An Introduction to Database Systems

Second Edition

C. J. DATE

An Introduction to Database Systems

Second Edition

C. J. DATE
IBM (UK) Laboratories Ltd.

ADDISON-WESLEY PUBLISHING COMPANY
Reading, Massachusetts • Menlo Park, California
London • Amsterdam • Don Mills, Ontario • Sydney

ISBN 0-201-14456-5
CDEFGHIJKL-HA-89876543210

Foreword

The field of systems programming primarily grew out of the efforts of many programmers and managers whose creative energy went into producing practical, utilitarian systems programs needed by the rapidly growing computer industry. Programming was practiced as an art where each programmer invented his own solutions to problems with little guidance beyond that provided by his immediate associates. In 1968, the late Ascher Opler, then at IBM, recognized that it was necessary to bring programming knowledge together in a form that would be accessible to all systems programmers. Surveying the state of the art, he decided that enough useful material existed to justify a significant publication effort. On his recommendation, IBM decided to sponsor The Systems Programming Series as a long term project to collect, organize, and publish principles and techniques that would have lasting value throughout the industry.

The Series consists of an open-ended collection of text-reference books. The contents of each book represent the individual author's view of the subject area and do not necessarily reflect the views of the IBM Corporation. Each is organized for course use but is detailed enough for reference. Further, the Series is organized in three levels: broad introductory material in the foundation volumes, more specialized material in the software volumes, and very specialized theory in the computer science volumes. As such, the Series meets the needs of the novice, the experienced programmer, and the computer scientist.

The Editorial Board

Preface to the First Edition

Computers have already had a considerable impact on many aspects of our society. Medicine, law enforcement, government, banking, education, transportation, planning—these are only some of the fields in which computers are already playing a highly significant role. Over the next few years we can expect a vast increase in the range of computer applications and a corresponding increase in the effect computers will have on our daily lives. The two areas of computer technology that will make the new applications possible—indeed, in most cases they are absolutely fundamental—are telecommunications and the integrated database.

In the years ahead, then, database systems will become increasingly widespread and increasingly important. At present, however, they represent a new and relatively unexplored field, despite the fact that the number of systems installed or under development is growing at considerable speed. There is a real need for a good basic textbook that covers the fundamental concepts of such systems in a clear and concise manner. This book represents an attempt to meet the need.

The reader is assumed to be professionally interested in some aspect of data processing. He or she may, for example, be a systems analyst or designer, an application programmer, a systems programmer, a student following a university or similar course in computer science, or a teacher of such a course. (The book is in fact based on an intensive course in the subject that the author has been teaching to professional staff within IBM over a considerable period of time.) Generally speaking, the reader is expected to have a reasonable appreciation of the capabilities of a modern computer system, with particular reference to the file-handling features of such a system. He or she should also have some knowledge of at least one high-level programming language. Since these prerequisites are not particularly demanding, I am hopeful that the book will prove suitable as an introductory text for anyone concerned with using or implementing a database system, or for anyone who simply wishes to broaden a general knowledge of the computer science field.

The book is divided into six major parts:

1. Database System Architecture
2. The Relational Approach
3. The Hierarchical Approach
4. The Network Approach
5. Security and Integrity
6. Review, Analysis, and Comparisons

Each part in turn is subdivided into a number of chapters. Part 1 provides a general introduction to the concepts of a database system, and in particular outlines three distinct approaches to the design of such a system, namely, the relational, hierarchical, and network approaches. Part 2 then examines the relational approach in considerable detail; Part 3 performs the same function for the hierarchical approach; and Part 4 does the same for the network approach. Part 5 presents a discussion of the problems of security and integrity in a database system. Part 6 draws together some of the more important themes introduced earlier in the book and considers them in somewhat more depth.

The structure just outlined requires some justification. As explained, Part 2 is concerned with the relational approach. In fact, it is largely devoted to an exposition of the ideas of Dr. E. F. Codd, the recognized authority in the relational database field. It is only fair to point out, however, that most commercial systems currently available (1977) are based on one of the other two approaches. Why, then, the emphasis on the relational approach? There are at least two answers to this question.

1. The relational approach may be viewed as the beginnings of a theory of data; as such, it provides an excellent basis for understanding and comparing the other two approaches, and a convenient measure or yardstick against which any existing system can be judged. The soundness and permanence of the theory would make it an ideal vehicle for tutorial purposes, even if it possessed no other advantages.

