Semantic Data Intergration:

An Introduction
--(semantic)description of data source

Hui Wang

Content:

- 1. Ontology-based Integration approach
  - Single Ontology approach
  - Multiple Ontologie approach
  - Ontology and Information System
- 2. Context-based Intergration approach
  - Context logic
  - Semantic value
- 3. Relations-based Integration approach

1. Ontology-based Integration approach

- Single Ontology approach
- Multiple Ontologie approach
- Ontology and Information System

Single Ontology approach

- Single ontology approach which supports expression of complex constraints as a part of the query.
- Prominent project: SIMS
- SIM projected every information source into a global ontology, while the schemaelemente of a structure inforamtion source were mapped on the elements. A user submit a query on the items of ontology,the SIM mediator translate it in the sub-query on informations source.

Single ontology approach

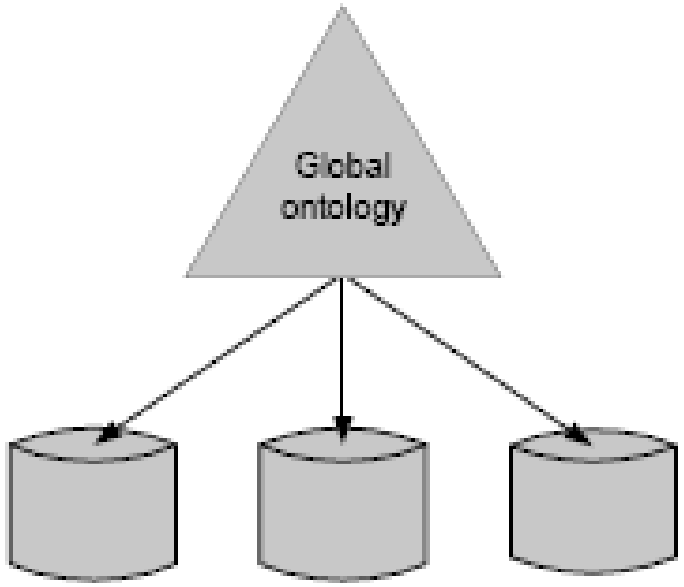


Figure 1a: Single ontology approach

The SIMS Architecture

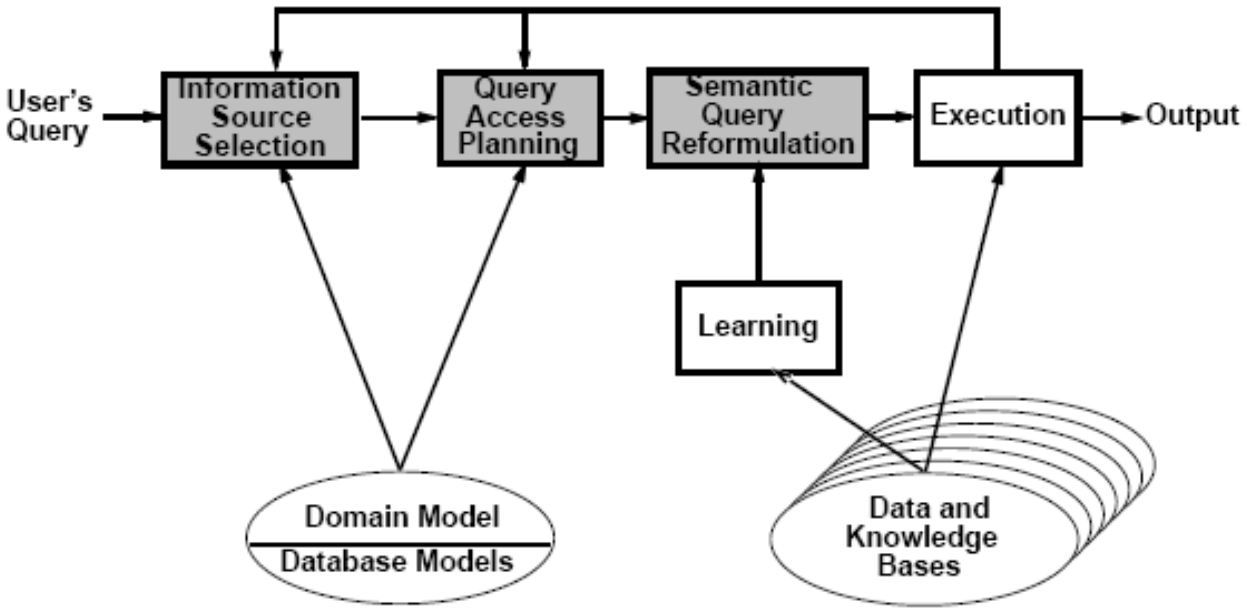


Figure 1: The SIMS Architecture

Multiple Ontologies approach

- Characteristic of it is that every information system are described through their own ontology(„source ontology“).it existed in ideal case,there are exactly many ontologies as much as informtion systems.
- Prominent project : Observer
- Observer: an achitecture for global information system.

