Chapter VI

Consistent Queries Over Databases With Integrity Constraints

Sergio Greco and Ester Zumpano
Università della Calabria, Italy

INTRODUCTION

Integrity constraints represent an important source of information about the real world. They are usually used to define constraints on data (functional dependencies, inclusion dependencies, etc.). Nowadays integrity constraints have a wide applicability in several contexts such as semantic query optimization, cooperative query answering, database integration and view update.

Often databases may be inconsistent with respect to integrity constraints, that is, one or more integrity constraints are not satisfied. This may happen, for instance, when the database is obtained from the integration of different information sources. The integration of knowledge from multiple sources is an important aspect in several areas such as data warehousing, database integration, automated reasoning systems and active reactive databases.

Since the satisfaction of integrity constraints cannot generally be guaranteed, in the evaluation of queries, we must compute answers which are consistent with the integrity constraints. Example 1 shows a case of inconsistency.

Example 1: Consider the following database schema consisting of the single binary relation Teaches (Course, Professor) where the attribute Course is a key for the relation. Assume there are two different instances for the relations Teaches as reported in Figure 1.
The two instances satisfy the constraint that Course is a key but, from their union we derive a relation which does not satisfy the constraint since there are two distinct tuples with the same value for the attribute Course.

In the integration of two conflicting databases simple solutions could be based on the definition of preference criteria such as a partial order on the source information or a majority criteria (Lin and Mendelzon, 1996). However, these solutions are not generally satisfactory and more useful solutions are those based on 1) the computation of repairs for the database, 2) the computation of consistent answers (Arenas et al., 1999).

The computation of repairs is based on the definition of minimal sets of insertion and deletion operations so that the resulting database satisfies all constraints. The computation of consistent answers is based on the identification of tuples satisfying integrity constraints and on the selection of tuples matching the goal.

For instance, for the integrated database of Example 1, we have two alternative repairs consisting of the deletion of one of the tuples (c2,p2) and (c2,p3). The consistent answer to a query over the relation Teaches contains the unique tuple (c1,p1) so that we don’t know which professor teaches course c2.

Therefore, it is very important, in the presence of inconsistent data, not only to compute the set of consistent answers, but also to know which facts are unknown and if there are possible repairs for the database. In our approach it is possible to compute the tuples that are consistent with the integrity constraints and answer queries by considering as true facts–those contained in every repaired database, false facts–those that are not contained in any repaired database and unknown–the remaining facts.

**Example 2.** Consider the integrated relation of Example 1 containing the tuples (c1, p1), (c2, p2) and (c2, p3). The database is inconsistent and there are two possible repairs which make it consistent: \( R1 = (\emptyset, \{\text{Teaches}(c2, p2)\}) \) and \( R2=(\emptyset, \{\text{Teaches}(c2, p3)\}) \) which delete, respectively, the tuples (c2, p2) and (c2, p3), from the relation Teaches. The set of consistent tuples in the relation Teaches consists of the singleton (c1, p1).

This chapter illustrates recent techniques for computing consistent answers and repairs for possibly inconsistent databases.
Related Content

FSQL: A Fuzzy SQL for Fuzzy Databases
www.irma-international.org/chapter/fsql-fuzzy-sql-fuzzy-databases/18764/

On Non-Constrained, Constrained and Mandatory Many-to-Many Relationship Types
www.irma-international.org/article/non-constrained-constrained-mandatory-many/51113/

Antecedents of the Closeness of Human-Avatar Relationships in a Virtual World
www.irma-international.org/chapter/antecedents-closeness-human-avatar-relationships/63666/

The Effect of User View Characteristics on Database Design Performance of Novices
www.irma-international.org/article/effect-user-view-characteristics-database/51182/

Maintaining Mappings between Conceptual Models and Relational Schemas
www.irma-international.org/chapter/maintaining-mappings-between-conceptual-models/63669/