Asian Journal of Information Technology 13 (10): 612-617, 2014
ISSN: 1682-3915

Tools and Techniques for Ontology Interoperability: A Survey

[1]R. Lakshmi Tulasi and [2]M. Srinivasa Rao
[1]Department of IT, QISCET, Ongole, India
[2]SIT, JNTUH, Hyderabad, India

Abstract: The idea of the semantic web is to add machine process able information to web-based data in order to realize interoperability. Ontology is a shared conceptualization of knowledge representation of particular domain. These are used for the enhancement of semantic information explicitly. Ontologies play a prominent role in the concept of the semantic web to provide semantic information for assisting communication among heterogeneous information repositories. Ontology interoperability provides the reusability of ontologies Different domain experts and ontology engineers create different ontologies for the same or similar domain depending on their data modelling requirements. These cause ontology heterogeneity and inconsistency problems. As increasing numbers of ontologies are developed by diverse communities, the demand for rapid ontology mapping is arising. For more better and precise results ontology mapping is the solution. As their use has increased, providing means of resolving semantic differences has also become very important. Studies on ontology interoperability report the results on different frameworks and this makes their comparison almost impossible. Therefore, the main focus of this study will be on providing some basics of ontology interoperability and briefly introducing its different approaches. In this study, researchers survey the approaches that have been proposed for providing interoperability among domain ontologies and its related techniques and tools.

Key words: Ontology interoperability, ontology mapping, ontology alignment, ontology merging, semantic heterogeneity, semantic web

INTRODUCTION

The WWW has become a vast resource of information. It is growing rapidly from last few decades. The problem is that finding the information and the individual desires are often quite difficult because of complexity in organization and quantity of the information stored. In traditional search engines, Information Retrieval (IR) is keyword based or with a natural language. Query entered by the users is not understandable, so it retrieves the large number of documents in the ranked order which have poor semantic relationships among the documents. This keyword based approach results poor precision-list of retrieved documents contain a high percentage of irrelevant documents and poor recall-list of relevant retrieved among possible relevant. To avoid the above problems semantic search engines are required.

Ontology is used to model knowledge representation of a particular domain (e-Learning, sports, medical, etc.). Ontologies are explicit specifications of the conceptualization and corresponding vocabulary used. Ontology is the fundamental factor for semantic web. So, users create different ontologies depending on their data modeling requirements for the same or similar domain. They are free to use vocabulary of their own. This leads to heterogeneity and inconsistency problems.

The basic operation, researchers perform to solve above problems among ontologies is “mapping” which interprets the sets of correspondences between similar concepts and among two or more ontologies of same or similar domains. This is prominent research area in the field of AI (Artificial Intelligence). These mappings support two other related operations ontology alignment and ontology merging. Ontology alignment process takes two or more input ontologies and produces a set of relationships between concepts that match semantically with each other. These matches are also called mappings. Ontology merging as its name implies merges two ontologies of same or similar domain in to one based on semantic similarity of concepts and produces unique ontology. Three important mismatches may exist between ontologies syntactic, semantic and lexical mismatches. The recent researchers developed several methods and techniques to identify these mismatches.

Corresponding Author: R. Lakshmi Tulasi, Department of IT, QISCET, Ongole, India

ONTOLOGY INTEROPERABILITY

This study describes several operations on ontologies like transformation and translation, merging, mapping, integration. These can be considered as an ontology reuse process (Kalfoglou and Schorlemmer, 2003a; Pinto and Martins, 2001).

Ontology transformation and translation: Ontology transformation (Chalupsky, 2000; Dou *et al*., 2005) is the process used to develop a new ontology to cope with new requirements made by an existing one for a new purpose by using a transformation function ‘t’. Many changes are possible in this operation including changes in the semantics of the ontology and changes in the representation formalism. Ontology translation is the function of translating the representation formalism of ontology while keeping the same semantic. In other words, it is the process of change or modification of the structure of ontology in order to make it suitable for purposes other than the original one. There are two types of translation. The first is translation from one formal language to another, for example from RDFS to OWL called syntactic translation. The second is translation of vocabularies, called semantic translation (Chalupsky, 2000). The translation problem arises when two web-based agents attempt to exchange information describing it using different ontologies.

