

Аналитический обзор традиционного подхода формирования реляционных таблиц с учетом использования существующей информации табличного вида

77-48211/425220

08, август 2012

Брешенков А. В., Мин Т. Т.

УДК 681.3.07

Россия, МГТУ им. Н.Э. Баумана

breshenkov@rambler.ru

Введение

За последние годы выполнен значительный объем научных исследований, посвященных проектированию реляционных баз данных (РБД). Среди них можно назвать работы Е.Ф. Кодда [1-5], К. Дж. Дейта [6-8, 9-14], Гэри Хансена, Джэймса Хансена [15], Ульмана Дж. [16-17], Тихомирова Ю.В. [18-19], Григорьева Ю.А., Ревункова Г.И. [20-21], Карповой Т.С. [22] и других.

Однако, несмотря на значительные успехи в области проектирования РБД, большинство задач проектирования формализовать не удается. Это, в частности, отмечает и Дейт К. Дж. [8]. Такое положение вещей связано в большинстве случаев с тем, что в процессе проектирования баз данных (БД) в основном используются не сами данные или информация, преобразуемая в данные, а их предполагаемые схемы отношений. В процессе эксплуатации БД может оказаться, что предположения ошибочны и схемы данных сформированы ошибочно. Действительно, не зная содержимого таблиц, а только отталкиваясь от их схемы отношений, далеко не всегда возможен правильный и оптимальный выбор ключевых полей, выявление функциональных зависимостей, решение вопросов нормализации, обоснованное формирование связей между таблицами. Ведь все эти вопросы решаются неформально на основании предполагаемого содержимого таблиц с данными, которых еще нет. В связи с этим БД приходится нередко перепроектировать, что связано с большими издержками различного характера.

Ситуация кардинально меняется, когда в качестве исходных сведений для проектирования используются существующая информация. Ее анализ можно

формализовать и в конечном итоге принять лучшие решения. В особенности это касается информации табличного вида (ИТВ).

С другой стороны, значительная часть информации, в том числе и ИТВ, находится вне баз данных и даже вне ЭВМ, хотя потребители этой информации, как правило, очень заинтересованы в возможности использования многочисленных возможностей БД [23, 24].

Эти два положения и определяют актуальность разработки методов и средств формирования реляционных таблиц на основе существующих ИТВ.

Конечно, традиционная методология проектирования РБД является мощным, а на сегодня лучшим подходом к разработке БД, использующих в качестве концептуальной модели реляционную модель данных. Поэтому было бы просто неразумно в данной работе не использовать инструмент создания целостных, непротиворечивых и неизбыточных систем.

С другой стороны неразумно не воспользоваться фактами повсеместного использования ИТВ и не решать задачи проектирования БД на основе анализа имеющейся информации. В связи с этим и возникает проблема теоретических и практических разработок, ориентированных на разработку метода формирования реляционных таблиц на основе информации табличного вида.

1. Понятие информации табличного вида

Далеко не всякую информацию можно представить в виде данных. Ведь данные - это информация, представленная в регламентированном виде. К сожалению, не всю информацию можно строго регламентировать. Введем понятие "информация табличного вида" (ИТВ). Представление такого рода информации близко к представлению данных в РБД, и поэтому в принципе процесс преобразования ИТВ в формат РБД можно формализовать. Но ИТВ по ряду признаков существенно отличается от данных.

ИТВ - это информация, которая воспринимается пользователем как двумерные таблицы.

По сути, это могут быть ведомости, прайс-листы, словари, списки и многое другое. Главная задача разработчиков такого рода таблиц - это обеспечение минимальной субъективной сложности восприятия информации. При этом:

- данные в столбцах могут не совпадать по типу;
- возможно дублирование записей;
- допустимо отражение семантики данных посредством цвета, фона, шрифта и т.п.;
- допустимо использование подзаголовков и внутренних заголовков;

- допустимо повторное использование заголовков и подзаголовков (суть их определяется посредством места в таблице, цвета, фона, шрифта и т.п.);
- возможны пустые строки;
- допустимо использовать первый столбец в качестве заголовков.

В этом случае о регламентации информации не может быть и речи. Проблема и состоит в преобразовании нерегламентированной информации к регламентированному виду. А регламентированный вид - это формат БД.

Естественно задаться вопросом, а нужны ли преобразования ИТВ в таблицы БД. Собственные исследования, работа с экспертами, участие в разработках показали, что, с одной стороны, БД исключительно редко создаются на пустом месте (чаще всего имеются значительные объемы информации вида ИТВ), а, с другой стороны, к настоящему времени накопилось множество ИТВ, которую просто необходимо обрабатывать средствами современных систем управления базами данных (СУБД).

2. Основы современной методологии проектирования реляционных баз данных

Проектирование реляционных баз данных в соответствии с традиционной методологией включает в себя 4-е этапа [11-13]. Как правило, выделяют следующие этапы проектирования: формулировка и анализ требований; инфологическое проектирование; датологическое проектирование; физическое проектирование.

Формулировка и анализ требований связаны с определением целей разработки, выделением информационных потоков, экспертированием и многими другими мероприятиями, связанными с анализом предметной области. Главной задачей этого этапа является формулировка требований к разрабатываемой БД. В связи с этим этот этап заканчивается техническим заданием (ТЗ) на разрабатываемую БД.

