Creating Databases from Ontologies

C. Forbes Dewey, Jr.
Massachusetts Institute of Technology
Cambridge, MA USA
Here’s our agenda

Review of database technology

*History*
*What is their role?*
*What do they do well?*
*What do they do poorly?*

Ontologies

*Why are they desirable?*
*What are their advantages?*
*What are their weaknesses?*

Relationship between DB and Ontologies

Exploiting the combination - OWLdb
Review of relational database technology

Started in mid-80’s

- Composed of tables with rows and columns
- Minimal language to construct and query them: SQL
- Generally fixed and unchanged relationships

Java-based interface JDBC a real breakthrough; common interface for all flavors. Can write programs!!!

Databases don’t talk to each other

Stove-piping, warehousing, federation

Updates, security, reliability, performance
OWL Ontologies

Describe meaning

- OWL is a language (set of relationships) that can be extended with additional definitions
- OWL is written in RDF, is machine-readable, and can be parsed and combined

Representing databases as ontologies

- Can represent a database schema by an ontology
- Will discuss state-of-the-art programs to do the conversion
- If we can convert, can have the best of both environments
OWL – Web Ontology Language

• An expressive and uniform way of defining meaning for terms used to transmit data and relationships

• Can be used for many key purposes
  ➢ Guarantee that two definitions are the same
  ➢ Discover that two terms are synonymous
  ➢ Encode complete object descriptions
  ➢ Define unambiguous database schema

• Comes in multiple flavors
  ➢ OWL Lite – OWL DL – OWL Full

http://www.w3.org/TR/2004/REC-owl-features-20040210/#s2.1
Owl Structure

• Classes
• Properties
• Types
• Meta-Data

Owl is written in RDF
OWL – Web Ontology Language (2)

RDF Schema Features:
Class (Thing, Nothing)
rdfs:subClassOf
rdf:Property
erdfs:subPropertyOf
erdfs:domain
rdfs:range
Individual

(In)Equality:
equivalentClass
equivalentProperty
sameAs
differentFrom
AllDifferent
distinctMembers

Property Characteristics:
Property Restrictions:
Restricted Cardinality:
Header Information:
Class Axioms:
Arbitrary Cardinality

Class Intersection:
Versioning:
Annotation Properties:
Datatypes:
Boolean Combinations:
Filler Information

http://www.w3.org/TR/2004/REC-owl-features-20040210/#s2.1
OWL – Web Ontology Language (3)

**RDF Schema Features:**

`rdfs:subPropertyOf`

`rdfs:subPropertyOf`: Property hierarchies may be created by making one or more statements that a property is a subproperty of one or more other properties. For example, hasSibling may be stated to be a subproperty of hasRelative. From this a reasoner can deduce that if an individual is related to another by the hasSibling property, then it is also related to the other by the hasRelative property.

One of 54 base language constructs

[http://www.w3.org/TR/2004/REC-owl-features-20040210/#s2.1](http://www.w3.org/TR/2004/REC-owl-features-20040210/#s2.1)
OWL – Web Ontology Language (2)

**RDF Schema Features:**
- Class (Thing, Nothing)
- rdfs:subClassOf
- rdfs:subPropertyOf
- rdf:Property
- rdfs:domain
- rdfs:range
- Individual

**Property Characteristics:**
- Property Restrictions:
- Restricted Cardinality:
- Header Information:
- Class Axioms:
- Arbitrary Cardinality

**(In)Equality:**
- equivalentClass
- equivalentProperty
- sameAs
- differentFrom
- AllDifferent
- distinctMembers

**Class Intersection:**

**Versioning:**

**Annotation Properties:**

**Datatypes:**

**Boolean Combinations:**

**Filler Information**

http://www.w3.org/TR/2004/REC-owl-features-20040210/#s2.1
XML Schema Datatypes used in OWL

http://www.w3.org/2001/XMLSchema#name

Our approach to creating and merging databases

- Create ontologies from database schema
- Edit and maintain the ontologies
- Utilize merged & aligned ontologies
- Create on-the-fly databases from ontologies
- Model scientific processes - use cases
- Allow for interoperability
Creating ontologies from database schema

D2RQ is a declarative language to describe mappings between relational database schemata and OWL/RDFS ontologies.

http://www4.wiwiss.fu-berlin.de/bizer/D2RQ/

Courtesy of Prof. Dr. Christian Bizer. Used with permission.
Options for Ontology Merging & Aligning


Architecture of MoA system.
Options for Ontology Merging & Aligning (2)

Options for Ontology Merging & Aligning

Ontology Merging for Federated Ontologies on the Semantic Web

Gerd Stumme
Institute for Applied Computer Science and Formal Description Methods (AIFB)
University of Karlsruhe
D-76128 Karlsruhe, Germany
www.aifb.uni-karlsruhe.de/WBS/gst

Alexander Maedche
FZI Research Center for Information Technologies
Haid-und-Neu-Strasse 10-14
D-76131 Karlsruhe, Germany
www.fzi.de/wim

\[ D: \text{set of natural language documents} \quad K: \text{formal context} \quad O: \text{ontology} \quad R: \text{set of relation names} \]
\[ B\{K\}: \text{concept lattice of } K \]


Figure by MIT OpenCourseWare.
Options for Ontology Merging & Aligning

An Algorithm for Merging and Aligning Ontologies: Automation and Tool Support

Natalya Fridman Noy and Mark A. Musen

Stanford Medical Informatics
Stanford University
Stanford, CA 94305-5479
{noy, musen}@smi.stanford.edu
OWLdb: A new paradigm for creating databases

Integration - Ontology

UniProt

BioModels

UniModels

External Ontologies

Internal Ontologies
The Species Link

UniProt

Protein

Protein Table

BioModels

Model

Protein
The OWLdb design
Architecture

- BioModels.net
- SBO/Bio PAX
- UniProt
- UniProt RDF
- Front End + Contextualization + Use Case
- UniModel MetaOntology in OWL (Handcrafted) + inference rules
- UniModel UniProt + BioModel DB (MySQL etc.)
- OWL db
- SPARQL
- SQL
## Key Components

<table>
<thead>
<tr>
<th>Key Component</th>
<th>UniModels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registry</td>
<td>LSID</td>
</tr>
<tr>
<td>Interface</td>
<td>UniModels Ontology</td>
</tr>
<tr>
<td>Nomenclature</td>
<td>UniModels / SBO / UniPROT RDF</td>
</tr>
<tr>
<td>Editor</td>
<td>Protege/SBMLEditor</td>
</tr>
<tr>
<td>Parser</td>
<td>Jena</td>
</tr>
<tr>
<td>Reasoner</td>
<td>Jena/OWLdb</td>
</tr>
<tr>
<td>Storage</td>
<td>Oracle / MySQL / PostgresSQL / Sesame</td>
</tr>
<tr>
<td>Query</td>
<td>SPARQL / SQL / D2RQ / SeRQL</td>
</tr>
</tbody>
</table>
Computational and Systems Biology