

Inhaltsverzeichnis

Vorwort zur ersten Auflage	xv
Vorwort zur zweiten Auflage	xvii
Vorwort zur dritten Auflage	xxi
Vorwort zur vierten Auflage	xxiv
I Einführung	1
1 Motivation. Historisches	3
1.1 Datenintensive Anwendungen und deren Anforderungen	3
1.2 Einordnung und historische Entwicklung des Gebiets „Datenbanken“ . .	5
1.3 Generische Sicht eines Datenbanksystems	9
1.4 Netzwerk- und hierarchische Datenbanken	11
1.5 Relationale Datenbanken	14
1.6 Bibliographische Hinweise und Ergänzungen	19
1.7 Übungen	19
2 Aufbau und Organisation von Datenbanken und Datenbank- systemen	21
2.1 Datenunabhängigkeit	21
2.2 Die ANSI/SPARC-Schemaebenen	22
2.3 Sprachebenen und -klassen	25
2.4 Ein Systemschichtenmodell	28
2.5 Zentrale Systemkomponenten	30
2.6 Bibliographische Hinweise, Ergänzungen, Querbezüge	36
2.7 Übungen	37
3 Physische Architekturkonzepte	39
3.1 Zentralisierte Datenbanksysteme	40
3.2 Datenbank-Serverarchitekturen	41
3.2.1 Das Client/Server-Konzept	41
3.2.2 Anfrageserver	43

3.2.3	Objekt- und Seitenserver	44
3.3	Verteilte Datenbanksysteme	47
3.3.1	Vor- und Nachteile	48
3.3.2	Klassifikationen	49
3.3.3	Homogene Verteilung	51
3.3.4	Heterogene Verteilung	54
3.4	Parallele Datenbanksysteme	57
3.5	Bibliographische Hinweise und Ergänzungen	62
3.6	Übungen	63
 II Datenmodelle und Datenbankentwurf		65
4	Statischer Datenbank-Entwurf mit dem Entity-Relationship-Modell	67
4.1	Der Datenbank-Lebenszyklus	68
4.2	Aspekte der Qualitätssicherung	71
4.3	Phasen des Entwurfsprozesses	72
4.3.1	Anforderungsanalyse und -spezifikation	74
4.3.2	Konzeptioneller Entwurf	75
4.3.3	Logischer Entwurf	77
4.3.4	Physischer Entwurf	77
4.3.5	Weitere Entwurfsschritte	78
4.4	Allgemeine Abstraktionskonzepte	79
4.5	Das Entity-Relationship-Modell	80
4.5.1	Entities und Attribute	81
4.5.2	Relationships	89
4.5.3	IS-A-Beziehungen	95
4.5.4	Zusammenfassung	101
4.6	Konzeptioneller Entwurf mit dem ER-Modell	102
4.7	Bibliographische Hinweise und Ergänzungen	109
4.8	Übungen	110
5	Das relationale Datenmodell	115
5.1	Relationen	115
5.2	Datenabhängigkeiten. Relationenschemata	118
5.2.1	Intrarelationale Abhängigkeiten	119
5.2.2	Relationale Datenbanken	121
5.2.3	Interrelationale Abhängigkeiten	122
5.3	Transformation eines ER-Diagramms in das Relationenmodell	125
5.3.1	Vorbereitende Schritte	125
5.3.2	Transformation	126
5.4	Datenbank-Definition mit SQL	129
5.4.1	Datenbankschema-Definition	130
5.4.2	Tabellen-Definition	131
5.4.3	Integritätsbedingungen	132
5.4.4	Weitere Definitionsaspekte	133
5.5	Fallstudie Buchverlag	135
5.6	Bemerkungen zum Reverse Engineering	142

