Towards Ontological Context Mediation for Semantic Web Database Integration: Translating COIN Ontologies Into OWL

Sumit Bhansali, Benjamin Grosof, Stuart Madnick

Massachusetts Institute of Technology, Sloan School of Management

50 Memorial Drive, Cambridge, MA 02142, USA

{bhansali, bgrosof, smadnick}@ mit.edu

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

Overview I

- **Context Interchange (COIN) is an approach to <u>information integration</u>. It uses <u>ontological mappings</u> and enables powerful <u>context-sensitive query mediation</u> for semantic integration of knowledge across multiple <u>heterogeneous database sources</u>.**

- **Existing COIN applications include:**
  - **<u>financial</u>: reporting and analysis**
  - **<u>travel</u>: airfare and car-rental aggregators**

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

Overview II

- **COIN's original development preceded the Semantic Web.**

- **How best to combine COIN's capabilities with those of OWL and RuleML?**
  - **This paper provides a big first step**

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

Overview III

- **We present a translation of COIN's ontology representation into OWL Description Logic.**

- **We identify at a high level how to use RuleML LP rules together with ontologies to perform COIN-type reasoning such as mapping of ontologies and mediation of queries.**

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

COIN Ontological Model's Translation to OWL I

- Semantic Type → OWL Class
- Attribute → OWL Property
- Source Relation → OWL Class
- Source Relation Column → OWL Property
- is-a inheritance link -> subClassOf axiom

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

COIN Ontological Model's Translation to OWL II

- **Context → instance of COINContext Class**
- **Modifier → OWL Property**

Actually, only RDF-Schema features are used:

- **rdfs:class**
- **rdfs:property**
- **rdfs:subClassOf**
- **rdfs:domain**
- **rdfs:range**

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

Kinds of Reasoning in COIN

COIN does several kinds of reasoning:

- **Ontological mapping**
- **Query mediation**
  - **Abduction**
  - **Constraint Handling & Equation Solving**

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

Use Rules for COIN Reasoning I

- **Description Logic *alone* is not well-suited for COIN reasoning – OWL is not enough**

- **Want Logic Program (LP) Rules – RuleML**

- **RuleML can express the DLP subset of OWL**
  - **DLP = Description Logic Programs knowledge representation**

- **Future Direction for Work: Use RuleML to express rules to do COIN-type reasoning**

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

Use Rules for COIN Reasoning II

- **Challenge for Future – how to treat abductive reasoning, constraint handling rules in the context of the Semantic Web**

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

MORE ABOUT COIN FOLLOWS

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

COIN Motivation

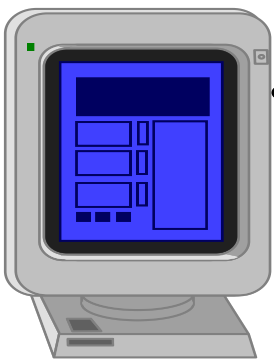
- Distributed databases makes many disparate sources available.
- The web is making even more semi-structured sources available.
  -- With XML and Web Wrapping, these can be treated as databases.
- Schema integration addresses the problem of syntactic inconsistencies.
  -- i.e., differing structures.
- How do we address semantic inconsistencies.
  -- i.e., differing meanings.
  (e.g., what does “price” really mean?)

