Tkanko O.V.
Institute for applied system analysis NTUU "KPI", Kyiv, Ukraine

Overview of existing approaches to semantic querying using ontology in relational databases

Nowadays human operator is required to perform numerous operations on a data stored in a Web, but it is possible for a computer to do them without any guidance. The Semantic Web aims to change Web presenting information in such way that machines understand it.

The traditional Web stores huge amount of data in relational databases that uses SQL for access information [1]. While the core of the Semantic Web consists of: Resource Description Framework (RDF), RDF Schema (RDFS) and Web Ontology Language (OWL) [2]. Converting available data stored in relational database into RDF format is tedious task and it is clearly better if ontology-based queries could directly retrieve the specific data required via SQL rather than first transforming potentially gigabytes of relational data into RDF. It means that integrating existing relational databases with ontology-based systems becoming one of the most important research problems for the Semantic Web.

All modern frameworks based on 2 layers: Ontology extraction and Syntax wrapping. There are few approaches investigating the transformation of relational schemas to ontologies. Frameword described in [1] use the semi-automatic process of extracting ontology based on primary and foreign keys. Later they adapt the extracted ontology to pre-defined one. While Premerlani [5] proposed a seven- step reverse engineering process and gave the guidelines to get mappings between semantic models and original schemas. A recent research described in [6] provided a description logic based ontology language, which captures features from ER and UML class diagrams. It is proven to preserve the semantics of the constraints in the relational databases.

After extracting and refining the ontology, the user can issue semantic queries based on extracted concepts and these queries will be mapped onto plain syntactic SQL queries. The semantic queries are based on SPARQL [7] as an RDF query language. It expresses queries across diverse data sources, because it contains capabilities for querying required and optional graph patterns along with their conjunctions and disjunctions.

There are many concepts involved in mapping process of the semantic query into SQL statements. Algoryth described in [1] replaces all predicates in the triples with corresponding columns name in relational databases tables. If the predicate is not in the columns name, then it will be in object property names from linked tables. Meanwhile, OntoGrate [3], uses an approach of deductive query answering, which rewrites original queries into a finite set of conjunctive queries in terms of the DB schema, which is only then converted to SQL.

As logical result all of them create sub-query clause, which consists of three parts: **SELECT** (columns), **FROM** (tables) and **WHERE** (join conditions) [6], but missing GROUP BY, LIMIT to retrieve set of data instead of getting explicit information. A client-server API implementation is a comprehensive task and considered for a future development for many modern frameworks.

References. **1.** Ayesha Banu, Syeda Sameen Fatima, Khaleel Ur Rahman Khan, "Semantic – Based Querying Using Ontology in Relational Database of Library Management System", Canadian Journal on Data, Information and Knowledge Engineering Vol. 2, No. 1, January 2011. **2.** OWL Web Ontology Language http://www.w3.org/TR/owl-ref/. **3.** Dejing Dou, Han Qin, Paea LePendu: Ontograte: towards Automatic Integration for Relational Databases and the Semantic Web through an Ontology-Based Framework. Int. J. Semantic Computing 4(1): 123-151 (2010) **4.** W. J. Premerlani and M. R. Blaha, An approach for reverse engineering of relational databases, Commun. ACM Journal 37(5) (1994) 42–49, 134. **5.** L. Lubyte and S. Tessaris, Extracting ontologies from relational databases, in Proceedings of Description Logics, 2007, pp. 122–126. **6.** A. Riazanov. Resolution-based Query Answering for Semantic Access to Relational Databases: A Research Note. Preprint 0901.0339, ArXiv, January 2009. **7.** SPARQL Query Language for RDF http://www.w3.org/TR/rdf-sparql-query/.

ISBN 978-966-2748-33-8, revision 1.0 (2013.04.29), available at http://sait.kpi.ua/books/sait2013.ebook.pdf