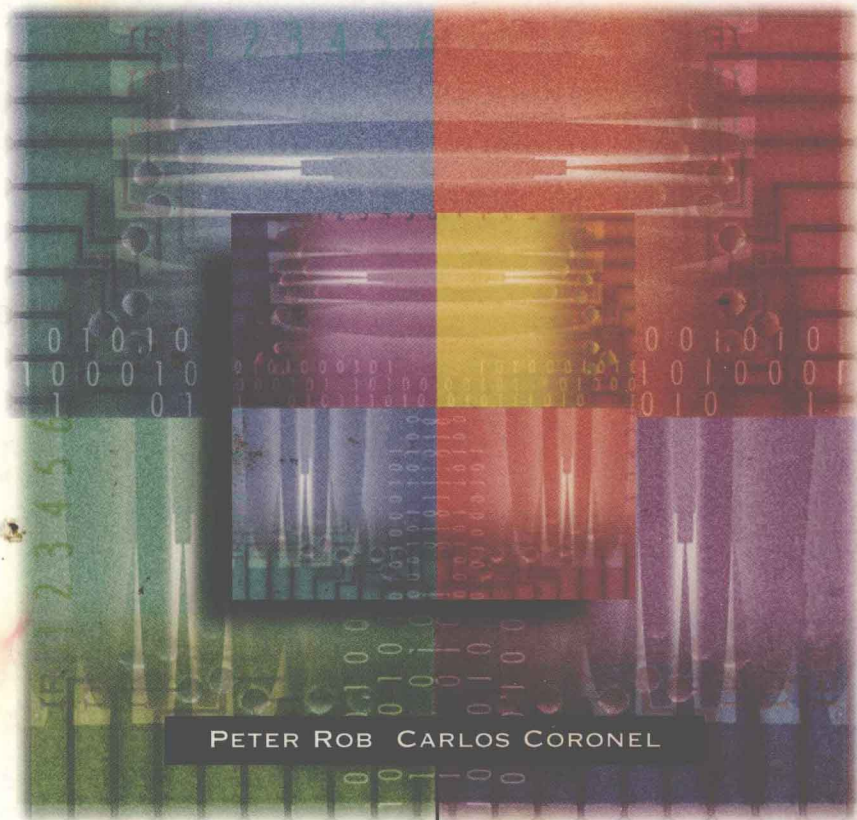THIRD EDITION

DATABASE SYSTEMS

DESIGN, IMPLEMENTATION, AND MANAGEMENT
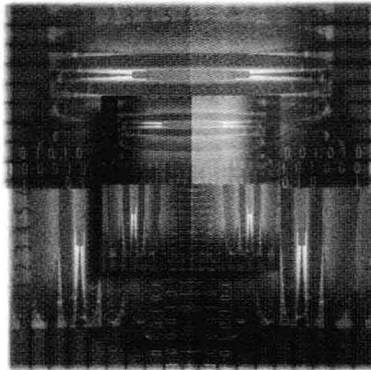
PETER ROB CARLOS CORONEL

THIRD EDITION

DATABASE SYSTEMS

DESIGN, IMPLEMENTATION, AND MANAGEMENT

PETER ROB CARLOS CORONEL

COURSE
TECHNOLOGY

an International Thomson Publishing company I T P®

CREDITS:

Managing Editor	Kristen Duerr	Composition House	GEX, Inc.
Product Manager	Jennifer Normandin	Text & Cover Design	Efrat Reis
Production Editors	Nancy Ray	Marketing Manager	Susanne Walker
	Christine Spillett	3-D Illustrator	Doug Goodman

FOR MORE INFORMATION CONTACT:

Trademarks

Disclaimer

To Anne, who remains my best friend after thirty-five years of marriage. To our son, Peter William, who turned out to be the man we hoped he would be and who proved his wisdom by making Sheena our treasured daughter-in-law. To Sheena, who stole our hearts so many years ago. To our grandson, Adam Lee, who is growing up to be as fine a human being as his parents are. To my in-laws, Han and Nini Fontein, whose life experiences in Europe and Southeast Asia would fill a history book, who trusted me with their daughter's future, and who are a treasured part of our lives. To the memory of my parents, Hendrik and Hermine Rob, who rebuilt their lives after WW II's horrors, who did it again after a failed insurgency in Indonesia, and who finally found their promised land in these United States. And to the memory of Heinz, who taught me daily lessons about loyalty, uncritical acceptance, and boundless understanding. I dedicate this book to you, with love.

