관계형 데이터베이스로부터 온톨로지 생성을 위한 변환 규칙에 관한 연구

XuNi[O] 정혜경 배우정 안동언 이용석
전북대학교 전자정보공학부
xuni1109@hotmail.com[O], sysj2@paran.com, topaz.inu@gmail.com
{duan, yslee}@chonbuk.ac.kr

A Study on Converting Rules for Generating Ontology from RDB

N. Xu[0] H.K. Jeong W.J. Bae D.U. An Y.S. Lee
Division of Electronics and Information Engineering, Chonbuk National University

요 약

Semantic Structuralization of data, not Semantic Web base on Ontology but software engineering on many points, can be applied to trace of state or requirements usefully. But it is required much effort that a man converts RDB(Relational DataBase) into Ontology. Hence this study propose rules to convert RDB into Ontology description language OWL automatically.

1. Introduction

At present, a lot of information is serviced innumerably enough to arrive in saturation through internet. Nevertheless, searched information from web is not secured certainty. The biggest cause of these problem is that a man recognizes relations of terms by brain activity, but computer can not recognize and reason relations and meaning of terms[1]. Accordingly, current web is in the entering step of semantic web and various research in Ontology are progressed. Also, W3C adopted OWL (Ontology Web Language) for automatic construction of Ontology into recommendation in 2004.

OWL includes various vocabulary and formal semantics to express information in the semantic web, and supports more expression way, so it shows good performance to make out understandable information by computer[1].

OWL language provides three increasingly expressive sublanguages OWL Lite, OWL DL(Description Logic), OWL FULL[2,3].

OWL Lite supports those users primarily needing a classification hierarchy and simple constraint features.

OWL DL supports those users who want the maximum expressiveness without losing computational completeness and decidability of reasoning systems. OWL DL includes all OWL language constructs with restrictions such as type separation. OWL DL was designed to support the existing Description Logic business segment and has desirable computational properties for reasoning systems.

OWL Full is meant for users who want maximum expressiveness and the syntactic freedom of RDF with no computational guarantees. 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 every feature of OWL Full[2].

The rest of this paper is organized as follows; in the next section, we present related works to our research. The third section deals with the Conversion Rules. This section illustrates how to convert RDB into OWL. The final section summarizes the overall discussion and present future works.

2. Related Works

In recent years, there are three approaches have been reported. The first one is based in the semi-automatic generation of ontology from Relational DB model[4,9]. Then mappings are defined between the DB and the generated ontology. The second approach, proposes the manual annotation of dynamic Web pages which publish DB content, with information about the underlying DB and how each content item in a page is extracted from the DB[4,10]. The third approach tries to map an existing DB to an appropriate existing ontology implemented in RDF(S) or OWL with the declarative language R2O[4]. But the first one does not allow the population of an existing ontology, which is a big limitation, the second approach does not deal neither with complex mapping situations and assumes we want to make our database schema public, which is not always the case although the third one R2O is automatically generating language, it does not solve the definition of degrees of similarity between DB elements and ontology very well and the readability of it can not be controlled easily by the human. Our approach is close to the first one.

As other research, T.S. Kim[6] proposed ontology generating algorithm, H.S. Kwak[7] proposed a conversion system HTML document into OWL ontology language, and S.H. Jang[8] proposed more detail approach than above.

Class hierarchies may be created by making one or more statements that a class is a subclass of another class. The Meaning of subclass in OWL is exactly the same: if the class description C1 is defined as a subclass of class description C2, then the set of individuals in the class extension of C1 should be a subset of the set of individuals in the class extension of C2[2].

3. Conversion Rules from RDB to OWL

In this section, we show a RDB schema(a company database) and relations of each tables, and present our converting rules.

3.1 Relation Database Schema

Figure 1 shows a database schema of relation database about company for our approach.

3.2 Relations of Tables

Via SQL, extract the table name from the RDB schema, and distinguish the entity table and relational table(see Figure 1).

