MSc Internship
Advances in Holistic Ontology Alignment

Topic description

One of the main challenges that the Semantic Web faces is the integration of a growing number of independently designed ontologies. PARIS [1] is an approach for the automatic alignment of ontologies. PARIS aligns not only instances, but also relations and classes. Alignments at the instance level cross-fertilize with alignments at the schema level. Thereby, the system provides a truly holistic solution to the problem of ontology alignment. The heart of the approach is probabilistic, i.e., degrees of matchings are measured based on probability estimates. This allows PARIS to run without any parameter tuning.

A number of problems are left open by PARIS, including:

1. Integration of subjective trust when matching multiple ontologies;
2. Matching within a single ontology and across multiple ontologies;
3. Capturing implicit classes of instances, in the spirit of [2, 3];
4. Indexing techniques for elaborate similarity metrics between literals.

The goal of this internship is to identify solutions to some of these problems, in order to make significant advances in the state-of-the-art of holistic ontology matching. The internship will combine documentation on related work, elaboration of models and algorithms, implementation, experimentation, and possibly theoretical results on the performance of the algorithms.

Supervision

This 5–6 month Master's internship will be supervised by Pierre Senellart at Télécom ParisTech. The research takes part in Serge Abiteboul's Webdam project on the Foundations of Web Data Management, and in the FP7 ARCOMEM project.

References

[1] Fabian M. Suchanek, Serge Abiteboul, and Pierre Senellart. PARIS: Probabilistic alignment of relations, instances, and schema. *Proceedings of the VLDB Endowment*, 5(3):157–168, December 2011. Presented at the VLDB 2012 conference, Istanbul, Turkey.

[2] Rahul Parundekar, Craig A. Knoblock, and José Luis Ambite. Linking and building ontologies of linked data. In *International Semantic Web Conference (1)*, pages 598–614, 2010.

[3] Rahul Parundekar, José Luis Ambite, and Craig A. Knoblock. Aligning unions of concepts in ontologies of geospatial linked data. In *Proceedings of the Terra Cognita 2011 Workshop in Conjunction with the 10th International Semantic Web Conference*, 2011.