Multiple ontologies approach

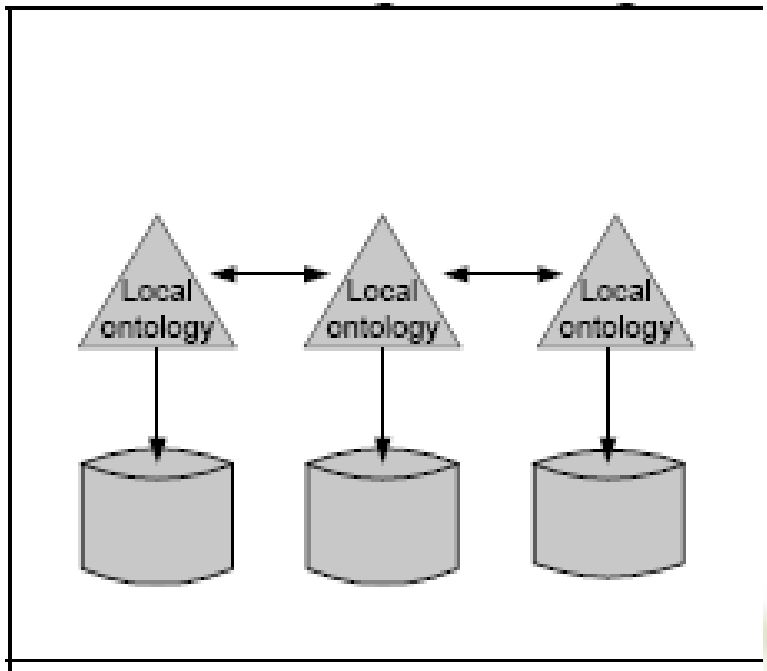
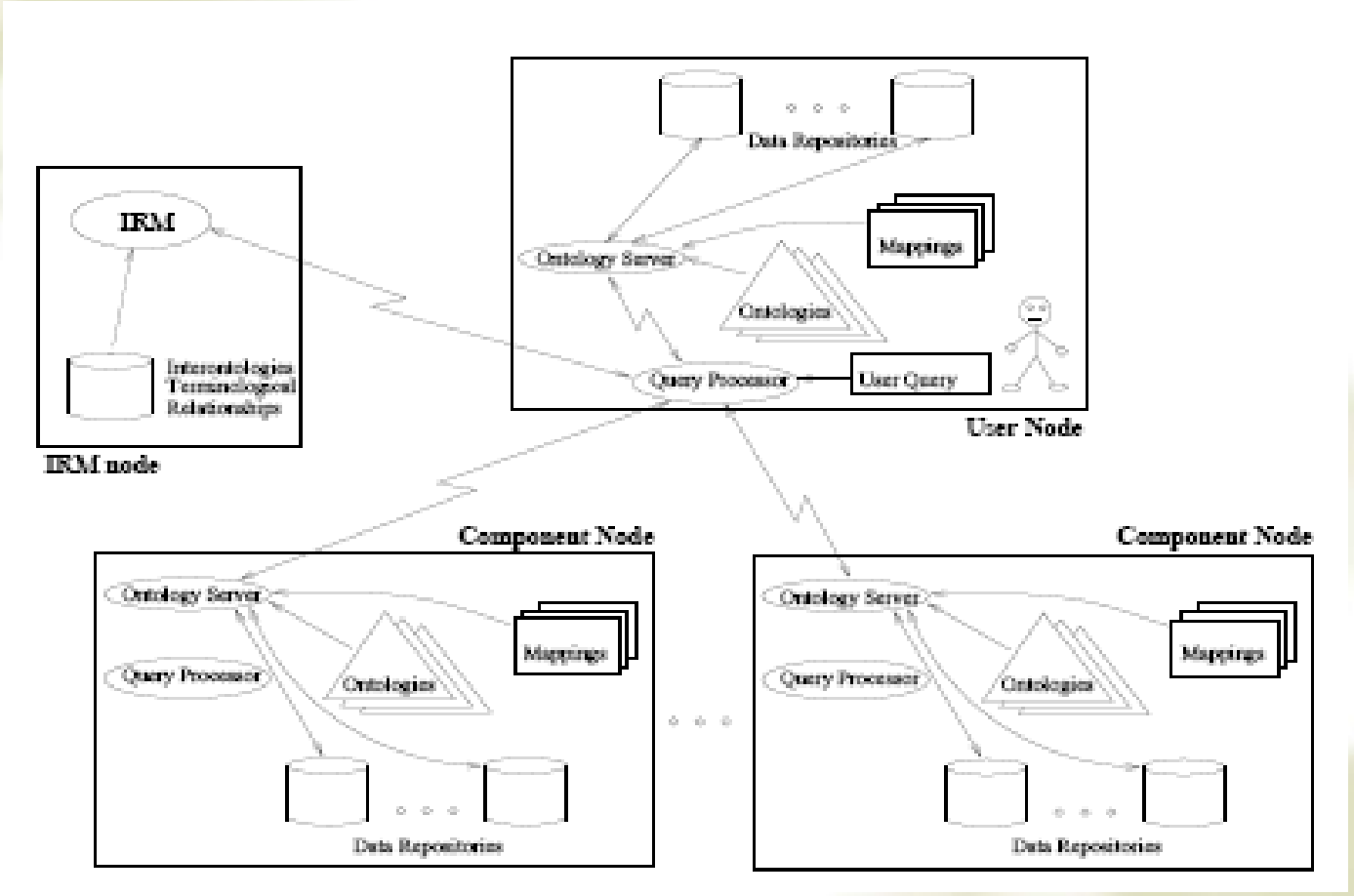