Ontology merging: Ontology merging (Ghidini and Giunchiglia, 2004; Klein and Fensel, 2001; Dou *et al*., 2005) is the process of creating a new single coherent ontology from two or more existing source ontologies related to the same domain. The new ontology will replace the source ontologies.

Ontology integration: Integration (Ghidini and Giunchiglia, 2004; Klein and Fensel, 2001) is the process of creating a new ontology from two or more source ontologies from different domains.

Ontology alignment: Ontology alignment (Ehrig and Staab, 2004; Euzenat and Shvaiko, 2007; Euzenat and Valtchev, 2004; Giunchiglia *et al*., 2005) is the process or method of creating a consistent and coherent link between two or more ontologies by bringing them into mutual agreement. This method is near to artificial intelligence methods: being a logical relation, ontology alignments are used to clearly describe how the concepts in the different ontologies are logically related. This means that additional axioms describe the relationship between the concepts in different ontologies without changing the meaning in the original ontologies. In fact the ontology alignment uses as a pre process for ontology merging and ontology integration. There are many different definitions for ontology alignment depending upon its applications and its intended outcome. Sample definitions include the following:

- Ontology alignment is used to establish correspondences among the source ontologies and to determine the set of overlapping concepts, concepts that are similar in meaning but have different names or structure and concepts that are unique to each of the sources (Dou *et al*., 2005)
- Ontology alignment is the process of bringing two or more ontologies into mutual agreement, making them consistent and coherent

Given two ontologies O1 and O2, mapping of one ontology in to another means that each entity (Concept C, Relation R, Instance I) in ontology is trying to find a corresponding entity which has the same intended meaning in ontology O2.

Formally, an ontology alignment function is defined as follows: an ontology alignment function, align based on the set E of all entities $e \in E$ and based on the set of possible ontologies O is a partial function:

$$\text{Align}: O1 \rightarrow O2$$
$$\text{Align}(eO1) = fO2 \text{ if } \text{Sim}(eO1, fO2) > \text{threshold}$$

Where:

Oi	=	Ontology
eOi, fOj	=	Entities of (Oi, Oj)
Sim(eO1, fO2)	=	Similarities function between two entities eO1 and fO2

The ontology alignment function is based on different similarity measures. A similarity measure is a real valued function. Sim(ei, fj): $O \times O \rightarrow [0, 1]$ measuring the degree of similarity between x and y. Ontology heterogeneity is shown in Fig. 1.

Ontology mapping: Ontology mapping (Giunchiglia *et al*., 2005; Kalfoglou and Schorlemmer, 2003a, b; Rahm and Bernstein, 2001; Shvaiko and Euzenat, 2005; Chalupsky, 2000) is a formal expression or process that defines the semantic relationships between entities from different ontologies. In other words, it is an important operator in many ontology application domains such as the semantic web and e-Commerce which are used to describe how to connect and from correspondences between entities across different ontologies. Ontology matching is the process of discovering similarities between two

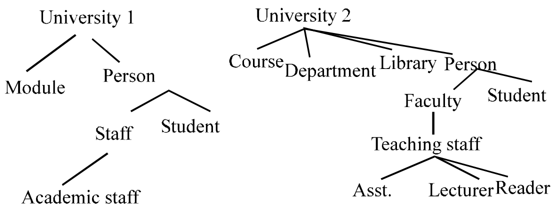

Fig. 1: Ontology heterogeneity among ontologies of same domain

ontologies. An entity ‘e’ is understood in an ontology O denoted by e|O is Concept C, Relation R or Instance I, i.e., $e|O \in C \cup R \cup I$. Mapping the two ontologies, O1 onto O2, means that each entity in ontology O1 is trying to find a corresponding entity which has the same intended meaning in ontology O2. The ontology mapping function “map” is defined based on the vocabulary, E of all terms $e \in E$ and based on the set of possible ontologies, O as a partial function:

$$\text{map}: E \times O \times O \rightarrow E, \text{ with}$$
$$e \in O1(f \in O2: \text{map}(e,O1,O2) = fz \in \text{map}(e,O1,O2) = \perp)$$

An entity is mapped to an other entity or none.

