FUNDAMENTALS OF DATABASE SYSTEMS

7TH Edition

ELMASRI • NAVATHE

FUNDAMENTALS OF

Database Systems

SEVENTH EDITION

This page intentionally left blank

FUNDAMENTALS OF

Database Systems

SEVENTH EDITION

Ramez Elmasri Department of Computer Science and Engineering The University of Texas at Arlington

Shamkant B. Navathe *College of Computing Georgia Institute of Technology*

PEARSON

Boston Columbus Indianapolis New York San Francisco Hoboken Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Vice President and Editorial Director, ECS:	Operations Specialist: Maura Zaldivar-Garcia
Marcia J. Horton	Cover Designer: Black Horse Designs
Acquisitions Editor: Matt Goldstein	Manager, Rights and Permissions: Rachel Youdelman
Editorial Assistant: Kelsey Loanes	Associate Project Manager, Rights and Permissions:
Marketing Managers: Bram Van Kempen, Demetrius Hall	Timothy Nicholls
Marketing Assistant: Jon Bryant	Full-Service Project Management: Rashmi Tickyani,
Senior Managing Editor: Scott Disanno	iEnergizer Aptara [®] , Ltd.
Production Project Manager: Rose Kernan	Composition: <i>iEnergizer Aptara</i> [®] , Ltd.
Program Manager: Carole Snyder	Printer/Binder: Edwards Brothers Malloy
Global HE Director of Vendor Sourcing	Cover Printer: Phoenix Color/Hagerstown
and Procurement: Diane Hynes	Cover Image: Micha Pawlitzki/Terra/Corbis
Director of Operations: Nick Sklitsis	Typeface: 10.5/12 Minion Pro
Director of Operations: Nick Skillsis	Typeface: 10.5/12 Minion Pro

Copyright © **2016**, **2011**, **2007** by Ramez Elmasri and Shamkant B. Navathe. All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright and permissions should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use materials from this work, please submit a written request to Pearson Higher Education, Permissions Department, 221 River Street, Hoboken, NJ 07030.

Many of the designations by manufacturers and seller to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed in initial caps or all caps.

The author and publisher of this book have used their best efforts in preparing this book. These efforts include the development, research, and testing of theories and programs to determine their effectiveness. The author and publisher make no warranty of any kind, expressed or implied, with regard to these programs or the documentation contained in this book. The author and publisher shall not be liable in any event for incidental or consequential damages with, or arising out of, the furnishing, performance, or use of these programs.

Microsoft and/or its respective suppliers make no representations about the suitability of the information contained in the documents and related graphics published as part of the services for any purpose. All such documents and related graphics are provided "as is" without warranty of any kind. Microsoft and/or its respective suppliers hereby disclaim all warranties and conditions with regard to this information, including all warranties and conditions of merchantability. Whether express, implied or statutory, fitness for a particular purpose, title and non-infringement. In no event shall microsoft and/or its respective suppliers be liable for any special, indirect or consequential damages or any damages whatsoever resulting from loss of use, data or profits, whether in an action of contract. Negligence or other tortious action, arising out of or in connection with the use or performance of information available from the services.

The documents and related graphics contained herein could include technical inaccuracies or typographical errors. Changes are periodically added to the information herein. Microsoft and/or its respective suppliers may make improvements and/or changes in the product(s) and/or the program(s) described herein at any time. Partial screen shots may be viewed in full within the software version specified.

Library of Congress Cataloging-in-Publication Data on File

10 9 8 7 6 5 4 3 2 1



ISBN-10: 0-13-397077-9 ISBN-13: 978-0-13-397077-7 To Amalia and to Ramy, Riyad, Katrina, and Thomas R. E.

To my wife Aruna for her love, support, and understanding and to Rohan, Maya, and Ayush for bringing so much joy into our lives S.B.N. This page intentionally left blank

Preface

his book introduces the fundamental concepts necessary for designing, using, and implementing database systems and database applications. Our presentation stresses the fundamentals of database modeling and design, the languages and models provided by the database management systems, and database system implementation techniques. The book is meant to be used as a textbook for a one- or two-semester course in database systems at the junior, senior, or graduate level, and as a reference book. Our goal is to provide an in-depth and up-to-date presentation of the most important aspects of database systems and applications, and related technologies. We assume that readers are familiar with elementary programming and data-structuring concepts and that they have had some exposure to the basics of computer organization.

New to This Edition

The following key features have been added in the seventh edition:

- A reorganization of the chapter ordering (this was based on a survey of the instructors who use the textbook); however, the book is still organized so that the individual instructor can choose to follow the new chapter ordering or *choose a different ordering of chapters* (for example, follow the chapter order from the sixth edition) when presenting the materials.
- There are two new chapters on recent advances in database systems and big data processing; one new chapter (Chapter 24) covers an introduction to the newer class of database systems known as NOSQL databases, and the other new chapter (Chapter 25) covers technologies for processing big data, including MapReduce and Hadoop.
- The chapter on query processing and optimization has been expanded and reorganized into two chapters; Chapter 18 focuses on strategies and algorithms for query processing whereas Chapter 19 focuses on query optimization techniques.
- A second UNIVERSITY database example has been added to the early chapters (Chapters 3 through 8) in addition to our COMPANY database example from the previous editions.
- Many of the individual chapters have been updated to varying degrees to include newer techniques and methods; rather than discuss these enhancements here,

we will describe them later in the preface when we discuss the organization of the seventh edition.