EMPLOYEE

FNAME	SSN	BIRTHDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNUMBER
John	ID123456789	JAN09-55	TX, 731 Fondren, Houston	M	$ 30000	ID333445555	D5
Franklin	ID333445555	DEC08-45	TX, 638 Voss, Houston	M	$ 40000	ID888665555	D5
Alicia	ID999887777	JUL19-58	TX, 3321 Castle, Spring,	F	$ 25000	ID987654321	D4
Jennifer	ID987654321	JUN20-31	TX, 291 Berry, Bellaire	F	$ 43000	ID888665555	D4
Ramesh	ID666884444	SEP15-52	TX, 975 Fire Oak, Humble	M	$ 38000	ID333445555	D5
Joyce	ID453453453	JUL31-62	TX, 5631 Rice, Houston	F	$ 25000	ID333445555	D5
Ahmad	ID987987987	MAR29-59	TX, 980 Dallas, Houston	M	$ 25000	ID987654321	D4
James	ID888665555	NOV10-27	TX, 450 Stone, Houston	M	$ 55000	NULL	D1

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research	D5	ID333445555	MAY22-78
Administration	D4	ID987654321	JAN01-85
Headquarters	D1	ID888665555	JUN19-71

DEPT_LOCATIONS

DNUMBER	DLOCATION
D1	Houston
D4	Stafford
D5	Bellaire
D5	Sugarland
D5	Houston

PROJECT

PNAME	PNUMBER	PLOCATION	DNUMBER
ProductX	P1	Bellaire	D5
ProductY	P2	Sugarland	D5
ProductZ	P3	Houston	D5
Computerization	P10	Stafford	D4
Reorganization	P20	Houston	D1
Newbenefits	P30	Stafford	D4

WORK_ON

ESSN	PNUMBER	HOURS
ID123456789	P1	H32.5
ID123456789	P2	H7.5
ID66884444	P3	H40.0
ID453453453	P1	H20.0
ID453453453	P2	H20.0
ID333445555	P2	H10.0
ID333445555	P3	H10.0
ID333445555	P10	H10.0
ID333445555	P20	H10.0
ID999887777	P30	H30.0
ID999887777	P10	H10.0
ID987987987	P10	H35.0
ID987987987	P30	H5.0
ID987654321	P30	H20.0
ID987654321	P20	H15.0
ID888665555	P20	NULL

DEPENDENT

ESSN	DEPENDENT_NAME	SEX	BIRTHDATE	RELATIONSHIP
ID333445555	Alice	F	APR05-76	DAUGHTER
ID333445555	Theodore	M	OCT25-73	SON
ID333445555	Joy	F	MAY03-48	SPOUSE
ID987654321	Abner	M	FEB29-32	SPOUSE
ID123456789	Michael	M	JAN01-78	SON
ID123456789	Alice	F	DEC31-78	DAUGHTER
ID123456789	Elizabeth	F	MAY05-57	SPOUSE

Figure 1. RDB schema of company database.

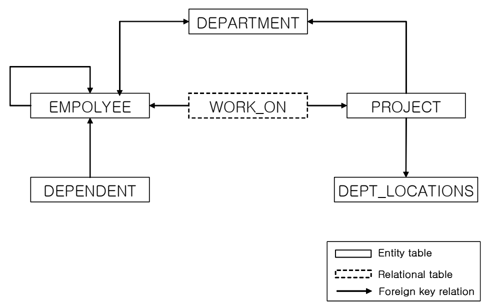

Figure 2. Table names with relation.

