A New Implementation for Ontology Mapping Based enterprise Semantic Interoperation

Xiaofeng Di, Yushun Fan

Department of Automation, Tsinghua university,Beijing 110004, China
dixf06@mails.tsinghua.edu.cn, fanyus@tsinghua.edu.cn

Keywords: Semantic, Ontology mapping, OWL, Interoperation

Abstract. In the interoperation among enterprises in eBusiness, there is a big problem that different data models and information description are used in different enterprises’ systems which blocking the ambient Semantic collaboration. Ontology is an important tool to overcome syntax and semantic misunderstanding. Our goal is to provide a user-friendly environment supporting syntax and neutral format data model for business information. In this paper, a unified description of data model which mapping from logical data of enterprise’s information system is developed to solve the gap in interoperation. It provides the methods to realize the mapping among serious types of data or information. Firstly, database and other types of information are transformed into neutral format which described by OWL. Secondly, the neutral format can be mapped into the semantic entities and semantic linking through the process of extraction and annotation and added into ontology and been described in a standard method which makes the collaboration could be understood. Lastly, a prototype is described and realized as an integration platform.

Introduction

Companies invariably need to exchange information electronically in the process of collaborating. However, due to the large number of diverse information systems, the data format (syntax) of messages usually differs from company to company. In such cases, users either have to agree on a common standard or map the data schema from one format to another.

Today, the schema mapping process focuses on syntax. The key drawback to this approach is that mappings must be performed by a technical person who has schema knowledge, whereas what is required is the business information mapping. We are wondering to have a friendly and easy instrumentality or agent which focuses on identifying semantic entities and enabling the mapping of those semantic pieces by business users so that the syntactical mapping process can be hidden completely from the users [1]. There are two Scenarios have been researched as follows:

Scenario1: When it comes to the exchange of business documents in the automotive sector, electronic data interchange (EDI) is heavily used by all members of the automotive supply chain. Over the last decades, four major standards were established which are namely ANSI X12, UN/EDIFACT, ODETTE and VDA.

Scenario 2: A machine to machine interface is not implemented and the automatic data flow is interrupted by a manual process when suppliers and dealers are interacting with the web-interface.

If enterprise wants to order mechanical components like transmissions from their first tier supplier, such order is generated in the SAP-system of enterprise. Via a fixed interface this order is
copied into the database and will be displayed in the web-front end to the receiving supplier. The supplier needs to copy this order into their own system manually for further processing.

Theory Research

Ontology is the effective method to deal with the syntax and semantic misunderstanding [2]. However, existing ontology construction methods mainly extract semantics from specific data schema directly. The current semantic interoperability environment handles multiple heterogeneous data schemas, such as RDB and XML [3]. It is time-consuming and labor-intensive even with an ontology editor. Thus, the collaboration process within and across enterprises should be as fast as possible in distributed, heterogeneous computing environments. The methods need to be improved to reduce the time and avoid the failure.

There is some research about ontology construction such as giving an approach including learning rules to automatically formulating OWL ontology directly from relational database or XML schema [4, 5]. Recently some approaches have been present to get ontology dynamically from the metadata and structural constraints of relational database systems [6]. There is also a method for semantic business which migrate data-intensive web pages into the ontology-based Semantic Web [7].

Preliminary Definitions. Some key definitions related to semantic interoperability in multi-enterprise business collaboration environments are presented in [8]. SSE (Semantic Entity) is an INSTANCE of concept; SSET (Terminological Semantic Entity) is a concept recognized by means of some semantic extraction or annotation; SNF is a non-semantic Neutral Format.

The architecture for whole mapping process. For above issues mentioned, a new method and its implementation were presented as follows.

Fig.1 Semantic interoperability process

The kernel of the method is to mapping the semantic entities derived from distributed users each other based on Neutral Format. For Scenario 1, the mapping between SQL database and ontology could be constructed by the method proposed in here, so the enterprise which has their own logical data model in information system could build domain ontology. For Scenario 2, the mapping between XML and ontology has been built and the enterprise depending on web could communicate with its supplier and dealer automatically.
Fig. 1 illustrates how this approach works in order to obtain the link specification reflecting the semantic correspondences of two business documents in a simple case. There are two organizations conducting electronic business with each other. The process was divided into four steps which will be described in the following section.

**Methods for some key steps**

**Import diversity local data.** Due to the variety of personal data format, especially data from relation database, or data from XML Schema, which need to be converted to a same format and described by a standard language. Of course, OWL is our first choice. This is a hard step for the whole mapping process and will not be explained in detailed here. After all, a data model described by OWL which come from users’ local data will be achieved. We call this as ‘OWL -description’.

