Ontological Mappings of Product Catalogues

Domenico Beneventano[1] and Daniele Montanari[2]

[1]Università di Modena e Reggio Emilia, Dipartimento di Ingegneria dell'Informazione
Via Vignolese 905, 41100 Modena, Italy
domenico.beneventano@unimore.it

[2]Eni S.p.A., ICT / Semantic Technologies Via Arcoveggio 74/2, 40129 Bologna, Italy
daniele.montanari@eni.it

Abstract. In this paper we built on top of recent effort in the areas of semantics and interoperability to establish the basis for a comprehensive and sustainable approach to the development and later management of bridging systems among a variety of corporate system that need to be interconnected without being individually modified. In particular, we collect some preliminary evidence that a sustainable approach exists to the definition of mappings which can withstand changes of the underlying classification schemes. This in turn adds evidence towards the feasibility of a dynamic interoperable infrastructure supporting a global adaptive electronic market place.

1 Introduction

The goal of this paper is to combine some results on the development of ontologies for products and services classification with other results in the area of system interoperability and ontology mapping to study the impact of evolution of the reference ontologies onto the catalogues/classification system annotated and then derivatively mapped w.r.t the ontologies. Slightly more formally, given comparable catalogues C1, C2, and assuming that OntologyA and OntologyB are reference ontologies which have been used to annotate the content of C1 and C2 respectively, given a mapping between OntologyA and OntologyB which provides a correspondence between concepts and relations in the two ontologies, a derived mapping can be defined at the level of the catalogues C1 and C2 (see Figure 1).

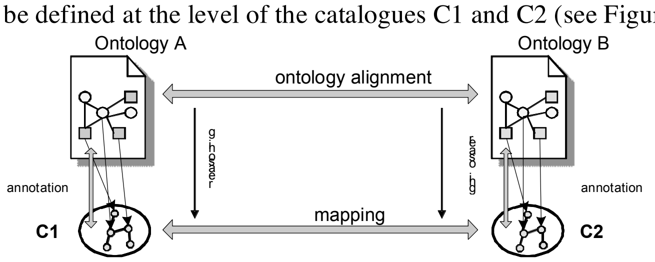

Figure 1 Mappings at reference and catalog levels [1]

It has been observed [7] that product and service ontologies exhibit a significant evolution of their content in time, due to changing market condition, and the evolving user sophistication and needs. This implies that the definition of the mapping between the catalogues will not be a one-time operation, but rather a repeated operation following the version cycles of the involved ontologies.

Being a heavy semi-automatic operation, the cost of the change of mapping must be carefully assessed, and understanding whether there are ways of minimizing the impact of these reviews can be a valuable information for people planning to position themselves, their products, and their services in an electronic market where they need to interoperate with other heterogeneous actors / systems.

This work was partially funded under the European Commission 6th Framework Programme under contract FP6-2005-IST-5-034980 (STASIS).

2 Ontologizing Product Catalogues

The first step when entering the semantic dimension of a field consists in providing a semantic reference for all the relevant entities in the domain. Several product and service classification schemes are available nowadays, both as in-house developments fulfilling the needs of their original users, and as more or less open public standards; some well known such schemes are UNSPSC [13], eCl@ss [6] and RosettaNet [11]. A good account of the subtleties lurking in the conversion from classification system into ontology is presented in [7]. One crucial point made is that the typical hierarchies of classification entities found in a classification system, being driven among other things, by the typical needs of purchasing departments in terms of searching, reporting, and classifying suppliers of goods and services.

Once we have reference ontologies derived from the standard classification systems, we can use them to annotate a given catalog of products and / or services. In [1] we describe a technique which derives an ontology for a specific database schema or semi-structured set of information (like web or XML pages); this technique was experimented in the STASIS project (http://www.stasis-project.net).

Let us introduce a real example of catalog by considering the eBay catalogue. This catalog is structured in three kinds of elements, called *categories*, *items* and *attributes*. Our Semantic Annotation of a Catalog with respect to a product ontology is based on the annotation of categories (called semantic entity in [1]) and is formally defined as follows. An annotation element is a tuple < *SE, AR, concept_description*> where *SE* is a semantic entity of the catalog; *concept_description* is a concept description of the *product ontology*; AR specifies the *Annotation Relation* which may hold between *SE* and *concept_description*; we consider *equivalence* (*AR_EQUIV*); *more general* (*AR_SUP*); *less general* (*AR_SUB*). Let us give some examples.

