

Предисловие к восьмому изданию

Настоящая книга представляет собой исчерпывающее введение в очень широкую в настоящее время область теории систем баз данных. С ее помощью читатель сможет приобрести фундаментальные знания в области технологии баз данных, а также ознакомиться с направлениями, по которым эта область, вероятно, будет развиваться в будущем. Книга предназначена для использования в основном в качестве учебника, а не справочника (но я надеюсь, что ее в какой-то мере можно будет использовать и в качестве справочного руководства). В книге сделан акцент на усвоении сути и глубоком понимании излагаемого материала, а не просто на его формальном изложении.

ДЛЯ КОГО ПРЕДНАЗНАЧЕНА ЭТА КНИГА

В целом, книга ориентирована на читателей, которые профессионально работают с компьютерами в той или иной области и хотят получить общее представление о теории и практическом использовании систем баз данных. Предполагается, что читатель имеет, по крайней мере, базовые знания в следующих областях:

- средства управления памятью и файлами (индексация и т.д.) в современных компьютерных системах;
- средства хотя бы одного из языков программирования высокого уровня (таких как Java, Pascal, PL/I и т.д.).

Что касается ознакомления с первой из указанных областей, то подробное учебное руководство по относящимся к ней темам приведено в приложении Г, “Структуры хранения и методы доступа”.

Структура книги

Автор хотел бы подчеркнуть свою обеспокоенность тем, что восьмое издание этой книги приобрело такой большой объем. Тем не менее, технология базы данных стала, несомненно, очень обширной областью знаний, поэтому невозможно рассмотреть всю относящуюся к ней проблематику в книге, которая занимает меньше 1000 страниц (и действительно, большинство книг, которые конкурируют на рынке компьютерной литературы с данной книгой, также имеют объем приблизительно 1000 страниц). Данная книга разделена на шесть частей.

- Часть I. Основные понятия.
- Часть II. Реляционная модель.
- Часть III. Проектирование базы данных.
- Часть IV. Управление транзакциями.

- Часть V. Дополнительные темы.
- Часть VI. Объекты, отношения и язык XML.

Каждая часть в свою очередь разделена на несколько глав.

- Часть I (четыре главы) — это обширное введение в теорию систем баз данных вообще и реляционных систем в частности. Здесь также излагаются основы стандартного языка баз данных **SQL**.
- Часть II (шесть глав) содержит подробное и весьма основательное описание **реляционной модели**, которая является теоретической основой не только для собственно реляционных систем, но фактически для всей сферы применения баз данных.
- Часть III (четыре главы) включает обсуждение основных проблем **проектирования баз данных**. Три главы посвящены теории проектирования, а в четвертой рассматриваются семантическое моделирование и модель “сущность–связь”.
- Часть IV (две главы) посвящена описанию средств **управления транзакциями** (т.е. средств восстановления и обеспечения параллельной работы).
- Часть V (восемь глав) охватывает довольно разнообразную тематику. Но в целом в ней показано, что с помощью реляционных понятий можно гораздо проще описать различные области применения технологии баз данных, включая **защиту данных, распределенные базы данных, хронологические данные, системы поддержки принятия решений** и т.д.
- Наконец, в части VI (три главы) показано, какие следствия вытекают из стремления применить **объектную технологию** в системах баз данных. Глава 25 посвящена исключительно описанию **объектных систем**, в главе 26 рассматривается возможность сближения между объектной и реляционной технологиями и обсуждаются принципы построения **объектно-реляционных систем**, а в главе 27 приведено описание того, как связаны между собой объектная технология и базы данных **XML**.

Кроме того, в книге имеется пять приложений. В приложении А приведен краткий обзор принципиально новой, полностью отличной от существующих технологии реализации, получившей название **модель TransRelational™**, в приложении Б представлена грамматика выражений SQL в **форме Бэкуса–Наура (БНФ)**, в приложении В приведен **гlossарий** наиболее важных сокращений и специальных символов, применяемых в тексте книги. Приложение Г, как уже было указано выше, включает учебное руководство по **структурам хранения и методам доступа**, а в приложении Д можно найти ответы к некоторым упражнениям, приведенным в главах книги.