При наличии информации табличного вида ТЗ в значительной степени уже сформулировано. Действительно, в общих чертах определены состав данных и их структура, известно смысловое назначение информации, имеются сведения о связях между сущностями. В связи с этим процесс формирования ТЗ при наличии ИТВ упрощается.

Инфологическое проектирование ориентировано на построение модели предметной области. В частности, выделяются сущности, формируются и оптимизируются локальные представления, назначаются первичные и внешние ключи, строятся диаграммы сущность-связь, таблицы приводятся к нормальным формам, формируются связи между таблицами. На этом этапе проектирования РБД разработчик

абстрагируется от инструментальных средств их разработки, от таких как Oracle, Clarion, Access и других. В этом и состоит одно из основных достоинств теории проектирования РБД – разделение логического и физического уровней разработок [15, 16] .

Именно на этом этапе проектирования РБД в наибольшей мере проявляется достоинство подхода формирования реляционных таблиц (РТ) на основе использования ИТВ. Оно обусловлена тем, что проектные задачи решаются не на основе анализа гипотетических схем отношений, а на основе анализа реальной информации, а это обеспечивает возможность формализации проектных процедур [28].

Датологическое проектирование связано с построением модели предметной области в терминах языков инструментальных средств разработки БД. При этом в качестве модели данных используются сами данные. Разработка датологической модели осуществляется на базе инфологической модели.

При проектировании РТ на основе существующих ИТВ достигается слияние инфологического и датологического этапов проектирования.

Физическое проектирование позволяет привязать датологическую модель к среде хранения. На этом этапе осуществляется выбор носителя данных, выбор внутренних форматов их хранения, выбор методов доступа к данным, методов сжатия данных, реализуются меры по безопасности данных

На этом этапе несущественно, использовалась ли ИТВ при проектировании РТ.

3. Реляционная модель данных (РМД)

РМД рассматривается практически во всех работах, посвященных проектированию РБД. В частности, она обсуждается в работах [25-27], [1-5]. Она является целевой моделью в контексте темы статьи, в связи с этим ее необходимо обсудить. К числу удачных определений основной компоненты РМД можно отнести определение, представленное в работе [29]. По определению авторов "реляционная модель данных (РМД) некоторой предметной области представляет набор отношений, изменяющихся во времени". Для однозначного понимания терминологии и сути данной работы оправдано привести таблицу элементов реляционной модели данных 1.1, которая представлена в работе [47].

Таблица 1

Элемент реляционной модели	Форма представления
Отношение	Таблица
Схема отношения	Строка заголовков столбцов таблицы (заголовков таблицы)
Кортеж	Строка таблицы
Сущность	Описание свойств объекта
Атрибут	Заголовок столбца таблицы
Домен	Множество допустимых значений атрибута
Значение атрибута	Значения поля в записи
Первичный ключ	Один из нескольких атрибутов
Тип данных	Тип значений элементов таблицы

Несмотря на детальное представление элементов РМД в таблицу по мнению автора можно добавить описательный атрибут – свойства и ограничения атрибута, внешний ключ – атрибут, используемый для связи между таблицами.

Основным понятием РМД является отношение, которое представляет собой подмножество декартового произведения доменов D_1, D_2, \dots, D_k вида:

$$D = D_1 \times D_2 \times \dots \times D_k, \text{ где}$$

$$D_1 = (d_{11}, d_{12}, \dots, d_{1i}, \dots, d_{1m1})$$

$$D_2 = (d_{21}, d_{22}, \dots, d_{2i}, \dots, d_{2m2})$$

...

$$D_k = (d_{k1}, d_{k2}, \dots, d_{ki}, \dots, d_{kmk})$$

Домен – множество элементов, типы которых могут не совпадать.

Отношение R представляется следующим образом:

$$R \subseteq D = D_1 \times D_2 \times \dots \times D_k \quad [1].$$

Следует отметить, что в таблице 1 описаны скорее элементы РТ (составной части РМД), а не РМД. РМД отражает связи между таблицами. А в рамках темы данной работы как раз и представляет основной интерес именно РТ.

РТ, соответствует отношению из k атрибутов и должна удовлетворять следующим свойствам:

- каждая строка представляет собой кортеж из k значений, принадлежащим k столбцам;
- каждый кортеж содержит точно одно значение (соответствующего типа) для каждого атрибута;

- порядок столбцов фиксирован ($1, 2, \dots, k$);
- порядок строк произволен;
- любые две строки различаются хотя бы одним элементом;
- строки и столбцы могут обрабатываться в любой последовательности, определяемой применяемыми операциями обработки;
- атрибуты не должны дублироваться [1-5].

Как правило, ИТВ, свойства которой рассмотрены выше, не удовлетворяет большинству из перечисленных требований. В изменении такого положения вещей и состоит основная задача разработчика реляционных таблиц баз данных на основе использования существующих РТ

4. Ключевые поля и обеспечение целостности данных

В работах, посвященных теории проектирования РБД, в частности в [6-8], дается определение ключевых полей, обосновывается их необходимость, формулируются требования к ним и определяются свойства внешних ключей.