5.7	Bibliographische Hinweise und Ergänzungen	143
5.8	Übungen	144
6	Integrität in relationalen Datenbanken	147
6.1	Arten von Integritätsbedingungen	147
6.2	Funktionale Abhängigkeiten	152
6.2.1	Definition. Implikation	152
6.2.2	Ableitung. Das Membership-Problem	155
6.2.3	Anwendungen des Membership-Algorithmus. Schlüssel	161
6.2.4	(Abhängigkeits-) Basen	164
6.3	Bemerkungen zur allgemeinen Implikation	167
6.4	Inklusionsabhängigkeiten	169
6.5	Assertions und Trigger	171
6.6	Aktive Datenbanken	177
6.7	Bibliographische Hinweise, Ergänzungen, Querbezüge	180
6.8	Übungen	184
7	Normalisierung und algorithmischer Schema-Entwurf	187
7.1	Selektion, Projektion und Verbundoperation	187
7.2	Universalrelationen und Update-Anomalien	191
7.3	Zweite, dritte und Boyce-Codd-Normalform	193
7.4	Dekomposition und Synthese	199
7.5	Mehrwertige und Verbund-Abhängigkeiten	209
7.6	Vierte und fünfte Normalform	212
7.7	Bibliographische Hinweise und Ergänzungen	215
7.8	Übungen	217
8	Semantische Datenmodellierung und Prozeßentwurf	219
8.1	Grenzen des ER-basierten Datenbank-Entwurfs. Semantische Datenmodellierung	219
8.2	Daten-, Funktions- und Prozeßentwurf	222
8.2.1	Revision des statischen Datenbankentwurfs	223
8.2.2	Geschäftsprozeßmodellierung und Workflow-Management	224
8.3	Das Generische Semantische Modell (GSM)	227
8.3.1	Lokale GSM-Konstrukte	227
8.3.2	Globale GSM-Aspekte	235
8.3.3	Zusammenfassung	236
8.4	Ein Modellierungswerkzeug für dynamische Aspekte	239
8.4.1	Petri-Netze	239
8.4.2	Verwendung von Petri-Netzen in der Modellierung	243
8.5	Einführung in IDEA	246
8.5.1	Übersicht	246
8.5.2	Eine Fallstudie zur Analyse	248
8.6	Bibliographische Hinweise und Ergänzungen	259
8.7	Übungen	260

9 Objektbasierte Modelle	261
9.1 Grundlagen	261
9.1.1 Modellierung stark strukturierter Information	261
9.1.2 Objektorientierung in höheren Programmiersprachen	264
9.1.3 Das Paradigma der Objektorientierung	268
9.2 Kennzeichen objektorientierter Datenmodelle	269
9.3 Einführung in den Object Database Standard	283
9.3.1 Grundzüge des ODMG-Modells	284
9.3.2 Die Objekt-Definitionssprache ODL	288
9.4 Grundzüge objektrelationaler Datenmodelle	291
9.5 Bibliographische Hinweise und Ergänzungen	297
9.6 Übungen	297
III Datenbanksprachen	299
10 Codd-vollständige (relationale) Sprachen	301
10.1 Die Relationenalgebra	301
10.1.1 Relationale Operationen und ihre Rechenregeln	302
10.1.2 Ausdrücke der Relationenalgebra	305
10.1.3 Eigenschaften der Relationenalgebra	310
10.1.4 Algebraische Optimierung	312
10.2 Relationenkalküle	315
10.2.1 Der Relationen-Tupelkalkül	316
10.2.2 Der Relationen-Domänkalkül	321
10.3 Datenbank-Updates. IDM-Transaktionen	322
10.4 View-Updates	328
10.5 Bibliographische Hinweise und Ergänzungen	333
10.6 Übungen	334
11 Der Sprachstandard SQL	339
11.1 Zur Evolution von SQL	339
11.2 Beispieldatenbank	340
11.3 Datenbank-Updates	342
11.4 Datenbank-Anfragen	343
11.4.1 Einführung	343
11.4.2 Zur Syntax von SELECT	345
11.4.3 Semantik von SELECT am Beispiel von RA-Ausdrücken	347
11.4.4 Verbund-Operatoren in SQL	349
11.4.5 Beispiele	353
11.4.6 Abschließende Bemerkungen	360
11.5 Rekursion in DB2 UDB	361
11.6 Bibliographische Hinweise und Ergänzungen	362
11.7 Übungen	363

