COLIN
RITCHIE

DATABASE PRINCIPLES AND DESIGN

THIRD EDITION

@ www.cengage.co.uk/ritchie

Database Principles and Design

Third Edition

Colin Ritchie

COURSE TECHNOLOGY
CENGAGE Learning™

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

COURSE TECHNOLOGY
CENGAGE Learning™

Database Principles and Design
3rd Edition
Colin Ritchie

Publishing Director: John Yates

Publisher: Gaynor Redvers-Mutton

Editorial Assistant: Matthew Lane

Content Project Editor: Alison Walters
Leonora Dawson-Bowling

Production Controller: Maeve Healy

Manufacturing Manager: Helen Mason

Marketing Manager: Jason Bennett

Typesetter: Newgen, India

Cover design: Nick Welch

Text design: Design Deluxe, Bath, UK

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978–1–84480–540–2

Cengage Learning EMEA
High Holborn House, 50-51 Bedford Row
London WC1R 4LR

Cengage Learning products are represented in Canada by Nelson Education Ltd.

For your lifelong learning solutions, visit
www.cengage.co.uk
Purchase e-books or e-chapters at:
http://estore.bized.co.uk

Printed by C & C Offset Printing Co Ltd, China
1 2 3 4 5 6 7 8 9 10 – 10 09 08

Preface

Overview

This text provides an introductory treatment of the principles and practice of relational databases, and is intended for students in HND and degree courses in computing or information technology in which the students are expected to develop competence in designing practical database systems.

No previous knowledge of databases is assumed but some knowledge of general principles of computers and a little programming experience would be helpful. The text is not a guide to any particular database system, although both Oracle, Microsoft Access and the (SQL) syntax of MySQL are used as exemplar systems.

Database design properly forms part of the broader process of systems analysis and design. However, data modelling and the consequent construction of a database are sufficiently distinct tasks to be worthy of treatment in a separate textbook. In addition to aspects of database design, this book also deals with other aspects of databases such as integrity and concurrency.

SQL is one of the most important elements of modern database technology. It is used very extensively by virtually all database systems and is a major vehicle in facilitating inter-database communication. Accordingly, it is important that students studying database principles obtain a good grounding in the language. To this end, this text contains a tutorial on Oracle SQL as well as guidance in using SQL in Access and MySQL. Other laboratory exercises on SQL are available from the text's support website www.cengage.co.uk/ritchie

Microsoft Access is used to demonstrate some of the aspects of database design. While Access is intended only for small scale database applications, it does provide a wide range of features such as forms and report generators that are found in larger systems and it is simple and very convenient to use – in short, it is ideal for student use.

Objectives

The objectives of the text are to provide the following:

- an introduction to the principles underlying relational databases
- instruction in the techniques used to design and develop practical database systems
- an introduction to some of the practical considerations with database systems, such security and integrity, distributed systems, etc.

- instruction in the use of SQL
- an introduction to other database technologies that are currently significant, such as object, object-relational and XML databases.

Summary of chapters

The text consists of thirteen chapters. The two Hands-On sections are somewhat distinct from the others: one is a tutorial on SQL, while the other is a simple introduction to Microsoft Access. The contents of the chapters are indicated below:

1 **Introduction and background**

This chapter provides a non-technical intuitive introduction to the subject.

2 **The relational data model**

The underlying principles of relational databases are introduced.

3 **Conceptual database design**

This chapter describes the process of developing a conceptual design for a proposed database application. Because of the size of this topic, the chapter has been divided into three sections, as follows:

Part 1: Entity-Relationship (ER) model

Part 2: Converting an ER model into a relational database model

Part 3: Normalisation of the relational model

4 **Physical database design**

Aspects of building a practical system based on the conceptual models discussed earlier.

Hands-On Section A: Learning SQL

A practical guide to learning SQL.

5 **Interfacing with the database**

How the database can be accessed using a range of programming tools.