Ряд специалистов в области реляционных баз данных не включают в качестве требований к РТ наличие первичных ключей. Это оправдано лишь отчасти.

Действительно, если эту проблему рассматривать в теоретическом аспекте, то отдельные первичные и внешние ключи могут быть назначены в процессе нормализации таблиц. Если эту проблему рассматривать в практическом аспекте, большинство инструментальных средств, ориентированных на проектирование РБД, позволяют описывать реляционные таблицы без указания первичных и внешних ключей.

В связи с этим ключевые поля могут быть просто не назначены, а это недопустимо. Ведь одно из центральных требований к РТ звучит следующим образом: любые две строки таблицы должны различаться хотя бы одним элементом. А в этом и состоит одно из назначений первичных ключей.

К основным требованиям к первичным ключам относятся их уникальность и минимальность.

В соответствии с требованием уникальности первичного ключа ни одно значение, соответствующее атрибуту первичного ключа не должно повторяться. Если первичный ключ включает в себя несколько атрибутов, то это требование звучит следующим образом: ни один кортеж значений, соответствующий атрибутам, входящим в первичный ключ, не должен дублироваться.

В соответствии с требованием минимальности первичного ключа, с одной стороны, суммарная длина атрибутов, входящих в первичный ключ, должна быть минимальной, а, с

другой стороны, ни один атрибут, входящий в первичный ключ, не должен быть уникальным, в том смысле, что его значения не должны повторяться.

При наличии только схемы отношения не просто выбрать атрибуты, удовлетворяющие данным условиям. При этом легко ошибиться, т.к. после заполнения таблицы реальная ситуация может отличаться от предполагаемой ситуации.

Используя ИТВ, разработчик РБД имеет возможность принимать решение на основе анализа существующих данных. Кроме того, процесс поиска уникальных и минимальных ключей можно формализовать.

Кроме первичных ключей в РБД широко используются внешние ключи. [15, 16]. Такие ключи используются для связи между таблицами. Определяют внешние ключи следующим образом: если в отношении R1 значения какого-либо атрибута, не входящие в его первичный ключ, равны значениям первичного атрибута отношения R, то такой атрибут называют внешним ключом для отношения R.

При наличии ИТВ возможна формализация назначения внешних ключей на основе анализа пар заполненных таблиц. Целостность данных в основном рассматривается в двух аспектах – целостность сущностей и целостность согласования. [8].

Требование целостности сущностей можно сформулировать так: сущности реального мира должны быть различимы. То есть ни одна запись в таблице не должна повторяться. В противном случае возникнет множество проблем в процессе обработки данных.

Целостность согласования: ссылаться можно только на те данные, которые существуют. Такую целостность еще называют ссылочной целостностью. Нарушение этого правила может привести к трудно разрешимым проблемам в процессе функционирования БД.

Обеспечение целостности данных связано с правильным использованием первичных и внешних ключей. В работах [38-40] достаточно полно рассмотрены вопросы обеспечения целостности данных при проектировании РБД на основе ИТВ.

5. Нормализация отношений

Нормализация отношений – аппарат ограничений на формирование отношений, который позволяет устранить избыточность БД, обеспечивает непротиворечивость хранимых данных.

“Концепции этой методологии являются не чем иным как соображениями здравого смысла, записанными в формальном виде. Сутью теории является поиск и формирование этих принципов здравого смысла, что, конечно же, является весьма непростой задачей.

Однако, если такая задача будет решена, найденные принципы могут быть положены в основу решений автоматизации, т.е. можно будет написать программу, позволяющую выполнять проектирование с помощью компьютера” [8].

Процесс нормализации отношений основывается на концепции нормальных форм. Переменная отношения находится в нормальной форме, если она удовлетворяет заданному для нее набору условий. В таком случае процедуру нормализации можно охарактеризовать как последовательное приведение данного набора отношений к некоторой более желательной форме. Процедура нормализации включает разбиение, или декомпозицию переменной отношения, не удовлетворяющей условиям нормальной формы [8]. Процесс декомпозиции на самом деле является операцией проекции исходной переменной отношения.

Важно отметить, что операции проекции имеют смысл, когда отношение имеет не только схему, но и тело, то есть при наличии ИТВ. А при традиционном проектировании тело отношения может иметь лишь гипотетический характер.

В работах [33, 41-49, 59] предложены и реализованы решения автоматизации, которые возможны для весьма часто встречающихся ситуаций, когда РБД проектируются на основе существующей ИТВ. В этих работах, в том числе, разработаны и методы нормализации таблиц ИТВ.

5. Семантическое моделирование.

Разработчики БД обычно обладают весьма ограниченными сведениями о смысле хранящихся в них данных. Однако в РМД присутствуют семантические (смысловые) аспекты. Их примерами являются домены, первичные и внешние ключи [8]. В связи с этим семантическое моделирование данных представляет теоретический и практический интерес. Этим проблемам посвящены работы как зарубежных [25-27], так и отечественных ученых [46-48].