**OWL -description transform to LDM_OWL(Neutral Format).** The transformation process is a process to transform a source schema S of a model M in an “equivalent” target schema S’ of a different model M’. The source schema could be RDB, EDIFACT, FlatFile, and XML-schema while the target schema is the logical data model LDM.

The transformation process from the relational structural database into ontology representation is described here as an important facet. The rules are given in the following table 1.

<table>
<thead>
<tr>
<th>Database</th>
<th>Neutral format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table/view</td>
<td>complexnode</td>
<td>All tables are represented as complexNode</td>
</tr>
<tr>
<td></td>
<td></td>
<td>For each x in rdb:Table, create ldm:ComplexNode(cn), i.e.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>create cn as a new instance of ldm:ComplexNode</td>
</tr>
<tr>
<td>Column</td>
<td>simplenode</td>
<td>All columns are represented as simpleNode</td>
</tr>
<tr>
<td></td>
<td></td>
<td>For each x in rdb:Column, create ldm:SimpleNode(sn), i.e.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>create sn as a new instance of ldm:SimpleNode</td>
</tr>
<tr>
<td>Keyconstraint</td>
<td>identification</td>
<td>All keyconstraints (primaryKeys and alternativeKeys) are represented as Identifications</td>
</tr>
<tr>
<td>foreignKey</td>
<td>reference</td>
<td>All foreignKeys are represented as references</td>
</tr>
<tr>
<td>hascolumns</td>
<td>containment</td>
<td>The relationship between a table and its columns is represented as a containment relationship</td>
</tr>
</tbody>
</table>

As a simple example to discuss transformation rules, the following simple relational schema with three tables related by two foreign keys is considered in Fig.2. And the LDM graph in Fig.3 is obtained with the transformation rules.

![Fig.2. Works-on association](image)

**Extract SSE from SNF.** Two methods are adopted:
Conceptual extraction. Conceptual extraction entails extracting from a schema (neutral format) the typical constructs of a conceptual model, such as relationships and constraints. The output is the schema of a conceptual model, which is expressed in OWL_DL.

Keyword extraction. The keyword extraction method is proposed to construct metadata that provides a synthesized view of the values of a SimpleNode directly extracted from the data.

**Annotation for SSET.** Semantic annotation method is proposed to identify the possible links of entities. Firstly, an automatic annotation method is presented for simple annotation with an external ontology and then this is extended to the multi-ontology environment using the concept of an ontology bridge. The basic strategy is grounded on terminological and structural similarity scoring. Finally, manual annotation is proposed to extend annotations manually.

The techniques for automatic disambiguation of schema elements are discussed as a part of automatic annotation using the following methods: Gloss similarity, WordNet Domains disambiguation and Structural disambiguation.

**Mapping.** Two organizations conducting electronic business with each other are interested in the semantic correspondences between their semantic entities and could establish corresponding semantic entities linking to express them.

Apart from the direct entity mapping, there is an indirect way to derive the semantic entity correspondences through reasoning: for this purpose the SSEs are annotated and linked with some existing ontology (for example expressing their semantic equivalence of the SSE and a concept in some ontology) and the logic relations between the concepts in the involved ontology are identified. In addition, possible discrepancies between the ontology have to be reconciled. The process is labeled “ontology alignment”. Based on the annotation made against the existing ontology and on the logic relations identified between these ontology, reasoning can identify correspondences on the SSE level and support the (semi-)automatic SSE mapping.

**Implementation and Validation**

In order to allow SME and corporation participate in the e-Economy, address semantic gap in integration, provide services and applications, a structure is provided and realized between information systems of different enterprises. The information between the enterprises could be understood automatically based on the ontology. All of the rules mentioned above are implemented into user-friendly software for the construction of ontology in this project.

---

**Fig.4 Import data file**

**Fig.5 Interaction relation model**
User can import their database or other format information into software as shown in Fig.4. All information will be turned into Neutral Format and mapped into semantic entities automatically. Fig.5 is the diagram of semantic entities gotten by importing a simple relational database directly. All semantic entities and semantic linking between entities are described with OWL.

Conclusion

In this paper, a unified description of data model which mapping from logical data of enterprise’s information system is developed to build ontology for understanding among enterprises in e-Business interoperation. It provide a user-friendly platform supporting syntax and neutral format data model for user to add their information into network and also provide a way to get entities from other enterprise despite of different information system which blocking the ambient Semantic collaboration between different enterprise. User can add their data into network automatically or manually and find the entities coming from other enterprises which correlative with user’s requirements.

Acknowledgements

Supported by the National 973 Project (2006CB705407), the Europeans Commission’s 6th Framework programmer (FP6-2005-IST-5-034980), the Country’s Natural Science Fund of China (60674080), and the 863 Plan of China (2007AA04Z150). The interested readers are able to access http://www.stasis-project.net for information.

References

[8] Information on http://www.stasis-project.net