Mapping Between RDBMS And Ontology: A Review

Vishal Jain, Dr. S. V. A. V. Prasad

Abstract: Today Semantic web is playing a key role in the intelligent retrieval of information. It is the new-generation Web that tries to represent information such that it can be used by machines not just for display purposes, but for automation, integration, and reuse across applications. It allows the representation and exchange of information in a meaningful way. Ontologies form the backbone of the Semantic Web; they allow machine understanding of information through the links between the information resources and the terms in the ontologies. Ontology describes basic concepts in a domain and defines relations among them. An ontology together with a set of individual instances of classes constitutes a knowledge base. An effort has been made by the Semantic Web community to apply its semantic techniques in open, distributed and heterogeneous Web environments, and for sharing the knowledge in the semantic web. For sharing the knowledge ontologies were introduced, and have grown considerably in number. Building ontologies, and for allow machine understanding of information through the links between the information resources and the terms in the ontologies. Information Retrieval mainly focuses on retrieval of unstructured documents (natural text language documents). These documents may include videos, photos and audios etc. IR addresses retrieval of documents from an organized well defined huge collection of documents available on net which may be email, maps, news etc. Various goals of IR are described below:

- IR aims on retrieving unstructured documents.
- IR engine may produce collection of relevant documents to user according to specified query entered by user.
- IR engine also arranges documents according to its rank which involves Page Rank algorithm. If a document ‘A’ has more effective results than document ‘B’, then ‘A’ will organized first. It has been discussed in further sections.

1 INTRODUCTION

Information Retrieval is the retrieval of information or data, either structured or unstructured. It retrieves in response to query statement which may be unstructured or structured also. Unstructured Query is like sentence which is written in common understandable language while structured query is in form of expression which is combination of equations and operands. IR deals with fusion of streams of output documents produced from multiple retrieval methods. They combined to form single ranked stream which is shown to user. There are two methods for solving queries:

a) By submitting a given query to multiple document collections.
b) By submitting a given query through multiple IR methods.

Traditional text search engines fails for finding optimal documents because of following reasons:

- Improper style of natural language: - These engines are not capable of understanding complex way of writing documents.
- High level unclear concepts: - Some concepts which are included in document but present search engines can’t find those words.
- Semantic Relations: - We can’t find relevant documents for word specified in part of document. E.g. If we have searched for Juice, then it will not find type or part of Juice.

Semantic Web (SW) is combination of SWD’s that are expressed in ontology languages (RDF, OWL). Ontology refers to categorization of concepts and relationships between terms in hierarchical fashion. Although SWD’s retrieves relevant information because they are characterized by semantic methods and ideas, but it is tedious job to find URL’s of SWD’s.

Index Terms: WWW, Semantic Web, Ontology, Ontology Mapping, OWL, RDBMS

Vishal Jain, Research Scholar, Computer Science and Engineering Department, Lingaya’s University, Faridabad, Haryana, INDIA, E-mail: vishaljain83@ymail.com

Dr. S. V. A. V. Prasad, Professor, Electronics and Communication Department, Lingaya’s University, Faridabad, Haryana, INDIA, E-mail: prasad svav@gmail.com

©IJSTR©2014
www.ijstr.org

307
3 WEB ONTOLOGY LANGUAGE (OWL)

W3C's Web Ontology Working Group defined OWL as three different sublanguages:

1. OWL Lite
2. OWL DL (includes OWL Lite)
3. OWL Full (includes OWL DL)

The W3C-endorsed OWL specification includes the definition of three variants of OWL, with different levels of expressiveness.

OWL Lite was originally intended to support those users primarily needing a classification hierarchy and simple constraints. For example, while it supports cardinality constraints, it only permits cardinality values of 0 or 1. It was hoped that it would be simpler to provide tool support for OWL Lite than its more expressive relatives, allowing quick migration path for systems utilizing thesauri and other taxonomies. In practice, however, most of the expressiveness constraints placed on OWL Lite amount to little more than syntactic inconveniences: most of the constructs available in OWL DL can be built using complex combinations of OWL Lite features. Development of OWL Lite tools has thus proven almost as difficult as development of tools for OWL DL, and OWL Lite is not widely used.