Figure 1b: Multiple ontologies approach

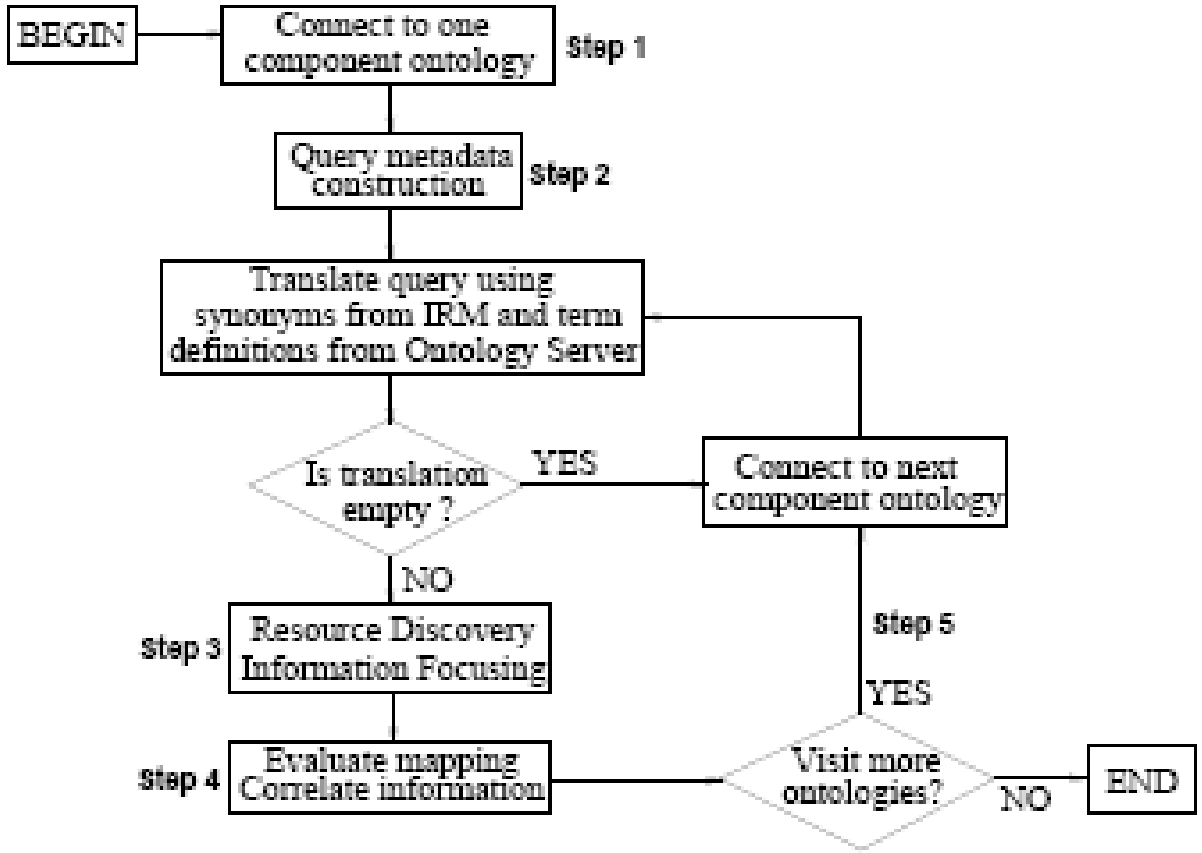
OBSERVER:
An architecture to support Query Processing



The basic elements of the architecture

The query processing approach in Observer:

- **Step 1:** Connection to the system.
- **Step 2:** Query metadata construction.
- **Step 3:** Resource Discovery and Information Focusing.
- **Step 4:** Information Access and Correlation.
- **Step 5:** Iteration.