ДОПОЛНИТЕЛЬНЫЕ МАТЕРИАЛЫ, ДОСТУПНЫЕ В ОПЕРАТИВНОМ РЕЖИМЕ

Представленные в этом разделе дополнительные материалы предназначены только для квалифицированных преподавателей. Для получения информации о том, как приобрести эти учебные материалы, обратитесь в местное торговое представительство издательства Addison-Wesley или отправьте по электронной почте письмо по адресу aw.cse@aw.com.

- В руководстве для преподавателей *Instructor's Manual* даны указания о том, как использовать эту книгу в ходе преподавания учебного курса, посвященного базам данных. В нем приведен ряд примечаний, советов и предложений по каждой части, главе и приложению, а также другой вспомогательный материал.
- Ответы к упражнениям (включены в руководство *Instructor's Manual*).
- Слайды в формате PowerPoint для демонстрации в ходе лекций.

Оригиналы следующих дополнений доступны всем читателям этой книги по адресу <http://www.aw.com/cssupport> (а перевод представлен в данной книге).

- Приложение Г, где представлен материал по структурам хранения и методам доступа.
- Ответы к отдельным упражнениям (приложение Д).

КАК ЧИТАТЬ ЭТУ КНИГУ

В целом, книга рассчитана на последовательное чтение, но можно пропустить последние главы и последние разделы внутри глав по своему усмотрению. Ниже приведен рекомендуемый план первого чтения.

- Бегло прочитать главы 1 и 2.
- Очень внимательно изучить главы 3 и 4 (возможно, кроме разделов 4.6 и 4.7).
- Бегло прочитать главу 5.
- Внимательно прочитать главы 6, 7, 9 и 10, но пропустить главу 8 (возможно, кроме раздела 8.6, посвященного языку SQL).
- Бегло прочитать главу 11.
- Внимательно прочитать главы 12 и 14¹, но пропустить главу 13.
- Внимательно прочитать главы 15 и 16 (возможно, кроме раздела 15.6 с описанием двухфазной фиксации).
- Прочитать следующие главы выборочно (но последовательно), по своему усмотрению.

Каждая глава начинается с введения и оканчивается заключением (резюме). Кроме того, в большинство глав включены упражнения. Ответы к некоторым упражнениям включены в приложение Д. Рекомендуется просматривать ответы к упражнениям, так как в них часто содержится дополнительная полезная информация по теме конкретной главы. Кроме того, в большинстве глав приведены обширные списки литературы, которые чаще всего дополнены комментариями. Такая структура книги позволяет усваивать материал на нескольких уровнях, поскольку наиболее важные понятия и результаты приведены в основном тексте, а дополнительные вопросы и более сложные аспекты исследований вынесены в соответствующие упражнения, ответы к ним и комментарии к литературе.

Примечание. Ссылки обозначаются в тексте двойным номером в квадратных скобках. Например, ссылка [3.1] означает первый пункт в списке литературы к главе 3, а именно статью Э.Ф. Кодда, опубликованную в феврале 1982 года в журнале *SACM*, том 25, № 2. (Объяснение сокращений, которые используются в ссылках, например, “*SACM*”, можно найти в приложении В.)

¹ При желании можно приступить к изучению главы 14 еще при первом чтении, сразу после главы 4.

СРАВНЕНИЕ С ПРЕДЫДУЩИМИ ИЗДАНИЯМИ

Ниже перечислены главные отличия настоящего издания от непосредственно предшествующего ему издания.