6 **Transactions**

This chapter describes the basic concepts of the transaction mechanism which is fundamental in maintaining system integrity.

7 **Integrity and security**

Considers techniques that are utilised in maintaining the integrity and security of database systems.

8 **Network and distributed systems**

Many if not most database systems now operate in a networked environment. We explore the implications of this for database systems.

Hands-On Section B: Microsoft Access

A practical guide to learning Access.

9 **Post relational databases**

This chapter describes database variants, object and object-relational systems, that provide object-oriented features.

10 **Web databases**

This chapter examines techniques used to provide database facilities within the context of a website.

11 **XML and databases**

XML is an increasingly significant factor in modern computing and this chapter indicates current developments regarding the storage of XML documents in conventional and purpose-built databases.

Appendix Answers to review questions.

Treatment

In order to make the text as readable as possible, each chapter, with the exception of the two Hands-On chapters, has the following structure:

- The chapter starts with a set of learning objectives that indicate what the student should learn from the chapter.
- After the main body of the chapter, a summary section recaps on the most significant points covered in that chapter.
- The summary is followed by a set of review questions. These are questions requiring a short answer and are intended simply to test the student's recall of the material covered. The review questions are answered in the appendix at the end of the book.
- A set of exercises is also provided. These are either appropriate practical exercises, and/or questions that might appear in an examination paper. Answers to exercises are provided on the text's support website.

The Hands-On chapters are intended to support the students' laboratory work:

1 Hands-On Section A: Learning SQL is a fairly extensive tutorial on SQL intended to be used in the laboratory to instruct the student in SQL. A set of exercises

based on the same database is available from the text's support website. These exercises can be attempted by the student after working through the corresponding part of the tutorial.

2 Hands-On Section B: Microsoft Access is a simple introduction to using Microsoft Access to create tables, queries, forms and reports. This might form the basis for project work in Access, in which a full application is generated consisting of tables, transaction forms and business reports.

Supporting material

Additional resources are available on the website, **www.cengage.co.uk/ritchie**. These include:

- Answers to the chapter exercises.
- Sets of presentation slides for lecture use.
- Sample databases corresponding to the case studies introduced in Chapter 1, implemented in Microsoft Access. SQL scripts are also provided to enable the tables for these databases to be constructed in Oracle and MySQL.
- Sample examination-type questions for some topics.
- Sample specifications for project assignments.

Brief contents

1 Introduction and background 1

2 The relational data model 35

3 Conceptual database design 61

4 Physical database design 133

Hands-On Section A: Learning SQL A1–A75

5 Interfacing with the database 157

6 Transactions 179

7 Integrity and security 197

8 Network and distributed systems 215

Hands-On Section B: Microsoft Access B1–B56

9 Post relational databases 231

10 Web databases 273

11 XML and databases 335

Appendix 363

Index 383

Contents

Preface xvii

1 Introduction and background 1

Learning objectives 1
Introduction 2
Information systems analysis and design 2
  Current methodologies 3
  Alternative life-cycle techniques 5
Data storage in computers 6
The database concept 8
Definition of 'database' 10
Data modelling 12
  General properties of data models 13
  Modelling concepts 14
Earlier forms of database 15
  Hierarchical model database 15
  Network model database 17
The relational database 18
Case studies 19
  Case study 1 – Employees 20
  Case study 2 – DVD hire shop 20
  Case study 3 – Training courses company 21
  Case study 4 – Job agency 21
  Case study 5 – Correspondence college 21
Tables in practice 22
A further example 26
  Elements of a practical database system 30
The current database software market 30
Summary 31
Review questions 32
References 33
  Textbooks 33

2 The relational data model 35

Learning objectives 35
The relational model 36
  Sets and relations 36
  Other relational concepts and terminology 39
  Relational algebra 47
  Relational views 55
Summary 56