Peter Rob

To my parents for my upbringing. To my wife, Victoria, who always brings out the best in me and who keeps me on the right track. And to my children, Carlos Anthony and Gabriela Victoria, whose smiles give me one more reason to be happy.

Carlos Coronel

BRIEF CONTENTS

PREFACE

This third edition continues the theme of the successful first two editions. That is, we continue to provide a solid and practical foundation for the design, implementation, and management of database systems. We build this foundation on the notion that, while good databases are very practical things, their successful creation depends on understanding the important concepts that define them. It's not easy to come up with the proper mix of theory and practice, but we are grateful that many of our second edition's adopter comments and many of our third edition's reviewer evaluations suggest that we have largely succeeded in our quest to provide the proper balance.

In this third edition, we continue to cover current topics in detail. The updated coverage of Client/Server and Object Oriented database systems has been augmented by an in-depth exploration of Data Warehousing. In fact, we not only examine the Data Warehouse's role as a new and powerful basis for decision support, we also provide the details of its design.

As its title suggests, *Database Systems: Design, Implementation, and Management* covers three broad aspects of database systems. However, we believe that the practical aspects of database system *design* and *implementation* deserve special attention for several important reasons:

- The availability of excellent database software enables even database-inexperienced people to create databases and database applications. Unfortunately, the "create without design" approach usually paves the road to any number of database disasters. In our experience, many, if not most, database system failures are traceable to poor design that cannot be solved with the help of even the best programmers and managers. Nor is better DBMS software likely to overcome problems created or magnified by poor design. Using an analogy, even the best bricklayers and carpenters can't create a good building from a bad blueprint.

- Most of the really vexing database system management problems seem to be triggered by poorly designed databases. It hardly seems worthwhile to use scarce resources to develop excellent and extensive database system management skills in order to exercise them on crises induced by poorly designed databases.

- Design provides an excellent means of communication, too. Clients are much more likely to get what they need when database system design is approached carefully and thoughtfully. In fact, clients may even discover how their organizations really function once a good database design is completed.

- Familiarity with database design techniques promotes one's understanding of current database technologies. For example, because Data Warehouses derive much of their data from operational databases, the Data Warehouse concepts, structures, and procedures make much more sense if the operational database's structure and implementation are understood.

In short, database design, far from being a theoretical nicety, is an eminently practical activity that justifies the in-depth coverage we have given it.

Because we stress the practical aspects of database design, we have covered design concepts and procedures in detail and we have made sure that the end-of-chapter problems are sufficiently challenging to let students develop real and useful design skills. We also make sure that students understand

the potential and actual conflicts between database design elegance, information requirements, and transaction processing speed. For example, it makes little sense to design databases that meet design elegance standards while they fail to meet end-user information requirements. Therefore, we explore the use of carefully-defined trade-offs to ensure that the databases are capable of meeting end-user requirements while still conforming to high design standards.

Our third edition reflects comments and suggestions made by the adopters of our second edition and by a host of reviewers who examined our extensive rewriting efforts. In addition, we have very important reviewers, known as students, who can be relied upon to provide useful feedback on what works in the classroom and what doesn't. As teachers, we continued to discover better ways to make a point and we discovered that, in spite of the careful reviewing and editing that preceded the second edition, there were some errors of omission and commission that required redress. The march of database technology required that we address new topics and that we approach some "old" topics differently. Finally, by practicing our database craft in the real world, we found better ways to develop and implement some of the designs and to better explain these ways to others. In short, experience dictated many changes in coverage and the addition of considerable support.

TOPICAL COVERAGE

Our book's title leads with DATABASE SYSTEMS. Therefore, we examine the database and design concepts covered in Chapters 1-5 as part of a larger whole by placing them within Chapter 6's systems analysis framework. We believe that database designers who fail to understand that the database is a part of a larger system are likely to overlook important database design requirements. In fact, Chapter 6 provides the map for the advanced database design we develop in Chapter 7 and 8. Within the larger systems framework, we can also explore such issues as Transaction Management and Concurrency Control (Chapter 9), Distributed Database Management Systems (Chapter 10), Data Warehousing (Chapter 13), and Database Administration (Chapter 14).

