The Design and Implementation of Minimal* RDFS Backward Reasoning in 4store

Manuel Salvadores[1], Gianluca Correndo[1], Steve Harris[2], Nick Gibbins[1], and Nigel Shadbolt[1]

[1] Electronics and Computer Science, University of Southampton, Southampton, UK
{ms8,gc3,nmg,nrs}@ecs.soton.ac.uk
[2] Garlik Ltd, UK
steve.harris@garlik.com

Abstract. This paper describes the design and implementation of *Minimal* RDFS semantics based on a backward chaining approach and implemented on a clustered RDF triple store. The system presented, called *4sr*, uses 4store as base infrastructure. In order to achieve a highly scalable system we implemented the reasoning at the lowest level of the quad store, the *bind* operation. The *bind* operation runs concurrently in all the data slices allowing the reasoning to be processed in parallel among the cluster. Throughout this paper we provide detailed descriptions of the architecture, reasoning algorithms, and a scalability evaluation with the LUBM benchmark. *4sr* is a stable tool available under a GNU GPL3 license and can be freely used and extended by the community[1].

Keywords: Triple Store, Scalability, Reasoning, RDFS, SPARQL, 4store.

1 Introduction

RDF stores - or triple stores - implement some features that make them very attractive for certain type of applications. Data is not bound to a schema and it can be asserted directly from RDF sources (e.g. RDF/XML or Turtle files) due to their native support of Semantic Web data standards. But the most attractive characteristic is the possibility of implementing an entailment regime. Having entailment regimes in a triple store allows us to infer new facts, exploiting the semantics of properties and the information asserted in the knowledge base. To agree on common semantics, some standards have arisen for providing different levels of complexity encoded in a set of inference rules, from RDF and RDFS to OWL and RIF, each of them applicable to different scenarios.

Traditionally, reasoning can be implemented via forward chaining (FC henceforth), backward chaining (or BC), or hybrid algorithms (a mixture of the two).

* Minimal RDFS refers to the RDFS fragment published in [8].
[1] Preliminary results were presented at the Web-KR[3] Workshop [10] and demoed at ISWC 2010 [9].

G. Antoniou et al. (Eds.): ESWC 2011, Part II, LNCS 6644, pp. 139–153, 2011.