Это моделирование осуществляется на этапе инфологического проектирования РБД. При этом выполняются следующие шаги:

- выявляется множество смысловых понятия;
- выделяются сущности предметной области;
- выполняется классификация сущностей;
- определяются свойства сущностей;
- формируются связи между сущностями;

Широкое распространение получил метод семантического моделирования “сущность-связь” или ER-модель. Связи в модели “сущность-связь” могут иметь тип

“один к одному”, “один ко многим”, “многие к одному” и “многие ко многим”. Одной из существенных проблем, которая возникает при формировании ER-диаграмм, является назначение связей.

В работах [38-41] предложены методы формирования всех типов связей между РТ, которые формируются на основе ИТВ.

Заключение

В свете вышесказанного можно сделать вывод о том, что в процессе проектирования БД при наличии соответствующих ИТВ оправдано использование формальных средств преобразования ИТВ в РТ. Это позволит свести к минимуму ошибки проектирования БД и снизить трудоемкость их создания.

Литература

1. Монографии, брошюры и т.п.:

1. Codd E.F. Data Base Sublanguage Founded on the Relational Calculus // Proc. 1971 ACM SIGFIDET Workshop on data Description, Access and Control. – San Diego, Calif. – November, 1971.

2. Codd E.F. A relational model data for larger shared data banks // Comm; ACM.- 1970. V.13.- № 6.- P. 377-387.

3. Codd E.F. Further normalization of the database relational model, in data base systems (R. Rustin, ed.). Prentice Hall, Endlewood Cliffs, NJ, 1972.

4. Codd E.F. Recent Investigations into Relational Data Base Eystems // Proc IFIP Congress.- Stockholm, Sweden, 1974

5. Codd E.F. The Relational Model For Database Management Version 2. Reading, Mass.: Addison-Wesley, 1990.

6. Дейт К. Дж. Введение в системы баз данных: Пер. с англ. - М.: Наука, 1980.- 464 с.

7. Дейт К. Дж. Введение в системы баз данных. 6-е изд.: Пер. с англ. - Киев: Диалектика, 1998. - 784 с.

8. Дейт К. Дж. Введение в системы баз данных. 7-е изд.: Пер. с англ. - М.: Вильямс, 2001. - 1072 с.

9. Date C. J. Why Quantifier Order Is Important // Date C. J. and Hugh Darwen. Relational Database Writings 1989 – 1991. – Reading, Mass.: Addison-Wesley, 1992.

10. Date C. J. What’s Wrong with SQL? // Date C. J. Relational Database Writings 1985 – 1989. – Reading, Mass.: Addison-Wesley, 1990.

11. Date C. J. How We Missed the Relational Boat // Date C. J. Relational Database

Writings 1991 – 1994. – Reading, Mass.: Addison-Wesley, 1995.

12. Date C. J. Why Relational? // C.J. Date. Relational Database Writings 1985 – 1989. – Reading, Mass.: Addison-Wesley, 1990.

13. Date C. J.: “There’s Only One Relational Model!”, <http://www.dbdebunk.com> (February 2001).

14. Date C. J. What Not How: The Business Rules Approach to Application Development.-Reading, Mass.: Addison-Wesley, 2000

15. Гэри Хансен, Джэймс Хансен. Базы данных: разработка и управление: Пер. с англ. - М.: Бином, 1999. – 699 с.

16. Ульман Дж. Основы систем баз данных: Пер. с англ. М.Р. Когаловского и В.В. Когутовского. - М.: Финансы и статистика, 1983. - 334 с.

17. Ульман Д., Уидом Д. Введение в системы баз данных: Пер. с англ. - М.: Лори, 2000. - 319 с.

18. Тихомиров Ю.В. Microsoft SQL Server 7.0. – СПб.: БХВ-Петербург, 1999. – 720с.

19. Тихомиров Ю.В. Microsoft SQL Server 7.0: разработка приложений. – СПб.: БХВ – Петербург, 1999. - 352 с.

20. Григорьев Е. А. Представление идентифицируемых сложных объектов в реляционной базе данных // Открытые системы. - 2000. - № 1-2.

21. Григорьев Ю.А., Ревунков Г.И. Базы данных: Учебник для вузов. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2002. - 320 с.

22. Карпова Т.С. Базы данных: модели, разработка, реализация. – СПб.: Питер, 2001. - 304 с.

23. Брешенков А.В., Балдин А.В. Анализ проблемы проектирования реляционных баз данных на основе использования информации табличного вида и разработка модели методики проектирования. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2007. - 150 с.

24. Брешенков А.В. Методы решения задач проектирования реляционных баз данных на основе использования существующей информации табличного вида. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2007. – 154 с.

25. Дейт К. Дж. Введение в системы баз данных: Пер. с англ. - М.: Наука, 1980.- 464 с.

26. Дейт К. Дж. Введение в системы баз данных. 6-е изд.: Пер. с англ. - Киев: Диалектика, 1998. - 784 с.

27. Дейт К. Дж. Введение в системы баз данных. 7-е изд.: Пер. с англ. - М.: Вильямс, 2001. - 1072 с.

28. Брешенков А.В. Базы данных. Проектирование баз данных на основе информации табличного вида. LAP LAMBERT Academic Publishing GmbH & Co. KG Dudweiler, rbr, 66123 Saarbrucken, Germany, 2011, 394 с.

