

A GRADUATE COURSE IN DATABASE MANAGEMENT

J. Gerry Purdy*

Visiting Industrial Professor
Department of Computer Science and Operations Research
Southern Methodist University

and

Director, Computer Technology
Institute for Aerobics Research
Dallas, Texas

Abstract

Database management is a new and fast growing discipline within computer science. Curriculum 68 did not include database management in its recommended courses. With the advent of large direct access secondary storage devices to hold large volumes of information referred to as a database and with the need to manage that storage efficiently and conveniently, there has developed the computer science discipline of database management. This paper describes a course in database management, including a course catalog overview, course outline, and bibliography. This paper could serve as a guideline for future ACM Curriculum Committee on Computer Science recommendations in database management.

When the ACM Curriculum Committee on Computer Science (C3S) published its recommendations for academic programs in computer science in Curriculum 68 (1), the discipline of database management was in its infancy. The 1973 ACM Turing Award winner, C.W. Bachman, had implemented an early version of Integrated Data Store (IDS), and a few others were beginning to publish descriptions of data management systems. The only Curriculum 68 course which included the subject of data management and databases was: A8. Large-scale Information Processing Systems. Since the course primarily concerned the design and implementation of large-scale software systems, data management and databases were studied as secondary tools to the main "systems design" function.

More recently, the ACM Curriculum Committee on Computer Education for Management (C3EM) published its Curriculum Recommendations for Graduate Professional Programs in Information Systems (2). Although these curriculum recommendations were primarily intended for students in management science and business professional

programs, two courses were included which discussed data management and databases: C3. File and Communication Systems and D2. System Design. Although neither of these courses was dedicated to the study of databases, they did introduce the student to such concepts as network data structures, direct access secondary storage devices, file systems, and design of integrated databases. The associated literature had also grown, which provided support for these courses. In the C3EM Curriculum Recommendations for Undergraduate Programs in Information Systems (3), course C3 was renamed UC3 and course D2 was included as part of UD9.

The purpose of this paper is to propose a course entirely in database management since the discipline has come into its own right, separate and apart from other topics such as system design and implementation or software engineering. This paper could serve as a guideline for future C3S recommendations.

The course is divided into eight parts. The first part, Introduction, defines a database and outlines why integrated databases are desirable. The second part, Logical Data Structures, is a review of basic data structures as well as a

*Present address: Arizona Heart Institute,
St. Joseph's Hospital & Medical Center,
350 West Thomas Road, Phoenix, AZ 85013

presentation of networks. The Bachman database structure diagram, the Codd relational notation, and the hierarchical structure are described. Physical Storage Structures is the third section and is concerned with storage devices and media such as memory, disc's, and drum's.

The fourth part, File Structures, concerns sequential, indexed sequential, direct access, and inverted file organizations. This is followed by the fifth part, File Systems, which surveys vendor generated file management systems. The sixth part, Development of Integrated Databases, describes the ingredients of all database management systems and provides a historical development of the CODASYL activities.

The seventh part, Survey of Database Management Systems, lists the criteria to evaluate database systems and surveys a number of generalized systems. The final part, Database Management in Perspective, reviews the goals of database management and projects the future of the discipline.

A "Course Description" has been included which could be published in a university bulletin. This is followed by a course outline and a bibliography. Entries in the bibliography have been cross-referenced in the outline to aid in the instruction of the material.

This course can be given adequately in one semester or two quarters. If a number of case studies are presented and the students are required to implement a database management system (DMS) or an application using an existing DMS, then the course could be given over an entire year (two semesters or three quarters).

At the present time there are no adequate texts for this course. However, a new book by Date and a forthcoming book by Wiederhold are steps in the right direction.

References

1. C³S, "Curriculum 68," Comm. ACM, Vol. 11, No. 3, March 1968, pp. 151-197.
2. C³EM, "Curriculum Recommendations for Graduate Professional Programs in Information Systems," Comm. ACM, Vol. 15, No. 5, May 1972, pp. 363-398.
3. C³EM, "Curriculum Recommendations for Undergraduate Programs in Information Systems," Comm. ACM, Vol. 16, No. 12, December 1973, pp. 727-749.