The first item in the book's subtitle is DESIGN, and our examination of database design is comprehensive. For example, Chapter 1 illustrates the need for design, Chapters 2 and 3 lay the foundation for the relational database model, while Chapter 4 provides extensive, in-depth, and practical database design coverage. Chapter 5 is entirely devoted to critical normalization issues that affect the database's efficiency and effectiveness. Chapter 6 examines the database design within the systems framework and maps the activities required to successfully design and implement the complex real-world database we develop in Chapters 7 and 8. This background also enables the student to understand Chapter 13's detailed Data Warehousing coverage.

Because database design is affected by the database model, we examine the dominant database models in depth. For example, our coverage of the dominant relational model begins at Chapter 1's end and is expanded in Chapter 2. Chapter 3's SQL coverage demonstrates the relational model's querying power, while Chapters 4 and 5 complete the design and quality control techniques that lead to the relational model's implementation. Chapter 9 addresses transaction management and concurrency control, while Chapter 10 focuses on distributed database management systems. We provide detailed coverage of Object-Oriented Databases in Chapter 11, while Client/Server Systems are covered in depth in Chapter 12. Chapter 13 explores the Data Warehouse in detail and examines how its functions affect design issues.

The second portion of the subtitle is IMPLEMENTATION. Because we used Chapters 7 and 8 to demonstrate the design of a database that was actually and fully implemented, we were forced to deal with a wide range of implementation issues. Naturally, we had to deal with conflicting design goals: Design elegance, information requirements, and operational speed. Therefore, we carefully audited Chapter 7's initial design to check its ability to meet end-user needs and to establish appropriate implementation protocols. The result of this audit yielded the final, implementable design developed in Chapter 8. (This operational database serves as the basis for the first of Chapter 13's Data Warehouse problems.)

The final portion of the subtitle is MANAGEMENT. We deal with database management issues in Chapter 9's Transaction Management and Concurrency Control, Chapter 10's Distributed Database Management Systems, and Chapter 14's Database Administration.

TEACHING DATABASE: A MATTER OF FOCUS

Given the wealth of detailed coverage, instructors may "mix and match" chapters to produce the desired coverage. Depending on where database courses fit into the curriculum, instructors may choose to emphasize database design, current database technologies, or database management tracks. For example, if the instructor wants to focus on database design and implementation issues, Figure 1 may be used as a basis for developing the course syllabus. Although all three tracks share the database and design concept foundation, the design track targets Chapters 6-8 in which a detailed and complex database design is created, verified, and implemented.

FIGURE 1 DATABASE DESIGN AND IMPLEMENTATIONS ISSUES TRACK
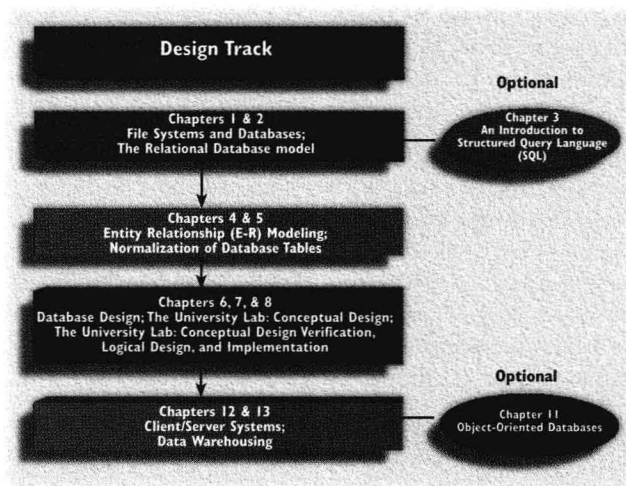
The hands-on nature of the design track lends itself particularly well to class projects in which students use instructor-selected software to prototype a student-designed system for the end user. Several of the end-of chapter problems are sufficiently complex to serve as projects, or the instructor may work with local businesses to give students hands-on experience.