BEGIN
Connect to one component ontology
Step 1
Query metadata construction
Step 2
Translate query using synonyms from IRM and term definitions from Ontology Server
Is translation empty ?
YES
Connect to next component ontology
NO
Step 3
Resource Discovery Information Focusing
Step 5
YES
Step 4
Evaluate mapping Correlate information
Visit more ontologies?
NO
END

Ontology and Information System

- Structure resemblance
- Definition of Terms
- Structure Enrichment
- Meta-Annotation

Structure resemblance

- The simplest possibility is the transmission of the schemata for the individual Information System

Definition of Terms

- For making the sematic of the data,just making a copy for the datastructure isn‘t enough.

Structure Enrichment:

- The most frequent case combined the two **categories** above-mentioned**.** It is formed a logic model, which was reproduced to a Data structure for the Information systems and enriched into another one with the additional Meta information.

Meta-Annotation

- An advanced possible is the semantic annotation, about it the information in Information systems was directly enriched with describing those Semantic.

2. Context-based Intergration approach

- Context logic
- Semantic Value

Context Logic

- The modality which McCarthy suggest as the formalization to such context-dependent statement is:

$$\bar{c}: \quad \mathbf{ist}(c, \phi)$$

Example:

- The relations description for the bourse in Frankfurt

boerse					
name:str	wkn:int	sym:str	art:char	preis:dbl	datum:str
Volkswagen AG	766400	VOW.FSE	A	62.05	08.05.2000
Bayrischen Motoren Werke AG Stammaktien EO 1	519000	BMW.FSE	A	44.26	08.05.2000
CITIBANK, N.A. NEW YORK, FIL.FFM KOS00/20.6.01 DAX 7400	752156	JSAK.FSE	O	6.46	08.05.2000
⋮	⋮	⋮	⋮	⋮	⋮

Example:

- the scenario in frankfurt bourse was represented in kontext logic:

$\bar{c}$: **ist**(fr, boerse(volkswagen-ag, 766400, VOW.FSE, A, 62.05, 08.05.00))
$\bar{c}$: **ist**(fr, boerse(bmw-ag, 519000, BMW.FSE, A, 44.26, 08.05.00))
...

Example:

- A semantic context for the Frankfurt exmaple could be expressed as the following:

$$\bar{c}: \text{ist}(\text{s-fr}, \text{boerse}(\text{volkswagen-ag}, 766400, \text{VOW.FSE}, \text{A}, 62.05, 08.05.00)) \wedge \text{currency}(62.05, \text{euro}) \wedge \text{scale}(62.05, 1) \wedge \ldots)$$

Semantic Value

- The „semantic value“ of the stock price in the frankfurt bourse:

  62.05(scale = 1, currency = euro)

The „semantic-value schema“ for the franfurt bourse could be expressed in parts as the following :

```
CREATE TABLE boerse (
    name char[30]
    preis double
       ( scale int
         currency char[24] )
    ... )
```

```
CREATE SCENE FOR boerse BY RULES
    IF true
    THEN scale = 1 AND
         currency = euro
```

Relations-based Integration approach

- Relations-based Integration approach based on the relations between the information Systems, which describe the semantic correlation of the information and structure elements respectively
- ISCA(inter-schema correspondence assertion)
- semPro(semantic proximity)

Relations-based Integration approach

- The semantic association between two elements form different information systems are represented as „inter-schema correspondence assertion(ISCA)“ with the following form:

$$\mathbf{Assert}[x, y](naming, abstraktion, heterogenity)$$

Relations-based Integration approach

- The another form of the spezification for the semantic association is semantic proximity.
- The semantic proximity semPro is defined as 4-tupel.

$$\mathbf{semPro}(O_F, O) = \langle C, M, (D_F, D), (S_F, S) \rangle$$

literatur

- Stand der Technik; Wache, 2003, S. 59-78; Akademische Verlagsgesellschaft Aka GmbH;Berlin
- Grundlagen und Anforderungsdefinition; Wache,2003, S. 13-20; Akademische Verlagsgesellschaft Aka GmbH; Berlin
- Query processing in the SIMS information mediator. Source, Year of Publication: 1997. ISBN:1-55860-495-2.Authors. Yigal Arens. Chun-Nan Hsu. Craig A. Knoblock. Publisher

literatur

- OBSERVER: An Approach for Query Processing in Global Information Systems based on Interoperation across Pre-existing Ontologies

  LSDIS Lab, http://lsdis.cs.uga.edu, Univ. of Georgia, Athens, GA30602

  Facultad de Inform´atica, Universidad del Pa´ıs Vasco. San Sebasti´an. Spain

  Dept. of Computer Science, Rutgers University, New Brunswick, NJ 08903

Thanks a lot !