Catalog Course Description

This course is primarily concerned with the definition, organization, and manipulation of a database. An overview of the goals of database management is examined. Next, the basic ingredients of all database management systems are presented, including logical data structures, physical storage structures, and file structures, with emphasis on direct storage access methods. The database management process is broken down into its four constituent parts: data definition, data manipulation, data retrieval, and report generation. Attempts toward standardization in database management are presented, with emphasis on the CODASYL activities. Security, integrity, and privacy measures are also discussed. The

concept of shared files and the deadlock problem are presented. A number of general database management systems are surveyed, and at least one case study is examined in detail.

Prerequisite: An undergraduate course in data structures.

Course Outline

1. Introduction [Texts: CAG 73, DAT 75, HOU 74, KLI 74, LEF 74, LYO 71, RUS 72 WEI 75]
 - A. Definition of a Database [LYO 71]
 - B. The Goal of Database Management [BAC 73]
 - C. Reasons for Developing the Integrated Database [BAC 73]
2. Logical Data Structures [CHI 68, DIM 69, DOD 69, EAR 71, MEA 67]
 - A. The Meaning of Data [LYO 71]
 1. Definition
 2. Attribute (Types)
 3. Values (Occurrences)
 - B. Hierarchical Representation of Data
 - C. Pointers [KNU 74]
 - D. Binding of Pointers with Data [KNU 74]
 - E. Sequential [KNU 74]
 - F. Linked Lists [KNU 74]
 - G. Ring Structures [KNU 74]
 - H. Doubly Linked Lists [KNU 74]
 - I. Trees [DEMA 71, LAN 63, NIE 72]
 - J. Strings
 - K. Hash Coding [ARO 69, BEL 70a, BEL 70b, BEY 68, BLO 70, COF 70, DAY 70, KNO 71, LAM 70, LUM 71, MAU 68, MOR 68, RAD 70, ULL 70]
 - L. Networks [BAC 65b, BAC 64]
 1. Definition
 2. Database Structure Diagram (Bachman) [BAC 69, ROB 73]
 3. Relational Structure (Codd) [CODD 70, CODD 71a, CODD 71b, CODD 72, CODD 74a, CODD 74b, DAT 74]
 4. Hierarchical Structures & Network Representation (à la IBM/IMS) [ELL 70, HEN 69]
 - M. Data Entry and Input Media [HOU 74]
3. Physical Storage Structures [DIM 69, HOU 74, TAY 71]
 - A. Hierarchy of Physical Structures [BAC 72a]
 - B. Types of Secondary Storage Physical Devices [IBM 74, HOU 74]
 1. Tape
 2. Disc
 3. Drum
 4. Data Cell
 - C. Comparison of Physical Structures
4. File Structures [CHA 68, CHA 69, CLI 66, COF 70, DEMA 71, DOU 66, HSI 71, HSI 70, LEF 69, MAD 70, MAR 68, MCG 68, MCG 69, NIE 72, PRI 71]
 - A. Definition
 - B. Sequential File Structures [CHO 69, IBM 74]
 1. Fixed
 2. Variable
 3. Blocked
 - C. Indexed Sequential File Structure [IBM 74, MUL 72, NIJ 71b]
 - D. Random (Direct Access) [HEI 63, IBM 74, JOH 61, NIJ 71a, NIJ 71b, OLS 69, PET 57]
 1. Randomization Techniques (See Hash Coding)