Instructors who want to focus on current database technologies may use Figure 3 as the basis for developing the course syllabus. Using this track, Chapters 6-8 are skipped or are used simply as background reading, while Chapters 9-13 become the focus of course coverage. Because it is difficult to understand the details of current database technologies without a solid foundation in database and design concepts, the coverage of Chapters 1-2 and 4-5 becomes the basis for subsequent course development.
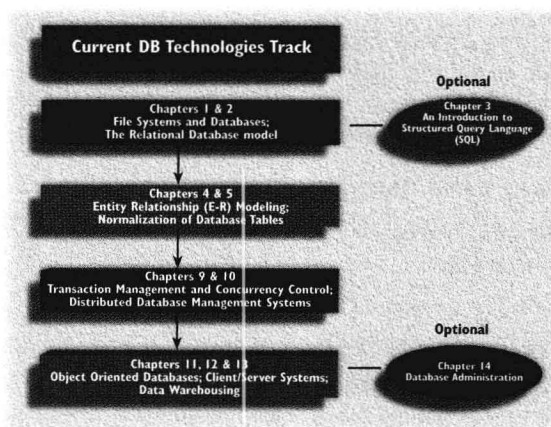

FIGURE 2 CURRENT DATABASE TECHNOLOGIES TRACK

If the instructor wants to focus on database management issues, Figure 3 may be used as a basis for developing the course syllabus.
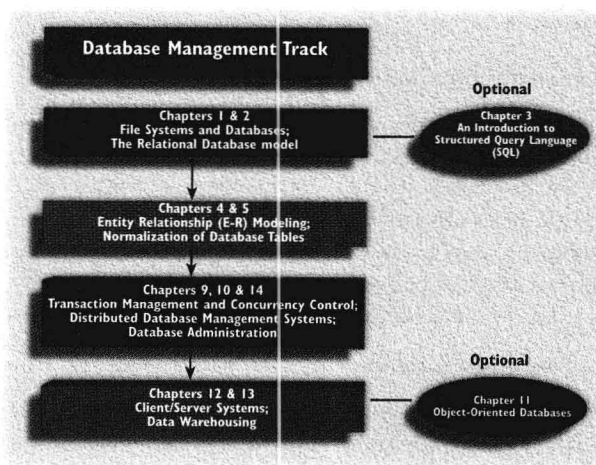

FIGURE 3 DATABASE MANAGEMENT ISSUES TRACK

Note that some elements of the database design and current database technologies tracks are also found in the database management track. This inclusion is dedicated to the notion that it is difficult to manage database technologies that are not understood.

SPECIAL FEATURES AND SUPPORT

We recognize that teaching database systems is hard work. If database coverage is broad, practical, and detailed, the work required includes the actual creation of the designs and the implementation of those designs through the creation of databases, tables, relationships, and so on. To make the teacher's life we have made sure that teachers are given lots of support. Such support includes:

1. DATABASES ARE INCLUDED

Unlike previous editions, this third edition includes all of the database structures and table contents on disk. This feature ensures that

- the database designs developed in the text meet all implementation requirements
- the designs are capable of meeting the stated information requirements
- the illustrations always match the actual database contents
- teachers do not waste their time creating the database structures and the table contents to illustrate the important design and information management principles

Although we have used Microsoft Access 2.0 screens to present the text's material more attractively, *Database Systems: Design, Implementation, and Management* is NOT an Access 2.0 book. Access 2.0 was used here because it is very flexible: the Access 2.0 tables can be exported in a wide variety of formats, including SQL. In fact, to demonstrate some of the SQL features in Chapter 3 we exported our Access 2.0 tables to ORACLE and then used ORACLE's SQL*Plus to produce the desired query results.

2. DETAILED INSTRUCTOR'S MANUAL

Because we tackle so many problems in depth, teachers will find the Instructor's Manual especially useful. Because our book emphasizes the practical aspects of the database environment, we believe it is appropriate to show you the details of the design solution process. Also, where appropriate, we present notes about selected design procedures or alternative approaches that may be used to solve a particular problem. Finally, we decided to reproduce the book's questions and problems together with their answers and solutions in this manual. We hope that this simple service helps make your teaching life simpler and more rewarding.

3. POWERPOINT PRESENTATIONS

To make sure that teachers do not need to create their own presentation materials, the support package includes PowerPoint presentations. These presentations follow each chapter's coverage precisely and include all appropriate illustrations.

4. VIDEO DEMONSTRATIONS

We have prepared video demonstrations in which we examine various aspects of database design, implementation, and applications development. We have covered four topics:

■ *Implementation of a Database Design and Basic Applications Development.* This video illustrates a database design from its E-R model to its implemention. It also examines the end-user view of the database through some sample applications, using a menu-driven interface. Because the presentation is based on two quizzes, one dealing with database design and the other dealing with implementation and applications development, this video may be used as a template for the development of hands-on quizzes and tests.

