Automated support for evaluating alignment and matching algorithms
Version 1.0

Ontology Alignment Evaluation Initiative[1]

http://oaei.inrialpes.fr

May 1, 2005

[1]Coordinator: Jérôme Euzenat (INRIA Rhône-Alpes)

Abstract

This document briefly consider supporting tools for running ontology alignment evaluation.

Executive Summary

Heterogeneity problems on the semantic web can be solved, for some of them, by aligning or matching heterogeneous ontologies. Aligning ontologies consists of finding the corresponding entities in these ontologies. Many techniques are available for achieving ontology alignment and many systems have been developed based on these techniques. However, few comparisons and few integration is actually provided by these implementations.

The present report describes valuable tools for automating the evaluation process: generating tests (§2), running tests (§3) and evaluating results (§4). These tools are in part already implemented.

Contents

Chapter 1

Introduction

Aligning ontologies consists of finding the corresponding entities in these ontologies. There have been many different techniques proposed for implementing this process. They can be classified along the many features that can be found in ontologies (labels, structures, instances, semantics), or with regard to the kind of disciplines they belong to (e.g., statistics, combinatorics, semantics, linguistics, machine learning, or data analysis) [Rahm and Bernstein, 2001; Kalfoglou and Schorlemmer, 2003; Euzenat *et al.*, 2004a]. The alignment itself is obtained by combining these techniques towards a particular goal (obtaining an alignment with particular features, optimising some criterion). Several combination techniques are also used. The increasing number of methods available for schema matching/ontology integration suggests the need to establish a consensus for evaluation of these methods.

Beside this apparent heterogeneity, it seems sensible to characterise an alignment as a set of pairs expressing the correspondences between two ontologies. We proposed, in [Bouquet *et al.*, 2004], to characterise an alignment as a set of pair of entities (e and e'), coming from each ontologies (o and o'), related by a particular relation (R). To this, many algorithms add some confidence measure (n) in the fact the relation holds [Euzenat, 2003; Bouquet *et al.*, 2004; Euzenat, 2004].

From this characterisation it is possible to ask any alignment method, given

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

to output an alignment. From this output, the quality of the alignment process could be assessed with the help of some measurement. However, very few experimental comparison of algorithms are available. It is thus one of the objectives of the Ontology Alignment Evaluation Initiative to run such an evaluation. We have organised two events in 2004 which are the premises of a larger evaluation event:

– The Information Interpretation and Integration Conference (I3CON), held at the NIST Performance Metrics for Intelligent Systems (PerMIS) Workshop, is an ontology alignment demonstration competition on the model of the NIST Text Retrieval Conference. This contest has focused "real-life" test cases and comparison of algorithm global performance.
– The Ontology Alignment Contest at the 3rd Evaluation of Ontology-based Tools (EON) Workshop, held at the International Semantic Web Conference (ISWC), targeted the characterisation of alignment methods with regard to particular ontology features. This contest defined a proper set of benchmark tests for assessing feature-related behavior.

These two events are described more thoroughly in [Sure *et al.*, 2004] and [Euzenat *et al.*, 2004b].

Both evaluations we carried out shown that the job of participants and of running the evaluation were greatly facilitated by providing tools for the evaluation. The tools also have the good features of providing the results to the participants without ambiguity.

We present below both what is already available and how it is desirable to develop tools for evaluating ontology alignment algorithms.

Chapter 2

Test generation framework

We did not use so far any test generation system. However, our competence benchmark would highly benefit from such systematic test generation facility. It is thus necessary to have some tools which, from one ontology, are able to discard any of the features and to generate both the obtained ontology and the corresponding alignment.

This generation facility could be relatively easy to provide for simple changes such as discarding entities or replacing labels by random strings. It is a bit more complicated when it must:

- replace by missspellings which would require a missspelling generator;
- translate terms which would require an automatic translation tool (some could be used for that);
- flatten subsumption and composition hierarchies which is however feasible;
- expand subsumption and composition hierarchies in a meaningful way which is far more difficult.

Such a generation tool would take some ontology as input and systematically generate directories corresponding to the combination of all the features considered by the competence benchmark and containing the altered ontology plus the corresponding reference alignment.

It could be useful to implement such a tool with interactive manipulations.

Chapter 3

Alignment framework

The I3CON Experiement Set Platform is a workbench under which the participants who wanted it could adapt their tools and plug them in for generating the results. It also provided formats in n3 notation for alignments and measures.