TYPES OF ONTOLOGY MAPPING

Based on the method of ontology mapping and how ontologies are created and maintained, it is divided in to three categories.

Ontology mapping between an integrated global ontology and local ontologies: In this case, ontology mapping is used to map a concept of one ontology into a view or a query over other ontologies (Doan *et al*., 2003a, b; Beneventano *et al*., 2003).

Ontology mapping between local ontologies: In this case, ontology mapping is the process that transforms the source ontology entities into the target ontology entities based on semantic relation. The source and target are semantically related at a conceptual level (Noy and Musen, 2001).

Ontology mapping in ontology merge and alignment: In this case, ontology mapping establishes correspondence among source (local) ontologies to be merged or aligned and determines the set of overlapping concepts, synonyms or unique concepts to that sources (Dou *et al*., 2005). This mapping identifies similarities and conflicts between the various source (local) ontologies to be merged or aligned.

CHALLENGES OF ONTOLOGY MAPPING

In this study, researchers discuss challenges of ontology mapping:

- Large-scale evaluation
- Performance of ontology-matching techniques
- Discovering missing background knowledge
- Uncertainty in ontology matching
- Matcher selection and self-configuration
- User involvement
- Explanation of matching results
- Social and collaborative ontology matching
- Alignment management: infrastructure and support
- Reasoning with alignments

TYPES OF MISMATCHES

Different types of mismatches may occur between different ontologies. Indeed different ontology designers opt for different representation languages and use different ontology editors to represent knowledge at different levels of granularity (detail). This explains the emergence of different forms of ontology mismatches. The identification of these types of mismatches is essential in order to solve them during the mapping, alignment or merging process.

Syntactic mismatches: Two ontologies are syntactically heterogeneous if they are represented by different representation languages such as OWL, KIF, etc. To resolve this type of mismatches, simply transform the representation language of one ontology to the representation language of the other ontology. Herein, researchers state that sometimes the translation is difficult and even impossible.

Lexical mismatches: It describes the heterogeneities among the names of entities, instances, properties or relations. In this type of mismatches, researchers may find four forms of heterogeneities: synonyms, homonyms, same name in different languages and same entities with the same name but with different syntactic variations.

Semantic mismatches: These kind of mismatches describe words belong to same synonym set. For example, ontology A has price and ontology B has cost. Then, both are said to be semantically equivalent or match otherwise it is a mismatched pair.

TOOLS AND TECHNIQUES FOR ONTOLOGY OPERATIONS

LSD (Learning Source Description): LSD semi automatically creates semantic mappings with a multi strategy learning approach (Doan *et al*., 2003a, b). This approach employs multiple learner modules with base learners and the meta-learner where each module exploits a different type of information in the source schemas or data. LSD uses the following base learners:

The name learner: It matches an XML element using its tag name.

The content learner: It matches an XML element using its data value and works well on textual elements.

Naive bayes learner: It examines the data value of the instance and does not work for short or numeric fields.

The XML learner: It handles the hierarchical structure of input instances. Multi-strategy learning has two phases: training and matching. In the training phase, a small set of data sources has been manually mapped to the mediated schema and is utilized to train the base learners and the Meta learner. In the matching phase, the trained learners predict mappings for new sources and match the schema of the new input source to the mediated schema.

