

Nadine CULLOT  
 University of Burgundy (Dijon)  
 Laboratory of Computer Science  
 LE2I (Laboratoire Electronique  
 Informatique et Image)

Cooperation of information systems  
 using ontologies

## Cooperation of Database Systems

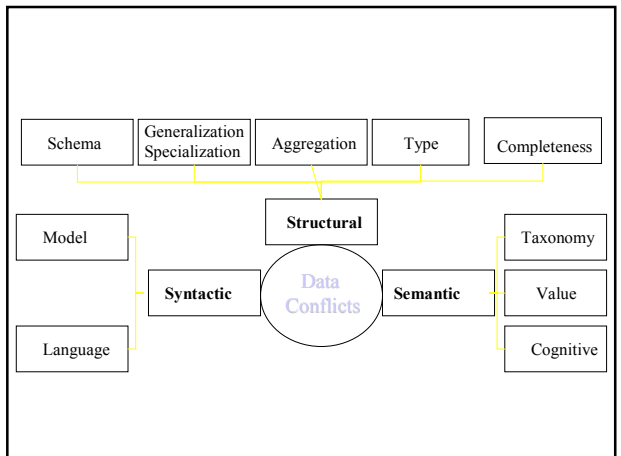
- How to access, exchange or share information from different databases in a transparent way to the end-users ?



## Heterogeneity problems

- Different levels of heterogeneity :
  - Hardware level,
  - Operating systems (OS) level,
  - Data level.

→ Data conflicts



## Tools to identify and solve data conflicts

- Metadata,
  - Metadata describe intrinsic properties of data.
- Contexts,
  - Contexts modelize the semantics of an application domain
- Ontologies.

## Outline of the talk

- Ontology motivation /challenges
- How to modelise semantics ?
- Ontologies classification (Guarino)
- Languages to express semantics (ontologies)
  - RDF (Resource Description Framework)
  - Description logics
  - DAML + OIL (I. Horrocks – University of Manchester)
- On using ontologies
  - DB-Globe Approach (D. Pfoser, E. Pitoura, N. Tryfona)
  - KAON Approach (B. Motik, A. Maedche, R. Volz) University of Karlsruhe (team of R. Studer)
  - DOGMA Approach (M. Jarrar, R. Meersman, University of Brussel)
- Integration using ontologies
  - Using Ontologies for Integrated GIS (F.T. Fonseca, M.J. Egenhofer, P. Agouris, G. Camara, NCGIA – University of Maine)

## Ontology motivation

- Enrich information with semantics
- Why ?
  - To be able to understand information over the Web
  - To share this information between users/applications
- How ?
  - With the definition of shared conceptualization of some domains

## Ontology challenges

- Ontology definition : need to modelize the real world (or some domains of) with unambiguous concepts/relationships
- Ontology use : need to find/ share/integrate information between end-users/applications
- Ontologies management : extension/evolution comparison/integration

## Classification of ontologies (Guarino 97)

- Top-level ontologies describe very general concepts.
- Domain ontologies describe the vocabulary related to a generic domain.
- Task ontologies describe a task or and activity;
- Application ontologies describe concepts depending of both a particular domain and a task and are usually a specialization of them.

## How to modelise semantics ?

- From the less to the more structured solutions
  - Natural language with some markups
  - Simple meta-data such as in XML-based languages
  - Data models such as RDF (Resource Description Framework)
  - Logical models

## Description Logics

- KL-One like systems (terminological logics)
  - CLASSIC (Brachman & al. 89), BACK (Peltason 91), LOOM (MacGregor 91) ...
  - DLP (Patel-Schneider 99), FACT (Horrocks 98), RACE (Haarslev-Moeller 99)
- Interest for describing ontologies
  - DAML + OIL

## BACK in few words (1)

- Basic notions are :
  - A concept represents a set of instances (intensional or extensional),
    - plant :< anything
    - product :< anything
    - biological\_plant :<plant
  - A role is a binary relation between instances of concepts
    - produces :< domain(plant) and range(product)
    - produced\_by :< inv (produces)

## BACK in few words (2)

- Value restrictions
  - water\_energy\_plant :< plant and all(produces,energy)
- Number restrictions
  - Toxic\_waste\_plant := plant and atleast(1,co\_produces,toxic-waste)
- A term is either a concept or a role
- An object is an instance of a concept
  - biograin :: biological\_plant
  - toxiplant :: atleast(1,produces) and produces : toxipharm

## BACK in few words (3)

- Non-definitional information (rules)
  - some (co\_produces, toxic\_waste) => all(pollution, risky)
- Type of queries
  - t1 ?< t2     subsumption mecanism
  - o ? : c     object classification
  - o1 ? : r:o2   are two objects related by a role
  - etc...