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

MIT Sloan COntext INterchange (COIN) Project

Appli-
cations

Receivers



Browsers

OUTPUT PROCESSING

ODBC Driver

Web - Publishing

CONTEXT MEDIATION

* Automatic conflict detection and conversion

- Source selection

TRUSTED AGENTS

INPUT PROCESSING

* Automatic web wrapping

- Semi-structured text

-Multi-source query plan and execution

Web Pages

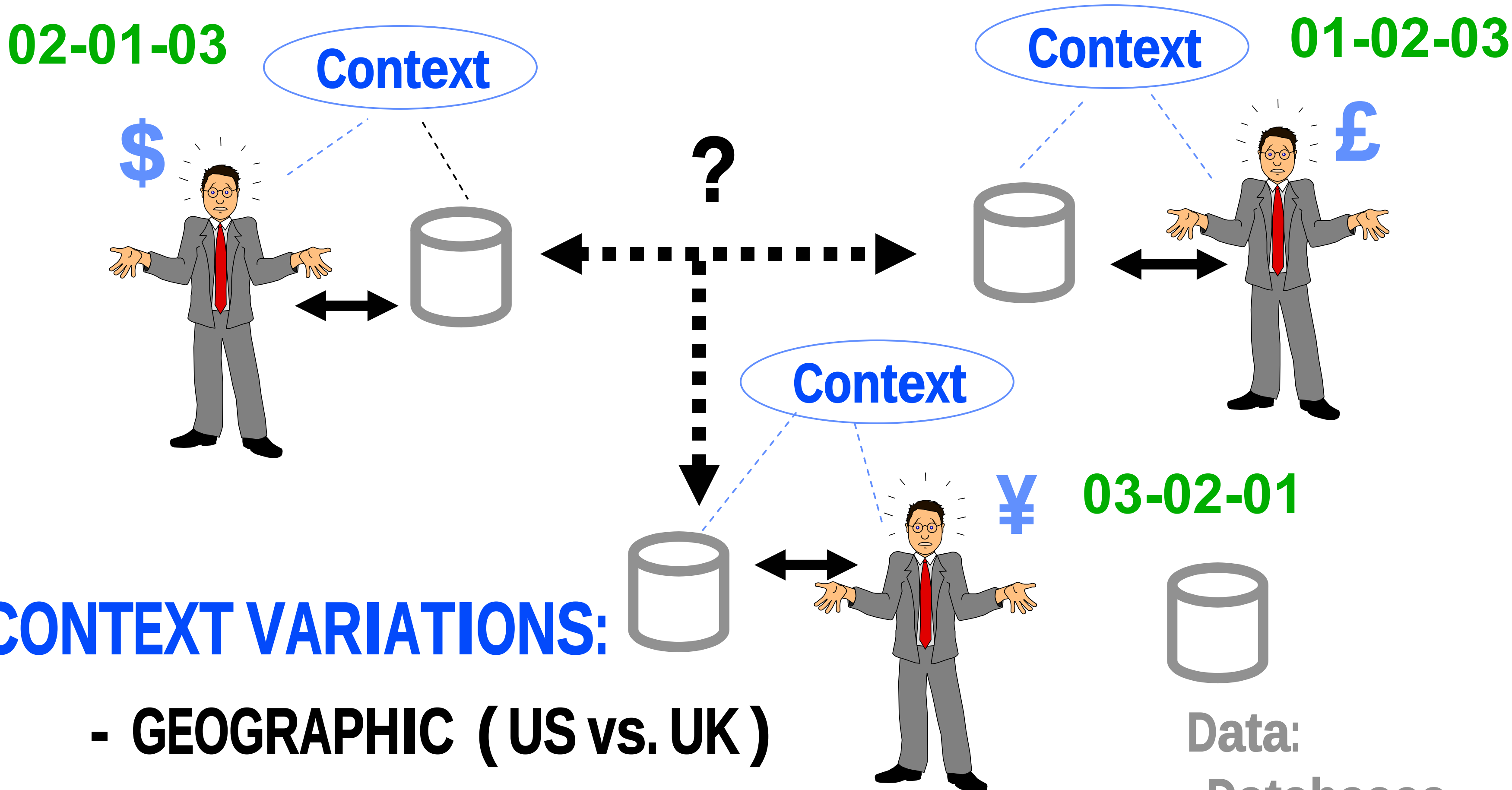
Sources

Data bases

APPLICATIONS: Financial services, electronic commerce

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

Role Of Context



- **CONTEXT VARIATIONS:**
  - GEOGRAPHIC (US vs. UK)
  - FUNCTIONAL (MARKETING vs. FINANCE)
  - ORGANIZATIONAL (CITIBANK vs. CHASE)

Data:
- Databases
- Web data
- E-mail

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

Types of Context

Surface **Ontological**

Temporal

	Example	Temporal
Surface	Currency: $ vs € Scale factor: 1 vs 1000	Francs before 2000, € thereafter
Ontological	Revenue: Includes vs excludes interest	Revenue: Excludes interest before 1994 but incl. thereafter

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

COIN - Summary

- Tremendous opportunity to gather and integrate information from many diverse sources
- But … need to overcome many context challenges
- Context-type "metadata" plays a critical role
- COIN technology can be an important aid for semantically meaningful information integration:
  - **- Scalable**
  - **- Extensible**
  - **- Application Domain Merging**
  - **- Reuse and extension of ontologies and contexts**

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

COIN ONTOLOGICAL MODEL I

Main components are –

- Semantic Types
- Attributes
- Modifiers
- Contexts
- Conversion Functions
- Elevated Relations

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

COIN ONTOLOGICAL MODEL II

Attributes and Modifiers

- **Properties of semantic types specified by attributes and modifiers.**

Attributes define the state of object or relationship between objects.

Example: "city" is an attribute of semantic type "location"

Modifiers are specialized attributes that take on different values in different contexts.

