Flexible Intensional Query-Answering for RDF Peer-to-Peer Systems

Zoran Majkić

University of Maryland, College Park, USA
zoran@cs.umd.edu

Abstract. We consider the Peer-To-Peer (P2P) database systems with RDF ontologies and with the semantic characterization of P2P mappings based on logical views over local peer's ontology. Such kind of virtual-predicate based mappings needs an embedding of RDF ontologies into a predicate first-order logic, or at some of its sublanguages as, for example, logic programs for deductive databases. We consider a peer as a *local* epistemic logic system with its own belief based on RDF tuples, independent from other peers and their own beliefs. This motivates the need of a semantic characterization of P2P mappings based not on the extension but on the *meaning* of concepts used in the mappings, that is, based on intensional logic. We show that it adequately models robust weakly-coupled framework of RDF ontologies and supports decidable query answering for the union of conjunctive queries.

1 Introduction

The notion of ontology has become widespread in fields such as intelligent information integration, cooperative information systems, information retrieval, electronic commerce and knowledge management. Using ontologies, semantic annotations on Web resources will allow structural and semantic definitions of documents, providing completely new possibilities: Intelligent search instead of keyword matching, query answering instead of information retrieval, document exchange between departments via ontology mappings, and *definition of views* on documents. The Semantic Web is a proposal to build an infrastructure of machine-readable semantics for the data on the Web.

RDF [1] follows the W3C design principles of interoperability, extensibility, evolution and decentralization. Particularly, the RDF model was designed with the following goals: simple data model and extensible URI-based vocabulary allowing anyone to make statements about any resource.

Data integration in RDF - Motivation
The new P2P data integration systems in Semantic Web [2] needs rich ontologies for the peer databases and efficient query-answering algorithms for expressive SQL-like query languages. So, the mapping between peer databases has to be based on the *views* (expressed by such SQL-like query languages), defined over peer ontologies, as in other cases of data integration systems [3, 4, 5, 6].

Thus, our attempt is to extend database technology of data integration based on rich ontologies and view-based P2P mappings also for semantic for the data on the Web

H. Legind Larsen et al. (Eds.): FQAS 2006, LNAI 4027, pp. 700–712, 2006.

expressed in the RDF syntax: principal motivation is that RDF is a reality in Web applications, so that P2P integration of RDF ontologies is necessarily very important issue.

Data integration in RDF - open problems and actual challenges
All such actual systems of view-based data integration systems, are based on some sublanguage of the FOL (First Order Logic), and also their query-rewriting algorithms are sound and complete w.r.t. the FOL semantics, differently from RDF. In fact, RDF semantics are given in terms of a non-standard model theory. This needs a bridge between RDFS and FOL theory. The embedding presented in this paper can be used to efficiently answer conjunctive queries in heterogeneous P2P settings.

Basically, RDF defines a data model for describing machine-understandable information on the Web. The basic data model consists of three object types: Resources, Properties and Statements. The modeling primitives of RDF are very basic: actually they correspond to binary predicates (RDF-properties) of ground terms (source and value), where, however the predicates may be used as terms so that RDF can not be embedded into the first order logic (FOL), which can be serious drawback in order to be fully integrated into the current logic based frameworks with *extensional equational* theory. Such problem will be explored in more details in the following (the use of *intensional* FOL [7] needs much more investigation). The RDF Schema (RDFS) [8] enriches RDF by giving an externally specified semantics to specific resources, e.g., to rdfs:subclassOf, to rdfs:Class, etc.. It is only because of this external semantics that RDFS is useful. RDFS is recognizable as an *ontology* representation language: it talks about classes and RDF-properties (binary relations), range and domain constraints (on RDF-properties), and subclass and subproperty (subsumption) relations.

So, all attempts to integrate RDFS into some more expressive FOL sublanguage with a built-in *extensional* equality theory (as OWL, largely based on Description Logic, or other interesting languages as Logic Programming, Deductive databases, or Modal Logic languages (e.g., epistemic logic), etc..) are unsuccessful [9, 10].

The difficulty comes from the fact that all FOL sublanguages have the model theory in which individuals are interpreted as *elements* of some domain, classes are interpreted as subsets of the domain, and RDF-properties are interpreted as binary relations on the domain; the semantics of RDFS, on the other hand, is given by a non-standard model theory, where individuals, classes and RDF-properties are *all* elements in the domain, RDF-property elements have extension which are binary relations on the domain, and class extensions are only implicitly defined by the rdf:type property.

A very big number of Web applications is based on simple data structures which actually do not need reification capability of RDFS, so that is really interesting to consider some FOL extensions of RDFS. Because of that we prefer to use directly logic expressions and logic connectives of FOL in a particular subset of RDFS language, which can be naturally embedded into decidable FOL sublanguages.

Data integration in RDF - Main contributions

1. We define the RDF sublanguage which can be embedded in the decidable FOL.
2. We extend the original syntax of such RDF sublanguge, which has only conjunction operator, by defining negation, disjunction and implication algebraic operators. Such