- **Часть I.** В главах 1–4 рассматриваются примерно те же темы, что и в главах 1–4 седьмого издания, но они были в значительной степени пересмотрены на уровне изложения конкретных сведений. В частности, в главе 4, в которой дано введение в SQL, вместо прежнего стандарта рассматривается современный стандарт SQL:1999, впрочем, как и везде в данной книге, где речь идет о языке SQL. (Сам этот факт стал причиной существенного пересмотра больше чем половины глав седьмого издания.)

Примечание. По мере необходимости упоминаются также средства, которые, по всей вероятности, будут включены в следующую версию стандарта, получившую условное название SQL:2003.

- **Часть II.** Главы 5–10, посвященные описанию реляционной модели, представляют собой полностью переработанные, значительно расширенные и существенно улучшенные версии глав 5–9 седьмого издания. В частности, сведения о типах (известных также как домены) были вынесены в отдельную главу (глава 5), а материал, который относится к проблеме целостности (глава 9), был заново продуман и полностью переработан. Автор спешит добавить, что изменения в этих главах являются следствием не пересмотра основополагающих концепций, а скорее пересмотра самого подхода, выбранного для их представления, на основании практического опыта преподавания этого материала для широкой аудитории.

Примечание. В связи с этим необходимо привести некоторые пояснения. В предыдущих изданиях данной книги в качестве основы изложения реляционных концепций использовался язык SQL в надежде на то, что студентам будет проще вначале освоить конкретный материал, а после этого приступить к изучению абстрактных понятий. Но, к сожалению, “пропасть” между языком SQL и реляционной моделью продолжает расти, поэтому в конечном итоге автор пришел к выводу, что дальнейшее применение языка SQL для описания реляционной модели способно лишь ввести читателей в заблуждение. К сожалению, теперь язык SQL настолько далек от того, чтобы быть истинным воплощением реляционных принципов (он страдает от слишком многих недоделок и переделок), что автор выражает свое искреннее пожелание вообще не вступать в обсуждение этой темы! Тем не менее, язык SQL, безусловно, остается важным с практической точки зрения. Поэтому любой профессионал в области баз данных должен иметь определенное представление об этом языке, кроме того, его нельзя полностью игнорировать в книге такого характера. Поэтому автор принял следующее общее решение: во-первых, включить главу по основам SQL в часть I этой книги, во-вторых, по мере необходимости вводить в главы других частей отдельные разделы с описанием тех средств SQL, которые относятся к теме рассматриваемой главы. Благодаря этому в данной книге удалось привести исчерпывающее описание тематики SQL (которое действительно занимает очень большой объем), но это описание дано в таком контексте, который автор считает наиболее приемлемым.

- **Часть III.** Главы 10–13 этой части представляют собой в основном откорректированные версии глав 9–12 из предыдущего, седьмого издания. Но на уровне изложения конкретных сведений в них внесены существенные изменения.

Примечание. И здесь необходимо дать некоторые дополнительные пояснения. В отдельных рецензиях к предыдущим изданиям были сделаны замечания, что вопросы проектирования базы данных в них рассматриваются слишком поздно. Но, по мнению автора, студенты смогут должным образом подготовиться к решению задач проектирования баз данных или полностью оценить важность проблем проектирования только после того, как будут иметь полное представление о том, что такое базы данных и как они используются; иными словами, автор считает важным уделить определенное внимание реляционной модели и связанным с ней вопросам и только после этого знакомить студентов с проблемами проектирования. Поэтому, по мнению автора, материал части III находится в должном месте. (Говоря об этом, следует признать, что многие преподаватели предпочитают излагать сведения о сущностях и связях гораздо раньше. С учетом этого автор попытался сделать главу 14 более или менее автономной, с тем чтобы преподаватели могли ввести ее в курс лекций, скажем, сразу же после главы 4.)

