

Содержание

Об авторах	27
Предисловие	28
Как работать с книгой	28
Предварительные условия	28
Упражнения	29
Ресурсы в World Wide Web	29
Благодарности	29
Глава 1. Мир баз данных	31
1.1. Эволюция систем баз данных	32
1.1.1. Первые СУБД	32
1.1.2. Системы реляционных баз данных	34
1.1.3. Уменьшение и удешевление систем	35
1.1.4. Тенденции роста систем	36
1.1.5. Системы “клиент/сервер” и многоуровневые архитектуры	37
1.1.6. Данные мультимедиа	38
1.1.7. Интеграция информации	38
1.2. Обзор структуры СУБД	39
1.2.1. Команды языка определения данных	40
1.2.2. Обработка запросов	40
1.2.3. Менеджеры буферов и хранения данных	41
1.2.4. Обработка транзакций	42
1.2.5. Процессор запросов	43
1.3. Обзор технологий СУБД	45
1.3.1. Проектирование баз данных	45
1.3.2. Программирование приложений баз данных	45
1.3.3. Реализация систем баз данных	47
1.3.4. Интеграция информации	48
1.4. Резюме	49
1.5. Литература	50
Глава 2. Модель данных “сущность-связь”	51
2.1. Элементы ER-модели	52
2.1.1. Множества сущностей	52
2.1.2. Атрибуты	53
2.1.3. Связи	53
2.1.4. Диаграммы сущностей и связей	53
2.1.5. Экземпляры ER-диаграммы	54
2.1.6. Множественность бинарных связей	55

2.1.7. Многосторонние связи	56
2.1.8. Связи и роли	57
2.1.9. Связи и атрибуты	59
2.1.10. Преобразование многосторонних связей в бинарные	60
2.1.11. Подклассы в ER-модели	61
2.1.12. Упражнения к разделу 2.1	63
2.2. Принципы проектирования	66
2.2.1. Достоверность	66
2.2.2. Отсутствие избыточности	67
2.2.3. Простота	67
2.2.4. Выбор подходящих связей	67
2.2.5. Использование элементов адекватных типов	69
2.2.6. Упражнения к разделу 2.2	71
2.3. Моделирование ограничений	74
2.3.1. Классификация ограничений	74
2.3.2. Ключи и ER-моделирование	75
2.3.3. Представление ключей в ER-модели	77
2.3.4. Ограничения уникальности	77
2.3.5. Ограничения ссылочной целостности	78
2.3.6. Ссылочная целостность и ER-диаграммы	79
2.3.7. Ограничения других видов	79
2.3.8. Упражнения к разделу 2.3	80
2.4. Слабые множества сущностей	81
2.4.1. Примеры использования слабых множеств сущностей	81
2.4.2. Требования к слабым множествам сущностей	82
2.4.3. Система обозначений слабых множеств сущностей	84
2.4.4. Упражнения к разделу 2.4	85
2.5. Резюме	85
2.6. Литература	86
Глава 3. Реляционная модель	87
3.1. Основы реляционной модели	87
3.1.1. Атрибуты	88
3.1.2. Схемы	88
3.1.3. Кorteжи	88
3.1.4. Домены	89
3.1.5. Формы представления отношений	89
3.1.6. Экземпляры отношения	89
3.1.7. Упражнения к разделу 3.1	90
3.2. От ER-диаграмм к реляционным схемам	91
3.2.1. От множеств сущностей к отношениям	91
3.2.2. От ER-связей к отношениям	93
3.2.3. Объединение отношений	95
3.2.4. Преобразование слабых множеств сущностей	96
3.2.5. Упражнения к разделу 3.2	99
3.3. Преобразование структур подклассов в отношения	100

3.3.1. Преобразование в стиле “сущность—связь”	101
3.3.2. Объектно-ориентированный подход	102
3.3.3. Значения “null” и объединение отношений	103
3.3.4. Сравнение стратегий	103
3.3.5. Упражнения к разделу 3.3	104
3.4. Функциональные зависимости	106
3.4.1. Определение функциональной зависимости	106
3.4.2. Ключи отношений	108
3.4.3. Суперключи	109
3.4.4. Выбор ключей для отношения	109
3.4.5. Упражнения к разделу 3.4	111
3.5. Правила использования функциональных зависимостей	112
3.5.1. Правила разделения/объединения	113
3.5.2. Тривиальные функциональные зависимости	114
3.5.3. Замыкание множества атрибутов	115
3.5.4. Обоснование алгоритма вычисления замыкания	116
3.5.5. Правило транзитивности	118
3.5.6. Замкнутые множества функциональных зависимостей	119
3.5.7. Проецирование функциональных зависимостей	120
3.5.8. Упражнения к разделу 3.5	121
3.6. Проектирование реляционных схем	123
3.6.1. Аномалии	123
3.6.2. Декомпозиция отношений	124
3.6.3. Нормальная форма Бойса—Кодда	126
3.6.4. Декомпозиция в BCNF	128
3.6.5. Реконструкция данных из отношений декомпозиции	132
3.6.6. Третья нормальная форма	134
3.6.7. Упражнения к разделу 3.6	136
3.7. Многозначные зависимости	137
3.7.1. Независимость атрибутов как причина избыточности	137
3.7.2. Определение многозначной зависимости	138
3.7.3. Правила использования многозначных зависимостей	139
3.7.4. Четвертая нормальная форма	141
3.7.5. Декомпозиция в 4NF	141
3.7.6. Взаимоотношения нормальных форм	143
3.7.7. Упражнения к разделу 3.7	143
3.8. Резюме	145
3.9. Литература	146
Глава 4. Другие модели данных	149
4.1. Обзор понятий объектно-ориентированного проектирования	150
4.1.1. Система поддержки типов	150
4.1.2. Классы и объекты	151
4.1.3. Идентификационный номер объекта	151
4.1.4. Методы	151
4.1.5. Иерархии классов	152
4.2. Введение в язык ODL	152

