INTRODUCTION

Nowadays, data management on the World Wide Web needs to consider very large knowledge databases (KDB). The larger is a KDB, the smaller the possibility of being consistent. Consistency in checking algorithms and systems fails to analyze very large KDBs, and so many have to work every day with inconsistent information.

Database revision—transformation of the KDB into another, consistent database—is a solution to this inconsistency, but the task is computationally untractable. Paraconsistent logics are also a useful option to work with inconsistent databases. These logics work on inconsistent KDBs but prohibit nondesired inferences. From a philosophical (logical) point of view, the paraconsistent reasoning is a need that the self human discourse practices. From a computational, logical point of view, we need to design logical formalisms that allow us to extract useful information from an inconsistent database, taking into account diverse aspects of the semantics that are “attached” to deductive databases reasoning (see Table 1). The arrival of the semantic web (SW) will force the database users to work with a KDB that is expressed by logic formulas with higher syntactic complexity than are classic logic databases.

BACKGROUND

Logic databases are based on the formalisms of first order logic (FOL); thus, they inherit a classical semantics that is based on models. Also, they can be interpreted within a proof–theoretic approach to logical consequence from the logic programming paradigm (Lloyd, 1987). The extended database semantics paradigm is developed to lay before the foundations of query-answering tasks and related questions (see Minker, 1999), but its aim is not to deal with inconsistencies. The data cleaning task may involve—in the framework of repairing logic databases—logical reasoning and automated theorem proving (Boskovitz, Goré, & Hegland, 2003).

On the other hand, new paradigms, such as SW, need new formalisms to reason about data. Description logics (DL) provide logic systems based on objects, concepts, and relationships, with which we can construct new concepts and relations for reasoning (Baader, Calvanese, McGuiness, Nardi, & Patel-Schneider, 2003). Formally, DL are a subset of FOL, and the classical problems on consistency remains, but several sublogics of DL provide nice algorithms for reasoning services. The ontology Web language (OWL; its DL-sublanguage) is a description logic designed for automated reasoning, not only designed for the classical ask–tell paradigm. With languages such as OWL, ontologies exceed their traditional aspects (e.g., taxonomies and dictionaries) to be essential in frameworks as data integration.

The classical notion of inconsistency in databases mainly deals with the violation of integrity constraints. This notion must be expanded because of the new notion of logic databases in SW, in which ontologies and data both play the same role in knowledge management. Therefore, there are several sources of inconsistency (see Table 2). This role is not only limited to the database but also includes the verification and validation task of knowledge-based systems (Bench-Capon, 2001). Inconsistency arises in the initial steps of ontology building due to several reasons and not only by the updating of data. In general, the repair of a logic database involves the study

Table 1. Semantics aspects to consider in logic databases

- Classical semantics for FOL
- Extended semantics for databases
- Reiter’s formalization of databases (Reiter, 1984). Closed world assumption
- Relations among a KDB, queries and integrity constraints
- Expressive power of recursive definitions
- Consistency checking versus intentional part of the KDB
- Multivalued semantics
- Contextualized semantics for ontologies or data
Table 2. A list of sources of inconsistencies from the practical knowledge management

<table>
<thead>
<tr>
<th>Source</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dirty data</td>
<td>Some kinds of data dirtiness give rise to fail of integrity constraints</td>
</tr>
<tr>
<td>Neglected development of the intentional database</td>
<td>The development of the intentional component part of the database produces an inconsistent theory (the ontology, in the SW paradigm. See, e.g., Backlawski, Kokar, Waldinger, &amp; Kogut, 2002).</td>
</tr>
<tr>
<td>Logical interpretation in data integration</td>
<td>The design of a data integration system—to provide uniform access to multiple and heterogeneous information sources—needs of query reformulation, ontology mapping, or integration and, in general, logical interpretation.</td>
</tr>
<tr>
<td>Procedural incompleteness</td>
<td>Incomplete query-answering algorithms do not produce any witness for some integrity constraint of existential character.</td>
</tr>
<tr>
<td>Conflict information in data integration</td>
<td>Special case in data integration: The information that is received from different consistent resources is inconsistent.</td>
</tr>
<tr>
<td>Deficient specification of the ontology language</td>
<td>The specification of the language for ontology representation is inconsistent (Fikes, McGuinness, &amp; Waldinger, 2002).</td>
</tr>
<tr>
<td>Inadequate data cleaning</td>
<td>Some criteria to take decisions in data cleaning make the KDB inconsistent.</td>
</tr>
<tr>
<td>Deficient maintenance of KDB</td>
<td>The kind of the selected method for preserving consistency is not robust under every sort of updates.</td>
</tr>
<tr>
<td>Wrong data mining</td>
<td>The output of data mining systems does not satisfy integrity constraints or ontology requirements. In multiagent data mining, the different outputs lead to a problem of data integration.</td>
</tr>
<tr>
<td>Expressiveness clashes with model theory</td>
<td>The logical syntax/semantics (from deductive database paradigm) does not allow new features to be used in the usual knowledge representation in the SW. This absence implies messy definitions that may be incorrect.</td>
</tr>
<tr>
<td>Bad design of the common knowledge shared by different users</td>
<td>The intentional component part does not describe the users intended requirements. The logical consistent KDB does not fit with users’s beliefs. Thus, new updates may produce inconsistencies.</td>
</tr>
<tr>
<td>Deficient ontology learning</td>
<td>The ontology acquisition is a tedious task that the user tends to finish before he or she thinks as advisable. A poor ontology associated to consistent data may produce inconsistency.</td>
</tr>
</tbody>
</table>

