The Sixth International Workshop on Ontology Matching (OM 2011)

Towards more Challenging Problems for Ontology Matching Tools

Ernesto Jiménez Ruiz - ernesto@cs.ox.ac.uk
Bernardo Cuenca Grau - berg@cs.ox.ac.uk

LogMap project:
http://www.cs.ox.ac.uk/isg/projects/LogMap/

Motivation

- OM tools have significantly improved in the last few years.
- More challenging and realistic problems are required.
- UMLS-Metathesaurus (UMLS) integrates large medical ontologies such as FMA (>70,000 concepts), SNOMED (>300,000 concepts) or NCI (>60,000 concepts).

Objectives

- Set up a new track within OAEI campaign with more challenging and large input ontologies.
- Consider UMLS as a "reference" alignment between FMA, SNOMED CT and NCI.
- Creation of a community-wide silver standard based on UMLS.

UMLS Metathesaurus

- Sets of mappings can be extracted from UMLS and represented as OWL axioms.
- The integration process in UMLS combines automated techniques, expert curation, and auditing protocols.
- The logic-based integration of FMA, NCI and SNOMED CT via UMLS mappings contains a huge amount of unsatisfiable classes (see Table 1).

Ontologies	Original UMLS Mappings	Unsat. Classes	Clean Mappings
FMA~NCI	3,024	655	2,898
FMA~SNOMED	9,072	6,179	8,111
SNOMED~NCI	19,622	20,944	18,322

Table 1 Refining UMLS mappings with LogMap's repair facility

Our proposed UMLS subset

- We present a non aggressive refinement of the UMLS mappings that does not lead to unsatisfiability (Table 1).
- Original and "clean" UMLS mappings are available and could be used as the basis of a new OAEI track.

Towards a Silver Standard

- To turn UMLS mappings into a gold standard additional (costly and time consuming) manual curation is required.
- A silver standard option would automate the process.
- The silver standard would include the harmonised outputs of different mapping-repair tools (e.g. our refined subset).

Feasibility and importance of the new track

- Significant leap in complexity w.r.t. existing OAEI tracks.
- Positive experiences with our matching tool LogMap.
- The outputs of such a new track would be of great value for the OM and bioinformatics research communities.

UNIVERSITY OF OXFORD

DEPARTMENT OF COMPUTER SCIENCE

Selected References

***LogMap**: Logic-based and scalable ontology matching.* In: 10th International Semantic Web Conference. 2011.

***LogMap** results for OAEI 2011.* In: 6th International Workshop on Ontology Matching. 2011