4.2.1. Объектно-ориентированное проектирование	153
4.2.2. Объявление класса	154
4.2.3. Атрибуты	154
4.2.4. Связи	155
4.2.5. Обратные связи	156
4.2.6. Множественность связей	158
4.2.7. Методы	159
4.2.8. Типы в ODL	161
4.2.9. Упражнения к разделу 4.2	163
4.3. Другие понятия ODL	164
4.3.1. Многосторонние связи	164
4.3.2. Подклассы	165
4.3.3. Множественное наследование	166
4.3.4. Экземпляры	167
4.3.5. Ключи	168
4.3.6. Упражнения к разделу 4.3	170
4.4. От ODL-проектов к реляционным схемам	171
4.4.1. От ODL-атрибутов к атрибутам отношений	172
4.4.2. Неатомарные атрибуты классов	172
4.4.3. Представление атрибутов типа множеств	173
4.4.4. Представление атрибутов других составных типов	175
4.4.5. Представление связей	177
4.4.6. А если ключей просто нет?	178
4.4.7. Упражнения к разделу 4.4	179
4.5. Объектно-реляционная модель	180
4.5.1. От отношений к объектам-отношениям	181
4.5.2. Вложенные отношения	182
4.5.3. Ссылки	183
4.5.4. Объектно-ориентированный и объектно-реляционный подходы: за и против	184
4.5.5. От ODL-проектов к объектно-реляционным решениям	185
4.5.6. Упражнения к разделу 4.5	186
4.6. Полуструктурированные данные	187
4.6.1. Обоснование модели полуструктурированных данных	187
4.6.2. Представление полуструктурированных данных	188
4.6.3. Интеграция информации с помощью модели полуструктурированных данных	189
4.6.4. Упражнения к разделу 4.6	191
4.7. Язык и модель данных XML	192
4.7.1. Семантические тэги	192
4.7.2. Правильные XML-документы	193
4.7.3. Определения типа документа	194
4.7.4. Использование DTD	195
4.7.5. Списки атрибутов	196
4.7.6. Упражнения к разделу 4.7	199
4.8. Резюме	199
4.9. Литература	200

Глава 5. Реляционная алгебра	203
5.1. Пример схемы базы данных	204
5.2. Алгебра реляционных операций	205
5.2.1. Основы реляционной алгебры	206
5.2.2. Теоретико-множественные операции над отношениями	207
5.2.3. Проекция	208
5.2.4. Выбор	209
5.2.5. Декартово произведение	210
5.2.6. Естественное соединение	211
5.2.7. Тета-соединение	212
5.2.8. Наборы операций и формирование запросов	214
5.2.9. Переименование атрибутов	215
5.2.10. Зависимые и независимые операции	216
5.2.11. Алгебраические выражения в линейном представлении	217
5.2.12. Упражнения к разделу 5.2	218
5.3. Реляционные операции над мультимножествами	225
5.3.1. Зачем нужны мультимножества	226
5.3.2. Объединение, пересечение и разность мультимножеств	227
5.3.3. Проекция мультимножеств	228
5.3.4. Выбор из мультимножеств	229
5.3.5. Декартово произведение мультимножеств	230
5.3.6. Операции соединения мультимножеств	230
5.3.7. Упражнения к разделу 5.3	231
5.4. Дополнительные операторы реляционной алгебры	232
5.4.1. Удаление дубликатов	233
5.4.2. Операторы агрегирования	234
5.4.3. Группирование	234
5.4.4. Оператор группирования	235
5.4.5. Оператор расширенной проекции	236
5.4.6. Оператор сортировки	238
5.4.7. Внешние соединения	238
5.4.8. Упражнения к разделу 5.4	241
5.5. Отношения и ограничения	241
5.5.1. Реляционная алгебра как язык описания ограничений	242
5.5.2. Ограничения ссылочной целостности	242
5.5.3. Другие ограничения	243
5.5.4. Упражнения к разделу 5.5	245
5.6. Резюме	246
5.7. Литература	247
Глава 6. Язык SQL	249
6.1. Простые запросы на языке SQL	250
6.1.1. Проекция в SQL	251
6.1.2. Выбор в SQL	253
6.1.3. Сравнение строк	254
6.1.4. Значения даты и времени в SQL	256
6.1.5. Значение NULL и операции сравнения с NULL	258

