

View-based Query Processing for Semistructured Data

Diego Calvanese
Dipartimento di Informatica e Sistemistica
Università di Roma “La Sapienza”
calvanese@dis.uniroma1.it

Abstract

View-based query processing is the problem of processing queries over a database based on a set of materialized views, rather than on the raw data in the database. It is a fundamental problem in several settings, such as data warehousing, mobile computing, query optimization, and data integration, and comes about in different forms. The basic task is view-based query answering, which amounts to directly computing the answer to a query by exploiting, besides the query and the view definitions, also the data in the materialized views. A second form of view-based query processing is query rewriting, where we are given a query and a set of view definitions and the task is to reformulate the query in terms of the views. The rewritten query can then be directly evaluated over the materialized views. Finally, query containment and view-based query containment are the essential tasks for evaluating the quality of the obtained rewritings, by allowing one to compare queries and rewritings. In the talk we discuss and analyze basic techniques for the different forms of view-based query processing in the setting of semistructured data, which are an abstraction for data on the web and XML. In such a setting flexible mechanisms for navigating the data are required, and regular path queries are considered the basic building blocks of more powerful query languages. The techniques we present apply to variants of regular path queries, and are based on automata, exploiting also a strong connection between view-based query answering and constraint satisfaction.