OWL DL was designed to provide the maximum expressiveness possible while retaining computational completeness (all conclusions are guaranteed to be computed), decidability (all computations will finish in finite time), and the availability of practical reasoning algorithms. OWL DL includes all OWL language constructs, but they can be used only under certain restrictions (for example, number restrictions may not be placed upon properties which are declared to be transitive). OWL DL is so named due to its correspondence with description logic, a field of research that has studied the logics that form the formal foundation of OWL.

OWL Full is based on a different semantics from OWL Lite or OWL DL, and was designed to preserve some compatibility with RDF Schema. For example, in OWL Full a class can be treated simultaneously as a collection of individuals and as an individual in its own right; this is not permitted in OWL DL. OWL Full allows an ontology to augment the meaning of the pre-defined (RDF or OWL) vocabulary. It is unlikely that any reasoning software will be able to support complete reasoning for OWL Full.

Each of these sublanguages is an extension of its simpler predecessor, both in what can be legally expressed and in what can be validly concluded. Usually, Ontologies are defined to consist of abstract concepts and relationships (or properties) only. In some rare cases, Ontologies are defined to also include instances of concepts and relationships. The following three types of Ontologies are common in literature and are classified on the basis of their generality:

Domain: they are domain-specific and are used to capture knowledge in a particular domain, e.g., engineering, medicine, e-commerce, etc.

Generic: they capture general, domain-independent knowledge (e.g., space and time). They are shared by large number of users across distinct domains. Examples are WordNet and CYC.

Application: they capture the knowledge necessary for a particular application, e.g., ontology representing the structure of a particular web site.

In enterprises, Google and Yahoo!, the major web search services, are using ontology-based approaches to find and organize the contents on the Web.
4 ONTOLOGY MAPPING

Michael Wick, Khashayar Rohanimanesh, Andrew McCallum, AnHai Doan [11] presented a fully supervised statistical model for ontology mapping based on conditional random fields. This model accounts for uncertainty in both the data and the data’s structure. Results on two domains and showed that our supervised model is able to generalize across them has been evaluated. Yuan An, Alex Borgida and John Mylopoulos [12] discussed about the different mapping methods from database to ontologies. Here, author focused on semi-automatic tool, called MAPONTO, that assists users to discover plausible semantic relationships between a database schema (relational or XML) and an ontology, expressing them as logical formulas/rules. Raji Ghawi and Nadine Cullot [13], focus on a component of the architecture which is a tool, called DB2OWL, that automatically generates ontologies from database schemas as well as mappings that relate the ontologies to the information sources. The mapping process starts by detecting particular cases for conceptual elements in the database and accordingly converts database components to the corresponding ontology components. A prototype of DB2OWL tool is implemented to create OWL ontology from relational database. Table 3 depicts the comparative study of various approaches to convert database management system to ontology. Mostafa E. Saleh [15], presented an approach for semantic query in traditional relational database by establishing ontological layer. In this paper, author has been described following rules for converting database to ontology:

- If the primary key of more than one relation is the same, then they should be merged in one ontological class, and their attributes should be merged.
- If the primary key of one relation is unique for that relation, and not contain the primary key in another relation, then that relation will be considered as one ontological class.
- If the foreign key in a relation Ri is a primary key in another relation Rj, then there is an object property (named by its name in Ri) from Ri to Rj, and the domain is Ri, and range is Rj.
- If the relation primary key consists of two other primary keys, then that relation is a property between two classes (resources), the classes are the two relations denoted by the two primary keys.