29. Хоменко А.Д., Цыганков В.М, Мальцев М.Г. Базы данных: Учебник для высших учебных заведений / Под ред. Проф. А.Д. Хомоненко – 6-е изд. – М.:Бином-Пресс.:СПб КОРОНА-Век, 2007. – 736 с.

30. Корнеев В.В. и др. Базы данных. Интеллектуальная обработка информации. - М.: Нолидж, 2000. - 162 с.

31. Розмахов О.Г. Основы проектирования баз данных. - М.: Московский авиационный институт, 1993. - 24 с.

32. Брешенков А.В. Методы решения задач проектирования реляционных баз данных на основе использования существующей информации табличного вида. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2007. – 154 с.

33. Брешенков А.В., Гудзенко Д.Ю., Казаков Г.И Проектирование реляционных баз данных на основе информации табличного типа - М.: Изд-во МГТУ им. Н.Э. Баумана, 2009 - 150с.

34. Берзтисс А.Т. Структуры данных. - М.: Статистика, 1974. - 408 с.

35. Armstrong W. W. Dependency Structures of data Base Relationships // Proc. IFIP Congress. – Stockholm, Sweden, 1974.

2. Периодические издания:

36. Брешенков А.В. Неформальная постановка проблемы преобразования информации табличного вида в файлы баз данных // Сб. трудов АУ МВД России "Актуальные вопросы технологий в деятельности органов внутренних дел". - М., 2004. - С. 55-70.

37. Брешенков А.В. Выявление и формирование связей один - к одному между заполненными реляционными таблицами. Сборник трудов №5 молодых ученых, аспирантов и студентов "Информатика и системы управления в XXI веке" (часть 2). – М. МГТУ им. Н.Э. – Баумана 2007. – С. 61– 65.

38. Брешенков А.В. Выявление и формирование связей один-ко многим в заполненных реляционных таблицах. Современные информационные технологии: Сб. трудов каф., МГТУ им. Н.Э. Баумана, фак "Информатика и системы упр.", каф. ИУ–6 – М.: Эликс+, 2010. Том 1– С. 48–51.

39. Брешенков А.В. Выявление и формирование связей многие-ко многим в заполненных реляционных таблицах. Современные информационные технологии: Сб. трудов каф., МГТУ им. Н.Э. Баумана, фак "Информатика и системы упр.", каф. ИУ–6 – М.: Эликс+, 2010. Том 2– С. 60–69.

40. Брешенков А.В. Избавление от сложных атрибутов в заполненных нереляционных таблицах // Сб. трудов каф. - М.: Эликс+, 2006. – С. 10-15.

41. Брешенков А.В. Разработка модели методики проектирования реляционных баз данных на основе использования информации табличного вида // Вестник Московского государственного технического университета им. Н.Э.Баумана. - М.: Изд-во МГТУ им. Н.Э.Баумана, 2007. - №2. - 15 с.

42. Брешенков А.В. Исследование методики проектирования реляционных баз данных на основе сетевой модели // Вестник Московского государственного технического университета им. Н.Э.Баумана. - М.: Изд-во МГТУ им. Н.Э.Баумана, 2007. - №3. - 15 с.

43. Брешенков А.В., Бараков Д.Д. Вопросы преобразования электронных таблиц в таблицы реляционных баз данных // Современные информационные технологии: Сб. трудов каф., посвященный 175-летию МГТУ им. Н.Э. Баумана. - М.: Эликс +, 2004. – С. 44-50.

3. Диссертации и авторефераты:

44. Брешенков А.В. Методология проектирования реляционных баз данных с использованием данных табличного вида. Дис. доктор техн. наук (05.25.05) – М., 2007

45. Бабанов А.М. Теория семантически значимых отображений и ее применение для проектирования реляционных баз данных: Дисс. ... канд. техн. наук (05.13.11). - М., 2005. - 182 с.

4. Электронные издания:

46. Брешенков А.В. Преобразование заполненных таблиц ко второй нормальной форме // (Наука и образование: Эл. науч. издание. Номер гос. регистрации 0420700025/0005.) Инженерное образование, 2007. - №2. - 16 с.

47. Брешенков А.В. Приведение заполненных таблиц к третьей нормальной форме // (Наука и образование: Эл. науч. издание. Номер гос. регистрации 0420700025/0016.) Инженерное образование, 2007. - №4. - 15 с.

48. Брешенков А.В., Белоус В.В. Преобразование заполненных таблиц к первой нормальной форме // (Наука и образование: Эл. науч. издание. Номер гос. регистрации 0420700025/0005.) Инженерное образование, 2007. - №2. – 14 с.

49. Брешенков А.В., Белоус В.В. Приведение заполненных таблиц к четвертой нормальной форме // (Наука и образование: Эл. науч. издание. Номер гос. регистрации 0420700025/0017.) Инженерное образование, 2007. - №4. - 15 с.

