QUERY BY DIAGRAM: A GRAPHICAL ENVIRONMENT FOR QUERYING DATABASES

Tiziana Catarci & Giuseppe Santucci

Dipartimento di Informatica e Sistemistica, Universita' degli Studi di Roma "La Sapienza"
Via Salaria 113, 00198 Rome, Italy
e-mail: catarci@infokit.dis.uniroma1.it, santucci@infokit.dis.uniroma1.it

1. Motivation

A central development in the database area concerns tools that allow non expert users to understand and easily extract information from a database. Fourth generation query languages, although non-procedural, are not friendly enough for a casual user who must know both the logical structure of the database and the syntax and semantics of the DBMS query language. Instead, recently proposed visual systems which allow a user to extract information by means of interactive graphical commands, have not yet been able to combine ease of use and high expressive power.

2. Prototype goals and functionalities

The QBD* (Query By Diagram) system tries to overcome the above limitations. It balances a high expressive power with a noticeable facility of use. Concerning the expressive power, it has been proved to be relationally complete; moreover it includes a significant class of recursive queries (transitive closure) and handles an extension of the relational algebra set-oriented operators that we call *generalized set-oriented operators*. Through those operators it is possible to compute set-oriented operations on any pair of relational tables sharing the same identifier. The easy of use has been reached through a fully graphical environment based on a diagrammatic representation of the Entity Relationship schema describing the underlying database. In that environment the user interacts with the system mainly with a mouse-like device, using the keyboard only when necessary. To increase the friendliness of the system a set of *helping facilities* has been devised. Among them, the most important ones are top down browsing and schema transformation. The former allows for browsing a library of top-down refinements documenting an ER schema at different levels of abstraction. While browsing through the top-down schemata, the user can easily locate the fundamental concepts and links. Top-down browsing may be also used to locate the subschema of interest. The schema transformation allows for converting a schema into a new version, that is "closer" to the query. As an example it is possible to contract a path in the schema into a new concept or to replace an ISA relationship with the child entity.

A systematic study has been performed to establish the effectiveness of the QBD* environment. The study involved 90 non experts in computer science (undergraduate students, employees, and professional persons). Two tutorials have been designed, introducing the SQL language and the QBD* environment, respectively. Moreover, a metric for measuring the complexity of the query has been studied, based on the idea of associating a weight to each relational algebra operator. First, the people read the tutorials, then they were asked to formulate a set of increasing complexity queries. During the experiment, a measure of both the answering time and the mistakes was carried out. The comparison between the SQL queries and the QBD* queries shows that using the QBD* environment both the answering time and the number of mistakes decrease.

The present prototype allows for retrieving and manipulating a stored ER schema and for graphically expressing a query on it. It is possible to browse the top-down refinement chain of the current schema, to extract a subschema, modify a concept or a portion of the schema, and save the resulting view of the schema. All types of queries can be expressed in the visual query language (including recursive queries). All the activities, but the schema transformation, can be fully intermixed, resulting in a very flexible user interaction. A metaschema, describing the structure of the whole set of the available schemata, can be queried with the same mechanisms. Such a meta-querying allows the user to quickly locate the schema s/he is looking for.

3. Current implementation state

An academic version of the QBD* system was developed in C (Microsoft 7.0) under MS-DOS environment. It includes all the functionalities described above. An engineered version of QBD*, called EQBD was developed under Microsoft Windows environment by an Italian Software House and it is actually a commercial product.

References

[CaSa88] T.Catarci, G.Santucci - GRASP : A Complete Graphical Conceptual Language for Definition and Manipulation of Statistical Databases. - Proc. of COMPSTAT-88 Conference, Copenaghen, 1988.

[CaSa88] T.Catarci, G.Santucci - QUERY BY DIAGRAM: A Graphic Query System - Proc. of 7th Conference on Entity-Relationship Approach, Roma, 1988.

[AnCaSa89] M.Angelaccio, T.Catarci, G.Santucci - QBD*: A Graphical Query Language with Recursion - Proc. of 3rd Human-Computer Interaction Conference, Boston, 1989.

[AnCaSa89] M.Angelaccio, T.Catarci, G.Santucci - QBD*: A Fully Visual System for E-R Oriented Databases - Proc. of Workshop on Visual Language 1989, Rome, 1989.

[AnCaSa90a] M.Angelaccio, T.Catarci, G.Santucci - QBD*: A Graphical Query Language with Recursion - *IEEE Transactions on Software Engineering*, vol. 16, No. 10, pp. 1150-1163, 1990

[AnCaSa90b] M.Angelaccio, T.Catarci, G.Santucci - QBD*: A Fully Visual Query System - *Journal of Visual Languages and Computing*, Vol. 1, No. 2 , pp. 255-273, 1990.

[SaSo93] G.Santucci, P. A. Sottile - Query By Diagram: a Visual Environment for Querying Databases - *Software Practice and Experience*, Vol. 23, No. 3, 1993.