Panel on Time and Databases

Chairpersons
Gad Ariav, Jim Clifford, Matthias Jarke
Graduate School of Business Administration
New York University

1. Introduction

Nowhere is the lag between database theory and database practice more apparent than with respect to the issue of time. Beginning long before the advent of database management systems, econometricians, industrial applications programmers and systems staff have been building programs and procedures for handling time-series data. It is one of the failures of the "three great data models" that none has addressed this issue. With the promise of several new technologies that will make feasible the storage of orders of magnitude more data on a single type of storage medium, it behooves us in the database research community to address the issue of databases and time. We have the opportunity to explore theoretical issues of time and its database properties and to develop disciplined models for its use, so as to guide and enhance practical attempts to deal with these issues, and to avoid merely ad hoc solutions to the growing demand for more time-oriented data and information.

We will use the generic term "historical database" (HDB) to refer to any database that has a built-in notion of time and its properties, for the purpose of archival and retrieval of data through time. There have been two trends to recent research in this area. In one, the focus has been on databases that record historical data as it changes over time, while the other has investigated systems that maintain an explicit notion of transitions from one database state to the next. [Bolour 1982] provides a useful survey of some of the literature in the field.

The panel addresses both these aspects of historical databases and the various research questions that such databases raise. For the purposes of discussion we might group these issues into the following four categories: time and database theory, the implementation of historical databases, the user interface to historical databases, and applications of historical databases. Each of these issues is briefly discussed below.

2. Time and Database Theory.

There have been several attempts recently to define theoretical or conceptual models for an historical database [Bubenko 1977], [Clifford & Warren 1981], [Klopprogge 1981]. These works have raised a number of important and interesting issues concerning a theory of time and its place in a database model.

Temporal Database Models: What is an appropriate formalism for an HDB model? For instance, is a first-order logic model appropriate, or should a temporal logic serve as a paradigm? How does the issue of time impact the already existing database models? Should we attempt to reformulate these models with an added temporal component, or should we rather be developing entirely new models for HDBs? Moreover, how does the issue of time relate to recent work in semantic data modelling, as for example in [Hammer & McLeod 1978]?

Temporal Query Languages: What kinds of query languages do we need to support questions that refer to this temporal dimension [Clifford 1982]? What is an appropriate definition of a "temporally complete" query language?

Temporal Consistency Checking: How can we retain or maintain the identity of data objects over time? Can we

develop a theory of temporal dependencies and constraints, to include the notions of "constant," "non-constant," and "monotonically in(de)creasing" attributes described in [Clifford & Warren 1981]? Finally, how does this concept of "data time" relate to the concept of "transaction time" that is central to the issue of database concurrency? [Rolland 1982].

3. Implementation of Historical Databases

To our knowledge there is no existing general-purpose DBMS in operational use which presents the notion of time as a primitive data-structuring device to the user. To be sure, there are some prototype systems under development, and moreover, systems which allow users to incorporate time-series data but which, having no built-in concept of time and its properties, provide no real support for this usage.

Structures & Operations: What are the file structures, indexing schemes, etc., appropriate for the implementation of HDBs? Can historical data and historical meta-data (concerning the parallel history of database definition and usage) be combined coherently into one operational unit? What kinds of language primitives are appropriate for a useful and consistent view of time-oriented data?

Efficiency: What are the appropriate efficiency concerns and trade-offs? For example, what design strategy will not "punish" the user of the latest version of the data for the expense of having the historical "data tail?" What kind of interface between an on-line historical database and an archival database for "long-past" information can be presented to the user?

Impact of New Technology: How can recent and future hardware developments (e.g., optical mass storage, associative processing, database machines) be exploited for the construction of effective HDBs?

4. The User Interface to Historical Databases

The notion of time-series data is very different from the notion of "the latest value" of an attribute. Flat-table presentations of dynamic data do not convey the notion of time in a rich way. Recent work [Ariav 1982] has explored alternative methods of presenting this data to the user in such a way as to provide a deeper sense of its temporal dimension. This work makes extensive use of graphics as a presentation device.

What are the appropriate forms of computer graphics for this interface task? Can we generalize about these graphic tools across different database domains? In what other ways can the temporal dimension to data be presented to the user, both on input and on output? Finally, what can we learn from psychological inquiries into the experience of time, and the growing literature on human interfaces, and even from the video games phenomenon, to make historical databases appear dynamic at the user interface?

5. Applications of Historical Databases

A major development in information systems research over the past decade has been the emergence of the concept of a decision support system (DSS) as an integrated set of tools to assist in the managerial decision-making process. Central to a DSS are a DBMS and a planning model component, and their integrated use for modelling future developments.

One major function that directly involves the concept of time is the forecasting of possible future conditions based upon performance in the past. How can an HDB with a sophisticated query language rich in the historical dimension support this predictive function?

A second task of a DSS is the planning of resource usage over time. How can an HDB be integrated with planning tools for modelling the future, such as Petri nets or CPM networks [Jarke 1982], [De Antonellis & Zonta 1981]? More generally, what can we learn from operations research and artificial intelligence planning concepts that will enable us to develop theoretically the concept of an HDB as a planning tool?

6. The Panel.

The panel brings together a number of people working on various issues involved in incorporating a temporal dimension into databases. We hope to stimulate a broad discussion of these issues at the conference, and thereby to encourage further work in this interesting area.

References

Ariav, Gad and Howard L. Morgan (1982). "MDM: Embedding the Time Dimension in Information Systems," TR #82-03-01, Dept. of Decision Sciences, The Wharton School, Univ. of Pennsylvania, Philadelphia.

Bolour, A., T.L. Anderson, L.J. Dekeyser, and H.K.T. Wong (1982). "The Role of Time in Information Processing: A Survey," ACM SIGMOD Record, 12,3.

Bubenko, Janis A. Jr. (1977). "The Temporal Dimension in Information Modelling," in Architecture and Models in Data Base Management Systems, ed. G. M. Nijssen, North Holland, Amsterdam.

Clifford, James and David S. Warren (1981). "Formal Semantics for Time in Databases," TR #81/025, Dept. of Computer Science, SUNY at Stony Brook, Stony Brook. To appear in TODS.

Clifford, James (1982). "A Logical Framework for the Temporal Semantics and Natural-Language Querying of Historical Databases," Ph.D. dissertation, Dept. of Computer Science, SUNY at Stony Brook, Stony Brook.

De Antonellis, V. and Zonta, B. (1981). "Modelling Events in Data Base Applications Design," Proc. 7th VLDB, Cannes, 23-31.

Hammer, Michael and Dennis Mcleod (1978). "The Semantic Data Model: A Modelling Mechanism for Data Base Applications," Proc. ACM SIGMOD 1978, Austin.

Jarke, Matthias (1982). "Developing Decision Support Systems: A Container Management Example," International Journal of Policy Analysis and Information Systems, 6,4, 351-372.

Klopprogge, Manfred R. (1981). "TERM: An Approach to Include the Time Dimension in the Entity-Relationship Approach," Proc. 2nd Intl. Conf. on Entity-Relationship Approach, Washington DC.

Rolland, C. and C. Richard (1982). "Transaction Modelling," Proc. ACM SIGMOD 1982,, Orlando.