6.1.6. Значение UNKNOWN	259
6.1.7. Упорядочение результатов	260
6.1.8. Упражнения к разделу 6.1	261
6.2. Запросы к нескольким отношениям	263
6.2.1. Декартово произведение и соединение в SQL	263
6.2.2. Как различать одноименные атрибуты нескольких отношений	264
6.2.3. Переменные кортежей и псевдонимы отношений	265
6.2.4. Интерпретация запросов к нескольким отношениям	266
6.2.5. Объединение, пересечение и разность запросов	268
6.2.6. Упражнения к разделу 6.2	270
6.3. Подзапросы	271
6.3.1. Подзапросы для вычисления скалярных значений	272
6.3.2. Условия уровня отношения	273
6.3.3. Условия уровня кортежа	273
6.3.4. Коррелированные подзапросы	275
6.3.5. Подзапросы в предложениях FROM	276
6.3.6. Выражения соединения в SQL	277
6.3.7. Естественное соединение	278
6.3.8. Внешние соединения	279
6.3.9. Упражнения к разделу 6.3	280
6.4. Операции над отношениями	283
6.4.1. Удаление кортежей-дубликатов	283
6.4.2. Дубликаты в объединениях, пересечениях и разностях отношений	284
6.4.3. Группирование и агрегирование в SQL	285
6.4.4. Операторы агрегирования	285
6.4.5. Группирование	286
6.4.6. Предложения HAVING	287
6.4.7. Упражнения к разделу 6.4	289
6.5. Модификация базы данных	291
6.5.1. Вставка кортежей	291
6.5.2. Удаление кортежей	293
6.5.3. Обновление данных	294
6.5.4. Упражнения к разделу 6.5	295
6.6. Определение схем отношений в SQL	296
6.6.1. Типы данных	296
6.6.2. Простые объявления схем отношений	297
6.6.3. Модификация реляционных схем	298
6.6.4. Значения по умолчанию	299
6.6.5. Индексы	299
6.6.6. Знакомство с технологиями выбора индексов	301
6.6.7. Упражнения к разделу 6.6	304
6.7. Виртуальные таблицы	305
6.7.1. Объявление виртуальной таблицы	305
6.7.2. Запросы к виртуальным таблицам	306
6.7.3. Переименование атрибутов	307

6.7.4. Модификация виртуальных таблиц	308
6.7.5. Интерпретация запросов к виртуальным таблицам	310
6.7.6. Упражнения к разделу 6.7	312
6.8. Резюме	314
6.9. Литература	315
Глава 7. Ограничения и триггеры	317
7.1. Ключи отношений	318
7.1.1. Объявление первичного ключа	318
7.1.2. Объявление ключа UNIQUE	319
7.1.3. Соблюдение ограничений ключей	320
7.1.4. Объявление внешнего ключа	321
7.1.5. Обеспечение ссылочной целостности	322
7.1.6. Отложенная проверка ограничений	324
7.1.7. Упражнения к разделу 7.1	327
7.2. Ограничения уровня атрибутов и кортежей	328
7.2.1. Ограничения NOT NULL	328
7.2.2. Ограничения CHECK уровня атрибута	329
7.2.3. Ограничения CHECK уровня кортежа	330
7.2.4. Упражнения к разделу 7.2	332
7.3. Модификация ограничений	334
7.3.1. Именованное ограничение	334
7.3.2. Модификация ограничений и команда ALTER TABLE	334
7.3.3. Упражнения к разделу 7.3	335
7.4. Ограничения уровня схемы и триггеры	336
7.4.1. Ограничения “общего вида”	336
7.4.2. Правила “событие—условие—действие”	339
7.4.3. Триггеры в SQL	339
7.4.4. Триггеры INSTEAD OF	343
7.4.5. Упражнения к разделу 7.4	344
7.5. Резюме	347
7.6. Литература	348
Глава 8. Системные аспекты SQL	349
8.1. SQL в среде программирования	350
8.1.1. Проблема согласования моделей данных	350
8.1.2. Интерфейс взаимодействия SQL и базового языка	352
8.1.3. Секция объявления	352
8.1.4. Использование общих переменных	353
8.1.5. Команды выбора единственного кортежа	354
8.1.6. Курсоры	355
8.1.7. Модификация данных посредством курсоров	358
8.1.8. Защита данных курсора от внешних воздействий	359
8.1.9. Изменение порядка просмотра содержимого курсора	360

8.1.10. Динамический SQL	361
8.1.11. Упражнения к разделу 8.1	362
8.2. Хранимые процедуры и функции	364
8.2.1. Создание хранимых процедур и функций	364
8.2.2. Простые формы выражений	366
8.2.3. Ветвления	367
8.2.4. Запросы	368
8.2.5. Циклы LOOP	369
8.2.6. Циклы FOR	370
8.2.7. Исключения	372
8.2.8. Использование хранимых процедур и функций	374
8.2.9. Упражнения к разделу 8.2	374
8.3. Среда SQL	376
8.3.1. Среды	377
8.3.2. Схемы	378
8.3.3. Каталоги	379
8.3.4. Клиенты и серверы в среде SQL	380
8.3.5. Соединения	380
8.3.6. Сеансы	381
8.3.7. Модули	381
8.4. Интерфейс уровня вызовов	382
8.4.1. Введение в SQL/CLI	383
8.4.2. Обработка команд	385
8.4.3. Извлечение данных	386
8.4.4. Передача аргументов в SQL-команды	388
8.4.5. Упражнения к разделу 8.4	389
8.5. Java Database Connectivity	390
8.5.1. Введение в JDBC	390
8.5.2. Создание объектов SQL-команд	390
8.5.3. Операции с курсорами	392
8.5.4. Передача аргументов в SQL-команды	393
8.5.5. Упражнения к разделу 8.5	393
8.6. Транзакции и SQL	393
8.6.1. Последовательное выполнение операций	394
8.6.2. Атомарность	396
8.6.3. Транзакции	398
8.6.4. Транзакции “только для чтения”	399
8.6.5. “Чтение мусора”	400
8.6.6. Другие уровни изоляции	403
8.6.7. Упражнения к разделу 8.6	404
8.7. Безопасность и авторизация пользователей в SQL	405
8.7.1. Привилегии доступа	406
8.7.2. Создание привилегий	407
8.7.3. Контроль привилегий	408
8.7.4. Назначение привилегий	409
8.7.5. Диаграммы назначения привилегий	411