- 2. Keyed Files
- E. Inverted Files [BAYER 70, BAYES 69, CAG 73]
 - 1. Partially Inverted Files (Directories)
 - 2. Fully Inverted Files
- 5. File Systems [BRY 69, LEF 69]
 - A. Basic Ingredients
 - 1. System Control Commands
 - 2. File Definition
 - 3. File Initialization
 - 4. File Update
 - 5. File Retrieval
 - 6. File Report Generation
 - 7. Multi-File Processing
 - B. MARS VI (CDC) [CON 74]
 - C. MANAGE (Xerox) [XER 70]
 - D. MARK IV (Informatics) [INF 71]
- 6. Development of Integrated Databases [KLI 74, WEI 75, MIN 69, RUS 72]
 - A. Motivation [BAC 73]
 - B. Standardization Efforts [BLE 70]
 - 1. Desirability
 - 2. CODASYL Activity [CODA 71a, CODA 74, OLL 72]
 - C. Ingredients of Database Management Systems
 - 1. Data Definition Facility [CODA 71a, CODA 74, DEMP 70, EAR 72, ELL 70, ENG 71, FRY 72, KOC 72, MCG 72, SIB 73, TAY 71]
 - 2. Data Manipulation Facility [CODA 71a, ENG 71, HEA 71]
 - 3. Database Query Facility [CODD 71a, CODD 74a, CODD 74b, GHO 68, LAN 63, MUL 71, SAL 68]
 - 4. Database Report Generator Facility
 - 5. Utilities (Journalization, Statistics, Initialization) [XER 74]
 - 6. Shared Databases [SHO 69]
 - 7. Database Restructuring [XER 74]
 - D. The Database Administrator [GUI 73]
- 7. Survey of Database Management Systems
 - A. Evaluation Criteria [CODA 69, CODA 71b, ENG 71]
 - 1. Type of System (Host vs. Total)
 - 2. Data Definition
 - 3. Data Manipulation
 - 4. Query Capability
 - 5. Report Generation
 - 6. Application Programming
 - 7. File Structures
 - 8. Documentation
 - 9. Communications
 - 10. Security, Integrity, and Privacy [BRO 71a, BRO 71b, DEA 71, SHE 72]
 - 11. Run Time Statistics
 - 12. Shared Files [CODD 70]
 - 13. Data Independence [ALT 72, AST 72, COL 72, DAT 71a, DAT 71b]
 - B. CODASYL DBTG Report [CODA 71a, CODA 74, ENG 71]
 - C. GUIDE/SHARE [GUI 73, EVE 71]
 - D. EDMS (Xerox) [XER 74]
 - E. DMS 1100 (Univac) [UNI 72, UNI 73]
 - F. IDS (Honeywell) [BAC 65a, HON 72]
 - C. IMS (IBM) [IBM 73]
 - H. IDMS (Cullinane Corp.) [CUL 73a, CUL 73b]
 - I. System 2000 (MRI Systems Corp.) [MRI 72]
 - J. TOTAL (Cincom Systems, Inc.) [CIN 74]
 - K. GIM (TRW, Inc.) [NEL 67]
 - L. TDMS (SDC) [BLE 67]
 - M. JOVIAL (SDC) [SYS 68]

- N. ADABAS (Software AG Corp) [SOF 73]
- O. Other Systems [ALT 72, AST 72, BAC 72b, BOB 72, CAN 74, DIX 66, FER 71, LIU 71, SEN 73]
- 8. Database Management in Perspective [BAC 73]
 - A. Review of the Goals of Database Management
 - B. Database Management System Evaluation
 - C. The Future of Database Management