Review questions 57
Exercises 57
References 59
Textbooks 59
Website 59

3 Conceptual database design 61

Learning objectives 61
Introduction 62
Part 1: Entity-Relationship (ER) model 62
Introduction 62
Entities and relationships 63
Optionality and participation 65
Visualisation of relationships 67
Attributes 68
Many-to-many relationships 71
Weak entities 72
More unusual relationships 73
Case studies 75
Case study 2 – DVD hire shop 76
Case study 4 – Job agency 80
UML system for ER modelling 83
Summary 85
Review questions 85
Exercises 86
Part 2: Converting an ER model into a relational database model 88
Part 2 Overview 88
Deriving a table design from an ER diagram 88
Summary of design process 88
Detailed process 88
Additional techniques 99
Case studies 106
Case study 2 – DVD rental 106
Case study 3 – Job agency 108
Summary 109
Review questions 109
Exercises 110
Part 3: Normalisation 111
Part 3 Overview 111
Overview of normalisation process 111
Normal forms 1NF, 2NF and 3NF 112
Boyce-Codd normal form 121
Fourth normal form (4NF) 123
Higher Forms: Fifth normal form (5NF) and DK/NF 125
Summary of normalisation 125
Case study 126

Summary 128
Review questions 128
Exercises 129
References 132

4 Physical database design 133

Learning objectives 133
Introduction 134
Database management systems 134
  Features of a DBMS 134
  Choice of database 137
  Sample DBMS systems 138
Design of tables 140
  Attribute design 140
  Data redundancy 143
  Indexing 146
Summary 152
Review questions 153
Exercises 153
References 154
  Textbooks 154
  Websites 155

Hands-On Section A: Learning SQL A1–A75

Introduction A2
  Scenario A3
  Implementations A5
SQL implementation differences A7
  Main differences in MySQL A7
  Main differences in Access A10
Session 1 – Simple queries A12
  Introduction A12
  Conditions A13
  Other forms of condition A15
  Logical operators A16
  Multiple logical operators A18
  ORDER BY clause A20
  Set operations A22
Session 2 – Calculations and functions A24
  Calculations A24
  Functions A26
Session 3 – Groups and group functions A33
  Introduction 33
  Group functions A34
  GROUP BY clause A36

HAVING clause A38
Session 4 – Joining tables A39
Introduction A39
Join definition A39
Inner and outer joins A44
SQL92 join formats A46
Session 5 – Subqueries A49
How SQL interprets a subquery A51
Subqueries yielding multiple rows A53
Subqueries producing multiple columns A55
More complex subquery constructions A56
Logical connectives A57
Session 6 – Data Definition Language (DDL) A58
Introduction A58
Creating tables A59
Creating a table from an existing table A60
Changing the structure of a table A61
Deleting a whole table A62
Adding rows to a table A62
Updating data in tables A64
Deleting rows from a table A65
Indexes A65
Constraints A66
Session 7 – Additional SQL features A68
Introduction A68
Views A68
DCL commands A71
Transactions A73
Constraints A74
References A74
Textbooks A74
Websites A74

5 Interfacing with the database 157
Learning objectives 157
Introduction 158
Programming the database 158
Applications of SQL 158
Standard database access interfaces 160
Fourth generation systems 167
User interfaces 171
Forms and report design – overview 171
Elements of forms 172
Report design 175
Summary 177
Review questions 177

Exercises 177
References 178
Textbooks 178
Websites 178

6 Transactions 179
Learning objectives 179
Introduction 180
Transactions 180
Concept and definition 180
ACID properties 181
Commit and rollback 182
Concurrent transactions 184
Problems of concurrency 184
Serialisation of transactions 187
Locking 188
Other locking variants 190
Deadlock 191
Introduction 191
Dealing with deadlocks 192
Summary 193
Review questions 193
Exercises 194
References 195
Websites 195

7 Integrity and security 197
Learning objectives 197
Introduction 198
Threats to the database 198
User erors 198
Software errors 198
Hardware failure 199
Malicious damage 199
Breach of confidentiality 199
Concurrency errors 199
Database integrity 200
Data validation 200
Type checking 201
Validation techniques 202
Assertions and triggers 206
Backup and recovery 208
Backup 208
Transaction logs 209
Checkpoints 209