of the soundness and perhaps completeness (i.e., the method output’s only correct solutions and all the relevant solutions). Semantics would support reasoning services such as self-consistency, checking the relations between concepts (as subsumption), and classification of objects according to the ontology.

Systems exist in which both paradigms, classical and SW logic databases, are conciliated under the extension of the former (see, e.g., Pan & Heflin, 2003). However, the relation between DL and database models may not be fruitfully formalized because of the limited expressiveness of the DL system selected to make the reasoning feasible (see chapter 4 in Baader et al., 2003).

INCONSISTENCY HANDLING

Solutions that are suggested to work in the presence of inconsistencies can be classified according to different views (see Table 3, where several references appear). The first aspect—and maybe the most important—is the compatibility between the original semantics of the KDB source and the logical formalism selected to handle inconsistency. From this point of view there exist paraconsistent logics that limit the inference power of FOL to avoid nondesired answers and also modal logics for representing different aspects of the information sources. These approaches manage semantics that are essentially different from the semantics of KDBs. On the other side we can find methods that classify, order, or both, interesting subsets according to the original semantics of the KDB, such as the argumentative approach or the integration of data by fusion rules, but they do not repair the KDB. Other methods also exist that propose how the KDB should be revised (e.g., integrity constraints of the extensional database). However, it is necessary to point out that the automated knowledge revision is an essentially different task in the case of ontologies, because the ontology source represents a key organisation of the knowledge of the owner and, as in every logical theory, minor changes may produce unexpected and dangerous anomalies.

Another point of view concerns the share of the KDB that is repaired when an anomaly is found. According to this, the methods based on arguments mentioned earlier can be used to repair only the anomalous argument. Due to the high complexity of consistency checking algorithms, to preserve consistency under updates is a better option than repairing. In the case of evolving ontologies, new systems such as KAON infrastructure are needed (for more information, see http://kaon.semanticweb.org/kaon).

There are methods dealing with the enforcement of consistent answers (i.e., answers that satisfies integrity constraints) from inconsistent databases; it is done by
Related Content

**Understanding the Role of Use Cases in UML: A Review and Research Agenda**
[www.igi-global.com/article/understanding-role-use-cases-uml/3256?camid=4v1a](www.igi-global.com/article/understanding-role-use-cases-uml/3256?camid=4v1a)

**A Cost Model for Path Expressions in Object-Oriented Queries**
[www.igi-global.com/article/cost-model-path-expressions-object/51166?camid=4v1a](www.igi-global.com/article/cost-model-path-expressions-object/51166?camid=4v1a)

**Assessing Business Process Modeling Languages Using a Generic Quality Framework**
Anna Gunhild Nysetvold and John Krogstie (2006). *Advanced Topics in Database Research, Volume 5* (pp. 79-93).
[www.igi-global.com/chapter/assessing-business-process-modeling-languages/4387?camid=4v1a](www.igi-global.com/chapter/assessing-business-process-modeling-languages/4387?camid=4v1a)

**Database Benchmarks**
[www.igi-global.com/chapter/database-benchmarks/7967?camid=4v1a](www.igi-global.com/chapter/database-benchmarks/7967?camid=4v1a)