8.7.6. Аннулирование привилегий	411
8.7.7. Упражнения к разделу 8.7	415
8.8. Резюме	416
8.9. Литература	418
Глава 9. Объектная ориентация и языки запросов	419
9.1. Введение в язык OQL	419
9.1.1. Пример объектно-ориентированной базы данных	420
9.1.2. Описание зависимости между объектами и их членами	420
9.1.3. Выражения “select–from–where”	421
9.1.4. Изменение типа результата запроса	423
9.1.5. Сложные типы выходных данных	424
9.1.6. Подзапросы	424
9.1.7. Упражнения к разделу 9.1	425
9.2. Дополнительные формы выражений OQL	429
9.2.1. Логические множественные условия	429
9.2.2. Выражения с операторами агрегирования	430
9.2.3. Выражения с оператором группирования	430
9.2.4. Предложения HAVING	432
9.2.5. Операторы объединения, пересечения и разности	433
9.2.6. Упражнения к разделу 9.2	434
9.3. Создание и присваивание объектов в OQL	435
9.3.1. Присваивание значений переменным базового кода	435
9.3.2. Извлечение элементов единичных коллекций	435
9.3.3. Получение элементов произвольных коллекций	436
9.3.4. Константы в OQL	437
9.3.5. Создание объектов	438
9.3.6. Упражнения к разделу 9.3	439
9.4. Типы данных SQL, определяемые пользователем	440
9.4.1. Определение типов в SQL	440
9.4.2. Методы в пользовательских типах	441
9.4.3. Объявление схем отношений с помощью пользовательских типов	442
9.4.4. Ссылочные типы	442
9.4.5. Упражнения к разделу 9.4	444
9.5. Операции над объектно-реляционными данными	445
9.5.1. Обращение по ссылке	445
9.5.2. Доступ к компонентам пользовательских типов	446
9.5.3. Генерирование объектов и модификация компонентов	446
9.5.4. Сопоставление объектов пользовательских типов	448
9.5.5. Упражнения к разделу 9.5	449
9.6. Резюме	451
9.7. Литература	452

Глава 10. Логические языки запросов	453
10.1. Логика отношений	453
10.1.1. Предикаты и атомы	453
10.1.2. Арифметические атомы	454
10.1.3. Правила и запросы Datalog	454
10.1.4. Интерпретация правил Datalog	455
10.1.5. Экстенциональные и интенциональные предикаты	458
10.1.6. Правила Datalog для множеств	459
10.1.7. Упражнения к разделу 10.1	460
10.2. От реляционной алгебры к языку Datalog	460
10.2.1. Пересечение	460
10.2.2. Объединение	461
10.2.3. Разность	461
10.2.4. Проекция	462
10.2.5. Выбор	462
10.2.6. Декартово произведение	464
10.2.7. Соединения	464
10.2.8. Представление наборов операций средствами Datalog	466
10.2.9. Упражнения к разделу 10.2	467
10.3. Рекурсивное программирование на языке Datalog	468
10.3.1. Рекурсивные правила Datalog	469
10.3.2. Вычисление рекурсивных правил Datalog	469
10.3.3. Отрицание в рекурсивных правилах	474
10.3.4. Упражнения к разделу 10.3	477
10.4. Рекурсия в SQL	478
10.4.1. Определение интенциональных отношений в SQL	479
10.4.2. Изолированная рекурсия в SQL	480
10.4.3. Проблемы реализации рекурсии в SQL	482
10.4.4. Упражнения к разделу 10.4	485
10.5. Резюме	486
10.6. Литература	487
Глава 11. Принципы хранения информации	489
11.1. Пример системы баз данных	490
11.1.1. Особенности реализации	490
11.1.2. Как СУБД выполняет запросы	491
11.1.3. Что плохого в нашей системе?	492
11.2. Иерархия устройств памяти	492
11.2.1. Кэш-память	493
11.2.2. Оперативная память	494
11.2.3. Виртуальная память	494
11.2.4. Вторичные устройства хранения	495
11.2.5. Третичные устройства хранения	497
11.2.6. Энергозависимые и энергонезависимые устройства памяти	499
11.2.7. Упражнения к разделу 11.2	500

11.3. Диски	500
11.3.1. Внутренние механизмы дисковых накопителей	500
11.3.2. Контроллер дисков	502
11.3.3. Общие параметры дисков	503
11.3.4. Параметры доступа	504
11.3.5. Запись блоков	508
11.3.6. Модификация содержимого блоков	508
11.3.7. Упражнения к разделу 11.3	509
11.4. Использование вторичных устройств хранения	510
11.4.1. Модель вычислений с функциями ввода-вывода	510
11.4.2. Сортировка данных во вторичных хранилищах	511
11.4.3. Сортировка слиянием	512
11.4.4. Сортировка двухфазным многокомпонентным слиянием	513
11.4.5. Многокомпонентное слияние и сверхбольшие отношения	516
11.4.6. Упражнения к разделу 11.4	517
11.5. Повышение эффективности дисковых операций	518
11.5.1. Группирование данных по цилиндрам диска	519
11.5.2. Использование нескольких дисковых устройств	521
11.5.3. Создание зеркальных копий дисков	522
11.5.4. Упорядочение дисковых операций и “алгоритм лифта”	523
11.5.5. Предварительное считывание и крупномасштабная буферизация данных	526
11.5.6. Приемы оптимизации дисковых операций: за и против	527
11.5.7. Упражнения к разделу 11.5	529
11.6. Отказы дисковых устройств	530
11.6.1. Перемежающиеся отказы	531
11.6.2. Контрольные суммы	531
11.6.3. “Устойчивые хранилища”	532
11.6.4. Устранение последствий ошибок	533
11.6.5. Упражнения к разделу 11.6	534
11.7. Восстановление данных при полном отказе диска	534
11.7.1. Модель отказов дисковых устройств	534
11.7.2. Зеркальные диски как средство резервирования	535
11.7.3. Блоки четности	536
11.7.4. Массивы RAID уровня 5	539
11.7.5. Восстановление данных после отказа нескольких дисков	540
11.7.6. Упражнения к разделу 11.7	543
11.8. Резюме	546
11.9. Литература	548
Глава 12. Представление элементов данных	549
12.1. Элементы данных и поля	549
12.1.1. Представление элементов реляционных баз данных	550
12.1.2. Представление объектов	550
12.1.3. Представление элементов данных	551