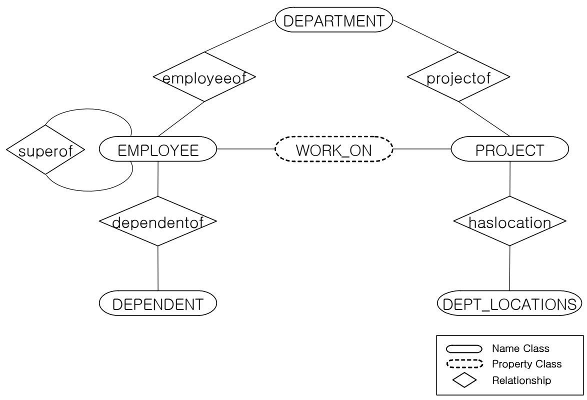

Figure 3. Generated Name class and Property class with relationship.

Figure 4 shows relationship between attributes of tables. we will explain of this figure in section 3.3 with Rule.

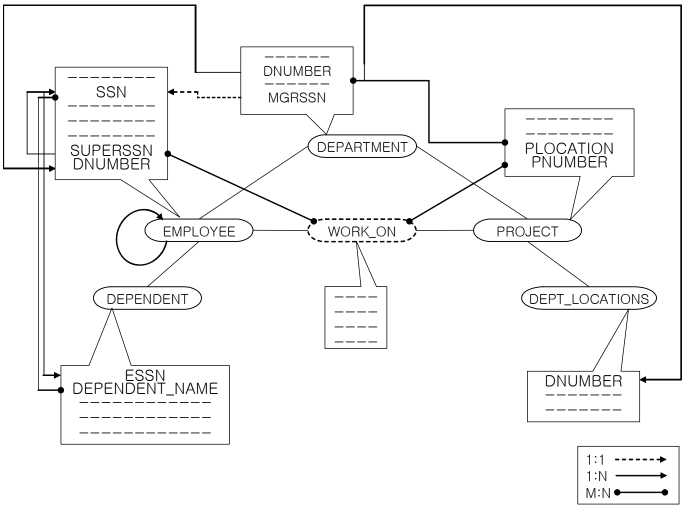

Figure 4 . Relationship between attributes of tables.

3.3 Conversion Rules

Step 1) Distinguish and select RDB table from the existing DB schema.
Via SQL extract the entity table and relational table from the DB schema, generating the entity table as Name Class with the same name of the corresponding table and generating the relational table as Property Class(see Figure 3).

Step 2) Generating subClass of the corresponding Class.

2.1) Via SQL extract the table which's name is implicated in the other table, generating this table as the subClass of the corresponding Class.

2.2) When the foreign key referenced to the table of itself, generating the foreign key attribute as a Class which is a subClassof the Class of the referenced table.

Step 3) Generating the instances of primary key attribute as the Individuals of the corresponding class.

3.1) when the primary key of a table is composed within 2 attributes, then the foreign key attribute in them generates the ObjectProperty(see Step4) the instances of other one generate the Individual of the table contains this kind of primary key.

Step 4) Generating the foreign key attribute as the ObjectProperty of the corresponding class.

4.1) when there are more than two primary key, primary key of other table is became ObjectProperty.

The domain of the ObjectProperty is the class of the table that contains this attribute, the range of it is the class of the referenced table.

Step 5) Generating the non-key attribute as the DatatypeProperty of the corresponding class.

The domain of the DatatypeProperty is the class of the table that contains the attribute, the range of it is one of the datatype of XML schema corresponding to the built-in datatype in the RDB schema.

Step 6) The relationship generating: (see Figure 4)

6.1) The 1:1 relationship generates the ObjectProperty with Cardinality=1.

6.2) The 1:n relationship generates the ObjectProperty with Cardinality=n.

Step 7) The n:m relationship generates the ObjectProperty. (see Figure 4)
when the table has the primary key composed within two attributes.

7.1) Each of the attributes is the primary of two different table, the domain of the property is the corresponding class of one table(e.q EMPLOYEE), the range of the property is the other table(e.q PROJECT) and vice versa.

7.2) Either of the attributes is the primary key of other table, the domain of the property is the corresponding class of the table(e.q DEPARTMENT) which has this primary key attribute the range of the property is the class of the table(e.q DEPT_LOCATIONS) that contains the two attribute and vice versa.

