The Fifteenth International Symposium on Artificial Life and Robotics 2010 (AROB 15th '10),
B-Con Plaza, Beppu,Oita, Japan, February 4-6, 2010

Semantic Query Processing Based on SQL

Hyun-Chang Lee

Division of Information and e-Commerce, WonKwang Univ., Institute of Information and Science
344-2 Sinyoun-dong, Iksan, Jeonbuk, 570-749, Korea
(Tel : 82-63-850-6260; Fax : 82-63-850-7302)
(hclglory@wku.ac.kr)

Abstract: According to the development of ubiquitous technology, the concept of ontology is widely used in the fields such as telematics and intelligent robots that require information processes for context. Web-based ontology languages like RDF/S and OWL represent resources in the web to a type of triple (subject, predicate, object). However, the use of languages is restrictive. Therefore, in this paper, we propose an OWL context relational model and semantic query processing mechanism based on SQL.

Keywords: SQL, semantic query, query processing.

I. INTRODUCTION

Ontologies are special forms of conceptual models whose primary role is not to serve documentation but to be machine interpretable and that capture the knowledge of a particular domain as computational artifacts based on principles of knowledge representation in information system. In various application contexts, and within different communities, ontologies have been explored from different points of view. The dominating definition of ontology is the following one, "an ontology is a formal explicit specification of a shared conceptualization of a domain of interest", by Gruber[1]. In addition, ontology concept is widely used in the fields of semantic web application, home networking, telematics, intelligent robot that require the processing of context awareness.

The definition of ontology captures several characteristics of ontology as a specification of domain knowledge, namely the aspects of formality, explicitness, being shared, conceptuality and domain-specificity. The principle constituents of ontology are concepts, relations and instances [2].

Concepts represent the ontological categories and classes of objects that are relevant in the domain of interest. The concepts and relations in an ontology can be intuitively grasped by humans, as they correspond to the elements in our mental model. An ontology tries to cover as many situations as possible that can potentially occur [3]. Relations semantically connect concepts, as well as instances, specifying their interrelations. Instances represent the named and identifiable concrete objects in the domain of interest. The most prominent insights have been published in [4] and are summarized in figure 1.

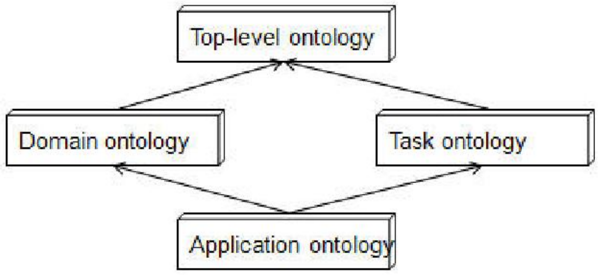

Fig.1. Types of ontologies

Top-level ontologies attempt to describe very abstract and general concepts that can be shared across many domains and applications. Prominent examples for top-level ontologies are DOLCE and SUMO [5,6]. The types of domain ontologies and task ontologies capture the knowledge within a specific domain of discourse, such as medicine or geography, or the knowledge about a particular task, such as diagnosis or configuration. Application ontologies provide the specific vocabulary required to describe a certain task enactment in a particular application context.

Owing to context-awareness systems provided by semantic web technology, most of applications accept the international standards because those have to serve the context awareness services [7]. As for an international standard, there are two types of standards: one is OWL that is a kind of web ontology language. The other is web service that enables to interoperate among software modules. [8] shows methodologies to store and manage ontology documents based on RDF/S to process context-awareness efficiently. In this paper, we propose a relational query process model based on OWL web ontology language.

The rest of this paper is organized as follows. Part 2 begins the discussion on related work. In part 3, it describes the semantic query processing models. And finally, part 4 is the conclusion.

II. RELATED WORKS

There are web based ontology languages such as RDF/S and OWL provided by W3C. In this paper, it is describing about query processes related to the above languages [9,10,11] and extended[12]. RDF has a triple form for framework consisting of subject, predicate and object as a description of resource. That documentation can be shown in form of graphs to be readable from machine. At that time, subject and object are shown as an oval and as a box respectively. The predicate is shown as an arc.