Bibliography*

- [ALT 72] Altman, E.B., Astrahan, M.M., Fehder P.L., Senko, M.E., "Specifications in a Data Independent Accessing Model", Proc. 1972 ACM-SIGFIDET Workshop on Data Description, Access and Control, ACM, New York, 1972, pp. 363-381.
- [ARO 69] Arora, S.R., Dent, W.T., "Randomized Binary Search Techniques", CACM, Vol. 12, No. 2, Feb. 1969, pp. 77-80.
- [AST 72] Astrahan, M.M., Altman, E.B., Fehder, P.L., Senko, M.E., "Concepts of a Data Independent Accessing Model", Proc. 1972 ACM-SIGFIDET Workshop on Data Description, Access and Control, ACM, New York, 1972, pp. 349-362.
- [BAC 72a] Bachman, C.W., "The Evolution of Storage Structures", CACM, Vol. 15, No. 7, July 1972, pp. 628-634.
- [BAC 69] Bachman, C.W., "Data Structure Diagrams", Data Base (ACM-SIGBDP Newsletter), Vol. 1, No. 2, Summer 1969.
- [BAC 65b] Bachman, C.W., "Software for Random Access Processing", Datamation, April 1965, pp. 36-41.
- [BAC 72b] Bachman, C.W., "Architecture Definition Technique: Its Objectives Theory, Process, Facilities and Practice", Proc. 1972 ACM-SIGFIDET Workshop on Data Description, Access and Control.
- [BAC 65a] Bachman, C.W., "Integrated Data Store", DPMA Quarterly, January 1965.
- [BAC 64] Bachman, C.W., et al, "A General Purpose Programming System for Random Access Memories", AFIPS, FJCC, 1964.
- [BAC 73] Bachman, C.W., "The Programmer as Navigator", CACM, Nov. 1973, pp. 653-658.
- [BAYER 70] Bayer, R., McCreight, E., "Organization and Maintenance of Large Ordered Indices," 1970 ACM-SIGFIDET Workshop on Data Description and Access, ACM, 1970.
- [BAYES 69] Bayes, A.J., "Retrieval Times for a Packed Direct Access Inverted File," CACM, Vol. 12, No. 10, October 1969, pp. 582-583.
- [BEL 70a] Bell, J.R., "The Quadratic Quotient Method: A Hash Code Elimination Secondary Clustering", CACM, Vol. 13, No. 2, February 1970, pp. 107-109.
- [BEL 70b] Bell, J.R., Kaman, C.H., "The Linear Quotient Hash Code", CACM, Vol. 13, No. 11, November 1970, pp. 675-677.
- [BEY 68] Beyer, J.D., "Comments on 'An Improved Hash Code for Scatter Storage'", CACM, Vol. 11, No. 5, May 1968, p. 378.
- [BLE 67] Bleier, R.E., "Treating Hierarchical Data Structures in the SDC Time-shared Data Management System (TDMS)", Proc. ACM 22nd

*This bibliography contains only the more important references. Ed Sibley of the University of Maryland has prepared a more comprehensive bibliography.