12.2. Записи	554
12.2.1. Конструирование записей постоянной длины	554
12.2.2. Заголовки записей	556
12.2.3. Группирование записей постоянной длины в блоках	557
12.2.4. Упражнения к разделу 12.2	558
12.3. Представление адресов записей и блоков	558
12.3.1. Системы “клиент/сервер”	559
12.3.2. Логические и структурированные адреса	560
12.3.3. “Подмена” указателей	561
12.3.4. Сохранение блоков на диске	565
12.3.5. “Закрепленные” записи и блоки	566
12.3.6. Упражнения к разделу 12.3	567
12.4. Элементы данных и записи переменной длины	569
12.4.1. Записи с полями переменной длины	569
12.4.2. Записи с повторяющимися полями	570
12.4.3. Записи переменного формата	571
12.4.4. Записи крупного объема	573
12.4.5. Объекты BLOB	574
12.4.6. Упражнения к разделу 12.4	575
12.5. Модификация записей	576
12.5.1. Вставка	576
12.5.2. Удаление	577
12.5.3. Обновление	579
12.5.4. Упражнения к разделу 12.5	579
12.6. Резюме	580
12.7. Литература	582
Глава 13. Структуры индексов	583
13.1. Индексы для последовательных файлов	584
13.1.1. Последовательные файлы	584
13.1.2. Плотные индексы	585
13.1.3. Разреженные индексы	586
13.1.4. Многоуровневые индексы	588
13.1.5. Дубликаты ключевых значений	589
13.1.6. Управление индексами при модификации данных	592
13.1.7. Упражнения к разделу 13.1	598
13.2. Вторичные индексы	599
13.2.1. Проектирование вторичных индексов	599
13.2.2. Применение вторичных индексов	600
13.2.3. Дополнительные уровни во вторичных индексах	601
13.2.4. Поиск документов и обращенные индексы	603
13.2.5. Упражнения к разделу 13.2	607
13.3. В-деревья	609
13.3.1. Структура В-дерева	609
13.3.2. Применение В-деревьев	612
13.3.3. Поиск в В-деревьях	613

13.3.4. Запросы в диапазонах значений	614
13.3.5. Вставка элементов в B-дерево	615
13.3.6. Удаление элементов из B-дерева	617
13.3.7. Оценки эффективности B-древовидных индексов	620
13.3.8. Упражнения к разделу 13.3	621
13.4. Хеш-таблицы	624
13.4.1. Хеш-таблицы для данных во вторичных хранилищах	624
13.4.2. Вставка записей в хеш-таблицу	625
13.4.3. Удаление записей из хеш-таблицы	626
13.4.4. Оценки эффективности хешированных индексов	626
13.4.5. Расширяемые хеш-таблицы	627
13.4.6. Вставка записей в расширяемую хеш-таблицу	627
13.4.7. Линейные хеш-таблицы	629
13.4.8. Вставка записей в линейную хеш-таблицу	631
13.4.9. Упражнения к разделу 13.4	633
13.5. Резюме	634
13.6. Литература	635
Глава 14. Многомерные и точечные индексы	637
14.1. Приложения модели многомерных данных	638
14.1.1. Географические информационные системы	638
14.1.2. Кубы данных	639
14.1.3. SQL-запросы к многомерным данным	640
14.1.4. Запросы в диапазонах значений с использованием традиционных индексов	641
14.1.5. Поиск ближайших соседних объектов с использованием традиционных индексов	643
14.1.6. Другие ограничения традиционных структур индексов	644
14.1.7. Обзор структур индексов для многомерных данных	645
14.1.8. Упражнения к разделу 14.1	645
14.2. Хеш-подобные структуры для многомерных данных	646
14.2.1. Сеточные файлы	647
14.2.2. Поиск записей в сеточном файле	647
14.2.3. Вставка записей в сеточный файл	648
14.2.4. Оценки эффективности структур сеточных файлов	650
14.2.5. Раздельные хеш-функции	652
14.2.6. Сравнение структур сеточных файлов и раздельных хеш-таблиц	654
14.2.7. Упражнения к разделу 14.2	654
14.3. Древовидные структуры для многомерных данных	657
14.3.1. Индексы с несколькими ключами	657
14.3.2. Оценки эффективности структур индексов с несколькими ключами	659
14.3.3. kd-деревья	659
14.3.4. Операции с kd-деревьями	660
14.3.5. kd-деревья и данные во вторичных хранилищах	663
14.3.6. Деревья квадрантов	664
14.3.7. R-деревья	666

