

The INA: A Simple Query Language with Only Attribute Names

A Video Presentation

Bruce I Blum and Ralph D Semmel

Applied Physics Laboratory
Laurel, MD 20707-6099
(301) 953-6235
bib@apl.vax.jhuapl.edu

Current query languages, such as SQL, assume that the user is familiar with the database schema including the attribute names, types, and relation associations. When a user has imperfect knowledge of this information (or when he balks at the data-processing orientation of the required statements), he normally asks an experienced analyst to perform his ad hoc query. The Intelligent Navigational Assistant (INA) was developed for the U S Army as a prototype query tool that permits the users to specify requests using only domain terms familiar to them. Once a request is made, it is converted into SQL for processing.^{1,2}

To facilitate query formulation, the INA supports an interface that allows the user to identify attributes without relation associations (i.e., it treats the data model as a universal relation). Because an attribute may appear in many relations, one of the principal tasks of the INA is the determination of the appropriate relation bindings. To aid in the selection of terms, the INA maintains a user vocabulary and provides facilities for browsing the vocabulary and examining term definitions. Thus, the INA has two primary functions: it provides an easy-to-use interface for query definition, and it converts a request into SQL.

The INA prototype has been implemented as a PC-resident knowledge-based system linked to a host-based DBMS. Its knowledge base is the logical schema of the target database, and the query transformation relies on the dependencies implicit in that schema. Supporting the knowledge-processing functions are the query definition interface, various tools to manage the target data model description, and facilities for communicating with other computers. The system was developed using TEDIUM[®],³ and the user interface and query resolution mechanism are extensions of earlier work with Tequila⁴ (which accessed the semantically-rich TEDIUM[®] data model).

Work on the INA began in 1987 and was terminated in 1988. The system was demonstrated as a prototype with an Army-supplied logical model consisting of approximately 40 relations and 200 attributes. After query definition, reformulation, and user acceptance, the SQL queries were submitted to the mainframe for processing. In those tests, the INA often produced better queries than those manually coded by analysts. The INA currently is undergoing a beta test with a much larger database schema. Its algorithms are described in reference 5, and reference 3 contains details regarding its implementation and semantic data model. Current research includes the development of improved query resolution algorithms based on an enriched semantic data model.

REFERENCES

- 1 Blum, B I, S D Diamond, M G Hammond, M E Perkins, and R D Semmel, An Intelligent Navigational Assistant for a Decision Resource Database, *Expert Systems in Government*, 1987, pp 19-25
- 2 Semmel, R D, B I Blum and S D Diamond, The Data Model as a Knowledge Base, in M L Emrich, A R Sadlowe and L F Arrowood (eds), *Expert Systems and Advanced Data Processing, NARDAC Washington-ORNL-DSRD Conference*, North Holland, 1988, pp 121-127
- 3 Blum, B I, *TEDIUM and the Software Process*, MIT Press, Cambridge, MA, 1990
- 4 Blum, B I, TEQUILA. An Intoxicating Query System, *14th MUMPS Users' Group*, 1985, p 87-94
- 5 Semmel, R D, A Knowledge-Based Approach to Query Formulation. Some Limits Based on the Logical Schema, Applied Physics Laboratory Research Center Report, RMI-90-004, January, 1990