Chapter XLV
Querying and Integrating P2P Deductive Databases

Luciano Caroprese
University of Calabria, Italy

Sergio Greco
University of Calabria, Italy

Ester Zumpano
*University of Calabria, Italy*

INTRODUCTION

Recently, there have been several proposals that consider the integration of information and the computation of queries in an open-ended network of distributed peers (Bernstein, Giunchiglia, Kementsietsidis, Mylopulos, Serafini, & Zaihrayen, 2002; Calvanese, De Giacomo, Lenzerini, & Rosati, 2004; Franconi, Kuper, Lopatenko, & Zaihrayeu, 2003) as well as the problem of schema mediation and query optimization in P2P (peer-to-peer) environments (Gribble, Halevy, Ives, Rodrig, & Suciu, 2001; Halevy, Ives, Suciu, & Tatarinov, 2003; Madhavan & Halevy, 2003; Tatarinov & Halevy, 2004).

Generally, peers can both provide or consume data and the only information a peer participating in a P2P system has is about neighbors, that is, information about the peers that are reachable and can provide data of interest. More specifically, each peer joining a P2P system exhibits a set of mapping rules, in other words, a set of semantic correspondences to a set of peers that are already part of the system (neighbors). Thus, in a P2P system, the entry of a new source, or peer, is extremely simple as it just requires the definition of the mapping rules. By using mapping rules as soon as it enters the system, a peer can participate and access all data available in its neighborhood, and through its neighborhood it becomes accessible to all the other peers in the system.

As stated before, the problem of integrating and querying databases in P2P environments has been investigated in Calvanese, De Giacomo, Lenzerini, et al. (2004) and Franconi et al. (2003). In both works, peers are modeled as autonomous agents that can export only data belonging to their knowledge, that is, data that are true in all possible

Figure 1. A P2P system
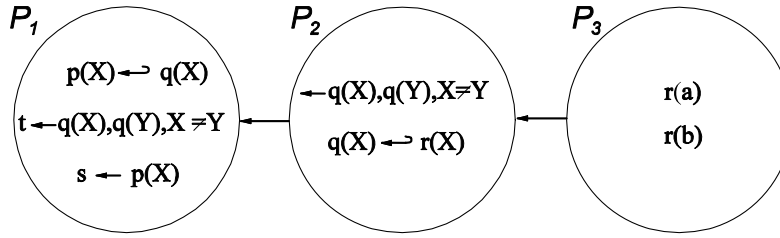
scenarios (models). In Calvanese, De Giacomo, Lenzerini, et al., new semantics for a P2P system, based on epistemic logic, is proposed.

The work also shows that the semantics is more suitable than traditional semantics based on FOL (first-order logic) and proposes a sound, complete, and terminating procedure that returns certain answers to a query submitted to a peer. In Franconi et al. (2003), a characterization of P2P database systems and a model-theoretic semantics dealing with inconsistent peers is proposed. The basic idea is that if a peer does not have models, all (ground) queries submitted to the peer are true (i.e., are true with respect to all models). Thus, if some databases are inconsistent, it does not mean that the entire system is inconsistent.

MOTIVATION

The motivation of this work stems from the observation that previously proposed approaches result in not being sound with respect to query answering when some peer is inconsistent.

Example 1. Consider the P2P system depicted in Figure 1 consisting of three peers P_1, P_2, and P_3 where

- P_3 contains two atoms, $r(a)$ and $r(b)$,
- P_2 imports data from P_3 using the (mapping) rule $q(X) \leftharpoondown r(X)$ (observe that a special syntax is used for mapping rules). Moreover, imported atoms must satisfy the constraint $\leftarrow q(X), q(Y), X \neq Y$ stating that the relation q may contain at most one tuple, and
- P_1 imports data from P_2 using the (mapping) rule $p(X) \leftharpoondown q(X)$. P_1 also contains the rules

$s \leftarrow p(X)$ stating that s is true if the relation p contains at least one tuple, and $t \leftarrow p(X)$, $p(Y), X \neq Y$, stating that t is true if the relation p contains at least two distinct tuples.

The intuition is that, with $r(a)$ and $r(b)$ being true in P_3, either $q(a)$ or $q(b)$ could be imported in P_2 (but not both, otherwise the integrity constraint is violated) and, consequently, only one tuple is imported in the relation p of the peer P_1. Note that whatever the derivation is in P_2, s is derived in P_1 while t is not derived. Therefore, the atoms s and t are, respectively, true and false in P_1. It is worth noting that the approaches above mentioned do not capture such a semantics. Indeed, the epistemic semantics proposed in Calvanese, De Giacomo, Lenzerini, et al. (2004) states that both the atoms $q(a)$ and $q(b)$ are imported in the peer P_2, which becomes inconsistent. In this case, the semantics assumes that the whole P2P system is inconsistent and every atom is true as it belongs to all minimal models. Consequently, t and s are true. The semantics proposed in Franconi et al. (2003) assumes that only P_2 is inconsistent as it has no model. Thus, as the atoms $q(a)$ and $q(b)$ are true in P_2 (they belong to all models of P_2), then the atoms $p(a)$ and $p(b)$ can be derived in P_1 and finally t and s are true.

The idea we propose in this work consists of importing in each peer maximal sets of atoms not violating integrity constraints.

BACKGROUND

We assume there are finite sets of predicate symbols, constants, and variables. A term is either

5 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/querying-integrating-p2p-deductive-databases/20726

Related Content

Information Technology Needs for Public Service Delivery in the Digital Era
 (2019). *Information Systems Strategic Planning for Public Service Delivery in the Digital Era (pp. 227-268).*
www.irma-international.org/chapter/information-technology-needs-for-public-service-delivery-in-the-digital-era/233409

Metaschemas for ER, ORM and UML Data Models: A Comparison
Terry Halpin (2002). *Journal of Database Management (pp. 20-30).*
www.irma-international.org/article/metaschemas-orm-uml-data-models/3277

The Second Wave ERP Market: An Australian Viewpoint
Andrew Steinand Paul Hawking (2003). *ERP & Data Warehousing in Organizations: Issues and Challenges (pp. 72-88).*
www.irma-international.org/chapter/second-wave-erp-market/18555

On the Computation of Recursion in Relational Databases
Yangjun Chen (2003). *Effective Databases for Text & Document Management (pp. 263-277).*
www.irma-international.org/chapter/computation-recursion-relational-databases/9215

Modeling Motion: Building Blocks of a Motion Database
Roy Gelbardand Israel Spiegler (2004). *Advanced Topics in Database Research, Volume 3 (pp. 327-345).*
www.irma-international.org/chapter/modeling-motion-building-blocks-motion/4366