Change paths in reasoning !

Raphael Volz

[1] FZI Forschungszentrum Informatik,
76131 Karlsruhe, Germany
volz@fzi.de, http://www.fzi.de/
[2] Institut AIFB, Universitaet Karlsruhe (TH) ,
D-76128 Karlsruhe, Germany

Abstract. Millions of research funding has been put down to develop - what I call - old forms - of reasoning that are characterized by strong focus on theoretical properties and strict adherence to the completeness properties of reasoning procedures. Despite the large amount of work and results that have been achieved, various benchmarks reinstate that the progress does not suffice for needs in the scale of many enterprise applications and the Web.

We believe that a fundamental change in research on reasoning is required, giving up basic assumptions such as completeness to gain performance required in many real-world applications. These new paths should start with a deriving a clear understanding of what types of questions should be solved, where reasoning can help, where properties like soundness and completeness are really required and what impact of departing from those properties is acceptable.

The standard ontology languages that have been standardized by W3C recently have varying but high complexity ranging from PTIME (RDFS) to NEXPTIME (OWL) complexity. Much research has focused on heavily optimizing reasoning techniques, e.g. [3], while other work such as [2] design new data structures and distribute computation on computing clusters.

While these results are interesting, the necessary scalability has not been achieved as of today. This state-of-the-art is derived from almost three decades of prior work along the same lines, therefore we argue that new paths in reasoning are needed to eventually achieve the scale that is needed in the age of the Web. We outline three potential paths in this position statement:

– Approximate reasoning - Trade off scalability to reasoning completeness
– Incremental reasoning - Leverage caching of prior results
– Tractable reasoning - Simplify ontology languages to a minimal core

The impact of pursuing these paths should be evaluated on top of a gold-standard - a collection of agreed upon ontologies, agreed upon reasoning tasks derived from reasoning steps utilized in real world applications and clear metrics. A consensus on what such a Gold-standard should look like is still to be found.

1 Baseline and Gold Standard Definition

An established baseline would be a strong benefit for research related to reasoning performance. A baseline should be based on a gold-standard ontology as well as standardized queries involving reasoning towards this ontology. The scope of these queries should involve queries typically involved in real-world applications. Therefore a characterization of reasoning potential in applications that we can see today is a sine qua non for us to identify typical queries. A standard data set for the performance evaluation should be established using currently available ontologies compiled from existing ontologies. The baseline for future reasoning performance should be established using publicly available state-of-the-art reasoners such as KAON2, RACER or PELLET by measuring response time, setup time and resource consumption involved in answering the standard queries on the standard data set.

2 Incremental Reasoning

Incremental reasoning leverages caching, i.e. precomputed data structures, to speed up reasoning at runtime. A trivial approach is to a pre-compute the set of all implicit information, which comes at the price of potentially large amounts of memory or space. Several other strategies should be investigated, such as caching prior queries or caching frequently accessed data or compiling declarative programs derived from translating logical programs via the logic-correspondence of the ontology language into executable programs. Another new path could be to depart from the ASK-TELL interface that we see with the reasoners of today to a PUBLISH-SUSCRIBE interface where answers are delivered incrementally and might be revised during the reasoning process.

3 Approximate Reasoning

The aforementioned theoretical bounds for existing ontology languages highly determine runtime behaviour of reasoners that usually aim to provide sound and complete reasoning procedures. Approximate reasoning tries to optimize the behaviour by relaxing (in a controlled manner) the requirement for complete reasoning and thereby decrease the complexity of the reasoning problem. An important focus of research should lie on controlling the approximation and understanding the effects of the approximation such that users can understand what information is potentially missing, as well to how acceptable missing this information is.

4 Tractable Reasoning

A paradigm shift in constructing ontology languages would be to push language features no longer towards the boundaries of expressivity but eliminating language features to obtain tractable languages. This path can already be observed in recent research such as [4, 1]. Other work that follows this lead is the layered design of the WSMO family of ontology languages and further simplify WSMO. Eventually we predict that also languages where sub-polynomial reasoning complexity can be shown will be of great interest, since this would eventually decouple reasoning from knowledge base size, i.e. the number of services, what we perceive to be key to Web scalability tasks.

5 Summary

While we pursue these new paths in reasoning, researchers should gather on a yearly basis at an established forum to repeatedly perform pair-wise independent and reliable performance evaluations[3], where we leverage the baseline and gold standard data set to continuously measure the improvements - and failures - achieved by the new techniques developed. This would ultimatively assess the viability of the research paths pursued.

References

1. Baader, F., Brandt, S., Lutz, C.: Pushing the EL Envelope, IJCAI 2005
2. Hogan, A., Harth, A., Umbrich, J., Decker, S.: Towards a scalable search and query engine for the web, WWW 2007: 1301-1302
3. Boris Motik: Reasoning in Description Logics using Resolution and Deductive Databases, PhD Thesis, Univesität Karlsruhe (TH), Karlsruhe, Germany, January 2006
4. Raphael Volz: Web Ontology Reasoning with Logic Databases, PhD Thesis, Univesität Karlsruhe (TH), Karlsruhe, Germany, February 2004

[3] A good idea is also to do a continuous evaluation on a set of unit tests based on the gold standard reasoning tasks where execution should be automatically scheduled and executed on top of the latest release of the current reasoning prototype