2. The fact that most existing systems are not relational may be viewed as the natural outcome of the way in which computing technology itself has developed. The comparatively small capacity and high access times of early direct-access devices, the traditional emphasis on sequential media such as tape and cards, the limited amount of storage available in the

computer itself—these and similar considerations had significant repercussions on the original design of most early systems. With modern hardware and techniques, however, it seems possible to design and build a system that does not have the shortcomings of earlier designs. To be more specific, many authorities now believe that the future will see the implementation of one or more large-scale systems based on the relational approach. (Since this was first written, in fact, a number of commercial systems incorporating relational concepts have begun to appear on the scene.)

From these remarks the reader will conclude, quite rightly, that the text is somewhat biased throughout in favor of the relational approach. Obviously the author believes that such a bias is justified; but it would be dishonest not to warn the reader that the bias is there.

Despite the views just expressed, however, the hierarchical and network approaches are obviously extremely important, and they possess the advantage of having several years' experience behind them. Parts 3 and 4 therefore deal with these approaches in some detail (and I hope fairly, in spite of my prejudices). Part 3 is based entirely on an existing system, IBM's Information Management System (IMS), which is already operating successfully in a number of computer installations. This system has been chosen as the basis for Part 3 both because it is a good example of the hierarchical approach and, of course, because it is an important system in its own right. For similar reasons, Part 4 is based on the proposals of the Data Base Task Group (DBTG) of the CODASYL Programming Language Committee. I hope, therefore, that Parts 3 and 4 will serve not only as a general introduction to the hierarchical and network approaches, but also specifically as a tutorial on the IMS and DBTG systems. However, teaching specific systems is not the major aim of the book; rather, the object is to describe some general concepts, using specific systems primarily for purposes of illustration. (For this reason many otherwise important systems are little more than just mentioned.) Even so, the descriptions of IMS and DBTG, in particular, do go into a fair amount of rather specific detail. The reader who is not too interested in the finer detail of these systems may omit certain portions of the text if he or she wishes—principally Chapters 16–18, and certain sections (appropriately indicated) of Chapters 20 and 22. In any case it may be as well to omit these portions on a first reading.

A note on terminology. Like many other new subjects, the field of database systems possesses as yet no commonly agreed nomenclature. In particular, the terminology of IMS differs in many respects from that of DBTG. This book attempts to reconcile the differences by relating both IMS and DBTG terminology to a "neutral" terminology defined in Parts 1 and 2. (Once this has been done, however, the "correct" terminology for each system is generally employed in subsequent discussions.) The terminology of Parts 1 and 2, in turn, is an amalgam derived from many sources.

A few further points about the structure of the book:

1. I have tried to write a textbook, not a reference work. Of course, these two objectives are not wholly incompatible—indeed, I hope that to a large extent both are achieved—but whenever they clash, I have aimed at the first rather than the second. To this end I have not hesitated to omit minor points in the interests of clarity, nor to simplify others for the same reason, although as a general rule I have attempted to be as thorough as possible. (The reader is referred elsewhere for further details where appropriate.)

2. Since it is a textbook, most chapters are followed by a set of exercises, of which the reader is strongly urged to attempt at least a few. Answers, sometimes giving additional information about the subject of the question, will be found at the end of the book.

3. Each chapter is followed by a list of references, many of them annotated. References are identified in the text by numbers in brackets. For example, [1.3] refers to the third item in the list of references at the end of Chapter 1, namely, a paper by the CODASYL Systems Committee published in the *BCS Computer Bulletin,* Vol. 15, No. 4, and also in *Communications of the ACM,* Vol. 14, No. 5.

There remains only the pleasant task of acknowledging the help I have received in writing this book. I am grateful, first, to Dr. Codd for a great deal of encouragement, for permission to make use of much of his published material, particularly in Part 2, and for his helpful comments on the initial draft. The following people also very kindly read that draft and produced many valuable criticisms and suggestions: Joel Aron, Jan Hazelzet, Roger Holliday, Paul Hopewell, Larry Lewis, Salah Mandil, Bill McGee, Herb Meltzer, John Nicholls, Terry Rogers, and Tom Work. I would also like to thank Professor Julius T. Tou, the organizer of the 4th International Symposium on Computers and Information Science (Miami Beach, Florida, 14–16 December 1972), and Plenum Publishing Corporation (publishers of the proceedings) for permission to use a paper I presented at that symposium as the basis for Chapters 3 and 11. Thanks should also be given to the many IBM students whose comments on the original course from which this book is derived have been most helpful. Finally, I am grateful to IBM for allowing much of the work of preparing the book to be done using company time and resources. I must emphasize, however, that I am entirely responsible for the contents of the book; the views expressed are my own and in no way represent an official statement on the part of IBM.