14.3.8. Операции с R-деревьями	667
14.3.9. Упражнения к разделу 14.3	669
14.4. Точечные индексы	671
14.4.1. Зачем нужны точечные индексы	672
14.4.2. Сжатые точечные индексы	673
14.4.3. Операции со сжатыми битовыми векторами	675
14.4.4. Управление точечными индексами	676
14.4.5. Упражнения к разделу 14.4	678
14.5. Резюме	679
14.6. Литература	680
Глава 15. Выполнение запросов	683
15.1. Знакомство с операторами физического плана запроса	684
15.1.1. Сканирование таблиц	685
15.1.2. Сортировка в процессе сканирования таблиц	686
15.1.3. Модель вычислений операторов физического плана	686
15.1.4. Критерии эффективности физических операторов	687
15.1.5. Оценки затрат на ввод-вывод для операторов сканирования	688
15.1.6. Реализация физических операторов с помощью итераторов	689
15.2. Однопроходные алгоритмы обработки данных	692
15.2.1. Однопроходные алгоритмы для унарных операций над отдельными кортежами	693
15.2.2. Однопроходные алгоритмы для унарных операций над полными отношениями	694
15.2.3. Однопроходные алгоритмы для бинарных операций над полными отношениями	697
15.2.4. Упражнения к разделу 15.2	700
15.3. Реализация соединений посредством вложенных циклов	702
15.3.1. Соединение посредством вложенных циклов с загрузкой отдельных кортежей	702
15.3.2. Итератор для соединения посредством вложенных циклов с загрузкой отдельных кортежей	702
15.3.3. Соединение посредством вложенных циклов с загрузкой блоков данных	702
15.3.4. Анализ алгоритмов соединения посредством вложенных циклов	704
15.3.5. Сводка характеристик однопроходных и циклических алгоритмов	705
15.3.6. Упражнения к разделу 15.3	705
15.4. Двухпроходные алгоритмы, основанные на сортировке	706
15.4.1. Алгоритм удаления кортежей-дубликатов	706
15.4.2. Алгоритм группирования и агрегирования	709
15.4.3. Алгоритм объединения	709
15.4.4. Алгоритмы пересечения и разности	710
15.4.5. Простой алгоритм соединения	712
15.4.6. Анализ простого алгоритма соединения	713

15.4.7. Более эффективный алгоритм соединения	714
15.4.8. Сводка характеристик алгоритмов, основанных на сортировке	715
15.4.9. Упражнения к разделу 15.4	716
15.5. Двухпроходные алгоритмы, основанные на хешировании	717
15.5.1. Распределение содержимого отношений по сегментам хеш-таблицы	718
15.5.2. Алгоритм удаления кортежей-дубликатов	718
15.5.3. Алгоритмы группирования и агрегирования	719
15.5.4. Алгоритмы объединения, пересечения и разности	719
15.5.5. Алгоритм соединения	720
15.5.6. Сокращение затрат на дисковый ввод-вывод	721
15.5.7. Сводка характеристик алгоритмов, основанных на хешировании	723
15.5.8. Упражнения к разделу 15.5	724
15.6. Алгоритмы, основанные на индексировании	725
15.6.1. Группирующие индексы	725
15.6.2. Алгоритм выбора	726
15.6.3. Алгоритм соединения с использованием индекса	728
15.6.4. Алгоритмы соединения посредством сортировки и индексирования	729
15.6.5. Упражнения к разделу 15.6	731
15.7. Управление буферизацией	732
15.7.1. Архитектура менеджера буферов	732
15.7.2. Стратегии управления буферизацией	734
15.7.3. Выбор физических операторов и проблемы управления буферизацией	735
15.7.4. Упражнения к разделу 15.7	737
15.8. Многопроходные алгоритмы	738
15.8.1. Многопроходные алгоритмы, основанные на сортировке	738
15.8.2. Характеристики многопроходных алгоритмов, основанных на сортировке	739
15.8.3. Многопроходные алгоритмы, основанные на хешировании	739
15.8.4. Характеристики многопроходных алгоритмов, основанных на хешировании	740
15.8.5. Упражнения к разделу 15.8	741
15.9. Параллельные алгоритмы для реляционных операторов	741
15.9.1. Архитектуры параллельных вычислительных систем	741
15.9.2. Параллельные алгоритмы для операций над отдельными кортежами	744
15.9.3. Параллельные алгоритмы для операций над полными отношениями	745
15.9.4. Характеристики параллельных алгоритмов	746
15.9.5. Упражнения к разделу 15.9	748
15.10. Резюме	749
15.11. Литература	751

Глава 16. Компиляция и оптимизация запросов	753
16.1. Синтаксический анализ	754
16.1.1. Синтаксический анализ и дерева разбора	754
16.1.2. Грамматика простого подмножества языка SQL	755
16.1.3. Препроцессор	759
16.1.4. Упражнения к разделу 16.1	759
16.2. Алгебраические законы и планы запросов	760
16.2.1. Коммутативный и ассоциативный законы	760
16.2.2. Законы выбора	762
16.2.3. Продвижение операторов выбора по дереву выражений	765
16.2.4. Законы проекции	766
16.2.5. Законы соединения и декартова произведения	769
16.2.6. Законы, касающиеся удаления кортежей-дубликатов	769
16.2.7. Законы группирования и агрегирования	770
16.2.8. Упражнения к разделу 16.2	772
16.3. От деревьев разбора к логическим планам запросов	774
16.3.1. От дерева разбора к дереву выражений реляционной алгебры	774
16.3.2. Изъятие подзапросов из условий	775
16.3.3. Улучшение логического плана запроса	780
16.3.4. Группирование ассоциативно-коммутативных операторов	781
16.3.5. Упражнения к разделу 16.3	782
16.4. Анализ стоимости операций	783
16.4.1. Оценка размеров промежуточных отношений	784
16.4.2. Оценка размера результата оператора проекции	785
16.4.3. Оценка размера результата оператора выбора	785
16.4.4. Оценка размера результата оператора соединения	788
16.4.5. Естественное соединение отношений с несколькими общими атрибутами	791
16.4.6. Соединение нескольких отношений	792
16.4.7. Оценка размеров результатов выполнения других операторов	794
16.4.8. Упражнения к разделу 16.4	796
16.5. Выбор планов с учетом их стоимости	797
16.5.1. Статистические характеристики данных	798
16.5.2. Вычисление статистик	801
16.5.3. Эвристические правила сокращения стоимости логических планов	802
16.5.4. Способы сопоставительного анализа физических планов	804
16.5.5. Упражнения к разделу 16.5	807
16.6. Выбор порядка соединения	809
16.6.1. Различия между левыми и правыми операндами соединения	809
16.6.2. Деревья соединения	809
16.6.3. Левосторонние деревья соединения	810
16.6.4. Динамическое программирование и выбор порядка соединения	814
16.6.5. Динамическое программирование и уточненные функции стоимости	819