- (eBay:ClassicToys *AR_SUB* UNSPSC:Toys) this annotation declares that the entity eBay:ClassicToys is less general than the concept UNSPSC:Toys
- (eBay:ClassicToys *AR_SUP* UNSPSC:ToyTrains) this annotation expresses the fact that all instances of UNSPSC:ToyTrains would be classified in eBay:ClassicToys

3 Derived Mappings between Ontologized Catalogues

Assume now that several catalogues ontologized with respect to some standard ontologies. We now want to establish correspondences among two or more such catalogues, so that e.g. our purchasing department will be able to see and compare the offer of different suppliers for the same class of goods. The plan is to align the relevant parts of the reference product and services ontologies used to annotate the catalogues, and then to derive a map on the underlying catalogues.

Ontology alignment

The basic expression of alignment mapping for ontologies modeled with description logic formalisms involves the use of a semantic (logic) constructs or evolved frameworks to express the existence and properties of similarities and then mappings [4][7][10][12]. In this paper we use a somewhat simplified setup. Let O1 and O2 be ontologies. Then, an *entity alignment mapping* between entities E1 in O1 and E2 in O2 is a tuple < E1, AM_R, E2> where AM_R specifies the semantic relation which holds between E1 and E2: equivalence (AM_EQUIV), subClass (AM_SUBS) and superClass (AM_SUP). The above notation then reads "E1 is a AM_R of E2".

For example < UNSPSC:ToyTrains, AR_SUBS, ECLASS:Toys>

The mapping process

We are in a position now to establish mappings between our (ontologized) catalogues at last, and the reader should keep in mind the picture in Figure 1. The idea is that the mappings at the ontology level will actually induce mappings at the lower level.

Let's begin with a simple example. Given the eBay catalog, and another catalog that we call C2. Suppose that eBay has been ontologized with respect to UNSPSC, C2 has been ontologized using eCl@ss, and an alignment mapping has been established between UNSPSC and eCl@ss. If the following three facts have been established:

1. (eBay:ClassicToys AR_SUB UNSPSC:Toys)
2. (C2:SE1 AR_SUP ECLASS:Toys)
3. <ECLASS:Toys, AM_EQUIV UNSPSC:Toys>

Then, from 1. and 3. we can deduce: (eBay:ClassicToys AR_SUB ECLASS:Toys); And from 2. and 4. we conclude <C2:SE1, SUP, eBay:ClassicToys> (A), which establishes a mapping at the ontologized catalog level (where SUP denotes moreGeneral at the ontologized catalogues level). This is a *derived mapping* from the ontology alignment, realizing the picture in Figure 1. We should note now that if we had 2'. (C2:SE1 AR_SUB ECLASS:Toys), then our reasoning would collapse and we would not be able to assert any mapping at the ontologized catalogues level. This is a common occurrence, since in real life conditions there is no guarantee that all of the mappings at the ontology level will actually induce mappings at the lower level.

The type of mapping should also be considered. The statement (A) above declares that a certain entity in the C2 catalog includes all the eBay:ClassicToys. This is a true fact, but it is not obvious that it is the fact we want. For example, we might want to have a stronger or stricter property. This may come as a further deduction from other mappings, but it may not. In the latter case, those in charge of the mapping need to enhance the annotation of the catalogues, refine the ontology alignment and finally, if all else fails, force the desired mappings by hand

At this point the discriminating reader may wonder whether this process actually returns some mappings at the ontologized catalogues level. The answer is affirmative, at least in some reasonable circumstances. In fact, we can state the following

Theorem. Assume that O1 and O2 are reference ontologies ontologizing the catalogues C1 and C2 via annotations A1, A2 resp. For all entities E1 in C1 and E2 in C2 with annotations (C1:E1 AR_SUB O1:o1), (C2:E2 AR_SUB O2:o2), if we have the mapping <M1, O1:o1, AM_SUB, O2:o2> and O2:o2 is the image of C2:E2 via the annotation A2, then M1 translates into a mapping <T1, C1:E1, SUB, C2:E2>

The proof of this statement follows immediately from the unfolding of the definitions. This theorem shows that mappings at the ontologized catalog level are generated indeed, provided that we can map all the entities in our classification schemes into entities in the reference ontologies, which is mostly the case if the reference ontologies are worth their salt.