The EON Ontology Alignment Contest made use of the Alignment API[1] for representing the resulting alignments. This API provide many different services (see [Euzenat, 2004]).

For using the Alignment framework, evaluation participants have to implement the Alignment API. The Alignment API enables the integration of the algorithms based on a minimal interface. Adding new alignment algorithms amounts to create a new `AlignmentProcess` class implementing the interface. Generally, this class can extend the proposed `BasicAlignment` class. The `BasicAlignment` class defines the storage structures for ontologies and alignment specification as well as the methods for dealing with alignment display. All methods can be refined (no one is final). The only method it does not implement is the one that implement the alignment algorithm: `align`. This method is invoked from the `Alignment` object which is already connected with the ontologies. I takes a `Parameters` structure enabling to communicate the parameters to the algorithms and must fill the `Alignment` object with the correspondence `Cells` that have been found by the algorithm.

Once this class (which can be thought of as a wrapper around the alignment algorithm) is implemented, it is used by creating an alignment object, providing the two ontologies, calling the `align` method which takes parameters and initial alignment as arguments. The alignment object then bears the result of the alignment procedure. It is thus possible to invoke it on a particular set of tests with particular parameters and to output the results on a variety of formats.

[1] `http://www.inrialpes.fr/exmo/software/ontoalign/`

This will be exploited by launching the `GroupAlign` facility of the Alignment API package to align all pairs of ontologies in a list of subdirectories and generate the result in the required format in this directory.

Chapter 4

Evaluation

The evaluation framework must enable the comparison of an alignment with another one and to generate a resulting evaluation. One of the available methods of the Alignment API (`PRecEvaluator`) directly provides precision, recall and F-measure in an extension of the format developed by Lockheed Martin.

Since the contest, the tools around the API have been improved. The first improvement consists of comparing the results of different algorithms simultaneously and generating a table. Other developments will consist in providing the opportunity to directly launch an algorithm to a full test bench (and even to optimise some parameter). We will try to merge both tools.

The evaluation framework is already implemented. It consists in gathering all the results in the same directory architecture and compare all of them to the reference alignment. This is implemented in the `GroupEval` class and has been used for the EON Ontology alignment contests (see Figure 4.1).

The Alignment API package provides a small utility (`GroupEval`) which allows to implement batch evaluation. It starts with a directory containing a set of subdirectories. Each subdirectory contains a reference alignment (usually called `refalign.rdf`) and a set of alignments (named `name1.rdf`...`namen.rdf`). These alignments can be provided directly by the `GroupAlign` facility.

Invoking `GroupEval` with the set of files to consider (-i argument) and the set of evaluation results to provide (-f argument with profm, for precision, recall, overall, fallout, f-measure as possible measures)

```
$ java -cp /Volumes/Phata/JAVA/ontoalign/lib/procalign.jar
    fr.inrialpes.exmo.align.util.GroupEval -f "pr" -c
    -l "karlsruhe2,umontreal,fujitsu,stanford"
```

returns an HTML file (which could also be other format) such as the one for Figure 4.1.

algo	karlsruhe2		umontreal		fujitsu		stanford	
test	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.
101	n/a	n/a	0.59	0.97	0.99	1.00	0.99	1.00
102	NaN	NaN	0.00	NaN	NaN	NaN	NaN	NaN
103	n/a	n/a	0.55	0.90	0.99	1.00	0.99	1.00
104	n/a	n/a	0.56	0.91	0.99	1.00	0.99	1.00
201	0.43	0.51	0.44	0.71	0.98	0.92	1.00	0.11
202	n/a	n/a	0.38	0.63	0.95	0.42	1.00	0.11
204	0.62	1.00	0.55	0.90	0.95	0.91	0.99	1.00
205	0.47	0.60	0.49	0.80	0.79	0.63	0.95	0.43
221	n/a	n/a	0.61	1.00	0.98	0.88	0.99	1.00
222	n/a	n/a	0.55	0.90	0.99	0.92	0.98	0.95
223	0.59	0.96	0.59	0.97	0.95	0.87	0.95	0.96
224	0.97	0.97	0.97	1.00	0.99	1.00	0.99	1.00
225	n/a	n/a	0.59	0.97	0.99	1.00	0.99	1.00
228	n/a	n/a	0.38	1.00	0.91	0.97	1.00	1.00
230	0.60	0.95	0.46	0.92	0.97	0.95	0.99	0.93
301	0.85	0.36	0.49	0.61	0.89	0.66	0.93	0.44
302	1.00	0.23	0.23	0.50	0.39	0.60	0.94	0.65
303	0.85	0.73	0.31	0.50	0.51	0.50	0.85	0.81
304	0.91	0.92	0.44	0.62	0.85	0.92	0.97	0.97

Figure 4.1: Precision and recall results for various alignment algoritms in HTML format.

