MAP-i: PhD Thesis Proposal

OTK – Ontology ToolKit

Adviser:

José João Antunes Guimarães Dias de Almeida (jj@di.uminho.pt)
Departamento de Informática, Universidade do Minho

1 Abstract

Study ontology representations approaches, manipulation techniques, problems and challenges. Implement an environment for providing general technologies and tools to use as building blocks for programming applications or tasks that deal with ontologies.

2 Motivations

Ontologies are a common approach used in nowadays for formal representation of concepts in a structured way. Natural language processing, translation tasks, or building blocks for the new web 2.0 (social networks for example) are instances of areas where the adoption of this approach is emerging and quickly growing. It is clear the importance of ontologies.

Many representations and tools already exist to do this kind of *ontology programming*. Unfortunately there is a lack of a powerful and compact, globally accepted way, of building and processing ontologies, and ontology information. There are many problems with current technologies: composition of smaller tools is hard, mainly because most of the times heterogeneous representations are used, existing tools lack expressiveness, it can get complicated to express operations or information, therefor these approaches lack some language efficiency, when we want to express concepts or operations in a simple and short way. For all these reasons building high order applications using these more simple tools is not possible most of the times. And, when possible it can be more complex than building the full application from scratch.

A good metaphor here is the Unix philosophy, we believe that small tools like: *grep*, *awk* or *sed* are simple and efficient enough to be used together, in elegant pipelines to build more complex applications or perform more complex tasks. What we would like to have is the same elegant and efficient components in an ontology context.

3 Purposed Work

In the context of this work we are proposing a set of tools and technologies to make it easier to design and implement ontology-aware tools and applications:

- Study of the current technologies and problems.

- A formal representation and an heterogeneous ontology algebra, that will help the definition of small tools. This representation needs to allow from simple ontologies to fully featured, scalable and multi-domain ontologies.
- A Domain Specific Language (DSL) that can be used to implement operations on ontologies. The formalism and languages developed should be linguistic efficient: semantics should be powerful enough but simple and concise.
- A set of small tools that perform simple operations on ontologies (*onto-grep*, *onto-awk*, *onto-open*, for example). These tools should be simple, modular, and easy to use in a composition way, like pipelines.
- This DSL and simple tools can then be used to build more complex tools: tools for publishing ontologies in different formats, or importing information from different formats and sources.
- Ultimately an environment for craft applications that create and manipulate ontologies in a simple and dynamic way.

While working on this technologies we also want to keep using generalist techniques as often as possible regarding domain specific languages typical problems (parsing, compilers, ...) to produce general approaches to solve similar problems in other contexts

References

[1] José João Almeida and Alberto Simões. T_2O — recycling thesauri into a multilingual ontology. In *Fifth international conference on Language Resources and Evaluation, LREC 2006*, Genova, Italy, May 2006.

[2] N. Carvalho. OML - Ontology Manipulation Language. Master's thesis, University of Minho, 2008.

[3] A. Gómez-Pérez, M. Fernández-López, and O. Corcho. Ontological Engineering. *AI Magazine*, 36:56, 1991.

[4] T.R. Gruber. A translation approach to portable ontology specifications. *KNOWLEDGE ACQUISITION*, 5:199–199, 1993.

[5] N. Guarino and Istituto (Roma) Consiglio nazionale delle ricerc. *Formal ontology in information systems*. IOS Press, 1998.

[6] I. Horrocks, P.F. Patel-Schneider, and F. van Harmelen. From SHIQ and RDF to OWL: the making of a Web Ontology Language. *Web Semantics: Science, Services and Agents on the World Wide Web*, 1(1):7–26, 2003.

[7] A. Maratheftis and N.I.T. Consulting. Knowledge Management.

[8] B. Smith, W. Kusnierczyk, D. Schober, and W. Ceusters. Towards a Reference Terminology for Ontology Research and Development in the Biomedical Domain. In *Proceedings of KR-MED*, pages 57–66, 2006.