- National Conference 1967, MDI Publications, Wayne, Pa. pp. 41-49.
- [BLE 70] Bleier, R.E., "Data Definition Standardization", 1970 ACM SIGFIDET Workshop on Data Description and Access, ACM, New York, 1970.
- [BLO 70] Bloom, B.H., "Space/time Trade-offs in Hash Coding with Allowable Errors", CACM, Vol. 13, No. 7, July 1970, pp. 422-426.
- [BOB 72] Bobrow, R.J., "An Experimental Data Management System", Data Base Systems, R. Rustin (Ed.), Prentice-Hall, 1972, pp. 125-142.
- [BRO 71a] Browne, P.S., "Data Privacy and Integrity: An Overview", Proc. 1971, ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1971, pp. 237-240.
- [BRO 71b] Browne, P.S., Steinauer, D., "A Model for Access Control," Proc. 1971 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1971, pp. 241-262.
- [BYR 69] Byrnes, C.J., Steig, D.B., "File Management Systems: A Current Summary", Datamation, November 1969, pp. 138-142.
- [CAG 73] Cagan, Carl, Data Management Systems, Melville, 1973.
- [CAN 74] Canaday, R.H., Harrison, R.D., Ivie, E.L., Ryder, J.L., and Weir, L.A., "A Back-End Computer for Data Base Management", CACM, Vol. 17, No. 10, October 1974, pp. 575-582.
- [CHA 68] Chapin, N., "A Comparison of File Organization Techniques", Proc. ACM Conference, 1968, pp. 631-638.
- [CHA 69] Chapin, N., "Common File Organization Techniques Compared", FJCC, Vol. 35, 1969, pp. 413-422.
- [CHI 68] Childs, D.L., "Feasibility of a Set Theoretical Data Structure - A General Structure Based on a Reconstituted Definition of a Relation", Proc. IFIP Congress 1968, North-Holland Pub. Co., Amsterdam, 1968, pp. 162-172.
- [CIN 74] Cincom Systems, TOTAL/7 Reference Manual, 2300 Montana Avenue, Cincinnati, OH 45211, Pub. No. PO2-1322-1, 1974.
- [CLI 66] Climenson, W.D., "File Organization and Search Techniques", Annual Review of Information Science and Technology, Vol. 1, John Wiley & Sons, Inc., 1966, pp. 107-135.
- [CODA 69] CODASYL Systems Committee, A Survey of Generalized Data Base Management Systems, ACM, New York, 1969.
- [CODA 71a] CODASYL, CODASYL Data Base Task Group Report, ACM, New York, April 1971.
- [CODA 71b] CODASYL Systems Committee, Feature Analysis of Generalized Data Base Management Systems, ACM, New York 1971.
- [CODA 74] CODASYL, Data Description Language, Journal of Development, U.S. Govt. Printing Office, SD No. C13.6/2:113, Jan. 1974.
- [CODD 70] Codd, E.F., "A Relational Model of Data for Large Shared Data Banks", CACM, Vol. 13, No. 6, June 1970, pp. 337-387.
- [CODD 71a] Codd, E.F., "A Data Base Sublanguage Founded on the Relational Calculus", Proc. 1971 ACM-SIGFIDET Workshop on Data Base Systems, ACM, New York, 1971, pp. 35-68.
- [CODD 71b] Codd, E.F., "Normalized Data Base Structure: A Brief Tutorial", Proc. 1971 ACM-SIGFIDET Workshop on Data Description, Access, and Control, ACM, New York, 1971, pp. 1-18. (Also IBM Research Report RJ 935, November 1971)