Conclusion

Providing formats has the advantage of being able to compute new measures when the consensus is latter made on a new evaluation method. The set of tools that we have presented help automating the generation of tests and evaluation of results. As can be seen from Figure 4.2, the three presented functions should be able to automate generation, processing and evaluation of the alignment algorithm f on the basis of ontology O. This has the advantage of decreasing the rate of errors and with it the risk of complaints. This also lowers the costs of generating a new set of tests or evaluating again the set of results. This helps managing the evaluations. But these tools also reduces the amount of work necessary from the participants to run the tests. They can thus concentrate on performing at best. Moreover, automation enables the participants to run the evaluation of their results easily which helps them to report problems early and to improve their algorithms against the actual benchmark results.

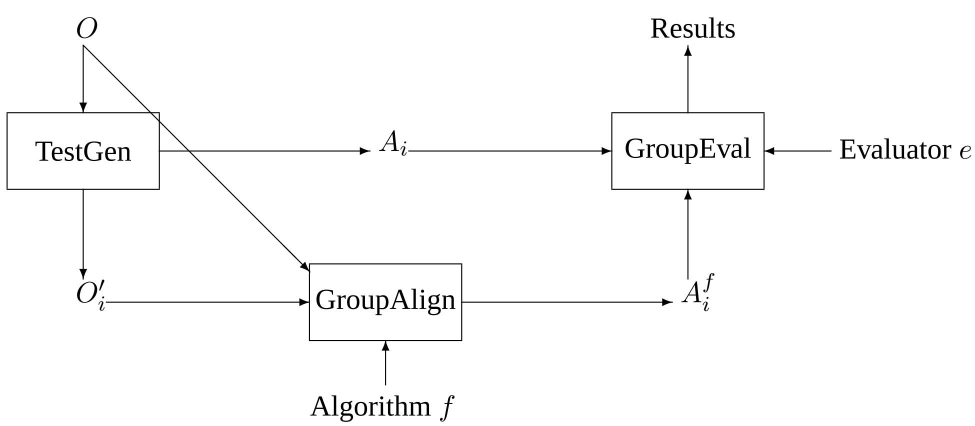

Figure 4.2: Process flow provided by the tool suite.

Bibliography

[Bouquet *et al.*, 2004] Paolo Bouquet, Jérôme Euzenat, Enrico Franconi, Luciano Serafini, Giorgos Stamou, and Sergio Tessaris. Specification of a common framework for characterizing alignment. deliverable D2.2.1, Knowledge web NoE, 2004.

[Euzenat *et al.*, 2004a] Jérôme Euzenat, Thanh Le Bach, Jesús Barrasa, Paolo Bouquet, Jan De Bo, Rose Dieng-Kuntz, Marc Ehrig, Manfred Hauswirth, Mustafa Jarrar, Rubén Lara, Diana Maynard, Amedeo Napoli, Giorgos Stamou, Heiner Stuckenschmidt, Pavel Shvaiko, Sergio Tessaris, Sven Van Acker, and Ilya Zaihrayeu. State of the art on ontology alignment. deliverable D2.2.3, Knowledge web NoE, 2004.

[Euzenat *et al.*, 2004b] Jérôme Euzenat, Marc Ehrig, and Raúl García Castro. State of the art on ontology alignment. deliverable D2.2.2, Knowledge web NoE, 2004.

[Euzenat, 2003] Jérôme Euzenat. Towards composing and benchmarking ontology alignments. In *Proc. ISWC-2003 workshop on semantic information integration, Sanibel Island (FL US)*, pages 165–166, 2003.

[Euzenat, 2004] Jérôme Euzenat. An API for ontology alignment. In *Proc. 3rd international semantic web conference, Hiroshima (JP)*, pages 698–712, 2004.

[Kalfoglou and Schorlemmer, 2003] Yannis Kalfoglou and Marco Schorlemmer. Ontology mapping: the state of the art. *The Knowledge Engineering Review*, 18(1):1–31, 2003.

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

[Sure *et al.*, 2004] York Sure, Oscar Corcho, Jérôme Euzenat, and Todd Hughes, editors. *Proceedings of the 3rd Evaluation of Ontology-based tools (EON)*, 2004.