Example: "tempUnit" is a modifier of semantic type "temperature". It has value "celsius" in metric ctxt and "fahrenheit" in imperial ctxt.

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

COIN ONTOLOGICAL MODEL III

Contexts and Elevation Axioms

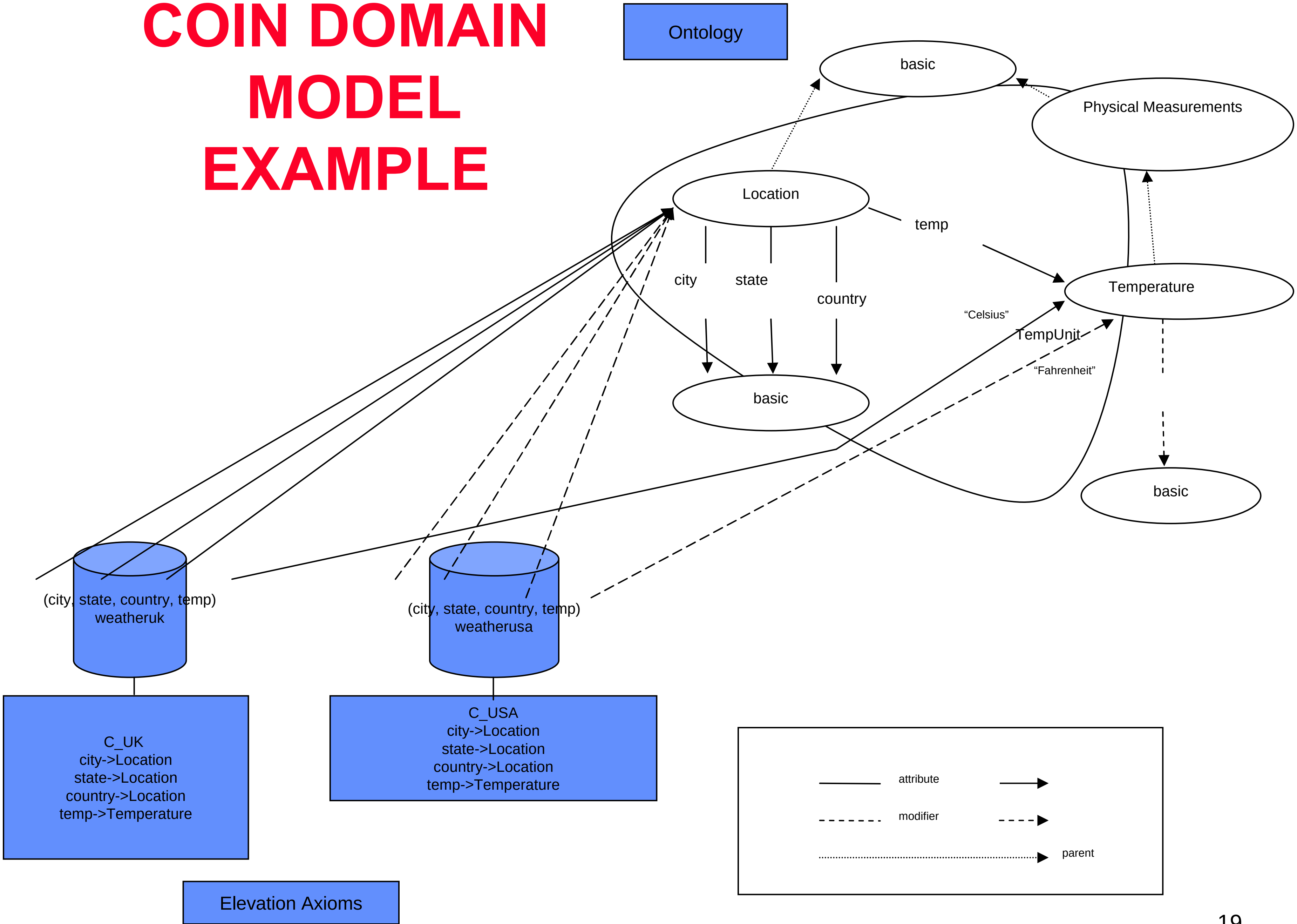
Elevation Axioms relate the source relations to the domain model. Each primitive source database relation is mapped to a semantic relation.

Semantic Relations are obtained by mapping each primitive column to a semantic type in a particular context.

Example: In the context metric units, the primitive source relation column "city" is mapped to the semantic type "location" in the domain model.

Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.

COIN DOMAIN MODEL EXAMPLE



Copyright 2004 by Sumit Bhansali, Benjamin Grosof, and Stuart Madnick. All Rights Reserved.