Wei Hu and Yuzhong Qu [16], propose a new approach to discovering simple mappings between a relational database schema and ontology. It exploits simple mappings based on virtual documents, and eliminates incorrect mappings via validating mapping consistency. Man Li, Xiao-Yong Du, Shan Wang [17], described the learning rules from relational database to OWL ontology. In this paper, an ontology learning approach has been proposed to construct OWL ontology.
automatically based on data in relational database. The related learning rules are discussed in detail. It can be seen that the approach is practical and helpful to the automation of ontology building. Guntar Bumans [18], demonstrated on a simple yet completely elaborated example how mapping information stored in relational tables can be processed using SQL to generate RDF triples for OWL class and property instances. Noredine Gherabi, Khoulia Addakiri [19], prototype has been implemented, which migrates a RDB into OWL structure, for demonstrate the practical applicability of approach by showing how the results of reasoning of this technique can help improve the Web systems. Authors have presented a new approach for mapping relational database into Web ontology. It captures semantic information contained in the structures of RDB, and eliminates incorrect mappings by validating mapping consistency. Secondly, we have proposed a new algorithm for constructing contextual mappings, respecting the rules of passage, and integrity constraints. Fuad Mire Hassan, Imran Ghani, Muhammad Faheem, Abdrahman Ali Hajji [20], in this authors has been reviewed and presented a number of articles for Human Resource Ontology in eRecruitment domain. The papers described the human resource ontology used within ontology matching approach, which provides means for semantic matching approach to match job seekers and job advertisements in a recruitment domain. Marc Ehrig and Steffen Staab [21], considered QOM, Quick Ontology Mapping, as a way to trade off between effectiveness (i.e., quality) and efficiency of the mapping generation algorithms and demonstrated that QOM has lower run-time complexity than existing prominent approaches. Jesús Barrasa, Óscar Corcho, Asunción Gómez-Pérez [22], in this paper authors has been illustrated “R2O, an Extensible and Semantically Based Database to-ontology Mapping Language”. Authors presented R2O, an extensible and declarative language to describe mappings between relational DB schemas and ontologies implemented in RDF(S) or OWL. R2O provides an extensible set of primitives with well defined semantics. Michal Laclavík [23], presented the approach for creating semantic metadata from relational database data. When building ontology based information systems, it is often needed to convert or replicate data from existing information systems such as databases to the ontology based information systems, if the ontology based systems want to work with real data. RDB2Onto converts selected data from a relational database to a RDF/OWL ontology document based on a defined template. Carlos Eduardo Pires, Damires Souza, Thiago Pachêco, Ana Carolina Salgado [24], has been presented a tool SemMatcher, for matching ontology-based peer schemas, combining different matching strategies (e.g., linguistic, structural, and semantic). SemMatcher allows the identification of semantic correspondences between two peer ontologies using domain ontology as background knowledge. Also, the tool determines a global similarity measure between the matching ontologies that can be used for peer clustering. Nikolaos Konstantinou, Dimitrios-Emmanuel Spanos, Michael Chalas, Emmanuel Solidakis and Nikolas Mitrou [25], presented a VisAVis, an approach to mapping relational database contents to ontologies. Authors has shown the key idea, instead of storing instances along with the ontology terminology, keep them stored in a database and maintain a link to the dataset.

5 NEW PROPOSED FRAMEWORK WHICH CAN BE DEVELOPED

In order to overcome the problems of the existing frameworks and tools, it is critical to design an appropriate framework. This paper proposes a framework similar to DB2OWL but has additional features that address the identified problems and deficiencies. This implies that the proposed framework can support many databases as possible and the most used programming languages. In addition, the new framework has the capacity to output information in different formats, which non-programmers can under re-use without the need for an expert. However, the proposed framework will not depend on particular table cases. It is a general framework that is applicable to all tables, whatever the case. The proposed mapping process involves converting tables into classes, which have several properties, as well as relationships. The conversion process will start when the user uses a well-designed user interface to send queries to the database. The proposed visualization service must be able to present the required queries in a suitable manner. In order to consider the requirements of different users, including those who do not have programming skills, the visualization service should have an interface that has select option for users to key-in commands in a desired language. This should consider all the available programming languages as well as human language, which diverse users can understand. Therefore, it is extremely critical to include a module that translates the input text and instruction into different programming languages. In addition, the new framework ought to incorporate a module that enables users to export information into different formats apart from the default format. Users must be able to output information in the form of text files, tables and datasets among others.
Table 3: “Features of different database-to-ontology mapping approaches” [13]