- **Часть IV.** Две главы этой части, 15 и 16, представляют собой полностью пересмотренные, расширенные и дополненные версии глав 14 и 15 из седьмого издания. В частности, теперь глава 16 включает тщательный анализ так называемых *свойств ACID* транзакций (а также некоторые довольно неожиданные выводы по этой теме).
- **Часть V.** Глава 20 о наследовании типов и глава 23 о хронологических базах данных были полностью пересмотрены с учетом новейших исследований и разработок в этих областях. Изменения, которые произошли в других главах, в основном представляют собой их корректировку, хотя внесены существенные уточнения в те пояснения и примеры, которые приведены в этих главах, а также повсеместно добавлен новый материал.
- **Часть VI.** Главы 25 и 26 представляют собой улучшенные и дополненные версии глав 24 и 25 из седьмого издания. Глава 27 является новой.

Наконец, новым также является приложение А, а приложения Б и В представляют собой, соответственно, пересмотренные версии приложений А и В из седьмого издания (материал из приложения Б седьмого издания включен в основную часть восьмого). Приложение Г — это существенно пересмотренная версия приложения А из шестого издания. Ответы к избранным упражнениям, которые в седьмом издании были приведены в соответствующих главах, теперь вынесены в отдельное приложение Д.

ОТЛИЧИТЕЛЬНЫЕ ОСОБЕННОСТИ ДАННОЙ КНИГИ

Каждая представленная на рынке книга по базам данных характеризуется своими собственными преимуществами и недостатками, и у каждого автора есть свой любимый конек. Одни сосредотачиваются на проблемах управления транзакциями, другие уделяют основное внимание моделям “сущность–связь”, третьи рассматривают все эту проблематику через призму языка SQL, четвертые принимают чисто “объектную” точку зрения, пятые рассматривают всю эту область знаний исключительно на основе определенного

коммерческого программного продукта и т.д. И, безусловно, автор данной книги не является исключением из этого правила — у него также есть любимый конек. Условно это можно было бы назвать *основаниями науки*. Автор сохраняет твердое убеждение в том, что вначале необходимо заложить прочные основания любой науки и полностью их освоить, прежде чем пытаться что-то создавать на этом основании. Именно эта убежденность автора позволяет объяснить, почему в настоящей книге такое внимание уделено описанию реляционной модели; в частности, это позволяет понять, почему такой большой объем имеет часть II (наиболее важная часть книги), где автор излагает свое понимание реляционной модели настолько тщательно, насколько он на это способен. Автора интересуют фундаментальные знания, а не преходящие увлечения и модные направления. Технические и программные продукты постоянно изменяются, а принципы остаются.

В связи с этим автор хотел бы привлечь внимание читателя к тому, что в данной книге нескольким важным (“фундаментальным”) темам посвящены целые главы (а в одном случае отдельное приложение), и в этом настоящая книга занимает совершенно особое положение среди других книг по данной тематике. Ниже перечислены основные указанные темы.

- Типы.
- Целостность.
- Представления.
- Отсутствующая информация.
- Наследование.
- Хронологические базы данных.
- Модель TransRelational™.

Развивая мысль о важности оснований науки, автор должен признать, что общая направленность этой книги с годами менялась. Первые несколько изданий по своему характеру в основном описывали текущее состояние дел; в них содержалось описание данной области в том виде, в каком она фактически реализовалась, со всеми своими сильными и слабыми сторонами. В отличие от этого, последующие издания становились все более *предписывающими*; в них говорилось о том, какой должна быть эта область и по какому пути она должна развиваться в будущем, если мы хотим, чтобы все в ней обстояло благополучно. А что касается настоящего издания, то в нем стремление определить правильное направление развития стало доминирующим (поэтому перед вами — не просто учебник, а книга с определенным подтекстом!). Тем не менее, добиться правильного положения дел можно лишь при условии полного понимания, в чем состоит это “правильное положение дел”, поэтому автор попытался сделать все от него зависящее, чтобы это новое издание могло способствовать успеху такого важного начинания.