■ *Exporting MS Access 2.0 Tables to Oracle.* Because database prototyping is often done with the help of microcomputer RDBMSs such as Microsoft Access, we show how database tables produced in Access can be exported to an "industrial strength" RDBMS such as ORACLE. Because the relationships are not exported, we also demonstrate how such relationships can be re-established in Oracle.

■ *Basic ORACLE SQL.* We show how Oracle SQL may be used to extract data from an Oracle database and how to do basic buffer and screen management. We then illustrate how SQL commands can be stored and edited with the help of a user-selected ASCII editor such as Notepad and how such stored SQL commands may be executed. We also demonstrate the use of a script file containing multiple SQL commands. A selection of Chapter 3's SQL problems is demonstrated. (All of Chapter 3's SQL problems are supplied as ASCII files, ready to be executed.)

■ *Using Access 2.0 as a Front End to Oracle.* Because many applications are prototyped through the use of microcomputer RDBMSs such as Microsoft Access, we show how Oracle tables may be attached to an Access database object. We then demonstrate how Access may be used to easily develop the forms, including form/subform formats, queries, and reports may be generated. This approach ensures that the data are maintained and managed at the server level, while the end user applications are quickly and easily developed at the client level.

5. MORE AND BETTER PROBLEMS

We have added many new problems and we have greatly improved their teaching potential. For example, just the first five chapters contain over 140 problems. Problem descriptions and illustrations have been strengthened considerably. As students work their way through the problem sets, the problems become progressively more complex while drawing on the lessons learned from the completion of preceding problems. The development of such comprehensive and carefully-crafted problem sets makes it possible for students to gain the necessary experience to tackle the real world design developed in Chapters 7 and 8. Many of the problems are sufficiently complex to serve as cases or as class projects, especially if they are carried through the implementation and applications development stages. These problems are marked with a T-square-like icon.
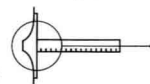
6. INCLUDES DATA WAREHOUSE AND DATA MINING COVERAGE

The new Chapter 13 introduces the Data Warehouse and how it may serve as the foundation for decision support activities. When done properly, Decision Support Systems are designed to provide a computerized interface to enable business decision makers to creatively approach, analyze, and understand business problems. In this chapter, we explore several approaches to the implementation of Decision Support Systems:

■ On-Line Analytical Processing using Relational DBMS

■ Multi-Dimensional Databases

Whatever the approach, it relies on operational data that are extracted, summarized, and stored in some type of Data Warehouse. Because we have covered the design and implementation of transaction-type (that is, operational) databases so thoroughly, we are able to provide practical coverage of the data warehousing environment without losing the student in the fog of the Data Warehouse's complexities.

Given the existence of the Data Warehouse, Data Mining activities become a growing component of new Decision Support Systems. Data Mining systems provide automated tools to perform data extraction and data analysis. The data analysis procedures are designed to identify data relationships, thus setting the stage for uncovering business problems and opportunities.

Because we focus on a hands-on approach to database coverage, we examine the criteria that form the basis for designing and building a data warehouse. Therefore, we are able to move beyond the basic concepts and we can show how the data warehouse "facts and dimensions" are used as the basis for the star schema and the multidimensional cube. Having provided practical examples, we offer students the opportunity to actually design and implement a small Data Warehouse.

7. DESIGNS ARE DEVELOPED TO REACH THE IMPLEMENTABLE LEVEL

As we developed database designs in our own classes, we discovered that too many students encountered problems when they tried to translate some of the designs into appropriate table structures and relationships. Those troubles were traceable to the fact that our relationships were sometimes displayed at the logical level, rather than at the implementation level. We have made sure that we corrected that shortcoming. Optional relationships within the M:N framework were particularly troublesome; although we decomposed the M:N relationship into 1:M relationships with the help of a composite entity, some students found it difficult to migrate the optionality to the new structure. Chapter 4's revised Entity Relationship illustrations, accompanied by examples, appear to have eliminated that hurdle. Teachers who choose to cover Chapter 3's SQL material will also be able to illustrate the impact of properly-defined optional relationships by examining the SQL delete and update requirements written in each table structure definition.

8. COMMONLY-USED GRAPHICAL MODELING DISPLAYS

Although Chen's Entity Relationship (E-R) model is an especially valuable classroom teaching device and end user communications tool, many computer-based modeling tools are based on alternate E-R presentation methods. Because several of our second edition's users suggested that we illustrate at least a few of these methods, we have added coverage of the Crow's Foot, the Rein85, and the IDEF1X methods in the new Chapter 4's Section 4.5. By using a familiar Chen model as the starting point for a simple invoicing system, students have found it easy to become used to the alternative modeling styles.

