

А. Н. Дробахина

МЕТОДИКА ОБУЧЕНИЯ ПРОЕКТИРОВАНИЮ БАЗ ДАННЫХ

При обучении студентов дисциплине «Информационные системы» особое внимание уделяется реализации содержательной линии «Проектирование баз данных».

Отметим, что нами предлагается классический подход, при котором весь процесс проектирования производится в терминах реляционной модели данных методом последовательных приближений к удовлетворительному набору схем отношений.

В основе процесса проектирования лежит метод нормализации - декомпозиции (разбиения) отношения, находящегося в предыдущей нормальной форме.

Процесс проектирования базы данных состоит из трех основных этапов: концептуальное, логическое и физическое проектирование.

Процесс создания информационной модели начинается с определения концептуальных требований ряда пользователей. Концептуальные требования могут определяться и для некоторых задач (приложений), которые в ближайшее время реализовывать не планируется. Это может несколько повысить трудоемкость работы, однако поможет наиболее полно учесть все нюансы функциональности, требуемой для разрабатываемой системы, и снизит вероятность ее переделки в дальнейшем. Требования отдельных пользователей интегрируются в едином «обобщенном представлении» (т.е. в концептуальной модели).

При проектировании концептуальной модели все усилия разработчика, как правило, направлены на структуризацию данных и выявление взаимосвязей между ними без рассмотрения особенностей реализации и вопросов эффективности обработки. Проектирование концептуальной модели основано на анализе решаемых на этом предприятии задач по обработке данных.

Концептуальная модель преобразуется затем в модель данных, совместимую с выбранной СУБД. Возможно, что отраженные в концептуальной модели взаимосвязи между объектами окажутся впоследствии нереализуемыми средствами выбранной СУБД. Это потребует изменения концептуальной модели.

1. Создание локальной концептуальной модели данных исходя из представлений о предметной области каждого из типов пользователей.

1.1. Определение объектов.

1.2. Определение атрибутов объектов.

1.3. Определение связей между объектами.

1.4. Определение доменов атрибутов (т.е. определение допустимых значений атрибута; сведений о размере и формате каждого из атрибутов).

1.5. Определение атрибутов, являющихся потенциальными и первичными ключами.

1.6. Проверка модели на отсутствие избыточности (например, удаление избыточных связей, производных атрибутов и т.д.).

1.7. Проверка соответствия локальной концептуальной модели конкретным пользовательским задачам.

1.8. Обсуждение локальных концептуальных моделей данных с конечными пользователями.

Версия концептуальной модели, которая может быть обеспечена конкретной СУБД, называется логической моделью.

Напомним, что логическая модель данных может быть реляционной, иерархической или сетевой.

2. Создание и проверка локальной логической модели данных на основе представления о предметной области каждого из типов пользователей.

2.1. Исключение особенностей, не совместимых с реляционной моделью данных.

2.2. Определение набора отношений (таблиц) на основе структуры локальной логической модели.

2.3. Проверка отношений (таблиц) с помощью правил нормализации.

2.4. Определение требований поддержки целостности данных.

2.5. Обсуждение разработанных локальных логических моделей данных с конечными пользователями.

3. Создание и проверка глобальной логической модели данных.

3.1. Слияние локальных логических моделей данных в единую глобальную модель данных.

3.2. Проверка глобальной логической модели данных.

3.3. Проверка возможностей расширения модели в будущем.

3.4. Обсуждение глобальной логической модели данных с пользователями.

Основная цель логического проектирования - устранение избыточности и дублирования информации. Для удовлетворения этих требований Э. Коддом предложен аппарат нормализации отношений.

Нормальная форма - это свойство отношения в реляционной модели данных, характеризующее его с точки зрения избыточности, потенциально приводящей к логически ошибочным результатам выборки или изменения данных. Нормальная форма определяется как совокупность требований, которым должно удовлетворять отношение.

Нормальная форма представляет собой ограничение на схему базы данных, которое избавляет базу данных от некоторых нежелательных свойств.