MOMIS (Mediator Environment for Multiple Information Sources): MOMIS creates a Global Virtual View (GVV) of information sources, independent of their location or their data's heterogeneity (Beneventano *et al*., 2003). MOMIS builds an ontology through five phases as follows:

- Extraction of local schema
- Local source annotation using Word Net (online dictionary)
- Common thesaurus generation: relationships of inter-schema and intra-schema knowledge about classes and attributes of the source schemas
- Generation of GVV: a global schema and mappings between the global attributes of the global schema and source schema are generated
- GVV annotation is generated by exploiting annotated local schemas and mappings between local schemas and a global schema

A framework for OIS (Ontology Integration System): Mappings between an integrated global ontology and local ontologies are expressed as queries and ontology as Description Logic (Calvanese *et al*., 2001). Two approaches for mappings are proposed as follows: concepts of the global ontology are mapped into queries over the local ontologies (global-centric approach) and concepts of the local ontologies are mapped to queries over the global ontology (local centric approach).

GLUE: It semi-automatically creates ontology mapping using machine learning techniques (Doan *et al*., 2003a, b). It consists of distribution estimator, similarity estimator, and relaxation labeler. It finds the most similar concepts between two ontologies and by using a multi-strategy learning approach calculates the joint probability distribution of the concept for similarity measurement. It has content learner, name learner and meta learner. Content and name learners are two base learners while meta learner combines the two base learners' prediction. The content learner exploits the frequencies of words in content of an instance and uses the Naive Bayes' theorem. The Name Learner uses the full name of the input instance. The Meta-Learner combines the predictions of base learners and assigns weights to base learners based on how much it trusts that learner's predictions.

ONION (ONtology compositION system): It resolves terminological heterogeneity in ontologies and produces articulation rules for mappings (Mitra and Wiederhold, 2002). The linguistic matcher identifies all possible pairs of terms in ontologies and assigns a similarity score to each pair. If the similarity score is above the threshold, then the match is accepted and an articulation rule is generated. After the matches generated by a linguistic matcher are available, a structure-based matcher looks for further matches. An inference-based matcher generates matches based on rules available with ontologies or any seed rules provided by experts. Multiple iterations are required for generating semantic matches between ontologies. A human expert chooses, deletes or modifies suggested matches using a GUI tool.

LOM (Lexicon-based Ontology Mapping): LOM finds the morphism between vocabularies in order to reduce human labor in ontology mapping using four methods: whole term, word constituent, synset and type matching (Li, 2004). LOM does not guarantee accuracy or correctness in mappings and has limitations in dealing with abstract symbols or codes in chemistry, mathematics or medicine.

QOM (Quick Ontology Mapping): QOM is an efficient method for identifying mappings between two ontologies because it has lower run-time complexity (Ehrig and Staab, 2004). In order to lower run-time complexity, light weight ontologies QOM uses a dynamic programming approach. A dynamic programming approach has data structures which investigate the candidate mappings, classify the candidate mappings into promising and less promising pairs and discard some of them entirely to gain efficiency. It allows for the ad-hoc mapping of large size, light-weight ontologies.

PROMPT: PROMPT is a semi-automatic ontology merging and alignment tool. It begins with the linguistic-similarity matches for the initial comparison but generates a list of suggestions for the user based on linguistic and structural knowledge and then points the user to possible effects of these changes (Noy and Musen, 2000).

Onto Morph: Onto Morph provides a powerful rule language for specifying mappings and facilitates ontology merging and the rapid generation of knowledge-base translators. It combines two powerful mechanisms for knowledge-base transformations such as syntactic rewriting and semantic rewriting. Syntactic rewriting is done through pattern-directed rewrite rules for sentence-level transformation based on pattern matching. Semantic rewriting is done through semantic models and logical inference (Chalupsky, 2000).

Anchor-PROMPT: Anchor-PROMPT takes a set of anchors (pairs of related terms) from the source ontologies and traverses the paths between the anchors in the source ontologies. It compares the terms along these paths to identify similar terms and generates a set of new pairs of semantically similar terms (Noy and Musen, 2001).

CMS (CROSI Mapping System): CMS is an ontology alignment system. It is a structure matching system on the rich semantics of the OWL constructs. Its modular architecture allows the system to consult external linguistic resources and consists of feature generation, feature selection, multi-strategy similarity aggregator and similarity evaluator (Kalfoglou and Hu, 2005).