<table>
<thead>
<tr>
<th>Approach</th>
<th>Ontology</th>
<th>Exploitation</th>
<th>Automatisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volz et al.</td>
<td>×</td>
<td>×</td>
<td>Semi Auto</td>
</tr>
<tr>
<td>DataGenie</td>
<td>×</td>
<td>×</td>
<td>Auto Auto</td>
</tr>
<tr>
<td>Relational.OWL</td>
<td>×</td>
<td>×</td>
<td>Auto Auto</td>
</tr>
<tr>
<td>KAON reverse</td>
<td>×</td>
<td>×</td>
<td>Semi Auto</td>
</tr>
<tr>
<td>vis A vis</td>
<td>×</td>
<td>×</td>
<td>Manual Auto</td>
</tr>
<tr>
<td>D2R map</td>
<td>×</td>
<td>×</td>
<td>Manual Auto</td>
</tr>
<tr>
<td>R2O</td>
<td>×</td>
<td>×</td>
<td>Manual Auto</td>
</tr>
<tr>
<td>DB2OWL</td>
<td>×</td>
<td>×</td>
<td>Auto Auto</td>
</tr>
</tbody>
</table>

Table 4: Ontologies based on newly-built approaches and its associated matching algorithms [19, 20]

<table>
<thead>
<tr>
<th>Author/Year</th>
<th>Title</th>
<th>Technique</th>
<th>Solution</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>L. Yahiaoui et al., 2006</td>
<td>Semantic Annotation of Documents Applied to e-recruitment</td>
<td>Used semantic matching computation coefficient as matching algorithm which is terminological matching technique.</td>
<td>Proposed scenario of automatic eRecruitment web which the applications and job descriptions matched through semantic annotation and indexing based on competency profiles.</td>
<td>Not implemented as a real world prototype to validate the proposed model.</td>
</tr>
<tr>
<td>Fazel-Zarandi, M.; Fox, M.S. 2009</td>
<td>Semantic Matchmaking for Job Recruitment: An Ontology-Based Hybrid Approach</td>
<td>Node-based semantic similarity measure which is terminological matching technique</td>
<td>HR ontology is presented in this work where they propose an ontology-based hybrid approach to efficiently match job seekers and job descriptions.</td>
<td>Not mentioned the composition of sub-ontologies or does not propose any model.</td>
</tr>
<tr>
<td>Chaoxiang Chen, et al., 2009</td>
<td>Design and Implementation of SMS Employment Agent Based on Ontology</td>
<td>Concept compatible matching algorithm which is terminological matching technique</td>
<td>Presented SMS recruitment web service using ontology and agent systems.</td>
<td>Not focused to talk about ontology components and their development.</td>
</tr>
<tr>
<td>Ly Hexion, Zhu Bin, 2010</td>
<td>Elastic Information Matching Technology and its Application in Electronic Recruitment</td>
<td>Used semantic similarity algorithm determined by distance of a concept tree structure, it is both terminological and structural</td>
<td>Presented an approach to develop ontology based framework, its semantic matching similarity algorithm</td>
<td>Not focused to talk about domain ontology and real world implementation prototype.</td>
</tr>
</tbody>
</table>
6 COMPARISON OF ALREADY DEVELOPED TOOLS AND FRAMEWORKS

<table>
<thead>
<tr>
<th>List of Tools</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. RDB To Onto</td>
<td>(a) Design and implements ontology based on relational database (b) User-oriented i.e. user can modify and access records.</td>
</tr>
<tr>
<td>2. Asio Semantic Bridge</td>
<td>(a) Creates ontology and represent rows of table as classes and columns as their properties. (b) Allows updating SPARQL queries to SQL and performs its execution.</td>
</tr>
<tr>
<td>3. Data Grid Semantic Web Kit</td>
<td>(a) Performs mapping as well as querying RDF triples. (b) User defined tool that uses GUI (visual interface) to define individual classes. (c) Generates SPARQL queries and translates them into SQL queries.</td>
</tr>
<tr>
<td>4. DB2OWL</td>
<td>(a) Automatically creates ontology by converting each component of databases into classes, properties and relations. (b) Represents developed ontology in OWL-DL (Description Logic) language. (c) Supports only MySql, Oracle databases.</td>
</tr>
<tr>
<td>5. SOAP (Simple object Access protocol)</td>
<td>(a) Predictive in Nature. (b) Uses classes to predict nature of ontologies.</td>
</tr>
<tr>
<td>6. R2O</td>
<td>(a) Uses XML for expressing elements of database and ontology. (b) Detects ambiguities between classes and their properties.</td>
</tr>
<tr>
<td>7. Triplify</td>
<td>(a) Represents data that is also present in other databases. (b) Generates SQL queries by linking and converting requests from various databases connected on remote hosts. (c) Does not support SPARQL and is easy to use in various applications.</td>
</tr>
</tbody>
</table>

