

This issue will be a valuable reference work for beginning students of data base. Also it should provide ample evidence, (for those who might still require it), that data base is clearly a well established discipline.

Charles J. Bontempo
IBM Systems Research Institute
219 East 42nd Street
New York, N. Y. 10017
(212) - 983-7216

[Those interested in reviewing papers or books of interest to SIGMOD readers should contact the Reviews Editor.]

Fry, J.P., Sibley, E.H. "Evaluation of Data-Base Management Systems"
ACM Computing Surveys 8 (1) (March 1976) 7-42.

Computing Reviews
Categories: CR 4.33, 4.34, 1.2

Data-base management systems provide generalized tools for the storage, retrieval, maintenance and manipulation of data-bases be they large or small. Since most organizations have data-bases that require manipulation, it is not surprising that a great deal of interest has been generated on the subject. The importance of data-base technology is evidenced by the large number of data-base management systems that have been developed, the numerous conferences that have been held on the subject, and the several journals that are devoted exclusively to the subject. Indeed, the article under review appears as the lead article in a special issue of Computing Surveys devoted exclusively to data-base management systems.

The intent of the article is to provide a brief introduction to data-base management, and to provide historical perspective upon the subject. Subsequent articles in the issue are devoted to the technological aspects of the various data-base models such as relational, network, and hierarchic.

The development of techniques for generalized systems arose in the mid 1950's primarily because of the needs of the military. Thus, the

earliest work in the field was sponsored by the Navy, the Air Force, the Army, and the Atomic Energy Commission. Generalized techniques arose because the dynamic nature of the military problem, even for a single data-base, forced programmers to continually modify their programs. To avoid this, the concept of a data description mechanism arose. This was but one of the many early developments. Indeed, most of the major concepts needed for implementing data-base management systems came about during the mid 1950's to early 1960's. This includes techniques for translating queries, data structures, indexing files (using indexed sequential files, although they were not so named), and other innovations. Some of the developments in software were not supported by hardware. Disks and drums needed for random access for large-scale on-line retrieval were not available until the mid 1960's

It was unfortunate that much of the work took place on classified projects and were not reported upon in the open literature. Because there were few papers written on the systems developed, it was necessary that many of the tools and techniques be rediscovered. This does not mean that some of the early insights gained were not used. Indeed, the article under review provides a very nice genealogical description of data-base technology. The authors trace generalized report generators that are in IBM systems and other computers back to the Hanford Washington operations of the Atomic Energy Commission. Work in the Navy in 1959 and work on the Strategic Air Ground Environment (SAGE) in the late 1950's led to the current Formatted File Systems in wide use throughout the military establishment and on current IBM systems. Work growing out of ADAM, developed for the Air Force in 1962, has led to the current S2000 data-base management system. These developments and others are carefully traced in a most readable and interesting manner by the authors.

In contrast to the military who led the way, manufacturers were

generally slow in recognizing the importance of the generalized techniques that their systems and programming staff were developing and being paid for under military contracts. Thus, when this reviewer suggested to RCA management in approximately 1962 that it could make an early retrieval system, the Retrieval Command Oriented Language (RECOL), by Doug Climenson, then of RCA, designed and implemented under work on an Army project, into an RCA product for general use with the RCA 501 with no additional effort, the suggestion went ignored. IBM, although working on data-base systems sponsored by the military in the late 1950's, did not recognize the importance of the technology until the late 1960's and early 1970's. It was also in the late 1960's that the CODASYL committee was formed to attempt to tie together work in data-base systems. Although RCA, IBM, UNIVAC, Honeywell, and most other manufacturers were not cognizant of the significance of the technology, this was not apparently the case at the General Electric Corporation (GE). Under the leadership of Charles Bachman, the Integrated Data Store (IDS) work was started in the early 1960's and has continued after Honeywell took over the GE computer activities. The network approach of the CODASYL Committee was influenced greatly by Bachman and his work.

Whereas manufacturers generally arrived late on the scene of data-base management systems, one of the software houses, Informatics, recognized its importance as early as 1962 when John Postley, the developer of a system termed GIRLS, joined Informatics and started work on the MARK series of data-base management systems. Since then, numerous software firms have implemented and are now marketing data-base management systems. It would not be incorrect to state that data-base management systems developed by software firms, rather than manufacturers, are in the forefront of the technology.

Although one may look back with nostalgia to the 1950's and 1960's and note that many new innovations were discovered then, and that nothing

tremendously new has been developed, this is a gross over-simplification. Most systems implemented during that time period were developed in an ad-hoc fashion without the designers having a clear concept of the entire system when it was completed. A major innovation of the past few years is the recognition that there are several different data models upon which data-base management systems can be built. Although some of the models may not be founded on mathematical formalisms, they are, nevertheless, models which can be understood, described, and around which techniques can be used to implement them. The understanding of the various models has led and will lead to more efficient implementations. The formal models further permit universities and colleges to present data-base management systems in a coherent manner rather than as a collection of isolated tools and techniques. Thus, computer scientists and business majors are being graduated who know a great deal about data-base management systems and need not undergo extensive training in industry before they become useful.

The article, unfortunately, does not trace the genealogy of relational data bases. The concepts of relational data-bases was discussed as early as 1962 by Manfred Kochen (1). At the Rand Corporation, under the direction of Roger Levien and M.E. (Bill) Maron (2) a sophisticated relational system, termed, Relational Data File was developed in the early 1960's. Furthermore, Larry Kuhns (3) wrote an interesting and important paper on the logical and philosophical aspects of the work. There were also a number of other important developments on relational systems in the 1960's. A description of several of these are provided in the survey by Minker and Sable (4). These developments preceded the work by Codd. However, Codd's achievement was to formalize the concept of a relational data-base system and to put it on a sound mathematical foundation. It was through Codd's work that relational data-bases have received such

wide interest.

The paper provides a very good description of many early developments in the area of data-base management systems. It is one of the few papers to appear in the published literature that treats the subject of data-base management systems from an historical view. "Evolution of Data-Base Management Systems" is important and should be read by all workers in the field.

References

- (1) M. Kochen "Adaptive Mechanisms in Digital Concept Processing" in Discrete Adaptive Processes - Symposium and Panel Discussion, AIEE, New York, 1962, pp. 50-58.
- (2) R.E. Levien and Maron, M.E. "A Computer System for Inference Execution and Data Retrieval" CACM 10 (11) (Nov. 1967), 715-721.
- (3) J.L. Kuhns "Logical Aspects of Question-Answering by Computer", Proc. Third International Symposium on Computer and Information Sciences (COINS IV), December 1969, Academic Press, New York 1969.
- (4) J. Minker and Sable, J.D. Relational Data Systems Study, Final Report, Auerbach Technical Report AVER-1776-TR-1 July 1970 submitted to Rome Air Development Center. Also RADC-TR-70-180. Final Technical Report September 1970 AD-720-263, ED-049-645.

Jack Minker

RELATIONAL DATA-BASE MANAGEMENT SYSTEMS

Donald D. Chamberlain

Computing Surveys, Vol 8, No 1. pp 43-66

Chamberlain has succeeded where many others have failed. The phrase "Relational Data-Base" has attained an aura of mystique about it, largely due to the new terminology introduced, and the inability on the part of early authors to simplify the concepts involved. The author, in this paper, has carefully defined the terminology in terms of common nomenclature, and has successfully conveyed the basic concepts of the relational model, and