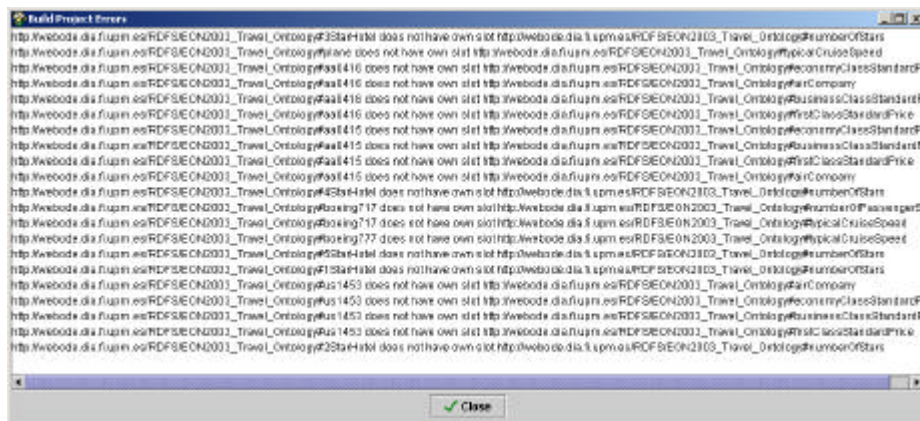

Figure 3. Own slots' import problem with Protégé-2000.

The result of the import process is shown in figure 4. There we can see the details of the concept *accommodation*, whose template slots are the same as those defined in WebODE (except for *hasRoom* and *placedIn*, which were defined as relations in WebODE). However there are some differences between these attributes and relations, which are related to their cardinalities and types. As a result of using RDFS as an exchange language, we have lost the cardinality information for template slots. Additionally, the types "integer", "Boolean", etc., have been transformed to "String" in Protégé-2000, since they were transformed by WebODE to *rdfs:Literal*. Finally, the type of the slot *ur*" is "Instance", of the class *:THING*, as it was transformed to a property whose range was *rdfs:Resource* by WebODE.

After the import process, we have compared the WebODE ontology and the Protégé-2000 ontology (shown in figure 4), finding the following differences:
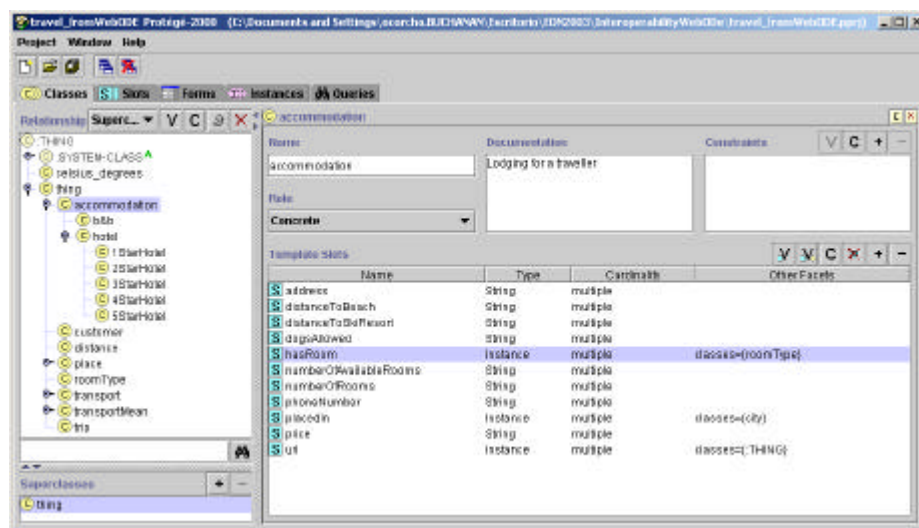
Figure 4. Travel ontology in Protégé-2000.

- Attributes whose type was "integer" or "Boolean" in WebODE have changed in Protégé-2000 to type "String". This is due to the fact that the RDFS file already contained a transformation of these basic types to *rdfs:Literal*.
- The cardinalities of attributes have changed. All of them have 0 as a minimum cardinality and N as a maximum cardinality (that is, they are defined as "multiple").
- The class attributes defined in WebODE have disappeared, because of the own slot problem described in the import process.
- The attributes with multiple documentations (multiple *rdfs:label* properties attached) have now one single documentation that joins all of them.
- The knowledge about disjoint and exhaustive decompositions, and partitions is lost in Protégé-2000, since it was not available in the RDF(S) files. The same applies to axioms, concept groups, constants, etc.