Next example shows how a property established in an ontology may propagate to the other ontology and both ontologized catalogues. Let O1, O2, be reference ontologies, and C1, C2, catalogues that have been ontologized with respect to O1, O2, E1i (i=1,2) entities in C1 and E2i (i=1,2) entities in C2, o1i (i=1,2) classes in O1 and o2i (i=1,2) classes in O2. Assume the following facts:

1. (C1:E1i AR_SUB O1:o1i) (i=1,2)
2. (C2:E2i AR_SUB O2:o2i) (i=1,2)
3. <M1i, O1:o1i, AM_SUB, O2:o2i> (i=1,2)
4. areDisjoint(O2:o21,O2:o22)

Then, a reasoner should be able to infer that E21 and E22 are disjoint, that O1:o11 and O1:o12 are disjoint, and finally that E11 and E12 are also disjoint. The nice outcome of this line of reasoning is therefore that any strong separation property established in O2 will propagate to O1 and both catalogues. This means that a comparison of ontologized catalogues can propagate qualifying properties and improve the quality of all the structures involved.

4 Conclusions and Future Work

This paper outlines the results of some scouting done in the area of effective and sustainable management of mappings among common industry tools like catalogues. While this exercise applies some general techniques in a specific context, it is suggestive of potential generalizations and difficulties to be tackled next.

The most interesting development should be to understand the relation between the mappings at the ontology level and derived mappings at the ontologized catalogues level in Figure 1, as modulations of the annotations of the catalogues using the ontologies.

Moreover, a more extensive approach including relations, instances, properties, rules, axioms, and constraints should be progressively pursued. This will enhance our

understanding of the properties that we should strive to identify a priori, in order to ease a forthcoming mapping process. More generally, the interplay of annotations and alignments could be investigated for general mappings between ontologies.
Finally, catalogues are one single area of interest. They are usually simply structured, yet large, occasionally idiosyncratic, evolving in time, reflecting real business needs. As such, they are a very relevant sandbox to try ideas for semantic applications. Eventually these techniques should migrate to other fields like EDI and general industrial and commercial operations of all kinds.

6 References

[1] Beneventano, D., Dahlem, N., El Haoum, S., Hahn, A., Montanari, D., and Reinelt, M., 2008. Ontology-driven Semantic Mapping, (I-ESA 2008) in Enterprise Interoperability III – Springer, Berlin, pp. 329-342.

[2] Beneventano, D., El Haoum, S., Montanari, D., 2007. Mapping of heterogeneous schemata, business structures, and terminologies, in DEXA Workshop on Database and Expert Systems Applications, IEEE Comper Society, pp. 412-418.

[3] D. Beneventano, D., Magnani, S., A 2004. Framework for the Classification and the Reclassification of Electronic Catalogs, 19th Annual ACM Symposium on Applied Computing Nicosia, Cyprus, March 14 -17.

[4] Choi, N., Song, I-Y., Han, H., 2006. A Survey of Ontology Mapping. In SIGMOD Record 35, Nr. 3.

[5] Corcho, O. and Gomez-Perez, A., 2001. Solving Integration Problems of E-commerce Standards and Initiatives through Ontological Mappings, Workshop on E-Business and Intelligent Web at the IJCAI2001 , Seattle, USA, August 5.

[6] eCl@ss web site: http://www.eclass-online.com/ (accessed on July 19, 2008).

[7] Euzenat, J., Mocan, A., Scharffe, F., 2008. Ontology alignments. In Hepp, M., De Leenheer, P., de Moor, A., and Sure, Y., Eds. 2008 Ontology Management. Springer. Semantic Web and Beyond: Computing for Human Experience.

[8] Hepp, M., 2006. Product and Service Ontologies: A Methodology for Deriving OWL ontologies from Industrial Categorization Standards. Int'l Journal on Semantic Web & Information Systems (IJSWIS), Vol. 2, No. 1, pp. 72-99.

[9] Klein, M., 2002. Web page on DAML+OIL and RDF Schema Representation of UNSPSC. http://www.cs.vu.nl/~mcaklein/unspsc/ (accessed on July 20, 2008).

[10] Ontology Alignment Evaluation Initiative (OAEI) web site: http://oaei.ontologymatching.org/ (accessed on July 21, 2008).

[11] RosettaNet http://www.rosettanet.org/cms/sites/RosettaNet/

[12] Shvaiko, P., Euzenat, J., 2005. A survey of schema-based matching approaches. Journal on Data Semantics 3730, pp. 146-171.

[13] UNSPSC web site: http://www.unspsc.org/ (accessed on July 19, 2008).