## Use of BACK

- TIME project (C. Nicolle PhD) to describe a hierarchy of metatypes and to be able to classify (introduce) a new metatype using the subsumption mecanism

SHB : A Strategic Hierarchy Builder for Managing Heterogeneous Database  
 Proc. of the International Database Engineering and Applications Symposium (IDEAS99) Montreal, Canada, August 2-4, 1999  
 Christophe Nicolle, Nadine Cullot, Kokou Yétongon

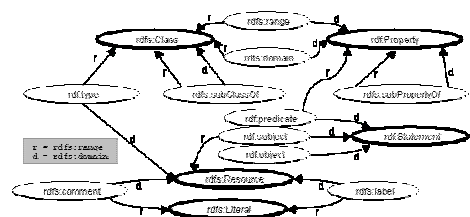
## Resource Description Framework (RDF)

- RDF is a general-purpose language for representing information in the Web.
- It proposes a vocabulary which is XML- based to describe resources.
- A RDF schema corresponds to the specification of some resources with their properties and their relationships.
- The used syntax is XML-based, and allows to describe different elements.
  - (resources, classes/subclasses, properties/subproperties, domain and range, etc ...

## RDF Syntax - Example

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <rdfs:Class rdf:about="http://www.w3.org/2000/01/rdf-schema#Resource">
    <rdfs:isDefinedBy rdf:resource="http://www.w3.org/2000/01/rdf-schema#" />
    <rdfs:label xml:lang="en">Resource</rdfs:label>
    <rdfs:comment>The class resource, everything.</rdfs:comment>
  </rdfs:Class>
  <rdfs:Property rdf:about="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">
    <rdfs:isDefinedBy rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#" />
    <rdfs:label xml:lang="en">type</rdfs:label>
    <rdfs:comment>Indicates membership of a class</rdfs:comment>
    <rdfs:range rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class" />
    <rdfs:domain rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource" />
  </rdfs:Property>
  ....
</rdf:RDF>
```

## RDF Properties



## DAML + OIL

(Darpa Agent Markup Language/ Our Ideas of a Language)

- DAML+OIL is an ontology language designed for use on the Web
- OIL has three roots :
  - description logics (for formal foundations and reasoning support/ subsumption mechanism)
  - frame-based systems (essential modeling primitives)
  - Web languages (XML/RDF based syntax)
- The connection with DL is done by defining a set of constructors as in DL

## DB-Globe Approach

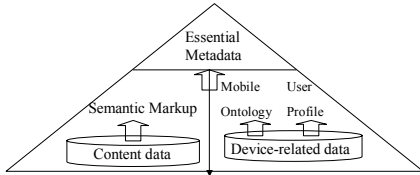
Metadata Modeling in a Global Computing Environment,  
D. Pfoser, E. Pitoura, N. Tryfona

DB-Globe Architecture deals with

- « PMO » (Primary Mobile Objects) which access via proxy to « Data Handler » linked with « DataStore »
- Different kinds of data are considered :
  - Content data (actual data stored in every PMO)
    - Descriptive data, spatially and temporally referenced information (where and when these data were stored)
    - Addition of semantic markup on content data
  - Device-related data
    - User profile, Device parameters, Movement data,
    - Description of a Basic Mobile Ontology which describe the properties of a trajectory relatively to an area of interest (stay within, bypass, leave, enter, cross // intersect, meet, equal, near, far)
    - An UML Schema of the Ontology is proposed*

## DB-Globe Approach

- Essential Data
  - It is used to create an image of the PMO in the DataStore
  - It contains
    - Part of device-related data
    - Abstractions of the content data



## DB-Globe Approach

- Positive points
  - Classification of the different kinds of data
  - Use of ontologies for GIS aspect and semantic information of content data
  - Keep the essential data on the Data Store to minimize the communications

## KAON (Karlsruhe Ontology) Approach

A Conceptual Modeling Approach for Semantics-Driven Enterprise Applications,  
B. Motik, A. Maedche, R. Volz

- Conceptual modeling approach suitable for business-wide applications
- Requirements
  - Unambiguous Semantics : to avoid diverging interpretations of intended meanings
  - Object Oriented Paradigm : successful and intuitive
  - Meta-Concepts : How to modelize an element ? As a concept or an instance, it is not always clear

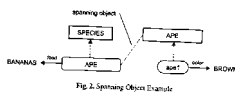


Fig. 2. Splitting Object Example

## KAON (Karlsruhe Ontology) Approach

- Requirements
  - Modularization : both concepts and instances may be subjected to modularization
  - Lexical information : on the entities of an ontology
  - Root Concept : hierarchy of concepts
  - Light-weight Inferences : predefined types of rules
- Definition of the modeling language
  - On a mathematical level (definitions of the structures of the language)
  - With its denotational semantics

### KAON (Karlsruhe Ontology) Approach

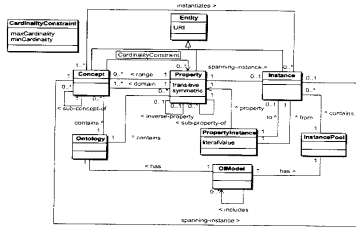
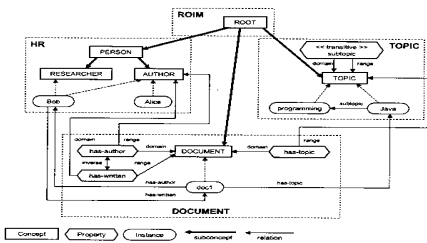


Fig. 4. UML View of KAON API

### KAON (Karlsruhe Ontology) Approach

- An ontology can be define using the proposed concepts,
- KAON propose an API, which is a set of interfaces to access and manipulate ontologies
- Different implementations are proposed for accessing RDF repository or any database
- A generic schema is given for database implementation

### KAON (Karlsruhe Ontology) Approach Example Domain ontology



### KAON (Karlsruhe Ontology) Approach Example Domain ontology

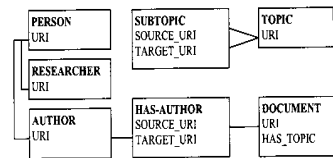


Fig. 6. Optimized KAON Schema

### KAON (Karlsruhe Ontology) Approach

- Positive points
  - Complete approach which define a modelling language but also API to give access to defined ontologies through different implementations

### DOGMA Approach

Formal Ontology Engineering in the DOGMA approach  
M. Jarrar, R. Meersman

- A database-inspired approach for engineering formal ontologies
- Knowledge are splitted into two groups : abstract contexts (set of lexons) and a layer of commitments
  - Lexons are binary facts (term1 role term2)
  - Commitments are rules, constraints ...

DOGMA Approach  
 Example : The Scientific Conference Domain  
 The ontology base