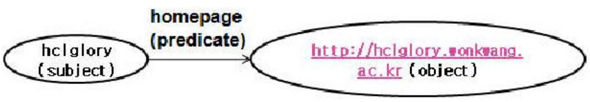

Fig.2. Graphical statement for RDF

As a semantic query language for RDF which is made by W3C, RDQL provides a similar type of the conventional SQL query model for relation database. The following table 1 shows syntax for RDQL briefly. An example for the syntax is in figure 3.

Table 1. The illustration of RDQL syntax

Keyword	Description
SELECT	Variables returned as result
FROM	URI for documentation to query
WHERE	Triples to query Variables for querying Patterns : ({subject|?var} {predicate|?var} {object|?var}) Enable Boolean operations
USING	Prefix to URI

Jena, Sesame, Parka and Tap are known for semantic information management system [13,14,15]. In case of Jena, it is an open source and operates on JAVA framework to build applications. It also provides JAVA API to store and manage documentations of OWL, RDF, RDFS and DAML+OIL[16], and RDQL for query language of ontology data.

III. SEMANTIC QUERY PROCESSING

In this section, we consider the steps for semantic query process and conversions.

First, the representative query language for relational database refers to SQL. The type of fundamental retrieve formula for SQL is "SELECT <field name> FROM <table name> WHERE <condition>;". Moreover, an example for conversion RDF into SQL statement to process semantic query language is shown in figure 3.

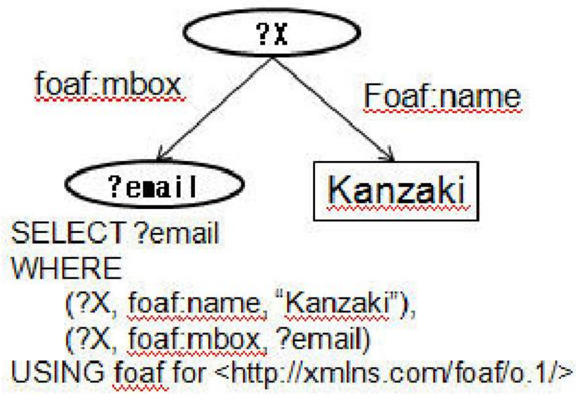

Fig.3. Graphical statement and SQL for RDF data retrieve

In figure 3, the query of graphical statement means that name(foaf:name) is Kanzaki and it has a node for email address(foaf:mbox). Nodes, such as "?X" and "?email", begin with character "?". These nodes play a role in variables that represent what nodes find in searching. In case of RDF, the meaning of graph is products of triples. Therefore, we do not use "AND" operator in triples of "WHERE" clause.

If we make a SQL statement using a theorem to convert variable name into field name, the statement is that "SELECT email FROM addressbook WHERE name = 'Kanzaki'". Therefore, we propose the following theorem as an algorithm to convert graphical RDF into SQL statement.

Lemma1. Let T be a set of triples consisting of RDF graph. T = { t_1, t_2, …, t_n }. Let S be a set of SQL statements. S= { s_1, s_2, …, s_n}. Each element of T can be converted to SQL statements in the following figure 4. For instance, t_1 is s_1, t_2 is s_2 etc.

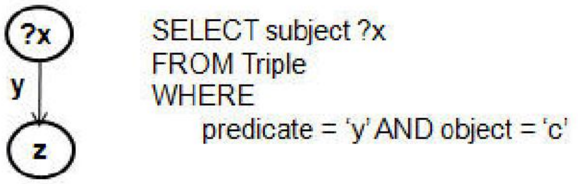

Fig.4. SQL query conversion in single triple

Lemma2. In multiple triples, literal operations are first processed.

For example, one of literals in figure 3 is "Kanzaki". Therefore, the node named "Kanzaki" is first computed and then an email address with foaf:mbox of that node is retrieved.

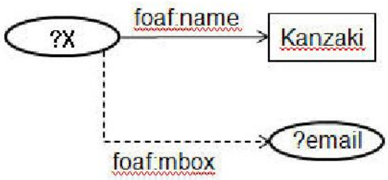

Fig.5. Order of literal and nodes

Lemma3. Graph having multiple triples, i.e., the case of a subject with several properties is joined with operation "Inner Join".

For example, in figure 5, if each node and literal is kind of nodes, then they are sharing the subject. In addition, in figure 6, a node (?X) has two other nodes (z, y). As for 'z' and 'y' nodes, they have same parent node. Then, we can make two SQL statements with same subject using key.