FCA-Merge: FCA-Merge is a method for ontology merging based on Ganter and Wille's formal concept analysis, lattice exploration and instances of ontologies to be merged. The overall process of ontology merging consists of three steps: instance extraction and generation of the formal context for each ontology, the computation of the pruned concept lattice by algorithm TITANIC29 and the non automatic generation of the merged ontology with human interaction based on the concept lattice (Stumme and Maedche, 2001).

CHIMAERA: CHIMAERA is an interactive ontology merging tool based on the ontolingual ontology editor. It makes users affect merging process at any point during merge process, analyzes ontologies to be merged and if linguistic matches are found, the merge is processed automatically otherwise, further action can be made by the user. It uses subclass and super class relationship (McGuinness *et al.*, 2000).

ConcepTool: This is an interactive and analysis tool that aims to facilitate knowledge sharing. It supports ontology alignment process where the ontologies are represented in Entity Relationship Model resulting from reasoning based on description logic. ConcepTool is based on heuristic and linguistic inferences to compare attributes of two entities belonging to the input ontologies. The analyst is then charged of identifying relevant information to resolve conflicts between overlapping entities. Overlapping entities are related to each other through semantic bridges. Each bridge provides a semantic transformation rule to solve the semantic mismatches between these entities. Summarizing, ConcepTool begins by analyzing the input models to derive taxonomic links and overlapping entities. Then, the analyst matches the common entities. The articulation ontology entities are automatically generated and the analyst defines mappings between the attributes of the matched entities. Finally, the articulation ontology is analyzed (Compatangelo and Meisel, 2002).

CONCLUSION

The ontology interoperability is a prominent issue in many application domains such as semantic query processing, data integration, data-warehousing, e-Commerce and e-Business. Issues of heterogeneity and inconsistency among the ontologies of same or similar domains will be resolved using ontology mapping.

Definitions of ontology matching, ontology merging, ontology integration are given. Researchers have presented a general framework situating ontology mapping. Kinds of ontology mapping are proposed. Ten challenges which we face while mapping ontologies are presented. Researchers have located three forms of mismatches that are usually studied in these processes, namely, lexical, syntactic and semantic mismatches. Because of the wide usage of ontology Interoperability techniques there is a need to consolidate different techniques and tools have been proposed to handle ontology alignment, ontology mapping and merging processes. In this study, researchers have surveyed the literature of these techniques and described the different criteria and approaches adopted by algorithms.

REFERENCES

Beneventano, D., S. Bergamaschi, F. Guerra and M. Vincini, 2003. Synthesizing an integrated ontology. IEEE Internet Comput., 7: 42-51.

Calvanese, D., G. de Giacomo and M. Lenzerini, 2001. A framework for ontology integration. Proceedings of the 1st international Semantic Web Working Symposium, July 30-August 1, 2001, Stanford University, California, USA., pp: 303-317.

Chalupsky, H., 2000. OntoMorph: A translation system for symbolic knowledge. Proceedings of the 7th International Conference on Principles of Knowledge Representation and Reasoning, April 11-15, 2000, Breckenridge, Colorado, USA., pp: 471-482.

Compatangelo, E. and H. Meisel, 2002. Intelligent support to knowledge sharing through the articulation of class schemas. Proceedings of the 6th International Conference on Knowledge-Based Intelligent Information and Engineering Systems, September 16-18, 2002, Italy.

Doan, A.H., P. Domingos and A. Halevy, 2003a. Learning to match the schemas of data sources: A multistrategy approach. Machine Learn., 50: 279-301.

Doan, A.H., J. Madhavan, P. Domingos and A. Halevy, 2003b. Learning to map between ontologies on the semantic web. VLDB J., 12: 303-319.

Dou, D., D. McDermott and P. Qi, 2005. Ontology translation on the semantic web. J. Data Semantics, 2: 35-57.