Database privileges or permissions 210
Facilities in MS Access 210
Facilities in SQL 211
Summary 211
Review questions 212
Exercises 212
Reference 213

8 Network and distributed systems 215

Learning objectives 215
Introduction 216
Client-server systems 216
Concept 216
Three-tier architecture 217
Advantages of client-server approach 218
Peer-to-peer systems 219
Distributed databases 219
Overview 219
Homogeneous and heterogeneous 220
Partitioned, horizontal and vertical 220
Replication 222
Transparency 223
Schema management 223
Query processing 224
Concurrency control 225
Date's 12 objectives 225
Advantages and disadvantages of distributed databases 226
Summary 227
Review questions 228
Exercises 228
References 229
Textbooks 229
Website 229

Hands-On Section B: Microsoft Access B1–B56

Introduction B2
Getting started in Access B2
Creating a new database B3
Creating a table and entering data B4
Creating forms B9
Creating a report B15
Creating a query B18
Summary B20
Exercises B21

More about Access B21
More about tables B21
More about forms B24
Example B26
More about reports B38
Even more Access B43
Overview B43
Menus B44
Sub-forms B45
Query facilities B47
Object references and events B52
References B56

9 Post relational databases 231
Learning objectives 231
Introduction 232
Advantages and limitations of relational databases 233
Advantages 233
Limitations of the relational database 234
The object data model 237
Basic principles 237
Object examples 239
Messages 240
Classes 240
Collections 241
Object identity 241
Inheritance 241
Polymorphism 243
Post-relational systems 244
Object-relational databases 244
SQL:1999 standard 245
Oracle's SQL:1999 245
Intersystem's Caché object-relational database 254
Object-oriented databases 258
Persistence for objects 258
Object identity and references 260
OODBMS standardisation 260
Comparison with RDBMS 261
Current practical systems 262
Object-relational mapping 265
Current significant products 265
Summary 267
Review questions 267
Exercises 268

References 269
Textbooks 269
Websites 270

10 Web databases 273

Learning objectives 273
Introduction 274
Historical background 274
Client-server architecture 275
HTML 276
Common HTML tags 278
Overview of interactive web technologies 281
CGI 282
ASP and ASP.NET 283
Sample ASP scripts 284
Overview of ASP objects 287
ADO – ActiveX Data Objects 294
The development of ASP.Net 303
Background 303
Main features in page design 304
Database access in ASP.Net 308
Maintaining state in web applications 311
Introduction 311
ASP sessions 311
Cookies 311
URL-encoded variables 312
Hidden form variables 312
Database storage 313
Using cookies 313
PHP 314
Introduction 314
Availability 314
General principles 314
PHP coding 315
PHP language 316
Java servlets and JSP 323
Java servlets 323
Java server pages 325
Summary 329
Review questions 330
Client-server architecture 330
HTML forms 330
CGI 330
ASP, ADO 330
ASP.NET 330

Cookies 331
State maintenance 331
PHP 331
JSP/Servelets 331
Exercises 331
General ASP 332
Database 332
Cookies 332
References 333
Textbooks 333

11 XML and databases 335

Learning objectives 335
Introduction 336
XML overview 336
Attributes 338
Associated XML technologies 339
DTD 339
XML Schema 340
XSL 341
XPath 343
XQuery 345
XPointer and Xlink 346
XML databases 347
Introduction 347
XML-enabled databases 348
SQL/XML functions – examples 351
Microsoft Access XML features 353
MySQL XML features 354
Native XML databases (NXD) 355
Overview 355
Advantages and disadvantages of NXD 356
Current XML software implementations 356
XML tools 357
Native XML databases 357
Summary 357
Review questions 358
Exercises 358
References 359
Textbooks 359
Websites 359

Appendix 363

Index 383