9. SELF-DOCUMENTING NAMING CONVENTIONS

Given our own experience as database practitioners and as classroom teachers, and given the comments from several of the second edition's users, we have revised all the table structures to improve their self-documentation. Therefore, we introduce attribute naming conventions that ensure the use of descriptive attribute names, that use a prefix that ties the attribute to a table,

and that make it obvious which attributes serve as foreign keys. By consistently adhering to these naming conventions, we make it easier to track the system's attributes during the design process. In addition, the use of such self-documenting naming conventions makes it much easier to develop end-user applications, especially when "drag-and-drop" techniques are used.

10. RELATIONAL SCHEMAS

Because Relational Schemas are an excellent and simple way to illustrate the nature of the relationships we model, we introduce them early to reinforce the modeling lessons. Students encounter their first Relational Schema in Chapter 1's Figure 1.9, see the development of a more complex set of relationships in Chapter 2's Figure 2.13, and graduate to their common use in subsequent chapters.

11. TABLE CONTENTS ARE SHOWN TO ILLUSTRATE THE MORE COMPLEX DESIGN ISSUES

Several users (and more than a few of our own students) indicated that the inclusion of table contents would make the more complex relationships more readily comprehensible. We have, therefore, included such examples: note particularly the expanded discussion of ternary relationships presented in Chapter 4.

12. UPDATED SQL COVERAGE

The feedback we've had indicates that Chapter 3's SQL coverage has become a popular option. But, because we recognize that many teachers do not want to cover SQL, we have made sure that the rest of the book is not dependent on Chapter 3's SQL coverage. Although we have occasionally used SQL to illustrate some topics (such as transaction management and distributed database implementations) in subsequent chapters, proper coverage and understanding of these topics does not require SQL knowledge.

13. COMPREHENSIVE TEST BANK

Because a database course requires hands-on design and implementation work, testing tends to follow that format. However, basic knowledge of concepts, implementation issues, and vocabulary may be tested with the help of a comprehensive test bank supplied to adopters. This test bank contains approximately 150 questions per chapter. The test bank includes graphics from the text and provides a mix of multiple choice, true-false, fill-in-the-blank, and essay questions.

ACKNOWLEDGEMENTS

former Boyd & Fraser Vice president and Executive Editor. Jim's decision to accelerate the book's second edition and to commit the necessary resources greatly contributed to its market presence. Thanks also to Chris Doran, Senior Project Manager, and Barb Worth, Production Editor, whose professional excellence still shines through that second edition.

In some respects, rewriting a book extensively is more difficult than writing the book in the first place. Fortunately, our third edition revision work received an incredibly valuable boost through the efforts of Dr. David Hatherly, a superb database professional who is a Senior Lecturer in the School of Information Technology, Charles Sturt University - Mitchell, Bathhurst, Australia. Dr. Hatherly first transmitted his detailed database class modules to us via the Internet, then used them as the basis for an in-depth review of our book, after first noting that he was "... generally quite pleased with both the book and the response of my students to it." During his review and several subsequent Email messages, Dr. Hatherly not only pointed out some of the errors of commission and omission that we noted earlier but, far better yet, suggested appropriate revisions. Best of all, Dr. Hatherly gave us several suggestions that made it much easier to improve our topical coverage. All of his help was given freely and without prompting on our part. His efforts are much appreciated and our thanks are heartfelt.

Every technical book receives careful scrutiny by several groups of reviewers selected by the publisher. While holding them blameless for any shortcomings, we owe these reviewers many thanks for their contributions to this third edition: Dr. Ahmad Abuhejleh, The University of Wisconsin/River Falls; Dr. Terence M. Baron, The University of Toledo; Dr. Juan Estava, Eastern Michigan University; Dr. Kevin Gorman, The University of North Carolina/Charlotte; Dr. Jeff Hedrington, The University of Wisconsin/Eau Claire; Dr. Herman P. Hoplin, Syracuse University, Dr. Sophie Lee, The University of Massachussetts/Boston; and Dr. Michael Mannino, the University of Washington.

