Multi-lingual Ontology Matching: A Case Study for Portuguese, English, and French

Cássia Trojahn[1], Paulo Quaresma[2], Renata Vieira[3]

[1] INRIA & LIG, Avenue de l'Europe 655, Montbonnot Saint Ismier, France
[2] Universidade de Évora, Évora, Rua Romão Ramalho 59, Évora, Portugal
[3] Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga 6681, Porto Alegre, Brazil

Abstract. We present an implementation for multi-lingual ontology matching that is based on two strategies, *direct translation-based* and *indirect*. The demonstration uses three ontologies written in English, French, and Portuguese, an extension of the OAEI benchmark test 206.

1 Introduction

Matching ontologies consists of finding corresponding entities (i.e., classes, properties, or instances) in different ontologies (usually one source ontology and one target ontology). However, existing matching systems are designed by assuming that the entities of both source and target ontologies are written in the same languages (English, for instance). Multi-lingual ontology matching is an open research issue.

We have proposed an API for multi-lingual matching [3] that specifies interfaces for two strategies, *direct translation-based* and *indirect*. The first strategy considers direct match between two ontologies (i.e., without intermediary ontologies), with the help of external resources (translators), while *indirect alignment*, proposed by [1], is based on composition of alignments. The notion of *direct translation-based alignment* is a simplification of our previous work [2], which uses external resources (WordNet and dictionaries) to obtain the set of synonyms for each term of one source ontology, composing the corresponding concept description in the target language.

In this demonstration we present an implementation that implements this API. We demonstrate how to use the tool by using as case study the matching between ontologies written in English, French, and Portuguese, an extension of the OAEI benchmark test 206.

2 The implementation

In direct translation-based strategy, one ontology source is translated into one translated ontology, using a URI translation strategy of name entities. It is an implementation for the *OWLEntityURIConverterStrategy* of OWL-API[4]. We use the Google-Translator-API[5] to provide the

[4] http://owlapi.sourceforge.net/
[5] http://code.google.com/p/google-api-translate-java

translations. Having the translated and target ontologies, they are used as input to a matcher. We reuse the matchers available in the Alignment API[6].
For the indirect strategy, two ontologies and two alignments (previoulsy computed) are loaded and the composition of the new alignment is performed. This implementation uses a maximisation to compute composed confidences and is restricted to equivalence relations.

3 The ontologies

The OAEI 2009 benchmark campaign[7] offers one open multi-lingual test case (Test 206), containing one alignment between one English (reference ontology) and one French ontologies. Based on this test case, we have created a Portuguese ontology and its corresponding alignments to French and English ontologies (reference alignments). We manually translated the labels of each entity of the reference ontology. We use these ontologies in the demonstration.

4 The demonstration

The demonstration will cover the following steps:

- How to translate one ontology;
- How to match two ontologies using a set of matchers (based on the Alignment API);
- How to match two ontologies using the indirect strategy;
- How to evaluate the alignments.

References

1. Jason J. Jung, Anne Håkansson, and Ronald Hartung. Indirect alignment between multilingual ontologies: A case study of korean and swedish ontologies. In *Proceedings of the 3rd International Symposium on Agent and Multi-Agent Systems: Technologies and Applications*, pages 233–241. Springer-Verlag, 2009.
2. Cássia Trojahn, Paulo Quaresma, and Renata Vieira. A framework for multilingual ontology mapping. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, may 2008.
3. Cássia Trojahn, Paulo Quaresma, and Renata Vieira. An API for Multilingual Ontology Macthing. In European Language Resources Association (ELRA), editor, *Proceedings of the 7th Edition of the Language Resources and Evaluation Conference (to appear)*, Malta, may 2010.

[6] http://alignapi.gforge.inria.fr/
[7] http://oaei.ontologymatching.org/2009/benchmarks/