Analytical review of traditional approach to the formation of relational tables considering the use of existent information in tabular form

77-48211/425220

08, August 2012

Breshenkov A.V., Min Thet Tin

Russia, Bauman Moscow State Technical University

breshenkov@rambler.ru

The authors consider features of traditional approach to design of relational tables in the presence of tabular style information. It was concluded that the presence of tabular style information allowed to formalize the development process. That, hence, allows to minimize the number of design errors and reduce laboriousness of formation of relational tables.

Publications with keywords: [data bases](#), [relational tables](#), [relational database](#), [information table form](#), [database design](#)

Publications with words: [data bases](#), [relational tables](#), [relational database](#), [information table form](#), [database design](#)

References

1. Codd E.F. A Data Base Sublanguage Founded on the Relational Calculus. *Proc. of the 1971 ACM SIGFIDET (now SIGMOD) Workshop on Data Description, Access and Control*. USA, New York, ACM, 1971, pp. 35-68. DOI: [10.1145/1734714.1734718](https://doi.org/10.1145/1734714.1734718)
2. Codd E.F. A relational model of data for larger shared data banks. *Communicstions of the ACM*.- 1970.- V.13.- № 6.- PP. 377-387. DOI: [10.1145/362384.362685](https://doi.org/10.1145/362384.362685)
3. Codd E.F. Further normalization of the database relational model. *Courant Computer Science Symposia. Ser. 6. Data Base Systems*. Prentice Hall, Endlewood Cliffs, NJ, 1972.
4. Codd E.F. Recent Investigations into Relational Data Base Eystems // *Proc. of IFIP Congress 74*, Stockholm, Sweden, 5-10 August 1974. N.Y., North-Holland, 1974, pp. 1017-1021.
5. Codd E.F. *The Relational Model For Database Management: version 2*. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1990. 567 p.

6. Date C.J. *An introduction to database systems*. 2nd ed. Reading, MA, Addison-Wesley, 1979. (Russ. ed.: Deit K.Dzh. *Vvedenie v sistemy baz dannykh*. Moscow, Nauka Publ., 1980. 464 p.).
7. Date C.J. *An introduction to database systems*. 6th ed. Reading, MA, Addison-Wesley, 1997. 839 p. (Russ. ed.: Deit K.Dzh. *Vvedenie v sistemy baz dannykh*. 6th ed. Kiev, Dialektika Publ., 1998. 784 p.).
8. Date C.J. *An introduction to database systems*. 7th ed. Reading, Mass.: Addison-Wesley, 2000. (Russ. ed.: Deit K.Dzh. *Vvedenie v sistemy baz dannykh*. 7th ed. Moscow, Vil'iams, 2001. 1072 p.).
9. Date C. J. Why Quantifier Order Is Important. In: Date C. J., Hugh Darwen. *Relational Database: Selected Writings 1989-1991*. Reading, Mass.: Addison-Wesley, 1992.
10. Date C. J. What's Wrong with SQL? In: Date C. J. *Relational Database: Selected Writings 1985-1989*. Reading, Mass.: Addison-Wesley, 1990.
11. Date C. J. How We Missed the Relational Boat. In: Date C. J. *Relational Database: Selected Writings 1991-1994*. Reading, Mass.: Addison-Wesley, 1995.
12. Date C. J. Why Relational? In: Date C.J. *Relational Database: Selected Writings 1985-1989*. Reading, Mass.: Addison-Wesley, 1990.
13. Date C.J. *There's Only One Relational Model!*. Available at: <http://www.dbdebunk.com> (February 2001).
14. Date C. J. *What Not How: The Business Rules Approach to Appllication Devolopment*. Reading, Mass., Addison-Wesley, 2000.
15. Hansen G.W., Hansen J.V. *Database Management and Design*. Englewood Cliffs, Prentice-Hall, 1995. 582 p. (Russ. ed.: Khansen G., Khansen Dzh. *Bazy dannykh: razrabotka i upravlenie*. Moscow, Binom Publ., 1999. 699 p.).
16. Ullman J.D. *Principles of Database Systems*. Computer Science Press, Rockville, MD, 1982. (Russ. ed.: Ul'man Dzh. *Osnovy sistem baz dannykh*. Moscow, Finansy i statistika, 1983. 334 p.).
17. Ullman J.D., Widom J. *First course in database systems*. Englewood Cliffs, Prentice-Hall, 1997. (Russ. ed.: Ul'man D., Uidom D. *Vvedenie v sistemy baz dannykh*. Moscow, Lori Publ., 2000. 319 p.).
18. Tikhomirov Iu.V. *Microsoft SQL Server 7.0*. St. Petersburg, BKhV-Peterburg, 1999. 720 p.
19. Tikhomirov Iu.V. *Micrjsoft SQL Server 7.0: razrabotka prilozhenii* [Micrjsoft SQL Server 7.0: development of applications]. – St. Petersburg, BKhV - Peterburg, 1999. 352 p.
20. Grigor'ev E.A. Predstavlenie identifitsiruemykh slozhnykh ob"ektov v reliaisionnoi baze dannykh [Presentation of the identifiable complex objects in a relational database]. *Otkrytye sistemy* [Open systems], 2000, no. 1-2.

