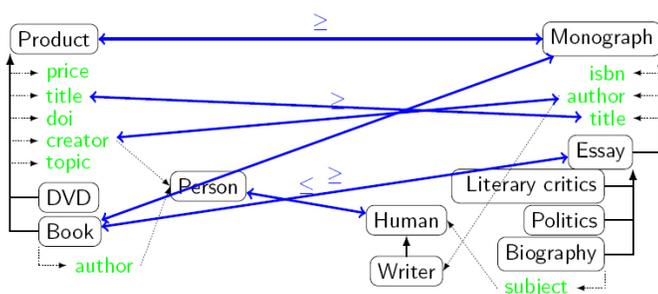# ONTOLOGY MATCHING & ALIGNMENT

Introduction

A realistic realisation of the Semantic Web will see thousands, if not millions, of ontologies describing domains, along with data marked up using those ontologies. Different ontologies will be used to describe the same domain, and if applications are to interoperate and share data, mechanisms for matching ontologies and processing alignments are crucial.

This briefing highlights work undertaken by members of the KnowledgeWeb network in the investigation of ontology matching and alignment techniques. The work covers underlying theoretical aspects, evaluation and methodologies.

Matching & Alignment

Ontology Matching is the process of determing correspondences between entities in different ontologies. For example one ontology may define the notion of *Person*, while another describes *Human*. Although the concepts have different names, they are intended to describe essentially the same collection of instances.

An Ontology Alignment

An Alignment of two ontologies describes a collection of correspondences between concepts in the ontologies. Correspondences may capture different relationships (e.g. one concept is narrower than the other, or the two concepts are an exact match) along with a measure of confidence. For example, the figure above shows an alignment between two ontologies used to described books and products.

Ontology Matching is needed to support a variety of tasks facing the semantic heterogeneity problem, including:

- Ontology engineering

- Information integration
- Peer-to-peer information sharing
- Web service composition
- Autonomous communication systems
- Navigation and query answering on the web

An example use case explored within KnowledgeWeb was within the Human Resources domain. Here, different ontologies may be used to describe job offers and the skill sets of potential employees. In order to match job seekers with job offers, alignment between the ontologies is required.

Approaches

Matching solutions have been proposed from various viewpoints, e.g., databases, information systems, artificial intelligence. They take advantage of various properties of ontologies including structures, data instances, semantics, or labels, and use techniques from fields such as statistics and data analysis, machine learning, automated reasoning, and linguistics. These solutions share some techniques and tackle similar problems, but differ in the way they combine and exploit their results. A uniform framework to compare and describe the various approaches has been proposed.

Representation of Alignments

A language has been defined that allows the description of alignments. This language is independent of the underlying ontology languages (providing greater independence) and is expressive enough to deal with complex cases.

Evaluating Matchers

Measuring the quality of alignments is important – users need confidence in the quality. Measures in traditional Information Retrieval are often based on notions of *precision* (how many of the returned results are relevant) and *recall* (how many of the potential relevant results have been returned). Such measures are less appropriate for measuring the quality of alignments as they do not differentiate sufficiently between good and bad alignments and lack semantics.

Alternatives generalising the classical approaches and providing *tolerant* and *semantic* precision and recall have been developed and used in the evaluation of systems.

Selecting Matchers

Selecting an ontology matching system given a particular application is a difficult problem. Even when the application needs are very precise, there are many criteria that can be used for choosing matchers and all criteria cannot be assessed in the same way.

An in-depth analysis of a number of application use cases (including recruitment, news aggregation and integrated access to biological data) along with the results of alignment evaluations has been performed. This has yielded guidelines that can be followed in order to facilitate the selection of an appropriate matching technique for a particular problem.

The Ontology Alignment Evaluation Initiative

The Ontology Alignment Evaluation Initiative (OAEI) is a coordinated international effort intended to establish a consensus for the evaluation and measurement of the quality of ontology alignments. A suite of ontology alignment benchmarks has been developed that facilitate comparison and evaluation of alignments.

The OAEI has organised a number of workshops at international conferences such as the International Semantic Web Conference (ISWC), where systems can be compared.

In the latest workshop, seven tracks investigated different matching problems drawn from use cases from domains including anatomy, recruitment, web directories, food and conferences. The tracks varied in the expressivity used, ranging from directories and thesauri through to richly expressive ontologies.

Results of the workshops can be found on the OAEI web site. These results show that more tools are becoming available, those tools are becoming more robust, and the overall quality of the results is improving in a measurable way.

Publications & Resources

J. Euzenat, P. Shvaiko, *Ontology matching*, Springer-Verlag, Heidelberg, 2007.

P. Shvaiko, J. Euzenat, *A survey of schema-based matching approaches* Journal of Data Semantics, no. IV, pp. 146-171, 2005.

L. Serafini, H. Stuckenschmidt, H. Wache *A Formal Investigation of Mapping Language for Terminological Knowledge*. IJCAI 2005

J. Euzenat. *Semantic Precision and Recall for Ontology Alignment Evaluation*. IJCAI 2007

M. Ehrig, *Ontology alignment*, Springer-Verlag, Heidelberg, 2007.

KnowledgeWeb Deliverable D1.2.2.2.1: *Case-based recommendation of matching tools and techniques*

KnowledgeWeb Deliverable D2.2.10 *Expressive alignment language and implementation*

OAEI Web Site
`http://oaei.ontologymatching.org/` *Ontology Matching Web Site*
`http://www.ontologymatching.org/`

Contact

For more information about the results presented here or the KnowledgeWeb Network of Excellence, please see the project web site

`http://knowledgeweb.semanticweb.org`

or contact the Network Manager.

Alice Carpentier
DERI, University of Innsbruck
Technikerstraße 13
Innsbruck 6020
AUSTRIA
`alice.carpentier@deri.org`

KnowledgeWeb was an EU-IST Framework 6 Network of Excellence.