12 Einschränkungen und Verallgemeinerungen Codd-vollständiger Sprachen	365
12.1 Universalrelationen–Anfragen	365
12.2 Geschachtelte Relationen und ihre Operationen	373
12.2.1 Geschachtelte Strukturen	373
12.2.2 Operationen auf geschachtelten Relationen	376
12.3 Grundlagen von Datalog	385
12.3.1 Motivationen	385
12.3.2 Logik als Anfragesprache	387
12.3.3 Intensionale und extensionale Datenbanken	388
12.3.4 Rekursion	389
12.3.5 Integritätsbedingungen	391
12.3.6 Datalog-Programme und deren Interpretation	391
12.3.7 Ausdruckskraft von Datalog	394
12.4 Weitere Vollständigkeitsbegriffe für Anfragesprachen	395
12.5 Ausdruckskraft und Komplexität von Anfragesprachen	397
12.6 Bibliographische Hinweise und Ergänzungen	401
12.7 Übungen	404
13 Sprachen für Objektdatenbanken	407
13.1 Spezielle Eigenschaften von objektorientierten Datenbanksprachen	407
13.1.1 Objekterhalt vs. Objekterzeugung	408
13.1.2 Zugriff auf komplexe Objekte. Pfad-Ausdrücke	410
13.1.3 Zugriff auf Mengen von Objekten	413
13.1.4 Update-Operationen	414
13.2 Beispiele objektorientierter Sprachen	415
13.2.1 O ₂	415
13.2.2 ObjectStore	420
13.3 Die ODMG-Sprache OQL	422
13.4 Beispiele objektrelationaler Sprachen	424
13.4.1 Postgres	424
13.4.2 UniSQL/X	429
13.5 Bibliographische Hinweise und Ergänzungen	431
13.6 Übungen	432
IV Datenbanksystemtechnik	433
14 Interne Datenbank- und Speicherorganisation	435
14.1 Plattenspeicher	435
14.1.1 Blöcke und Blockzugriffe	437
14.1.2 RAID-Architekturen	438
14.2 Pufferverwaltung	441
14.3 Files	442
14.3.1 Sequentielle Files	443
14.3.2 Effizienter Filezugriff durch Indexierung	445
14.4 Spezielle Indexstrukturen	446
14.4.1 ISAM	446

14.4.2	Baumstrukturen für eindimensionale Punktdaten	448
14.4.3	Baumstrukturen für mehrdimensionale Punktdaten	453
14.4.4	Hash-Organisationsformen	458
14.4.5	Index-Strukturen zur Auswertung von Pfad-Ausdrücken	462
14.5	Beispiel: Speicherorganisation bei DB2	464
14.6	Bibliographische Hinweise und Ergänzungen	467
14.7	Übungen	469
15	Verarbeitung und Optimierung von Anfragen	471
15.1	Anfrageverarbeitung im Überblick	472
15.2	Anfragen in homogen verteilten Datenbanksystemen	473
15.3	Anfrage-Optimierung	476
15.3.1	Ziele und Ansätze	476
15.3.2	High-Level-Optimierung	477
15.3.3	Low-Level-Optimierung	481
15.4	Implementierungstechniken für den Verbund	482
15.4.1	Nested-Loop-Join	482
15.4.2	Sort-Merge-Join	483
15.4.3	Hash-Join	484
15.5	Spezielle QEP-Parameter	487
15.6	Anfrage-Prozessoren	490
15.7	Der Systemkatalog am Beispiel DB2	492
15.8	Bibliographische Hinweise und Ergänzungen	495
15.9	Übungen	495
16	Physischer Datenbankentwurf und Tuning	499
16.1	Grundlagen des physischen Datenbankentwurfs	499
16.2	Fragmentierung in verteilten Datenbanken	503
16.3	Datenbank-Tuning	504
16.3.1	Grundlegende Prinzipien	505
16.3.2	Tuning von Indexen	507
16.3.3	Tuning in relationalen Systemen	507
16.4	Datenbank-Benchmarks	508
16.4.1	Der Wisconsin-Benchmark	509
16.4.2	Die TPC-Benchmarks	510
16.4.3	Der OO7-Benchmark	513
16.5	Bibliographische Hinweise und Ergänzungen	516
16.6	Übungen	516
V	Transaktionsverarbeitung	519
17	Transaktionen und deren korrekte Verarbeitung	521
17.1	Das Transaktionskonzept	522
17.2	Beispiele für Synchronisationsprobleme	525
✓17.3	Das ACID-Prinzip	527
17.4	Das Read-Write-Modell für Transaktionen	529
17.5	Serialisierbarkeit von Schedules	531

