Explicit Representation of Constrained Schema Mappings for Mediated Data Integration

C. Altenschmidt, J. Biskup
Fachbereich Informatik, Universität Dortmund, D–44221 Dortmund
{altensch|biskup}@ls6.cs.uni-dortmund.de

Abstract

In an environment of heterogeneous data sources it may be necessary to integrate these in order to provide a single global view to the data. Nowadays this problem is solved by mediators, which are tolerant not only to heterogeneity of the sources, but also of their availability and of structural changes. For some mediation problems it is reasonable to assume the existence of a fixed structured target schema as the global view. In these cases, mismatches of target concepts and source concepts can occur, which make it impossible for a mediator to interpret the data correctly and completely at the same time. We will show how to enforce correct interpretations by imposing constraints on the mappings between the target schema and the source schemas. The strength of such constraints can be decreased in a flexible and controlled way, for the sake of exploiting more sources, and at the cost of potentially loosing assurance in correctness. Additionally, we treat interpretation completeness of sources. A careful specification of data structures and algorithms allows for using mappings of this kind in a generic mediation system. The data structures represent mappings explicitly by linking structural descriptions of source data to the target schema expressed in an object oriented data model.