Ehrig, M. and S. Staab, 2004. QOM-Quick Ontology Mapping. In: The Semantic Web, McIlraith, S.A., D. Plexousakis and F. van Harmelen (Eds.). Springer, USA., ISBN: 978-3-540-23798-3, pp: 683-697.

Euzenat, J. and P. Shvaiko, 2007. Ontology Matching. 1st Edn., Springer-Verlag, Berlin, Germany, ISBN-13: 9783540496120, Pages: 333.

Euzenat, J. and P. Valtchev, 2004. Similarity-based ontology alignment in OWL-Lite. Proceedings of the 16th European Conference on Artificial Intelligence, August 22-27, 2004, Valencia, Spain, pp: 333-337.

Ghidini, C. and F. Giunchiglia, 2004. A semantics for abstraction. Proceedings of the 16th European Conference on Artificial Intelligence, August 22-27, 2004, Valencia, Spain, pp: 343-347.

Giunchiglia, F., P. Shvaiko and M. Yatskevich, 2005. Semantic schema matching. Proceedings of the OTM Confederated International Conferences, October 31-November 4, 2005, Agia Napa, Cyprus, pp: 347-365.

Kalfoglou, Y. and B. Hu, 2005. CROSI mapping system (CMS) results of the ontology alignment contest. Proceedings of the K-CAP Workshop on Integrating Ontologies, October 2, 2005, Banff, Canada.

Kalfoglou, Y. and M. Schorlemmer, 2003a. Ontology mapping: The state of the art. Knowl. Eng. Rev., 18: 1-31.

Kalfoglou, Y. and W.M. Schorlemmer, 2003b. IF-Map: An ontology-mapping method based on information-flow theory. J. Data Semantics, 1: 98-127.

Klein, M.C.A. and D. Fensel, 2001. Ontology versioning on the semantic web. Proceedings of the 1st Semantic Web Working Symposium, July 30-August 1, 2001, Stanford University, California, USA., pp: 75-91.

Li, J., 2004. LOM: A lexicon-based ontology mapping tool. Proceedings of the Performance Metrics for Intelligent Systems Workshop, August 24-26, 2004, Gaithersburg, MD., USA.

McGuinness, D.L., R. Fikes, J. Rice and S. Wilder, 2000. The chimaera ontology environment. Proceedings of the 17th National Conference on Artificial Intelligence and 12th Conference on Innovative Applications of Artificial Intelligence, July 31-August 2, 2000, Austin, Texas, pp: 1123-1124.

Mitra, P. and G. Wiederhold, 2002. Resolving terminological heterogeneity in ontologies. Proceedings of the ECAI-02 Workshop on Ontologies and Semantic Interoperability, July 22, 2002, Lyon, France.

Noy, N. and M. Musen, 2001. Anchor-PROMPT: Using non-local context for semantic matching. Proceedings of the Workshop on Ontologies and Information Sharing at the Seventeenth International Joint Conference on Artificial Intelligence, August 4-10, 2001, Seattle, USA., pp: 63-70.

Noy, N.F. and M.A. Musen, 2000. PROMPT: Algorithm and tool for automated ontology merging and alignment. Proceedings of the 17th National Conference on Artificial Intelligence and 12th Conference on Innovative Applications of Artificial Intelligence, July 31-August 2, 2000, Austin, Texas, pp: 450-455.

Pinto, H.S. and J.P. Martins, 2001. A methodology for ontology integration. Proceedings of the 1st International Conference on Knowledge Capture, October 21-23, 2001, Victoria, BC, Canada, pp: 131-138.

Rahm, E. and P.A. Bernstein, 2001. A survey of approaches to automatic schema matching. VLDB J., 10: 334-350.

Shvaiko, P. and J. Euzenat, 2005. A survey of schema-based matching approaches. J. Data Semantics, 4: 146-171.

Stumme, G. and A. Maedche, 2001. FCA-merge: Bottom-up merging of ontologies. Proceedings of the 17th International Joint Conference on Artificial Intelligence, Volume 1, August 4-10, 2001, Seattle, Washington, USA., pp: 225-230.