Because this third edition has been carefully constructed on the solid base of its predecessors, we also continue to owe a debt of gratitude to the careful and in-depth reviews provided by Dr. Carol Chrisman, Illinois State University; Dr. Timothy Heintz, Marquette University; Dr. Herman Hoplin, Syracuse University; Dr. Dean James, Embry-Riddle University; Dr. Constance Knapp, Pace University; Dr. Mary Ann Robbert, Bentley College; Dr. Francis J. Van Wetering, The University of Nebraska; Dr. Joseph Walls, The University of Southern California; Dr. Stephen C. Solosky, Nassau Community College; Dr. Robert Chiang, Syracuse University; Dr. Crist Costa, Rhode Island College; Dr. Sudesh M. Duggal, Northern Kentucky University; and Dr. Chang Koh, The University of North Carolina/Greensboro.

In some respects, writing books resembles building construction: When 90% of the work seems done, 90% of the work remains to be done. This observation seems especially appropriate with respect to the third edition. The logistics of complex graphics, screen captures, actual databases, the use of color, and so on, are complex. That is why we gratefully acknowledge the "behind the scenes" work done by Kristen Duerr, managing editor for the MIS list at Course Technology. Authors seldom have much contact with managing editors, yet their writing fortunes are to no small extent dependent on what managing editors do: organize the team that produces the book and help smooth out the inevitable bumps in the road. We owe special thanks to Jennifer Normandin, our MIS product manager, who had ample opportunity to exercise her excellent managing and editing abilities. She also was our all-important and frequently-used pipeline to the rest of the team. In addition, Jennifer spread cheer to occasionally grumpy, often tired, and sometimes discouraged authors. In short, Jennifer is a class act.

Laura Ellingson, development editor extraordinaire, managed to translate the many and detailed reviewer comments into very readable summaries. These absolutely first-rate summaries yielded the necessary patterns that helped guide our many rounds of revisions. Thanks for your upbeat and "can do" approach to the whole process, not to mention your unfailing good humor that frequently stood between us and writer purgatory. And thanks for trusting our professional judgement when it became time to make the final rewrite decisions. Because production editors manage the many and often complex implementation issues, they deal most directly with the authors' concerns. Given the need to combine scanned second edition text, our many inserts and corrections, a large infusion of new material, numerous screen captures, and bountiful illustrations, this third edition turned out to be a much more complex and frequently frustrating endeavor than we had anticipated. Added to our woes was the fact that Christine Spillett, who started our production journey, moved on to manage another project. Fortunately, Nancy Ray stepped into this complex production picture. (We are told she actually *volunteered* to take on this difficult task.) We quickly learned why Nancy's production editing skills are highly regarded. As the work progressed, we discovered that we really hit the production jackpot: quite aside from the fact that Nancy obviously knew her job very well, she also managed to retain her sense of humor while she invested untold overtime hours in this third edition. Many thanks for your fine work, Nancy. You are clearly "top-of-the-line." And how could we fail to notice Cynthia Anderson's truly superb copyediting? Given the incedible numbers of inserts and corrections, it took a really good copyeditor to help ensure that the material came together seamlessly and that the all-important details were preserved. Cynthia managed to get the job done so well that many of our return comments were "good eye!" and "thanks!" Finally, every database professional probably realizes that end users value the interface that connects them to the database. In short, the book's quality of content must be reflected in the book's formatting, its art work, its crucial supplements, and its cover. That's why we acknowledge with thanks the contributions made by Lisa Ayers, editorial assistant, who took care of coordinating the many supplements, Efrat Reis, who designed the book cover, and Angel Thibault at Gex, Inc., who was the book's production liason and who managed the team of scanners, artists, compositors, page formatters, and technology specialists.

We also thank our students for their comments and suggestions. They are the reason for writing this book in the first place. In fact, it may be reasonable to argue that, without the stimulus of students who reasonably expect their instructors to provide well-prepared and up-to-date coursework, writing books constitutes compelling evidence of masochism. Extensive university duties mean that writing is done at night, during week-ends, and during "vacations." Yet there is, of course, a bonus: our writing efforts are based on much research and practice, thus ensuring that our work yields better course materials for our students. We hope that is the reason why universities tend to view writing efforts favorably.

Last, and certainly not least, we thank our families for the solid home support. They graciously accepted that, during more fifteen months' worth of rewriting, there would be no vacation, no free week-ends, rare free nights, and even rarer free days. We owe you much and the dedication we wrote to you is but a small reflection of the important space you occupy in our hearts.

Peter Rob

Carlos M. Coronel

PART I

DATABASE

CONCEPTS

CONTENTS