- The values of the attribute *url* for two of the instances have been transformed to instance themselves, as instances of the class *:THING*. In WebODE and RDF(S) they were just URIs.

Since Protégé-2000 is not able to work with different instance sets at the same time, we have been only able to import one of the instance sets that could be exported by WebODE.

2.3 Step 3. Export the Protégé-2000 ontology to RDF(S)

Finally, we have exported the Protégé-2000 ontology to RDF(S) and we have obtained two files, one for the classes and another one for the instances. There are many differences (mainly syntactic) between the original RDF(S) files and the target RDF(S) files generated, as can be seen by simply comparing the four files.

2.4 Step 4. Import the RDF(S) ontology generated by Protégé-2000 into WebODE

In order to import the ontology into WebODE, we have had to join the two files generated by Protégé-2000 into only one file that contains both the ontology conceptualization and the instances. This file is called *Travel_fromProtegetoWebODE.rdf_s.*

In this import process we have found the following problems:
- Protégé-2000 uses a namespace for the RDFS KR ontology that comes from an old specification: *http://www.w3.org/TR/1999/PR-rdf-schema-19990303#*. This causes the WebODE RDF(S) import function to not correctly detect the concepts defined in the ontology. Consequently, we have edited the file manually so as to change this namespace by the following: *http://www.w3.org/2000/01/rdf-schema#*.
- The concepts whose identifier starts with a digit have not been imported correctly. As a consequence, we had to rename manually the terms *1StarHotel*, *2StarHotel*, *3StarHotel*, *4StarHotel*, and *5StarHotel*.
- The same applies to the instances whose identifier starts with a digit. In this case, the WebODE import function notifies the following error:
  "Error importing RDFS ontology: Error occurred in server thread; nested exception is: com.hp.hpl.mesa.rdf.jena.model.RDFError: org.xml.sax.SAXParseException: An invalid second ':' was found in the element type or attribute name."
  which is not much descriptive about the problem in the source RDF(S) file.

In this case, we have compared the original ontology built in WebODE and the resulting ontology of importing the RDF(S) of Protégé-2000 in WebODE (shown in figure 5).

We have found the following differences in our comparison:
- A new concept is generated in WebODE (*rdfs:Resource*) which is used as the root concept of the ontology.

- New relations, which did not exist in the original ontology, appear in the imported ontology. These relations were represented as attributes of type *URL* in the original ontology. Since they were transformed into slots with range *:THING*, and transformed back to RDFS as properties with range *rdfs:Resource*, they have not been recovered as originally during the last import process.

Figure 5. Travel ontology imported from Protégé-2000 RDF(S).

- The concept *b&b* (bed and breakfast) has been transformed to *b*, because of the symbol &.
- The documentation of concepts, attributes, relations, etc., now have more text: they include the term label (as defined in the Protégé-2000 RDF(S) files) and the comment, which was the original documentation.
- The cardinalities and types of the instance and class attributes are different from those that were originally present in WebODE. This knowledge was lost in the first step.
- All the information that was already lost in the first export process is, of course, missing: disjoint and exhaustive decompositions, partitions, axioms, etc.
- Relations with the same name represented in WebODE (e.g., *usesTransportMean*) are now transformed into a unique relation whose domain is *rdfs:Resource*.
- Class and instance attributes with the same name represented in WebODE (e.g., *airCompany*) are now transformed into a unique relation whose domain and range is *rdfs:Resource*. This is due to the fact that their domain was not exported to RDF(S) in step 1.

3 Conclusions

The table 1 summarizes the main conclusions of this circular import/export processes, with the number of ontology components that can be found in each of the ontologies generated during the process. We do not care about other issues, such as differences in the domains, ranges, cardinalities, term names, etc.

