Towards composing and benchmarking ontology alignments

Jérôme Euzenat
INRIA Rhône-Alpes
Jerome.Euzenat@inrialpes.fr

Formal resources on the web will be expressed in the framework of an ontology. Integrating such resources requires finding agreement between ontologies (ontology alignment). Many methods have been put forward for aligning ontologies. They involve different techniques (linguistic, statistical, structural) and are based on various features of ontologies (names, internal structure, external structure, extension or semantics). It is necessary to allow these methods to cooperate and to be able to compare them in order to stimulate research on ontology alignment. We propose here a first attempt toward this based on a format for expressing the alignment independently of the methods used for building them and a first benchmarking framework for alignment methods. This sketch of foregoing work in the framework of the European FP6 'Knowledge web' is intended to be discussed and improved.

1 What's in an ontology alignment

Like the web, the semantic web will have to be distributed and heterogeneous. As such, the integration of resources found on this semantic web is its main problem. For contributing solving this problem, data will be expressed in the framework of ontologies. However, ontologies themselves can be heterogeneous and some work will have to be done for restoring interoperability.

Semantic interoperability can be grounded on ontology reconciliation: finding relationships between concepts belonging to different ontologies. We call this process "ontology alignment". The alignment result can be used for various purposes such as displaying the correspondances, transforming one source into the other, creating a set of bridge axioms between the ontologies or buiding query wrappers which rewrite queries for reaching a particular source.

The ontology alignment problem can be described in one sentence: given two ontologies which describe each a set of discrete entities (which can be classes, properties, rules, predicates, etc.), find the relationships (e.g., equivalence or subsomption) holding between these entities. Hence, in first approximation, an alignment is a set of pairs of elements from each ontology.

However, there are other aspects of alignments that can be added to this first approximation:

- There is not only equivalence/subsumption but more sophisticated operators (e.g., concatenation of firstname and lastname for instance considered in [2]).
- Another relevant point is to define the kind of alignment sought. Usual notations are 1:1, 1:m, n:1 or n:m. We prefer to note if the mapping is injective, surjective and total or partial on both side. We then end up with more alignment arities (noted with, 1 for injective and total, ? for injective, + for total and * for none and each sign concerning one mapping and its converse): ?:?, ?:1, 1:?, 1:1, ?:+, +:?, 1:+, +:1, +:+, ?:*, *:?, 1:*, *:1, +:*, *:+, *:*. These partial alignments (i.e. they align only one part of each ontology) could be provided as input (or constraints) of the alignment problem. This would allow iterative alignment: starting with a first alignment, followed by user feed-back, subsequent alignment rectification, and so on.
- Last, since many alignment methods compute a strength of the relation between entities, this strength can be provided as a normalized measure.

Then the alignment description can be stated as follows:

a set of pairs with characterization of the relation (default "=") and strength (default 1);

a statement of arity (default 1:1) and even more generaly a statement of the properties of the alignment when this can be provided by the alignment method (e.g., a subsumption preservation assertion), This is simpler than the alignment representation of [1], but is supposed possible to produce by most alignment tools.

Having alignment results in a standardized format can be very useful for taking advantage of these alignments in various contexts (transformations, queries, etc.). It can be used for modularizing alignments; for instance, by first using terminological alignment methods for labels, having this alignment agreed or amended by a user and using it as input alignment for a structural alignment. But it can also be used for benchmarking alignment methods. To that extent, we will need a measure of the distance between such an alignment and an expected target alignment.

2 Benchmarking alignment methods

There are various methods for computing alignments, however, it seems sensible to ask alignment methods for an output alignment, given:

- two ontologies to be aligned;
- an input partial alignment (possibly empty);
- a characterization of the wanted alignment (1:+, ?:?, etc.).

A measure of the distance between alignments would allow to evaluate alignment methods. There are two kinds of benchmarks which seems useful: competence benchmarks and performance benchmarks.

2.1 Competence benchmark

Competence benchmarks aim at characterising the kind of task each method is good at. There are many different areas in which methods can be evaluated. One of them is the kind of features they use for finding matching entities (this complements the taxonomy provided in [2]):

terminological (T) comparing the labels of the entities trying to find those which have similar names;

internal structure comparison (I) comparing the internal structure of entities (e.g., the value range or cardinality of their attributes);

external structure comparison (S) comparing the relations of the entities with other entities;

extensional comparison (E) comparing the known extension of entities, i.e. the set of other entities that are attached to them (in general instances of classes);

semantic comparison (M) comparing the interpretations (or more exactly the models satisfying the entities).

A set of reference benchmarks, targetting one type of feature at a time can be defined. These benchmarks would caracterize the competence of the method for one of these particular features of the languages.

2.2 Performance benchmarks

Performance benchmarks are aimed at evaluating the overall behaviour of alignment methods in versatile real-life examples. It can be organised as a yearly or bi-annual challenge (à la TREC) for comparing the best compound methods. Such benchmarks should yield as a result the distance between provided output and expected result as well as traditional measures of the amount of resource consumed (time, memory, user input, etc.).

References

[1] Jayant Madhavan, Philip Bernstein, Pedro Domingos, and Alon Halevy. Representing and reasoning about mappings between domain models. In *Proc. 18th National Conference on Artificial Intelligence (AAAI 2002), Edmonton (CA)*, pages 122–133, 1998. http://citeseer.nj.nec.com/milo98using.html.

[2] Erhard Rahm and Philip Bernstein. A survey of approaches to automatic schema matching. *VLDB Journal*, 10(4):334–350, 2001.