Palo Alto, California C. J. D.
November 1974 (revised January 1977)

Preface to the Second Edition

Many changes have occurred in the field of database development since the first edition was written. The DBTG data manipulation language and sub-schema data description language have been accepted by the CODASYL Programming Language Committee for incorporation into COBOL, and a number of DBTG-based systems are now commercially available. Secondary indexing and various other features have been added to IMS. Commercial systems based on relational concepts have started to become available. Various standardization activities are under way. Perhaps most significant of all, universities and similar institutions throughout the world are displaying an unprecedented level of interest in the subject. The present edition is an attempt to reflect some of this activity. It includes a great deal of new material, which is of course its primary raison d'être; however, the opportunity has also been taken to correct some errors from the first edition and to improve the presentation in many places. Many new references have also been included, most of them with annotation.

Some of the more significant differences between this edition and the previous one are summarized below.

Part 1: The overall systems architecture (Chapter 1) has been revised to incorporate ANSI/SPARC terminology. The comparative presentation of the three approaches (Chapter 3) has been unified and extended.

Part 2: Treatment of the relational data model has been expanded into a separate chapter (Chapter 4); chapters on SEQUEL and Query By Example have been added (Chapters 7 and 8); and Chapter 9 (Further Normalization) has been totally rewritten and includes an improved treatment of third normal form and the new fourth normal form. All other chapters have been considerably revised.

Part 3: Secondary indexing has been introduced (Chapter 18). Other chapters have been revised in accordance with the latest version of IMS.

Part 4: All chapters have been revised to incorporate changes made by the Data Description Language and Programming Language Committees of CODASYL.

Part 5: Both chapters have been revised to incorporate IMS and DBTG changes. Additional relational systems have been included. The treatment of integrity constraints and concurrency (Chapter 24) has been greatly expanded.

Part 6: This section is completely new.

Once again it is a great pleasure to acknowledge the assistance I have received in producing this book. I am particularly pleased to have the chance to thank the many people who commented favorably on the first edition and encouraged me to expand it into its present form. In this regard, I would especially like to mention Ted Codd, Frank King, Ben Shneiderman, and Mike Stonebraker. I am also deeply indebted to the following people for helping me over numerous technical questions and for reviewing and criticizing various portions of the draft of this edition: David Beech, Don Chamberlin, Rod Cuff, Bob Engles, Ron Fagin, Peter Hitchcock, Roger Holliday, Bill Kent, Bill Lockhart, Ron Obermarck, Vern Watts, and Moshe Zloof. As with the previous edition, I am extremely grateful to IBM for supporting me in this work. I would also like to thank Technical Publishing Company, publishers of *Datamation,* for permission to base the revisions in Chapter 3 on an article that appeared in that journal in April 1976; and ACM, for permission to base portions of Chapter 25 on material from three papers (references [22.2], [25.3], and [25.8]) for which ACM holds the copyright. Last but not least, I would like to express my appreciation to the staff of Addison-Wesley for the tremendous enthusiasm, encouragement, and patience they have shown throughout the production of both editions.

San Jose, California
June 1977

C. J. D.

Contents

PART 1
DATABASE SYSTEM ARCHITECTURE

CHAPTER 1
BASIC CONCEPTS

CHAPTER 2
STORAGE STRUCTURES

CHAPTER 3
DATA MODELS AND DATA SUBLANGUAGES

PART 2
THE RELATIONAL APPROACH

CHAPTER 4
THE RELATIONAL DATA MODEL

CHAPTER 5
A DATA SUBLANGUAGE BASED ON RELATIONAL CALCULUS

CHAPTER 6
A DATA SUBLANGUAGE BASED ON RELATIONAL ALGEBRA

CHAPTER 7
THE DATA SUBLANGUAGE SEQUEL

CHAPTER 8
QUERY BY EXAMPLE

CHAPTER 9
FURTHER NORMALIZATION

CHAPTER 10
THE EXTERNAL MODEL

CHAPTER 11
SOME RELATIONAL SYSTEMS

PART 3
THE HIERARCHICAL APPROACH

CHAPTER 12
THE ARCHITECTURE OF AN IMS SYSTEM

CHAPTER 13
THE IMS DATA MODEL

CHAPTER 14
THE IMS EXTERNAL MODEL

CHAPTER 15
THE IMS DATA SUBLANGUAGE

CHAPTER 16
IMS STORAGE STRUCTURES

CHAPTER 17
IMS LOGICAL DATABASES

CHAPTER 18
IMS SECONDARY INDEXING

PART 4
THE NETWORK APPROACH