16.6.6. “Жадный” алгоритм выбора порядка соединения	819
16.6.7. Упражнения к разделу 16.6	821
16.7. Завершение формирования физического плана запроса	822
16.7.1. Поиск метода выбора	822
16.7.2. Поиск метода соединения	824
16.7.3. Организация каналов и материализация: за и против	825
16.7.4. Каналы и унарные операции	826
16.7.5. Каналы и бинарные операции	827
16.7.6. Система представления физических планов запросов	830
16.7.7. Упорядочение физических операторов	832
16.7.8. Упражнения к разделу 16.7	833
16.8. Резюме	834
16.9. Литература	836
Глава 17. Профилактика системных отказов и устранение их последствий	837
17.1. Модели живучести систем баз данных	838
17.1.1. Режимы отказов	838
17.1.2. Подробнее о транзакциях	839
17.1.3. Условия выполнения транзакций	841
17.1.4. Базовые операции транзакций	842
17.1.5. Упражнения к разделу 17.1	845
17.2. Протоколирование в режиме “undo”	846
17.2.1. Записи протокола	846
17.2.2. Правила протоколирования в режиме “undo”	847
17.2.3. Восстановление с применением протокола “undo”	850
17.2.4. Введение контрольных точек	852
17.2.5. Динамическое введение контрольных точек	854
17.2.6. Упражнения к разделу 17.2	857
17.3. Протоколирование в режиме “redo”	858
17.3.1. Правило протоколирования в режиме “redo”	859
17.3.2. Восстановление с применением протокола “redo”	860
17.3.3. Введение контрольных точек в режиме “redo”	861
17.3.4. Восстановление на основе протокола “redo” с контрольными точками	863
17.3.5. Упражнения к разделу 17.3	864
17.4. Протоколирование в режиме “undo/redo”	865
17.4.1. Правила протоколирования в режиме “undo/redo”	865
17.4.2. Восстановление с применением протокола “undo/redo”	866
17.4.3. Введение контрольных точек в режиме “undo/redo”	867
17.4.4. Упражнения к разделу 17.4	869
17.5. Защита от отказа дискового устройства	870
17.5.1. Архив	871
17.5.2. Динамическое архивирование	872
17.5.3. Восстановление с применением архива и протокола	874
17.5.4. Упражнения к разделу 17.5	875
17.6. Резюме	875
17.7. Литература	876

Глава 18. Управление параллельными заданиями	879
18.1. Последовательные и условно-последовательные расписания	880
18.1.1. Расписания	880
18.1.2. Последовательные расписания	881
18.1.3. Условно-последовательные расписания	882
18.1.4. О важности конкретной семантики транзакций	884
18.1.5. Система обозначений для представления транзакций и расписаний	885
18.1.6. Упражнения к разделу 18.1	886
18.2. Условно-последовательное упорядочение с учетом конфликтов	887
18.2.1. Конфликты	887
18.2.2. Графы предшествования и расписания, условно- последовательные с учетом конфликтов	888
18.2.3. Обоснование корректности процедуры проверки графа предшествования	891
18.2.4. Упражнения к разделу 18.2	892
18.3. Последовательные расписания и механизмы блокирования	894
18.3.1. Блокировки	894
18.3.2. Архитектура простого планировщика с блокированием	897
18.3.3. Двухфазное блокирование	897
18.3.4. Обоснование корректности условия двухфазного блокирования	898
18.3.5. Упражнения к разделу 18.3	900
18.4. Системы с несколькими режимами блокирования	901
18.4.1. Общие и монопольные блокировки	902
18.4.2. Матрицы совместимости	904
18.4.3. Повышение уровня блокирования	905
18.4.4. Обновляемые блокировки	906
18.4.5. Инкрементные блокировки	908
18.4.6. Упражнения к разделу 18.4	910
18.5. Архитектура планировщика с блокированием	912
18.5.1. Планировщик заданий и вставка инструкций блокирования	912
18.5.2. Таблица блокировок	915
18.5.3. Упражнения к разделу 18.5	918
18.6. Управление иерархиями элементов базы данных	918
18.6.1. Блокировки с множеством степеней детализации	918
18.6.2. Предупреждающие блокировки	919
18.6.3. Фантомные кортежи и корректная обработка операций вставки	923
18.6.4. Упражнения к разделу 18.6	924
18.7. Протокол блокирования древовидных структур	924
18.7.1. Модель блокирования древовидных структур	924
18.7.2. Правила доступа к элементам древовидных структур	925
18.7.3. Обоснование корректности протокола блокирования древовидных структур	927