Step 8) Generating the subProperty of the corresp-

onding property
When the attribute name of a table is included in another attribute of other table, it generates the subProperty of the corresponding property.

In the Figure 1, underlined attribute means the primary key.

Table name EMPLOYEE, DEPARTMENT, DEPT_LOCATIONS, PROJECT, DEPENDENT are generated as each class name by Step1). Because table WORK_ON has the primary key composed within 2 attributes, and each of the attributes is the primary key of other table. so this table can not be a entity table, it is a relational table, it presents the relationship between two table which contain the attributes separately.

Step2) is a rule for when all of the table name are data of another table.

Attribute SSN in EMPLOYEE table, DNUMBER of DEPARTMENT, PNUMBER of PROJECT, PLOCATION of DEPT_LOCATIONS, DEPENDENT_NAME of DEPENDENT are generated as individual by step3).

DNUMBER of DEPT_LOCATIONS and ESSN of DEPENDENT are became ObjectProperty in Step3.1).

In the Step4), DNUMBER of EMPLOYEE is became ObjectProperty of EMPLOYEE class and MGRSSN of DEPARTMENT, DNUMBER of PROJECT, DNUMBER of DEPT_LOCATIONS and ESSN of DEPEDENT are became ObjectProperty for each class.

The other attributes, except primary key and foreign key, are became DatatypeProperty by Step5).

4. Conclusions and Future Works

We proposed several converting rules to generate ontology from RDB. From these rules, we could generate basic elements such as Class, subClass, individual, ObjectProperty, DatatypeProperty, and subProperty in OWL. But there are many other elements in OWL, so we need to extend our rules for generation of the other elements in OWL.

5. References

[1] Whee Seo, “A Theoretical Study of Using Methods for OWL Vocabulary and Syntactics to Ontology Automatic Construction”, Journal of Korean Library and Information Science Society, Vol. 37, No. 2, (2006) 191-216

[2] http://www.w3.org/TR/owl-guide/

[3] Alan Rector, Nick Drummond, Matthew Horridge, et al., "OWL Pizzas:Practical Experience of Teaching OWL-DL:Common Errors & Common Patterns", 14th International Conference on Knowledge Engineering and Knowledge Management (EKAW), Whittlebury Hall, UK, (2004)

[4] Jesús Barrasa, et al., "R2O, an Extensible and Semantically Based Database-to-ontology Mapping Language", In Proceedings of the 2nd Workshop on Semantic Web and Databases, Toronto, (2004)

[5] Jesús Barrasa, Oscar Corcho, Asuncion Gomez-Perez, "Fund Finder: A case study of database-to-ontology mapping", Semantic Integration Workshop, ISWC 2003. Sanibel Island, Florida. Sept (2003)

[6] Tae-Suk Kim, Jin-Hyuk Yang, In-Jeong Chung, "Efficient Creation of Domain Ontology Using DataBase", The 24th KIPS Autumn Conference, Vol. 12, No. 2, (2005) 485-488

[7] Hyoun-Soo Kwak, Su-Kyoung Kim, Yeong-Geun Kim, Kee-Hong Ahn, "A Conversion System of HTML Document into OWL Ontology language", The 22th KIPS Autumn Conference, Vol. 11, No. 2, (2004) 539-542

[8] Sohee Jang, Insuk Park, Soon J. Hyun, "Converting Relational Database to Ontology in OWL", Koran DataBase Conference 2004 KDBC (SIGDB-KISS), (2004) 37-43

[9] Stojanovic L., Stojanovic N., Volz R., "A Reverse Engineering Approach of Migrating Data-intensive Web Sites to the Semantic Web", Intelligent Information Processing, Montreal (2002)

[10] Handschuh S, Staab S, Volz R, "On Deep annotation", 12th International World Wide Web Conference, (2003)