- [CODD 72] Codd, E.F., Relational Completeness of Data Base Sublanguages, Data Base Systems, R. Rustin (Ed.), Prentice-Hall, 1972, pp. 65-98.
- [CODD 74a] Codd, E.F., "Seven Steps to Rendezvous with the Casual User," Data Base Management, J.W. Klimbie and K.L. Koffeman (Eds.), North-Holland Pub. Co., 1974, pp. 179-200.
- [CODD 74b] Codd, E.F., "Recent Investigations in Relational Data Base Systems," Proc. IFIP Congress 74, North-Holland Pub. Co., 1974, pp. 1017-1021.
- [COF 70] Coffman, Jr., E.G., Eve, J., "File Structures Using Hashing Functions", CACM, Vol. 13, No. 7, July 1970, pp. 427-436.
- [COL 70] Collmeyer, A.J., Shemer, J.E., "Analysis of Retrieval Performance for Selected File Organization Techniques", Proc. Fall Joint Computer Conference, 1970, p. 206.
- [COL 72] Collmeyer, A.J., "Implications of Data Independence on the Architecture of Data-base Management Systems," Proc. 1972 ACM-SIGFIDET Workshop on Data Description, Access and Control, ACM, New York, 1972, pp. 307-322.
- [CUL 73a] Cullinane Corp., Integrated Data-base Management System Data Definition Language and Utilities Reference Guide (IDMS DDL), Boston, MA, 1973.
- [CUL 73b] Cullinane Corp., Integrated Data Management System COBOL Data Manipulation Language Programmer's Reference Guide (IDMS DML), Boston, MA, 1973.
- [CON 74] Control Data Corp., Multi-Access Retrieval System: (MARS VI) for Partial Inversion, Reference Manual, Version 2.1, #60385900, April 1974.
- [DAT 71a] Date, C.J., Hopewell, P., "File Definition and Logical Data Independence", Proc. 1971 ACM-SIGFIDET Workshop-Data Description, Access and Control, pp. 117-168.
- [DAT 71b] Date, C.J., Hopewell, P., "Storage Structure and Physical Data Independence", Proc. 1971 ACM-SIGFIDET Workshop-Data Description, Access and Control, pp. 139-168.
- [DAT 74] Date, C.J., Codd, E.F., "The Relational and Network Approaches: Comparison of the Application Programming Interfaces," IBM Report RJ 1401 (#21706), June 6, 1974.
- [DAY 70] Day, A.C., "Full Table Quadratic Searching for Scatter Storage," CACM, Vol. 13, No. 8, August 1970, pp. 481-482.
- [DEA 71] Dean, Jr., A.L., "Data Privacy and Integrity Requirements for On-line Data Management Systems", Proc. 1971, ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1971, pp. 279-298.
- [DEMA 71] DeMaine, P.A.D., Rotwitt, Jr., T., "Storage Optimization of Tree Structured Files", Proc. 1971 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1971, pp. 207-218.
- [DEMP 70] Dempsey, J.W., Mullin, J.K., "Problems of Building a Hybrid Data Definition Facility," 1970 ACM-SIGFIDET Workshop on Data Description and Access, ACM, New York, 1970.
- [DIM 69] D'Imperio, M.E., "Data Structures and Their Representation in Storage", Annual Review of Automatic Programming, Pergamon Press, 1969.