С этим связана еще одна важная мысль: как может быть известно читателю, автор недавно опубликовал вместе со своим коллегой Хью Дарвенем еще одну “*предписывающую*” книгу, *Foundation for Future Database Systems: The Third Manifesto* (в настоящей книге она указана в списке литературы как [3.3])². Эта книга, сокращенно называемая ее авторами

² Кроме того, Третьему Манифесту посвящен Web-узел <http://www.thethirdmanifesto.com>. Большой объем дополнительного материала можно также найти по адресу <http://www.dbdebunk.com>.

Третьим Манифестом, или просто *Манифестом*, содержит подробные технические предложения по созданию будущих систем баз данных на основе реляционной модели; она явилась результатом многолетнего опыта преподавания и размышления на эти темы, который в дальнейшем был обобщен и изложен самим Хью и автором данной книги. Поэтому не удивительно, что идеи *Манифеста* присутствуют и в настоящей книге. Но не следует думать, что без изучения *Манифеста* нельзя приступать к чтению данной книги — это не так; но *Манифест* непосредственно касается многих вопросов, изложенных в этой книге, и в нем часто можно найти важную дополнительную информацию.

Примечание. В [3.3] в иллюстративных целях решено использовать язык Tutorial D, и то же самое принято в данной книге. Язык Tutorial D разрабатывался таким образом, чтобы его синтаксис и семантика в основном не требовали дополнительных пояснений (сам этот язык можно неформально охарактеризовать как “подобный Паскалю”). Но когда отдельные его средства используются впервые, дается их описание, если в этом есть необходимость.

ЗАКЛЮЧИТЕЛЬНОЕ ЗАМЕЧАНИЕ

В завершение своих вступительных замечаний, автор хотел бы привести следующую немного отредактированную выдержку из собственного предисловия Бертрانا Рассела (Bertrand Russell) к его словарю *The Bertrand Russell Dictionary of Mind, Matter and Morals* (ed. Lester E. Denonn) — Citadel Press, 1993 (любезное разрешение на ее перепечатку получено).

Меня обвиняли в привычке менять свои суждения... Но разве мог бы физик, работающий с 1900 года, например, похвастаться в середине двадцатого века тем, что его суждения не изменились за последние полстолетия?... Та философия, которую я ценю и которой стараюсь следовать, научна в том смысле, что мы должны всегда стремиться получить неопровержимые знания, но новые открытия могут выявить прежние ошибки, неизбежные для любого беспристрастного разума. Когда бы и что бы я ни говорил, сейчас или в прошлом, я никогда не утверждал, что это — окончательная истина. Я утверждаю лишь то, что в свое время высказанное мной мнение было вполне обоснованным... Я был бы очень удивлен, если бы дальнейшие исследования не показали, что его необходимо пересмотреть. К тому же я никогда не высказывал свое мнение как окончательный вердикт, а просто подчеркивал, что это — лучшее, что я мог сделать в то время для достижения ясного и точного понимания. Моей целью была прежде всего полная ясность во всем.

Если читатели предыдущих изданий будут изучать настоящее, восьмое издание, то обнаружат, что его автор также меняет свое мнение по многим вопросам (и, несомненно, будет продолжать это делать). И он надеется, что приведенная выше цитата может служить достаточным оправданием такого положения дел. Автор во всем разделяет взгляды Бертрана Рассела на то, что представляет собой любая область научных исследований, но признает, что просто не мог бы выразить эти взгляды более красноречиво.

БЛАГОДАРНОСТИ

Хочу еще раз исполнить свой приятный долг и поблагодарить всех, кто прямо или косвенно принимал участие в работе над этой книгой.

