Intensional P2P Mappings Between RDF Ontologies

Zoran Majkić

University of Maryland, College Park, USA

We consider the Peer-To-Peer (P2P) database system 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. The approach to use conventional first order logic (FOL) as the semantic underpinning for RDF has many advantages: FOL is well established and well understood. We will consider an RDF-ontology as finite set of triples $< r, p, v >$, where r is a *resource* name (for class, an instance or a value), p is a *property* (InstanceOf or Property in RDF, or Subclass or Property in RDFS), and v is a value (which could also be a resource name). We denote by $\mathcal{T}$ the set of all triples which satisfy such requirements.

Definition 1 (Logic embedding of RDF-ontology). *We define the logic embedding as a mapping $\mathcal{E} : \mathcal{T} \rightarrow \mathcal{L}$, where $\mathcal{L}$ is FOL with a domain D, such that:*

1. Case when $p = Subclass$: $\mathcal{E}(< r, Subclass, v >) = r'(x) \leftarrow v'(x)$
where r', v' are unary predicates for the subject and object of a triple respectively, and x is a variable over a domain D.
2. Case when $p = InstanceOf$: $\mathcal{E}(< r, InstanceOf, v >) = r'(v) \leftarrow$
where r' is an unary predicate for the resource of the RDF triple.
3. Case when $p \notin \{InstanceOf, Subclass\}$:

$\mathcal{E}(< r, p, v >) = (r'(x) \wedge v'(y)) \leftarrow p'(x, y)$, *if r, v are classes;*
$= r'(x) \leftarrow p'(x, v)$, *if r is a class and v is an instance or value;*
$= v'(y) \leftarrow p'(r, y)$, *if v is a class and r is an instance or value;*
$= p'(r, v) \leftarrow$, *if r and v are instances or values;*

where p' is a binary predicate assigned to the property name p, and r', v' are unary predicates assigned to the subject and the object classes of a triple.

The peer database in this framework is just the logic theory, defined as union of the RDF-ontology FOL-embedding $\mathcal{E}(\mathcal{O})$ and a number of views defined over it (they constitute a virtual user-type interface). Such embedding of an RDF-ontology, together with its view-extension can be used as mean for intensional mapping with other peer databases in a Web P2P networks, and has the following nice properties:

M. Kitsuregawa et al. (Eds.): WISE 2005, LNCS 3806, pp. 592–594, 2005.

1. The FOL embedding of standard RDF-ontology together with views corresponds to the definite logic program, thus the database model of such RDF peer database has a *unique* Herbrand model: consequently, the query answering from peer databases with RDF-ontology and views is very efficient (polynomial complexity).
2. The defined views can be materialized and there are efficient algorithms for maintenance of RDF views when a new RDF triples are inserted in a peer database, when some of them are deleted, or modified.
3. The possibility to transform original RDF based peer database into the more expressive, but *decidable*, FOL sublanguages, is important if we want to add also integrity constraints over a peer database ontology (for example, in the simplest case, the key constraints over a view). In that case we are able to parse the original RDF structures into a deductive predicate-based database. In such way we are able to rich the full expressive power of standard relational Data Integration Systems.

Let P_i and P_j be the two different peer databases, denominated by 'Peter' and 'John' respectively, and $q_1(\mathbf{x})$, $q_2(\mathbf{x})$ be the concepts of "the Italian art in the 15'th century" with attributes in **x**.

The with*weakly-coupled* semantics mappings between peers [1, 2, 3], where each peer is completely independent entity, which has not to be directly, externally, changed by the mutable knowledge of other peers, is based on the *meaning* of the view-based concepts:

1. The knowledge of other peers can not be directly transferred into the local knowledge of a given peer.
2. During the life time of a P2P system, any local change of knowledge must be independent of the beliefs that can have other peers. The *extension* of knowledge which may have different peers about the same type of real-world concept can be very different.

'John' can answer only for a part of his own knowledge about Italian art, and not for a knowledge that 'Peter' has. Thus, when 'query-agent' asks 'John' about this concept, 'John' responds only by the facts known by himself (*known*) answers, and can tell to query-agent that, probably, also 'Peter' is able to answer. It is the task of the *query-agent* to reformulate the request to 'Peter' in order to obtain other (*possible*) answers.

We can paraphrase this by the kind of *belief-sentence*-mapping '*John believes that also Peter knows something about Italian art in the 15'th century*'. Such belief-sentence has *referential* (i.e., extensional) *opacity*. In this case we do not specify that the knowledge of 'John' is included in the knowledge of 'Peter' (or viceversa) for the concept 'Italian art in the 15'th century', but only that this concept, $q_1(\mathbf{x})$, for 'John' (expressed in a language of 'John') *implicitly corresponds* to the 'equivalent' concept, $q_2(\mathbf{x})$, for 'Peter' (expressed in a language of 'Peter'). More about intensional equivalence of concepts, and query-answering in such P2P systems, can be find in [4].

References

1. Z. Majkić, "Weakly-coupled P2P system with a network repository," *6th Workshop on Distributed Data and Structures (WDAS'04), July 5-7, Lausanne, Switzerland*, 2004.
2. Z. Majkić, "Massive parallelism for query answering in weakly integrated P2P systems," *Workshop GLOBE 04, August 30-September 3,Zaragoza, Spain*, 2004.
3. Z. Majkić, "Weakly-coupled ontology integration of P2P database systems," *1st Int. Workshop on Peer-to-Peer Knowledge Management (P2PKM), August 22, Boston, USA*, 2004.
4. Z. Majkić, "Intensional logic and epistemic independency of intelligent database agents," *2nd International Workshop on Philosophy and Informatics (WSPI 2005), April 10-13, Kaiserslautern, Germany*, 2005.