17.5.1 Schedules	532
17.5.2 View-Serialisierbarkeit	535
17.5.3 Konflikt-Serialisierbarkeit	537
17.5.4 Commit-Serialisierbarkeit	541
17.6 Fehlersicherheit von Schedules	543
17.6.1 Motivation	544
17.6.2 Rücksetzbarkeit	545
17.6.3 Vermeidung kaskadierender Aborts	545
17.6.4 Striktheit	546
17.7 Korrektheit von Schedules	547
17.8 Bibliographische Hinweise und Ergänzungen	548
17.9 Übungen	548
18 Concurrency Control	551
18.1 Überlegungen zum Scheduler-Entwurf	551
18.2 Sperrende Scheduler	555
18.2.1 Einführung	555
18.2.2 Das 2-Phasen-Sperrprotokoll	557
18.2.3 Varianten des 2-Phasen-Sperrprotokolls	561
18.2.4 Das MGL-Protokoll	563
18.2.5 Das TL-Protokoll	567
18.2.6 Tuning in Gegenwart sperrender Scheduler	569
18.3 Nicht-sperrende Scheduler	572
18.3.1 Zeitmarken-Verfahren	573
18.3.2 Serialisationsgraphen-Tester	574
18.4 Zur Ausnutzung semantischer Information	576
18.5 Bibliographische Hinweise und Ergänzungen	580
18.6 Übungen	581
19 Grundlagen des Recovery	583
19.1 Einführung	583
19.2 Organisation eines Data-Managers	585
19.3 Der Recovery-Manager	587
19.4 Ein Undo/Redo-Protokoll	591
19.5 Alternativen zum Undo/Redo-Protokoll	593
19.6 Recovery-Tuning	594
19.7 Bibliographische Hinweise und Ergänzungen	595
19.8 Übungen	595
20 Transaktionsverarbeitung bei homogener Verteilung	597
20.1 Verteilte Transaktionen und Schedules	597
20.2 Verteiltes Concurrency Control	601
20.2.1 Verteiltes TO	601
20.2.2 Verteiltes 2PL	602
20.2.3 Verteiltes SGT	603
20.2.4 Deadlock-Erkennung und -Auflösung	607
20.3 Verteilte Recovery-Protokolle	609
20.3.1 Atomare Freigaben	609

20.3.2	Das Zwei-Phasen-Commitprotokoll	611
20.3.3	Das Drei-Phasen-Commitprotokoll	616
20.4	Transaktionen auf replikativen Datenbanken	617
20.4.1	Ein-Kopien-Serialisierbarkeit	618
20.4.2	Concurrency Control	622
20.5	Bibliographische Hinweise	623
20.6	Übungen	623
21	Transaktionsverarbeitung bei heterogener Verteilung	627
21.1	Einführung	627
21.2	Ein MDBS-Modell	628
21.3	Transaktionen und Schedules in einem MDBS	630
21.4	Globale Serialisierbarkeit	633
21.5	Gewährleistung globaler Serialisierbarkeit	636
21.5.1	Forcierte Konflikte	636
21.5.2	Rigorousität	637
21.5.3	Freigabe-Ordnung	638
21.6	Alternative Korrektheits-Begriffe	640
21.6.1	Quasi-Serialisierbarkeit	640
21.6.2	Zweistufige Serialisierbarkeit	641
21.7	Bibliographische Hinweise	642
21.8	Übungen	643
VI	Neuere Entwicklungen	645
22	Daten-Integration	647
22.1	Integrierter Datenzugriff	647
22.2	Integration von Datenbank- und anderen Systemen	651
22.2.1	Verteilte Objektverwaltung	652
22.2.2	Die Object Management Architecture	655
22.2.3	CORBA	656
22.3	Integration von Daten aus heterogenen Quellen (im Internet)	660
22.3.1	Datenbanken und Internet	660
22.3.2	Datenbank-Web-Anbindungen	662
22.3.3	Das Web als Datenbanksystem	665
22.3.4	Semistrukturierte Daten	667
22.4	Bibliographische Hinweise und Ergänzungen	669
22.5	Übungen	669
23	Daten-Analyse	671
23.1	Datenwarenhäuser	671
23.1.1	Warenhaus-Anwendungen und -Architektur	673
23.1.2	Warenhaus-Wartung	677
23.2	Grundlagen des OLAP	682
23.2.1	Mehrdimensionale Daten und Datenwürfel	682
23.2.2	Schemaformen für das OLAP	686
23.2.3	Der relationale Cube-Operator	689

23.2.4	Der Cube-Operator bei DB2 UDB	693
23.2.5	Effiziente Speicherung von Faktentabellen	702
23.3	Data Mining	703
23.3.1	Ziele, Motivationen und Anwendungen des Data Mining	703
23.3.2	Klassifikation	705
23.3.3	Assoziationsregeln	706
23.3.4	Berechnung von Assoziationsregeln	709
23.3.5	Analyse von Ereignissequenzen	713
23.3.6	Sprachkonstrukte für das Data Mining	716
23.4	Bibliographische Hinweise und Ergänzungen	718
23.5	Übungen	719
24	Ausblicke	721
24.1	Temporale Datenbanken	721
24.2	Räumliche Daten und GIS	725
24.3	Multimedia-Datenbanksysteme	728
24.4	Stationäre Server mit mobilen Clients	729
24.5	Bibliographische Hinweise und Ergänzungen	731
	Literaturverzeichnis	733
	Sachindex	774