

# DIRECT: A Query Facility for Multiple Databases

Ulla Merz  
Roger King  
Campus Box 430  
University of Colorado  
Boulder, Colorado 80309

DIRECT<sup>1</sup> is an interactive software system that provides a uniform interface for retrieving data from heterogeneous databases. This is done by specifying a multi-database query. The databases and their definitions can differ in their data models, names, types and encoded values. Multi-database queries integrate data values based on semantically equivalent data elements that are stored in the different databases.

This research effort has focused on: (1) Providing flexibility to adjust to changing data retrieval requirements (2) Accommodating different data models and (3) Focusing on the user's role in defining the solution. Flexibility is achieved by integrating the data values through a multi-database query rather than the database schemas using a semantic data model. Allowing different database schemas and their data models to coexist is at the center of this solution. No global schema is created. DIRECT provides direct-manipulation interaction techniques to reduce the user's effort of creating a multi-database query and assists users in identifying semantically equivalent data elements.

The two major components of DIRECT are a browser for the different data definitions and a graphical, direct-manipulations query editor for specifying a multi-database query. The purpose of the browser is to assist users in identifying compatible data elements. It stores the data names and constructs of the individual databases and makes them accessible and visible to the user. The browser unifies the different data definitions visually through a standard presentation. DIRECT provides a query editor for retrieving data as needed. Only the data required for a particular retrieval request has to be integrated. A user selects the names of the data elements that determine the query result and chooses the query functions that relate semantically equivalent data elements. DIRECT unifies the different local query lan-

guages by representing a multi-database query as a data definition diagram.

Given the design considerations and their realization DIRECT supports the following features. *Flexibility*: Queries only integrate data values that are needed for a retrieval request. Queries may contain arithmetic expressions to convert data values and define views that provide alternative schema representations. *Autonomy*: Coexistence of the different database schemas and their data models preserves the autonomy of the individual databases. Existing applications do not have to be changed. *Extensibility*: Coexistence is a prerequisite for extensibility. Existing database schemas do not have to be mapped into the constructs and notation of a semantic data model. *Uniformity*: A uniform interface is made possible through a uniform, graphical presentation of the schema descriptions and the query specification.

DIRECT has been implemented on an IBM RISC System/6000. The graphical user interface has been developed using X Windows, the Motif toolkit and the user interface builder UIM/X<sup>2</sup>.

The video sets the stage by describing the problem addressed by DIRECT and by providing an overview of the proposed solution. This is followed by an example that illustrates the use of DIRECT to retrieve data stored in IMS and DB2 databases. The example uses existing, operational databases that are part of an automated import and export system. The scenario shows the different steps the user takes to obtain a report that lists all imported parts and their descriptions.

---

<sup>1</sup>This work was funded by an IBM Shared University Research Grant

---

<sup>2</sup>UIM/X is a product of Visual Edge Software Limited.