The following are key features of the book:

- A self-contained, flexible organization that can be tailored to individual needs; in particular, *the chapters can be used in different orders* depending on the instructor's preference.
- A companion website (http://www.pearsonhighered.com/cs-resources) includes data to be loaded into various types of relational databases for more realistic student laboratory exercises.
- A dependency chart (shown later in this preface) to show which chapters depend on other earlier chapters; this can guide the instructor who wants to tailor the *order of presentation of the chapters*.
- A collection of supplements, including a robust set of materials for instructors and students such as PowerPoint slides, figures from the text, and an instructor's guide with solutions.

Organization and Contents of the Seventh Edition

There are some organizational changes in the seventh edition as well as improvement to the individual chapters. The book is now divided into 12 parts as follows:

- Part 1 (Chapters 1 and 2) describes the basic introductory concepts necessary for a good understanding of database models, systems, and languages. Chapters 1 and 2 introduce databases, typical users, and DBMS concepts, terminology, and architecture, as well as a discussion of the progression of database technologies over time and a brief history of data models. These chapters have been updated to introduce some of the newer technologies such as NOSQL systems.
- Part 2 (Chapters 3 and 4) includes the presentation on entity-relationship modeling and database design; however, it is *important to note* that instructors can cover the relational model chapters (Chapters 5 through 8) *before Chapters 3 and 4* if that is their preferred order of presenting the course materials. In Chapter 3, the concepts of the Entity-Relationship (ER) model and ER diagrams are presented and used to illustrate conceptual database design. Chapter 4 shows how the basic ER model can be extended to incorporate additional modeling concepts such as subclasses, specialization, generalization, union types (categories) and inheritance, leading to the enhanced-ER (EER) data model and EER diagrams. The notation for the class diagrams of UML are also introduced in Chapters 7 and 8 as an alternative model and diagrammatic notation for ER/EER diagrams.
- Part 3 (Chapters 5 through 8) includes a detailed presentation on relational databases and SQL with some additional new material in the SQL chapters to cover a few SQL constructs that were not in the previous edition. Chapter 5

describes the basic relational model, its integrity constraints, and update operations. Chapter 6 describes some of the basic parts of the SQL standard for relational databases, including data definition, data modification operations, and simple SQL queries. Chapter 7 presents more complex SQL queries, as well as the SQL concepts of triggers, assertions, views, and schema modification. Chapter 8 describes the formal operations of the relational algebra and introduces the relational calculus. The material on SQL (Chapters 6 and 7) is presented before our presentation on relational algebra and calculus in Chapter 8 to allow instructors to start SQL projects early in a course if they wish (it is possible to cover Chapter 8 before Chapters 6 and 7 if the instructor desires this order). The final chapter in Part 2, Chapter 9, covers ER- and EER-to-relational mapping, which are algorithms that can be used for designing a relational database schema from a conceptual ER/EER schema design.

- Part 4 (Chapters 10 and 11) are the chapters on database programming techniques; these chapters can be assigned as reading materials and augmented with materials on the particular language used in the course for programming projects (much of this documentation is readily available on the Web). Chapter 10 covers traditional SQL programming topics, such as embedded SQL, dynamic SQL, ODBC, SQLJ, JDBC, and SQL/CLI. Chapter 11 introduces Web database programming, using the PHP scripting language in our examples, and includes new material that discusses Java technologies for Web database programming.
- Part 5 (Chapters 12 and 13) covers the updated material on object-relational and object-oriented databases (Chapter 12) and XML (Chapter 13); both of these chapters now include a presentation of how the SQL standard incorporates object concepts and XML concepts into more recent versions of the SQL standard. Chapter 12 first introduces the concepts for object databases, and then shows how they have been incorporated into the SQL standard in order to add object capabilities to relational database systems. It then covers the ODMG object model standard, and its object definition and query languages. Chapter 13 covers the XML (eXtensible Markup Language) model and languages, and discusses how XML is related to database systems. It presents XML concepts and languages, and compares the XML model to traditional database models. We also show how data can be converted between the XML and relational representations, and the SQL commands for extracting XML documents from relational tables.
- Part 6 (Chapters 14 and 15) are the normalization and relational design theory chapters (we moved all the formal aspects of normalization algorithms to Chapter 15). Chapter 14 defines functional dependencies, and the normal forms that are based on functional dependencies. Chapter 14 also develops a step-by-step intuitive normalization approach, and includes the definitions of multivalued dependencies and join dependencies. Chapter 15 covers normalization theory, and the formalisms, theories,