18.7.4. Упражнения к разделу 18.7	929
18.8. Хронометраж действий транзакций	930
18.8.1. Хронологические признаки	931
18.8.2. Физически неосуществимые расписания	931
18.8.3. Операции “чтения мусора”	932
18.8.4. Правила упорядочения действий с помощью механизма хронометража	934
18.8.5. Множественные версии элементов и хронологических признаков	936
18.8.6. Хронометраж и блокирование: за и против	938
18.8.7. Упражнения к разделу 18.8	939
18.9. Проверка достоверности действий транзакций	939
18.9.1. Архитектура планировщика с функциями проверки достоверности действий транзакций	939
18.9.2. Правила проверки достоверности действий транзакций	940
18.9.3. Упражнения к разделу 18.9	943
18.10. Сравнение схем управления параллельными заданиями	944
18.11. Резюме	945
18.12. Литература	947
Глава 19. Дополнительные аспекты управления транзакциями	949
19.1. Последовательное упорядочение операций и восстановление данных	949
19.1.1. Проблема “чтения мусора”	950
19.1.2. Каскадный откат	952
19.1.3. Восстановимые расписания	952
19.1.4. Расписания без каскадного отката	953
19.1.5. Управление процедурами отката с помощью схем блокирования	954
19.1.6. Групповая фиксация	956
19.1.7. Логическое протоколирование	957
19.1.8. Восстановление данных с помощью логического протокола	959
19.1.9. Упражнения к разделу 19.1	961
19.2. Условно-последовательное упорядочение с учетом источников данных	962
19.2.1. Эквивалентность расписаний с учетом источников данных	962
19.2.2. Теоретико-графовая модель расписаний, условно-последовательных с учетом источников данных	964
19.2.3. Обоснование корректности процедуры проверки графа предшествования	967
19.2.4. Упражнения к разделу 19.2	968
19.3. Разрешение взаимоблокировок	968
19.3.1. Обнаружение взаимоблокировок по истечении лимита времени	968
19.3.2. Графы ожидания	969

19.3.3. Предотвращение взаимоблокировок за счет упорядочения элементов данных	971
19.3.4. Выявление взаимоблокировок с помощью механизма хронометража	972
19.3.5. Сравнение методов обнаружения и предотвращения взаимоблокировок	975
19.3.6. Упражнения к разделу 19.3	976
19.4. Распределенные базы данных	977
19.4.1. Концепция распределенных данных	977
19.4.2. Распределенные транзакции	979
19.4.3. Репликация данных	979
19.4.4. Распределенная оптимизация запросов	980
19.4.5. Упражнения к разделу 19.4	981
19.5. Распределенная фиксация	981
19.5.1. Поддержка распределенной атомарности	982
19.5.2. Двухфазная фиксация	982
19.5.3. Восстановление распределенных транзакций	985
19.5.4. Упражнения к разделу 19.5	986
19.6. Распределенное блокирование	987
19.6.1. Централизованное блокирование	988
19.6.2. Модель трудоемкости алгоритмов распределенного блокирования	988
19.6.3. Блокирование реплицированных элементов	989
19.6.4. Режим блокирования “главной копии”	990
19.6.5. Глобальные блокировки на основе локальных блокировок	991
19.6.6. Упражнения к разделу 19.6	992
19.7. Длинные транзакции	993
19.7.1. Проблемы использования длинных транзакций	993
19.7.2. Хроники	995
19.7.3. Уравновешивающие транзакции	996
19.7.4. Обоснование корректности модели уравновешивающих транзакций	997
19.7.5. Упражнения к разделу 19.7	998
19.8. Резюме	999
19.9. Литература	1001
Глава 20. Интеграция информации	1003
20.1. Обзор технологий интеграции информации	1003
20.1.1. Проблемы интеграции информации	1004
20.1.2. Федеративные базы данных	1005
20.1.3. Хранилища данных	1007
20.1.4. Медиаторы	1009
20.1.5. Упражнения к разделу 20.1	1010
20.2. Компоненты-оболочки	1012
20.2.1. Шаблоны запросов	1012
20.2.2. Генераторы оболочек	1013

20.2.3. Фильтры	1014
20.2.4. Другие функции оболочек	1015
20.2.5. Упражнения к разделу 20.2	1017
20.3. Медиаторы и оптимизация с учетом возможностей источников данных	1017
20.3.1. Проблема ограниченных возможностей источника данных	1018
20.3.2. Классификация возможностей источника данных	1019
20.3.3. Выбор плана запроса с учетом возможностей источника данных	1020
20.3.4. Дополнительная оптимизация планов с учетом их стоимости	1022
20.3.5. Упражнения к разделу 20.3	1022
20.4. Оперативная аналитическая обработка данных	1023
20.4.1. Приложения OLAP	1024
20.4.2. Многомерная модель данных OLAP	1025
20.4.3. Схема “звезды”	1026
20.4.4. Рассечение и расслоение куба исходных данных	1028
20.4.5. Упражнения к разделу 20.4	1030
20.5. Кубы данных	1031
20.5.1. Оператор CUBE	1031
20.5.2. Реализация куба на основе материализованных представлений	1033
20.5.3. Сетки представлений	1036
20.5.4. Упражнения к разделу 20.5	1037
20.6. Разработка данных	1039
20.6.1. Приложения разработки данных	1039
20.6.2. Отыскание часто встречающихся наборов объектов	1042
20.6.3. Априорный алгоритм анализа рыночной корзины	1043
20.6.4. Упражнения к разделу 20.6	1045
20.7. Резюме	1046
20.8. Литература	1047
Послесловие редактора перевода	1049
Именной указатель	1053
Предметный указатель	1057