21. Grigor'ev Iu.A., Revunkov G.I. *Banki dannykh* [Databanks]. Moscow, Bauman MSTU Publ., 2002. 320 p.
22. Karpova T.S. *Bazy dannykh: modeli, razrabotka, realizatsiia* [Databases: model, design, implementation]. St. Petersburg, Piter, 2001. 304 p.
23. Breshenkov A.V., Baldin A.V. *Analiz problemy proektirovaniia reliatsionnykh baz dannykh na osnove ispol'zovaniia informatsii tablichnogo vida i razrabotka modeli metodiki proektirovaniia* [Analysis of the problem of designing a relational database using tabular form information, and development of the model for design techniques]. Moscow, Bauman MSTU Publ., 2007. 150 p.
24. Breshenkov A.V. *Metody resheniia zadach proektirovaniia reliatsionnykh baz dannykh na osnove ispol'zovaniia sushchestvuiushchei informatsii tablichnogo vida* [Methods for solving problems of designing relational databases using the existing tabular form information]. Moscow, Bauman MSTU Publ., 2007. 154 p.
25. Date C.J. *An introduction to database systems*. 2nd ed. Reading, MA, Addison-Wesley, 1979. (Russ. ed.: Deit K.Dzh. *Vvedenie v sistemy baz dannykh*. Moscow, Nauka Publ., 1980. 464 p.).
26. Date C.J. *An introduction to database systems*. 6th ed. Reading, MA, Addison-Wesley, 1997. 839 p. (Russ. ed.: Deit K.Dzh. *Vvedenie v sistemy baz dannykh*. 6th ed. Kiev, Dialektika Publ., 1998. 784 p.).
27. Date C.J. *An introduction to database systems*. 7th ed. Reading, Mass.: Addison-Wesley, 2000. (Russ. ed.: Deit K.Dzh. *Vvedenie v sistemy baz dannykh*. 7th ed. Moscow, Vil'iams, 2001. 1072 p.).
28. Breshenkov A.V. *Bazy dannykh. Proektirovanie baz dannykh na osnove d tablichnogo vida* [Databases. Database design based on the tabular form information]. Saarbrucken, Lambert Academic Publ. GmbH & Co., 2011. 394 p.
29. Khomenko A.D., Tsygankov V.M, Mal'tsev M.G. *Bazy dannykh: Uchebnik dlia vysshikh uchebnykh zavedenii. 6-e izd.* [Databases: A textbook for higher educational institutions. 6th ed.]. Moscow, Binom-Press Publ., St. Petersburg, Korona-Vek Publ., 2007. 736 p.
30. Korneev V.V., Gareev A.F., Vasiutkin C.B., Raikh V.V. *Bazy dannykh. Intellektual'naia obrabotka informatsii* [Databases. Intelligent processing of information]. Moscow, Nolidzh, 2000. 162 p.
31. Rozmakhov O.G. *Osnovy proektirovaniia baz dannykh* [Fundamentals of database design]. Moscow, MAI Publ., 1993, 24 p.
32. Breshenkov A.V. *Metody resheniia zadach proektirovaniia reliatsionnykh baz dannykh na osnove ispol'zovaniia sushchestvuiushchei informatsii tablichnogo vida* [Methods for solving problems of designing relational databases using the existing tabular form information]. Moscow, Bauman MSTU Publ., 2007. 154 p.
33. Breshenkov A.V., Gudzenko D.Iu., Kazakov G.I. *Proektirovanie reliatsionnykh baz dannykh na osnove informatsii tablichnogo tipa* [Relational

- database design based on the tabular type information]. Moscow, Bauman MSTU Publ., 2009. 150 p.
34. Berztiss A.T. *Data structures: Theory and practice*. Academic Press, New York, 1971. 442 p. (Russ. ed.: Berztiss A.T. *Struktury dannykh*. Moscow, Statistika, 1974. 408 p.).
 35. Armstrong W. W. Dependency Structures of data Base Relationships. *Proc. of IFIP Congress74*, Stockholm, Sweden, 5-10 August 1974. N.Y., North-Holland, 1974, pp. 580-583.
 36. Breshenkov A.V. Neformal'naia postanovka problemy preobrazovaniia informatsii tablitsnogo vida v faily baz dannykh [The informal setting of the problem of transformation of tabular form information into database files]. *"Aktual'nye voprosy tekhnologii v deiatel'nosti organov vnutrennikh del"*. Sb. tr. AU MVD Rossii ["Topical issues of technology in the activities of the internal affairs bodies". Academy of Management of the Interior Ministry of Russia. Coll. works]. Moscow, 2004. pp. 55-70.
 37. Breshenkov A.V. Vyiavlenie i formirovanie svyazei odin-k-odnomu mezhdz zapolnennymi relatsionnymi tablitsami [Detection and formation of connections of a one-to-many between the completed relational tables]. *Sb. trudov №5 molodykh uchennykh, aspirantov i studentov "Informatika i sistemy upravleniia v 21m veke"* [Coll. Of works no. 5 of young scientists, post-graduate students and students "Information Science and Control Systems in the 21st century"]. Moscow, Bauman MSTU Publ., 2007, ch. 2, pp. 61– 65.
 38. Breshenkov A.V. Vyiavlenie i formirovanie svyazei odin-ko-mnogim v zapolnennykh relatsionnykh tablitsakh [Detection and formation of connections of a one-to-many in the completed relational tables]. *Sovremennye informatsionnye tekhnologii: Sb. trudov kaf. IU-6 MGTU im. N.E. Baumana, fak. "Informatika i sistemy upravleniia"* [Modern information technology. Coll. of works of the Department of IU-6 of the Bauman MSTU, faculty of "Informatics and control systems"]. Moscow, Elikst Publ., 2010, vol. 1, pp. 48-51.
 39. Breshenkov A.V. Vyiavlenie i formirovanie svyazei mnogie-ko-mnogim v zapolnennykh relatsionnykh tablitsakh [Detection and formation of connections of a many-to-many in the completed relational tables]. *Sovremennye informatsionnye tekhnologii: Sb. trudov kaf. IU-6 MGTU im. N.E. Baumana, fak. "Informatika i sistemy upravleniia"* [Modern information technology. Coll. of works of the Department of IU-6 of the Bauman MSTU, faculty of "Informatics and control systems"]. Moscow, Elikst Publ., 2010, vol. 2, pp. 60-69.
 40. Breshenkov A.V. Izbavlenie ot slozhnykh atributov v zapolnennykh nerelatsionnykh tablitsakh [Getting rid of the complex attributes in the filled non-relational tables]. *Sb. trudov kaf. IU-6 MGTU im. N.E. Baumana, fak. "Informatika i sistemy upravleniia"* [Coll. of works of the Department of IU-6 of the Bauman MSTU, faculty of "Informatics and control systems"]. Moscow, Elikst Publ., 2006. pp. 10-15.