- [DIX 66] Dixon, P.J., Sable, J., "DM-1 A Generalized Data Management System", Proc. AFIPS 1966 Fall Joint Computer Conference, Vol. 29, pp. 677-684.
- [DOD 69] Dodd, G.E., "Elements of Data Management Systems", Computing Surveys, June 1969.
- [DOU 66] Douglas, C.W., "File Organization and Search Techniques", Annual Review of Information Technology, Interscience, 1966.
- [EAR 71] Earley, Jay, "Towards an Understanding of Data Structures", CACM, Vol. 14, No. 10, October 1971, pp. 617-627.
- [EAR 72] Earley, Jay, "On the Semantics of Data Structures in Data Base Systems", R. Rustin (Ed.), Prentice-Hall, 1972, pp. 23-32.
- [ELL 70] Ellis, Max, Nelson, Kenneth, "A Data Description Language for Hierarchical Data Files", 1970 ACM-SIGFIDET Workshop on Data Description and Access, ACM, New York, 1970.
- [ENG 71] Engles, R.W., "An Analysis of the April 1971 Data Base Task Group Report," Proc. 1971 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1971, pp. 69-92.
- [EVE 71] Everest, G.C., Sibley, E.H., "A Critique of the GUIDE - SHARE DBMS Requirements", Proc. 1971 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1971, pp. 113-116.
- [FER 71] Ferrari, D., "On the Architecture of Data Base Systems", Proc. 1971 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1971, pp. 113-116.
- [Fry 72] Fry, J.P., Smith, D.C.P., Taylor, R. W., "An Approach to Stored Data Definition and Translation", Proc. 1972 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1972, pp. 13-56.
- [GHO 68] Ghosh, S.P., "On the Problem of Query-Oriented Filing Schemes Using Discrete Mathematics", Proc. IFIP Congress 1968, Edinburgh, Scotland, pp. 74-79.
- [GHO 69] Ghosh, S.P., Senko, M.E., "File Organization: On the Selection of Random Access Index Points for Sequential Files", JACM, October 1969, pp. 569-579.
- [GUI 73] Guide International, Basic Requirements.
- [HEA 71] Heath, I.J., "Unacceptable File Operations in a Relational Data Base," Proc. 1971 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1971, pp. 19-34.
- [HEI 63] Heizing, W.P., "Note on Random Addressing Techniques", IBM Systems J., June 1963, pp. 112-116.
- [HEN 69] Henry, W.R., "Hierarchical Structure for Data Management", IBM Systems Journal, Vol. 8, No. 1, 1969.
- [HON 72] Honeywell, Integrated Data Store Series 600/6000 Software, BR 69, Rev. 1, May 1972.
- [HOU 74] House, W.C. (Ed.), Data Base Management, Petrocelli Books, Mason and Lipscomb Pub., 1974.
- [HSI 70] Hsiao, D.K., Harary, Frank, "A Formal System for Information Retrieval from Files", CACM, Vol. 13, No. 2, February 1970, pp. 67-73.
- [HSI 71] Hsiao, D.K., "A Generalized Record Organization," IEEE Trans. on Computers, C-20, No. 12, Dec. 1971, pp. 1490-1495.
- [IBM 74] IBM, Introduction to IBM/360 Direct Access Storage Devices and Organization Methods, GC20-1649-6, White Plains, New York, February 1974.
- [IBM 73] IBM, Information Management System/360, Version 2 System/Application Design Guide, IBM Corporation, SH20-0910-4, September 1973.
- [INF 71] Informatics, MARK IV File Management System, Technical System Description, April 1971.
- [JOH 61] Johnson, L.R., "An Indirect Chaining Method for Addressing on Secondary Keys", CACM Vol. 4, No. 5, May 1961, pp. 218-222.
- [KLI 74] Klimbie, J.W. and Koffeman, K.L., (Eds.) Data Base Management, North-Holland Pub. Co., 1974.
- [KNO 71] Knott, G.D., "Expandable Open Addressing Hash Table Storage and Retrieval", Proc. of Symp. on Data Structures in Programming Languages (ACM-SIGPLAN), 1971.
- [KNU 74] Knuth, D.E., The Art of Computer Programming, Vol. 1: Fundamental Algorithms, Chapter 2, Information Structures, Addison-Wesley, Reading, Massachusetts, Second Edition, 1974.
- [KOC 72] Koch, H.S., Gotterer, M.S., "An Independent Data Definition Facility for COBOL and FORTRAN," Proc. 1972 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1972, pp. 57-66.
- [LAM 70] Lamport, L., "Comments on Bell's Quadratic Quotient Method for Hash Code Searching", CACM, Vol. 13, No. 9, September 1970, pp. 573-574.
- [LAN 63] Landauer, W.I., "The Balanced Tree and Its Utilization in Information Retrieval", IEEE Trans. on Electronic Computers, Vol. EC-12, No. 6, December 1963.
- [LEF 69] Lefkovitz, D., File Structures for On-Line Systems, Spartan Books, New York, 1969.
- [LEF 74] Lefkovitz, D., Data Management for On-Line Systems, Hayden Book Co., Rochelle Park, NJ, 1974.
- [LIU 71] Liu, H.N., "A Data Base System Defined to Support a Large Information System", Proc. 1971 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1971, pp. 349-372.
- [LUM 71] Lum, V.Y., Yuen, P.S.T., Dodd, M., "Key-to-Address Transform Techniques: A Fundamental Performance Study on Large Existing Formatted Files," CACM, Vol. 14, No. 4, April 1971, pp. 228-239.
- [LYO 71] Lyon, J.K., An Introduction to Data Base Design, Wiley-Interscience, 1971.
- [MAD 70] Madnick, S.E., "Design Strategies for File Systems", MIT Project MAC M.S. Thesis, Report TR-78, October 1970.
- [MAR 68] Martin, L.D., "A Model for File Structure Determination for Large On-Line Data Files", in Proc. of the File 68 International Seminar on File Organization, Copenhagen, 1968, pp. 793-834.
- [MAU 68] Maurer, W.D., "An Improved Hash Code for Scatter Storage", CACM, Vol. 11, No. 1, January 1968, pp. 35-38.
- [MCG 68] McGee, W.C., "File Structures for Generalized Data Management", Proc. IFIP Congress 68, North-Holland Pub. Co., 1968, F68-F73.