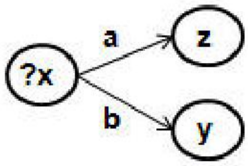

Fig.6. Multiple triples

Lemma4. If an object of a node in a triple is a subject of another node, then add property element to the node belonging to object.

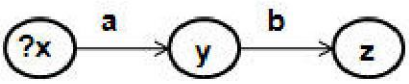

Fig.7. Multiple level nodes

The node (y) has a role in a value as an object and subject. Therefore, in that case of the node, we can make a SQL statement using the key in inner SQL statement.

Until now, we considered the model to convert graphical RDF model to SQL statement based on lemmas.

IV. CONCLUSION

The paper addresses the issues converting graphical RDF to SQL statements. The algorithm is used for applying to the ubiquitous technology requiring information processes for context.

Recently, owing to the development of web technology, web-based ontology languages like RDF/S and OWL are emerged. Those languages represent resources in the web to a type of triple which is consisting of subject, predicate and object. However, the use of languages is restrictive. Therefore, we have proposed and addressed the issues converting to SQL statements. First, we have briefly considered the representative query language for relational database. Next, we have shown the results as an example to convert RDF into SQL statement. To do those works, we have suggested several lemmas.

As a further works, we are investigating effective methods for converting RDF into SQL by taking an experience using data.

REFERENCES

[1] T. Gruber. A Translation Approach to Portable Ontologies. Knowledge Acquisition, 5(2):199-220, 1993.

[2] S. Grimm, R. Volz. Semantics At Work Ontology Management – Tools and Techniques, FZI Forshungszentrum Informatik an der Universitat Karlsruhe(TH), 2007.

[3] N. Guarino. Formal Ontology and Information Systems, Preface. In N. Guarino, editor, Proceedings of the 1st International Conference on Formal Ontologies in Information Systems, FOIS-98, Trento, Italy, pages 3-15. IOS Press, 1998.

[4] N. Guarino. Semantic Matching: Formal Ontological Distinctions for Information Organization, Extraction, and Integration. In M. Pazienza, editor, Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology, number 1299 in LNCS, pages 139-170. Springer-Verlag, 1997.

[5] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. Sweetening Ontologies with DOLCE.

In EKAW-02: Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web, pages 166-181. Springer-Verlag, 2002

[6] I. Niles and A. Pease. Towards a Standard Upper Ontology. In C. Welty and B. Smith, editors, Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), 2001.

[7] Z. X. Pan and J. Heflin: DLDB: Extending relational databases to support semantic web queries, Workshop on Practical and Scalable Web Systems, pp. 109-113, 2003.

[8] J. Broekstra, A. Kampman, F. van Harmelen: Sesame: A Generic Architecture for Storing and Querying RDF and RDF Schema. In The Semantic Web-ISWC 2002, volume 2342 of Lecture Notes in Computer Science, pp. 54-68, 2002. http://openrdf.org/

[9] Dave beckett and et al.: RDF/XML Syntax Specification (Revised), W3C Recommendation 10 February 2004. http://www.w3.org/TR/rdfsyntax-grammer/.

[10] Dan Brickely, R.V. Guha and et al.: RDF Vocabulary Description 1.0: RDF Schema, W3C Recommendation 10 February 2004. http://www.w3.org/TR/rdf-schema/.

[11]Deborah L. McGuinness, Frank van Harmelen: OWL Web Ontology Language Overview W3C Recommendation. Feb 2004. http://www.w3.org/TR/owlfeatures

[12] H.S.Kim, J.H.Son, SQL-based Semantic Query Processing in the OWL-aware Relational Model. In the Journal of KIISE Vol35. Pp44-53. Feb 2008.

[13] Jeremy Carrol, Brian McBride: The Jena Semantic Web Toolkit. HP-Labs, Bristor, 2001. http://jena.sourceforge.net/

[14] J. Broekstra, A. Kampman, F. van Harmelen: Sesame: A Generic Architecture for Storing and Querying RDF and RDF Schema. In The Semantic Web-ISWC 2002, volume 2342 of Lecture Notes in Computer Science, pp. 54-68, 2002. http://openrdf.org/

[15] Kilian Stoffel, Merwyn Taylor, James handler: Efficient management of very large ontologies. Proc. 14th Nat'l Conference AAAI, AAAI press, pp. 442-227, 1997.

[16] Ian Horrocks, Frank van Harmelen and et al.: DAML+OIL (March 2001). http://www.daml.org/2001/03/daml+oil-index.html.