41. Breshenkov A.V. Razrabotka modeli metodiki proektirovaniia reliatsionnykh baz dannykh na osnove ispol'zovaniia informatsii tablichnogo vida [Development of Model of Technique for Designing Relational Databases Based on Information of Tabular Form]. *Vestnik MGTU im. N.E. Baumana. Ser. Priborostroenie* [Herald of the Bauman MSTU. Ser. Instrumentation], 2007, no. 4, pp. 53-62.
42. Breshenkov A.V. Issledovanie metodiki proektirovaniia reliatsionnykh baz dannykh na osnove setevoi modeli [Study of Technique of Relational Databases Design Based on Network Model]. *Vestnik MGTU im. N.E. Baumana. Ser. Priborostroenie* [Herald of the Bauman MSTU. Ser. Instrumentation], 2008, no. 2, pp. 104-113.
43. Breshenkov A.V., Barakov D.D. Voprosy preobrazovaniia elektronnykh tablits v tablitsy reliatsionnykh baz dannykh [Problems of spreadsheets converting into tables of relational databases]. *Sovremennye informatsionnye tekhnologii. Sb. trudov kaf. IU-6, posviashchennyi 175-letiiu MGTU im. N.E. Baumana* [Modern information technology. Coll.of works of the Department of IU-6, on the 175th anniversary of the Bauman MSTU]. Moscow, Eliks+ Publ., 2004. pp. 44-50.
44. Breshenkov A.V. *Metodologiya proektirovaniia reliatsionnykh baz dannykh s ispol'zovaniem dannykh tablichnogo vida. Diss. dokt. tekhn. nauk* [The methodology of designing relational databases using tabular type data. Dr. tech. sci. diss.]. Moscow, 2007.
45. Babanov A.M. *Teoriia semanticheski znachimyykh otobrazhenii i ee primeneniia dlia proektirovaniia reliatsionnykh baz dannykh. Kand. diss.* [The theory of semantically meaningful mappings and its application for designing of relational databases. Cand. diss.]. Moscow, 2005. 182 p.
46. Breshenkov A.V. Preobrazovanie zapolnennykh tablits ko vtoroi normal'noi forme [Completed tables conversion to the second normal form]. *Nauka i obrazovanie MGTU im. N.E. Baumana* [Science and Education of the Bauman MSTU], 2007, no. 2. Available at: <http://technomag.edu.ru/doc/63872.html> , accessed 14.07.2012.
47. Breshenkov A.V. Privedenie zapolnennykh tablits k tret'ei normal'noi forme [Completed tables conversion to the third normal form]. *Nauka i obrazovanie MGTU im. N.E. Baumana* [Science and Education of the Bauman MSTU], 2007, no. 3. Available at: <http://technomag.edu.ru/doc/63920.html> , accessed 14.07.2012.
48. Breshenkov A.V., Belous V.V. Preobrazovanie zapolnennykh tablits k pervoi normal'noi forme [Completed tables conversion to the first normal form]. *Nauka i obrazovanie MGTU im. N.E. Baumana* [Science and Education of the Bauman MSTU], 2007, no. 2. Available at: <http://technomag.edu.ru/doc/63845.html> , accessed 14.07.2012.
49. Breshenkov A.V., Belous V.V. Privedenie zapolnennykh tablits k chetvertoi normal'noi forme [Completed tables conversion to the fourth normal form]. *Nauka i obrazovanie MGTU im. N.E. Baumana* [Science and Education of the Bauman MSTU], 2007, no. 3. Available at: <http://technomag.edu.ru/doc/63956.html> , accessed 14.07.2012.