```
(Organization-ContextID)
Person    IsMemberOf    Committee
Person    Chairs        Committee
Committee ChairsBy     Person
Reviewer  SubtypesOf   Person
Author    SubtypesOf   Person
Reviewer  Reviews       Paper
Paper     ReviewedBy    Reviewer
Paper     WrittenBy     Author
Author    Presents      Paper
Paper     Has            PaperTitle
Paper     Has            PaperNumber
[...]
(ResearchAreas-ContextID)

Representation_and_Storage SuperAreaOf Ontology_Languages
Representation_and_Storage SuperAreaOf Semi-Structured_Data
Applications_and_Evaluation SuperAreaOf Semantic_Web
Applications_and_Evaluation SuperAreaOf Media_Archives
[...]
```

DOGMA Approach  
 Example : The Scientific Conference Domain  
 Commitments

```
(ConferenceAdmin-Commitment)
<Each Committee ChairsBy at most one Person>
<Each Person who chairs a Committee must also IsMemberOf that Committee>
<Each Reviewer Reviews at least one Paper>
<Each Paper which is WrittenBy a Person must not ReviewedBy with that Person>
[...] (Rules are verbalized in a suitable pseudo-NL syntax)
```

DOGMA Approach

- The notion of context is introduced to assure the consistency of an ontology.
- Different but plausible representations of the real world may be defined in an ontology but not in the same context which can be view as an interpretation.

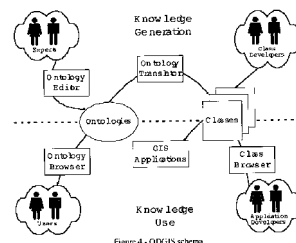
DOGMA Approach

- Positive points
  - The chosen approach is pragmatic and linked to database engineering.

Using Ontologies for Integrated GIS  
 F.T. Fonseca, M.J. Egenhofer, P. Agouris

- An architecture for an ODGIS (Ontology-Driven GIS) is proposed.
- The main idea is to propose tools to describe ontologies in a formal representation but also to translate these formal definitions into computing languages (as java, ...).
- These « classes/functions/... » can then be used in applications.

Using Ontologies for Integrated GIS  
 F.T. Fonseca, M.J. Egenhofer, P. Agouris



Using Ontologies for Integrated GIS  
F.T. Fonseca, M.J. Egenhofer, P. Agouris

- Components of an ODGIS architecture :
  - Ontology server : has to make ontologies available for applications.
  - Ontologies : specifications (made by experts) and classes generated by translation (software component)
  - Information sources : can be geographic databases which communicate with mediators.
  - Mediators extract the pieces of information necessary to generate an instance of an entity of an ontology.
  - Applications : for example information retrieval, ...

Using Ontologies for Integrated GIS  
F.T. Fonseca, M.J. Egenhofer, P. Agouris

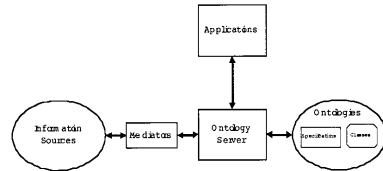


Figure 8 - Basic architecture of an ODGIS

Using Ontologies for Integrated GIS  
F.T. Fonseca, M.J. Egenhofer, P. Agouris

- Positive points
  - A complete architecture is proposed, with tools which allow to have software components from the formal ontologies
  - An example is developed in the paper for image retrieval.

## Conclusion

- There are, in a way, two approaches to modelize ontologies
  - Logic models (Description logics, Frame Based logics, ...).
  - More databases like approaches (Conceptual models).
- In the first case, there are well founded theories, with tools to make proofs, classifications, etc ...
- In the second case, there is all the engineering knowledge of databases (store, update, query)