RQL: A Declarative Query Language for RDF

Contributed by:
Greg Karvounarakis, Vassilis Christophides, and Dimitris Plexousakis
Institute of Computer Science, FORTH
Vassilika Vouton, P.O.Box 1385, GR 711 10
Heraklion, Greece
<gregkar, christop, dp @ics.forth.gr>

In the next evolution step of the Web, termed the Semantic Web,vast amounts of information resources (i.e., data, documents, programs) will be made available along with various kinds of descriptive information (i.e., metadata). This evolution opens new perspectives for Digital Libraries (DLs). Community Web Portals, E-Marketplaces, etc. can be viewed as the next generation of DLs in the Semantic Web era. Better knowledge about the meaning, usage, accessibility or quality of web resources will considerably facilitate automated processing of available Web content/services. The Resource Description Framework (RDF), http://www.w3.org/TR/REC-rdf-syntax , enables the creation and exchange of resource metadata as any other Web data. To interpret metadata within or across user communities, RDF allows the definition of appropriate schema vocabularies (RDFS), http://www.w3.org/TR/rdf-schema . The most distinctive feature of the RDF model is its ability to superimpose several descriptions for the same Web resources in a variety of application contexts (e.g., advertisements, recommendations, copyrights, content ratings, push channels, etc.), using different DL schemas (many of which are already expressed in RDF/RDFS. See http://139.91.183.30:9090/RDF/Examples.html . Yet, declarative languages for smoothly querying both RDF resource descriptions and related schemas, are still missing.

This ability is particularly useful for next generation DLs that require the management of voluminous RDF description bases, and can provide the foundation for semantic interoperability between DLs. For instance, in knowledge-intensive Web Portals, various information resources such as sites, articles, etc. are aggregated and classified under large hierarchies of thematic categories or topics. These descriptions are exploited by push channels aiming at personalizing Portal access (e.g., on a specific theme), using standards like the RDF Site. Summary http://purl.org/rss/1.0/spec . Furthermore, the entire catalog of Portals can be exported in RDF, as in the case of Open Directory http://dmoz.org The same is true for DL services (e.g., notification/filtering, annotation/personalization, recommendation based on user/group profiles). There is an ongoing effort to express service descriptions and schemas in RDF (e.g., see the RDF version of WSDL http://www106.ibm.com/developerworks/library/ws-rdf) to benefit from existing RDF support (e.g., query engines) in service matchmaking (i.e., matching service offers with service requests).

Motivated by the above issues, we have designed RQL, a declarative query language for RDF descriptions and schemas. RQL is a typed language, following a functional approach (as in ODMG OQL or W3C XQuery). RQL relies on a formal graph model (as opposed to other triple-based RDF query languages) that captures the RDF modeling primitives and permits the interpretation of superimposed resource descriptions by means of one or more schemas. The novelty of RQL lies in its ability to smoothly combine schema and data querying. In this way, DL applications have to specify in a high-level language only the schema and/or data resources needed for access (e.g., to implement browsing, personnalization, etc.), leaving the task of determining how to efficiently store or access their descriptions to the underlying RDF database engine. More precisely, RQL has been implemented on top of a persistent RDF Store that exploits, as much as possible, available RDF schema information in order to efficiently load and query resource descriptions in an Object-Relational DBMS (SQL3), whilst it preserves the flexibility of RDF in refining schemas and enriching descriptions at any time. The results of our experiments (using the Open Directory catalog) illustrate that our approach yields considerable performance gains in query processing and storage volumes as compared to other triple-based RDF Stores.

RDF Suite has been partially supported by EU projects C-Web (see < http://cweb.inria.fr > [IST-1999-13479]) and MesMuses (see < http://cweb.inria.fr/Projects/Mesmuses > [IST-2001- 26074]). For more information, related publications, downloads and the online RQL demo, see < http://139.91.183.30:9090/RDF/RQL/>