

Position Paper: Relaxing the Basic KR&R Principles to Meet the Emergent Semantic Web*

Vít Nováček

DERI, National University of Ireland, Galway
IDA Business Park, Galway, Ireland
E-mail: vít.novacek@deri.org

Abstract. The paper argues for an alternative, empirical (instead of analytical) approach to a Semantic Web-ready KR&R, motivated by the so far largely untackled need for a feasible emergent content processing.

1 Revisiting the Prevalent KR&R Trends

Since the onset of AI, the knowledge representation and reasoning (KR&R) field has been largely an analytical (in the early Wittgenstein sense) endeavour aimed at producing sound and complete results by algorithmic manipulation of rigorously defined symbol sets (knowledge bases). This works pretty well when the respective domain of interest is closed, deterministic and amenable for complete, indubitable formalisation. Unfortunately, the Semantic Web is not such a neat environment. As has been widely acknowledged in the community, the data one has to manage generally have one or more of the following qualities to them: they are dynamic, noisy, inconsistent, incomplete, intractably abundant, too inexpressive, uncertain and/or context-dependent.

Approaches extending the traditional analytical KR&R accordingly have been investigated recently, however, they seldom take the problem of the actual content acquisition into account as a primary design consideration. To illustrate the issue, we can think of the current RDF/OWL experience – substantially more people generate and use the rather relaxed OWL Full than the rigorous OWL DL flavour. Yet, much larger number of users employ the even simpler RDF(S). It seems to be quite risky to assume that future Semantic Web developers and users will eagerly and happily adopt complex uncertain, paraconsistent or contextualised extensions of the rather OWL-ish (analytical) approach to KR.

Therefore we argue that a truly Semantic Web-ready KR&R should natively tackle noisiness, uncertainty, etc., but also sensibly redefine and/or relax the rigorous assumptions and theoretical groundwork of the analytical approaches in order to follow the WWW success instead of the vapour-ware Xanadu path.

2 Towards the Relaxed, Empirical KR&R

The informatic universe we have to represent within the Semantic Web is very similar (yet simpler) to the perceptual reality of human beings – namely concerning its openness, noisiness and lack of complete, sufficiently formalised data.

* This work has been supported by the EU IST FP6 project ‘Nepomuk’ (FP6-027705) and by Science Foundation Ireland under Grant No. SFI/02/CE1/I131.

Therefore it can be quite useful to draw inspirations from the features of the human mind. These are, however, in many respects exact opposites of the traditional KR&R basic notions (e.g., entailment or model theory) [1]. Conversely, the high-performance and robust (although quite likely unsound and incomplete) natural reasoning abundantly employs similarity-based incorporation and retrieval of data to and from the memory [2]. The respective reasoning is much rather empirical than analytical then [1].

Expanding on these rough considerations, the proposed alternative KR&R conceptualisation can be described by three general canons: (1) **empirical** nature – everything shall be allowed to a degree once it is supported by an empirical evidence; (2) **relaxed** KR principles – the representation shall be as simple as possible so that even AI-illiterates can safely and efficiently contribute to the empirical knowledge refinement if need be; (3) **similarity-based** reasoning – any inference service shall employ soft analogical concept unification enabling to yield sufficient conclusions even from the relaxed representations. Moreover, we suggest that the particular implementations of these canons should maximally reduce the knowledge acquisition and maintenance burden imposed on the users. An obvious way is to support and reasonably employ automatically extracted knowledge as well as legacy resources, while minimising the necessary amount of modelling to be done by the users themselves.

We have recently started to implement our vision in a respective framework, with which we have already attained promising initial results in integration and “analogical closure” of automatically learned ontologies using a biomedical legacy resource [3]. We address the canon (1) by a mechanism of continuous conceptual change based on ordered weighted operators. The canon (2) is reflected by an intuitive, yet expressive basic knowledge representation (essentially compatible with RDF(S), adding heuristic uncertainty and negation). We support also simple, but already quite powerful user-defined uncertain conjunctive rules and queries. Eventually, the canon (3) is addressed by defining an ordered class of universal metrics on the set of basic KR units, which supports granular analogical concept retrieval and a well-founded soft rule and query evaluation. The implementation of these metrics allows for both closed and open world assumptions (can be chosen according to application needs at will). We are currently developing a packaged Python module comprising the framework (a public release is planned for December, 2008 at latest). Apart of that, we are going to further refine and disseminate the “philosophical” and theoretical principles among the relevant research communities.

References

1. Frith, C.: Making Up the Mind: How the Brain Creates Our Mental World. Blackwell Publishing (2007)
2. Gentner, D., Holyoak, K.J., Kokinov, B.K., eds.: The Analogical Mind: Perspectives from Cognitive Science. MIT Press (2001)
3. Nováček, V.: Empirical KR&R in action: A new framework for the emergent knowledge. Technical Report DERI-TR-2008-04-18, DERI, NUIG (2008) Available at <http://140.203.154.209/~vit/resources/2008/pubs/aerTR0408.pdf>.