Table 5: “Comparative Study of Tools”

7 CONCLUSION

This paper emphasis on the concept of Ontology Mapping, discuss various approaches for converting relational database to ontology and vice-versa. It is evident the conversion of relational databases to ontology is a diverse process and the frameworks and tools used are diverse. These frameworks and tools have their merits and demerits. Data presentation and output formats and languages are crucial concerns. The proposed frameworks will ensure that there is maximum data integrity in after conversion. In addition, it offers users the ability to customize queries depending on their literacy level. Automation is also a critical part of the proposed framework.

ACKNOWLEDGMENT

I, Vishal Jain, would like to give my thanks to Mr. Gagandeep Singh Narula for helping me in comparison of these tools and Dr. M. N. Hoda, Director, Bharati Vidyapeeth’s Institute of Computer Applications and Management (BVICAM), New Delhi, for giving me opportunity to do Ph.D from Linagay’s University, Faridabad, Haryana.

REFERENCES


[12]. Yuan An, Alex Borgida and John Mylopoulos, “Building Semantic Mappings from Databases to Ontologies”, American Association for Artificial Intelligence, Year 2006.


[14]. Nadine Cullot, Raji Ghawi, and Kokou Yétongnon, " DB2OWL: A Tool for Automatic Database-to-Ontology Mapping", Laboratoire LE2I, Université de Bourgogne, Dijon, FRANCE


[16]. Wei Hu and Yuzhong Qu, “Discovering Simple Mappings Between Relational Database Schemas and Ontologies”, School of Computer Science and Engineering, Southeast University, Nanjing 210096, P.R. China


[22]. Jesús Barrasa, Óscar Corcho, Asunción Gómez-Pérez, “R2O, an Extensible and Semantically Based Databaseto-ontology Mapping Language”, Ontology Engineering Group, Departamento de Inteligencia Artificial, Facultad de Informática, Universidad Politécnica de Madrid, Spain


Authors

Vishal Jain has completed his M.Tech (CSE) from USIT, Guru Gobind Singh Indraprastha University, Delhi and doing PhD from Computer Science and Engineering Department, Lingaya’s University, Faridabad. Presently he is working as Assistant Professor in Bharati Vidyapeeth’s Institute of Computer Applications and Management, (BVICAM), New Delhi. His research area includes Web Technology, Semantic Web and Information Retrieval. He is also associated with CSI, ISTE.

Dr. S. V. A. V. Prasad has completed M.Tech., Ph.D. Presently working as professor and Dean (Academic Affairs), Dean (R&D, IC), Dean School of Electrical Sciences. Sir has actively participated and organized many refresher courses, seminars, workshops on ISO, ROHS, Component Technology, WEEE, Organizational Methods, Time Study, Productivity enhancement and product feasibility etc. Sir has developed various products like 15 MHz dual Oscilloscope, High Voltage Tester, VHF Wattmeter, Standard Signal, Generator with AM/FM Modulator, Wireless Becom, High power audio Amplifier, Wireless microphone and many more in the span of 25 years (1981 – 2007). Sir has awarded for excellence in R&D in year 1999, 2004 and National Quality Award during the Year 1999, 2000, 2004 and 2006. Sir is Fellow member of IEEE and life member of ISTE, IETE, and Society of Audio & Video System. Sir has published more than 90 research papers in various National & International conferences and journals. Sir research area includes Wireless Communication, Satellite Communication & Acoustic, Antenna, Neural Networks, and Artificial Intelligence.