- [MCG 69] McGee, W.C., "Generalized File Processing", An Annual Review in Automatic Programming, Pergamon Press, New York, 1969, pp. 77-149.
- [MCG 72] McGee, W., "Some Current Issues in Data Description", Proc. 1972 ACM-SIGFIDET Workshop on Data Description, Access and Control, ACM, New York, 1972, pp. 1-12.
- [MEA 67] Mealy, G., "Another Look at Data", FJCC, Vol. 31, 1967, pp. 525-534.
- [MIN 69] Minker, J., "Generalized Data Management Systems - Some Perspectives", National Bureau of Standards, Technical Report, pp. 69-101, 1969.
- [MOR 68] Morris, R., "Scatter Storage Techniques", CACM, Vol. 11, No. 1, January 1968, pp. 38-44.
- [MRI 72] MRI Systems Corp., System 2000 General Information Manual, Austin, TX, 1972.
- [MUL 71] Mullin, J.K., "Retrieval-Update Speed Tradeoffs Using Combined Indices", CACM, Vol. 14, No. 12, December 1971, pp. 775-776.
- [MUL 72] Mullin, J.K., "An Improved Index Sequential Access Method Using Hashed Overflows", CACM, Vol. 15, No. 5, May 1972, pp. 301-307.
- [NEL 67] Nelson, D.B., Pick, R.A., Andrews, K.B., "GIM-1 - A Generalized Information Management Language and Computer System", Proceedings Spring Joint Computer Conference, 1967.
- [NIE 72] Nievergelt, J., "Binary Search Trees and File Organization", Proc. 1972 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1972, pp. 165-188.
- [NIJ 71a] Nijssen, G.M., "Efficient Batch Updating of a Random File", Proc. 1971 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1971, pp. 173-186.
- [NIJ 71b] Nijssen, G.M., "Indexed Sequential Versus Random", IFIP IAG Journal, Vol. 4, 1971, No. 1, pp. 29-37.
- [OLL 72] Olle, T.W., "Recent CODASYL Reports on Data Base Management", in Data Base Systems, R. Rustin (Ed.), Prentice-Hall, 1972, pp. 175-184.
- [OLS 69] Olson, C.A., "Random Access File Organization for Indirectly Addressed Records", Proc. ACM 24th National Conference 1969, pp. 539-549.
- [PET 57] Peterson, W.W., "Addressing for Random Access Storage", IBM J. Res. Devel., April 1957, pp. 130-146.
- [PRI 71] Price, C.E., "Table Look-up Techniques", Computing Surveys, Vol. 3, No. 2, June 1971, pp. 49-66.
- [RAD 70] Radke, C.E., "The Use of Quadratic Residue", CACM, Vol. 13, No. 2, February 1970, pp. 103-105.
- [ROB 73] Robertson, W., "Design Flexibility with EDMS and IDP," Proc. 20th Inter. Meeting, EXCHANGE (Xerox User's Group), Dec. 1973, pp. 147-181.
- [RUS 72] Rustin, R. (Ed.), Data Base Systems, Prentice-Hall, 1972.
- [SAL 68] Salton, Gerard, Automatic Information Organization and Retrieval, McGraw-Hill Book Company, New York, 1968.
- [SEN 73] Senko, M.E., Altman, E.B., Astrahan, M.M., and Fehder, P.L., "Data Structures and Accessing in Data-Base Systems", IBM Systems Jour., Vol. 12, No. 1, 1973, pp. 30-93.
- [SHE 72] Shemer, J., Collmeyer, A.J., "Data Base Sharing: A Study of Interference, Roadblock and Deadlock", Proc. 1972 ACM-SIGFIDET Workshop-Data Description, Access and Control, ACM, New York, 1972, pp. 147-164.
- [SHO 69] Shoshani, A., Bernstein, A.J., "Synchronization in a Parallel-Accessed Data Base," CACM, Vol. 12, No. 11, Nov. 1969, pp. 604-607.
- [SIB 73] Sibley, Edgar H., Taylor, Robert W., "A Data Definition and Mapping Language", CACM, Dec. 1973, pp. 750-759.
- [SOF 73] Software AG Corp., ADABAS Reference Manual, 11800 Sunrise Valley Dr., Reston, VA 22091, AM-1519, November 1973.
- [SYS 68] System Development Corp., JOVIAL Programming Language, Santa Monica, 1968.
- [TAY 71] Taylor, R.W., "Generalized Data Base Management System Data Structures and Their Mapping to Physical Storage", University of Michigan, Ph.D., Dissertation, 1971.
- [ULL 70] Ullman, J.D., "The Design of Hashing Functions", Technical Report No. 85, Electrical Engineering Department, Princeton University, September 1970, 19 pages.
- [UNI 73] Univac, Data Management System (DMS 1100), Schema Definition, Sperry-Univac, #UP-7007, Rev. 1, 1973.
- [UNI 72] Univac, Data Management System (DMS 1100), Data Manipulation Language, Sperry-Univac, #UP-7908, 1972.
- [WEI 75] Wiederhold, G.O., Database Management and Schemas, McGraw-Hill, In Press.
- [XER 70] Xerox, Xerox MANAGE, #90 16 10B, July 1970.
- [XER 74] Xerox, Xerox Extended Data Management System (EDMS), Reference Manual, #90 30 12B, Feb. 1974.
- Late addition:
- [DAT 75] Date, C.J., An Introduction to Database Systems, Addison-Wesley, 1975.