- Прежде всего, я должен поблагодарить моих друзей Дэвида Макговерна (David McGoveran) и Ника Тиндола (Nick Tindall) за их важный вклад в подготовку настоящего издания; Дэвид подготовил первый вариант главы 22 по поддержке принятия решений, а Ник — первый вариант главы 27 по XML. К тому же я очень признателен моему другу и коллеге Хью Дарвену (Hugh Darwen), который оказал мне большую помощь (в разных формах) в работе со всеми теми частями рукописи, которые касались языка SQL. Наградж Алур (Nagraj Alur) и Фабиан Паскаль (Fabian Pascal) предоставили мне большой объем технического вспомогательного материала. Особую благодарность я хочу выразить Стиву Тарену (Steve Tarin) за то, что он изобрел технологию, описанную в приложении А, и оказал мне значительную помощь в достижении ее полного понимания.
- Кроме того, текст книги был заметно улучшен благодаря замечаниям слушателей семинаров, которые я веду на протяжении последних нескольких лет. Кроме того, на нее чрезвычайно положительное влияние оказали комментарии многих друзей и рецензентов, а также мои дискуссии с ними. В их числе следует прежде всего назвать Хью Дарвена (Hugh Darwen), компания IBM; Ги де Тре (Guy de Tré), Гентский Университет; Карла Экберга (Carl Eckberg), Университет штата Сан-Диего; Ченга Хсу (Cheng Hsu), Политехнический институт Ренселлера; Абдул-Рахмана Итани (Abdul-Rahman Itani), Университет Мичигана-Дирборна; Виджая Канабара (Vijay Kanabar), Бостонский Университет; Брюса О. Ларсена (Bruce O. Larsen), Технологический институт Стевенса; Дэвида Ливингстона (David Livingstone), Нортумберлендский университет в Ньюкасле; Дэвида Макговерена (David McGoveran), компания Alternative Technologies; Стива Миллера (Steve Miller), IBM; Фабиана Паскаля (Fabian Pascal), независимая консультационная компания; Мартина К. Соломона (Martin K. Solomon), Флоридский Атлантический университет; Стива Тарена (Steve Tarin), компания Required Technologies; и Ника Тиндола (Nick Tindall), IBM. Каждый из них просмотрел, по крайней мере, часть рукописи этого издания, предоставил технический материал или как-то иначе помог мне найти ответы на многие технические вопросы, поэтому я всем им очень благодарен.
- Я хотел бы также поблагодарить мою жену Линди (Lindy) за то, что она снова участвовала в оформлении обложки книги, а в течение многих лет оказывала мне помощь при подготовке этого и других проектов, связанных с базами данных.
- И наконец, как всегда, выражаю свою признательность всем сотрудникам издательства Addison-Wesley, особенно Майте Суарес-Ривас (Maite Suarez-Rivas) и Кэтрин Харутуниан (Katherine Harutunian) за их поощрение и поддержку во время работы над проектом, а также моему редактору Элизабет Беллер (Elisabeth Beller) за ее безукоризненную работу.

*К. Дж. Дейт
Хилдсбург, Калифорния
2003*

ЖДЕМ ВАШИХ ОТЗЫВОВ!

Вы, читатель этой книги, и есть главный ее критик и комментатор. Мы ценим ваше мнение и хотим знать, что было сделано нами правильно, что можно было сделать лучше и что еще вы хотели бы увидеть изданным нами. Нам интересно услышать и любые другие замечания, которые вам хотелось бы высказать в наш адрес.

Мы ждем ваших комментариев и надеемся на них. Вы можете прислать нам обычное или электронное письмо, либо просто посетить наш Web-сервер и оставить свои замечания там. Одним словом, любым удобным для вас способом дайте нам знать, нравится или нет вам эта книга, а также выскажите свое мнение о том, как сделать наши книги более интересными для вас.

Посылая письмо или сообщение, не забудьте указать название книги и ее авторов, а также ваш обратный адрес. Мы внимательно ознакомимся с вашим мнением и обязательно учтем его при отборе и подготовке к изданию последующих книг. Наши координаты:

E-mail: info@williamspublishing.com
WWW: <http://www.williamspublishing.com>

Адреса для писем из:

России: 115419, Москва, а/я 783
Украины: 03150, Киев, а/я 152