	WebODE	RDF(S) (step 1)	Protégé-2000 (step 2)	RDF(S) (step 3)	WebODE (step 4)
#concepts	62	62	62[3]	62	**63**
#subclass of	24	**61**	63	63	63
#disjoint decompositions	6	**0**	0	0	0
#exhaustive decompositions	0	**0**	0	0	0
#partitions	3	**0**	0	0	0
#attributes/relations	69	**43**	43	43	**44**
#axioms	8	**0**	0	0	0
#constants	1	**0**	0	0	0
#instances	20	20	**22**	22	**20**

Table 1. Summary of knowledge preserved and lost during the circular import/export process

The most relevant comments that can be extracted from the previous table are the following:

- WebODE creates a new concept when importing ontologies from RDF(S). This class is *rdfs:Resource*, which is used as the root concept of all the ontology concepts, and is also used as the domain and/or range of several ad hoc relations for which the domain/range has not been defined explicitly in the RDFS file.
- With regard to the taxonomic relationships between concepts, we have two comments:
  - WebODE is able to represent disjoint and exhaustive knowledge in its concept taxonomies. However, with RDFS we cannot represent this kind of knowledge, and consequently it is transformed into simple subclass of relationships. This is the reason why there are 24 subclass of relationships defined in the original ontology, and they are transformed into 61 in the RDF(S) file and successive transformations.
  - Besides, when importing the ontology from RDF(S) to Protégé-2000 two new subclass of relationships appear. These are related to the use of the class *:THING* as the root class of any Protégé-2000 ontology. As a consequence, the classes *thing* and *celsius_degrees* from the original ontology are explicitly declared as subclasses of *:THING*.

---

[3] This figure does not include the system classes that are always generated by Protégé-2000

- The number of attributes and relations that are present in the original ontology is quite different than that of the ontology generated in RDF(S) and obtained in the subsequent processes. The reason for this is that WebODE allows representing different attributes and relations for different concepts with the same name. This is not allowed neither in RDF(S) nor in Protégé-2000. Consequently, in the transformation, attributes and relations with the same name are transformed into only one attribute/relation.
- We have discovered an error in the import process of WebODE with the RDF(S) property *url*, whose range is *rdfs:Resource*. This property is transformed into an attribute of type *URL* and a relation between the concept accommodation and the concept *rdfs:Resource*.
- Axioms and constants are lost in the transformation to RDF(S), since they cannot be represented in this language.
- Finally, the number of instances is constant, except for the import to Protégé-2000, in which instances are created for two resources that appear as the range of the property *url* (holidayInn hotels' URLs), and except for the import to WebODE, where these instances are lost since they are instances of *rdfs:Resource*.

Acknowledgments

References

1. Arpírez JC, Corcho O, Fernández-López M, Gómez-Pérez A (2003) *WebODE in a nutshell*. AI Magazine 24(3):37:48
2. Brickley D, Guha RV (2003) *RDF Vocabulary Description Language 1.0: RDF Schema*. W3C Working Draft. http://www.w3.org/TR/PR-rdf-schema
3. Corcho O, Fernández-López M, Gómez-Pérez A (2002) *Evaluation experiment for the editor of the WebODE ontology workbench*. In: Angele J, Sure Y (eds) EKAW02 Workshop on Evaluation of ontology-based tools (EON2002). Sigüenza, Spain
4. Corcho O, Fernández-López M, Gómez-Pérez A, Vicente O (2002) *WebODE: an Integrated Workbench for Ontology Representation, Reasoning and Exchange*. In: Gómez-Pérez A, Benjamins VR (eds) 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW'02). Sigüenza, Spain. (Lecture Notes in Artificial Intelligence LNAI 2473) Springer-Verlag, Berlin, Germany, pp 138–153
5. Lassila O, Swick R (1999) *Resource Description Framework (RDF) Model and Syntax Specification*. W3C Recommendation. http://www.w3.org/TR/REC-rdf-syntax/
6. Noy NF, Fergerson RW, Musen MA (2000) *The knowledge model of Protege-2000: Combining interoperability and flexibility*. In: Dieng R, Corby O (eds) 12th International Conference in Knowledge Engineering and Knowledge Management (EKAW'00). Juan-Les-Pins, France. (Lecture Notes in Artificial Intelligence LNAI 1937) Springer-Verlag, Berlin, Germany, pp 17–32