Potential Applications of Description Logics for Data Integration

Richard Hull
Bell Labs, Lucent Technologies
hull@research.bell-labs.com

This talk discusses the research challenges raised by two potential applications of Description Logics (DLs) for data integration.

The first application concerns the generation of mediators, i.e., software components that support a read-only integrated view of selected data from multiple databases. A mediator specification consists in a family of "local" database schemas, a single integrated "global" database schema, and descriptions of the relationship between the local and global schemas, e.g., by defining the global schema as a (multi-relation) view over selected portions of the local databases. In the current state-of-the-art, mediators can be automatically generated from such mediator specifications. We study the problem of generating a mediator specification. This might be possible if ontologies for various application domains were expressed using a DL, and the semantic relationships between databases and their relevant ontologies were specified.

The second application concerns maintaining consistency when updating multiple databases. Previous work in this area has focused on the use of essentially local rules that enforce individual interdatabase constraints, and a distributed rule execution engine that ensures the constraints will be maintained. A significant issue in that work is the potential for cycles between the local rules. We propose to express interdatabase constraints using a DL. Information about the DL specification is then captured in a graph-based data structure. Depending on the DL used, construction of the graph may require more than polynomial time. But once the graph is constructed, maintenance of the constraints may lie within polynomial time.