В теории реляционных баз данных обычно выделяется следующая последовательность нормальных форм: первая нормальная форма (1NF); вторая нормальная форма (2NF); третья нормальная форма (3NF); нормальная форма Бойса-Кодда (BCNF); четвертая нормальная форма (4NF); пятая нормальная форма, или нормальная форма проекции-соединения (5NF или PJ/NF).

Нормальные формы нумеруются последовательно, по мере ужесточения требований (1 НФ→...5НФ).

Основные свойства нормальных форм:

- каждая следующая нормальная форма в некотором смысле лучше предыдущей;
- при переходе к следующей нормальной форме свойства предыдущих нормальных форм сохраняются.

Первая нормальная форма требует атомарности данных в таблицах, т.е. данные в таблицах должны быть представлены в виде минимально возможных и далее неделимых кусочков информации.

Вторая нормальная форма требует, чтобы данные во всех неключевых столбцах полностью зависели от первичного ключа.

Третья нормальная форма требует, чтобы структура базы данных удовлетворяла требованиям первой и второй нормальной формы. Все неключевые поля таблицы должны зависеть от первичного ключа, но быть независимыми друг от друга.

База данных считается нормализованной, если ее таблицы представлены как минимум в третьей нормальной форме.

Последний этап - создание физической модели. Как правило, основной целью физического проектирования базы данных является описание способа физической реализации логического проекта базы данных. В случае реляционной модели данных под этим подразумевается следующее:

4. Создание набора реляционных таблиц и ограничений для них на основе информации, представленной в глобальной логической модели данных;
5. Определение конкретных структур хранения данных и методов доступа к ним, обеспечивающих оптимальную производительность СУБД;

6. Разработка средств защиты создаваемой системы.

Для приведения модели к требуемому уровню нормальной формы необходимо выполнить следующие действия:

Ограничение первой нормальной формы - значения всех атрибутов отношения должны быть атомарными. Данное требование является базовым требованием классической реляционной модели данных, поэтому любая реляционная таблица, по определению, уже находится в первой нормальной форме.

Чтобы перейти от первой нормальной формы ко второй, нужно выполнить следующие шаги:

1. Определить, на какие части можно разбить первичный ключ, так чтобы некоторые из неключевых полей зависели от одной из этих частей (причем эти части могут содержать несколько атрибутов).
2. Создать новую таблицу для каждой такой части ключа и группы зависящих от нее полей и переместить их в эту таблицу. Часть бывшего первичного ключа станет при этом первичным ключом новой таблицы.
3. Удалить из исходной таблицы поля, перемещенные в другие таблицы, кроме тех из них, которые станут внешними ключами.

Для анализа зависимостей между атрибутами удобно применять ER-диаграммы.

Чтобы перейти от второй нормальной формы к третьей, нужно:

1. Определить все поля (или группы полей), от которых зависят другие поля.
2. Создать новую таблицу для каждого такого поля (или группы полей) и группы зависящих от него полей и переместить их в эту таблицу. Поле (или группа полей), от которого зависят все остальные перемещенные поля, станет при этом первичным ключом новой таблицы.
3. Удалить перемещенные поля из исходной таблицы, оставив лишь те из них, которые станут внешними ключами.

Для выработки у студентов устойчивого навыка проектирования баз данных на лабораторных занятиях предлагаются индивидуальные задания по проектированию баз данных в различных предметных областях, например:

1. Университетская база данных, которая должна включать информацию о студентах, факультетах, преподавателях, курсах, а также о том, какие студенты слушают какие курсы и какие преподавателя преподают эти курсы, какие курсы читаются на том или ином факультете, об отметках студентов и лучших достижениях студентов по каждому курсу.

2. База данных «Рынок труда», которая должна включать информацию о безработных, зарегистрированных на бирже труда: фамилия, имя, отчество, пол, дата рождения, адрес, образование, учебное заведение, которое закончили, специальность, стаж работы, дополнительные возможности (владение иностранным языком, знание компьютера и т.д.), причина безработицы (сокращение, переезд, болезнь и т.д.). Кроме того, информацию о каждом предприятии, предоставляющем работу: название предприятия, адрес, перечень специальностей, имеющих вакансии. Для каждой специальности указаны критерии отбора: образование, стаж, пол, возраст, умения и условия труда на предприятии: рабочий день, выходные, отпуск, заработная плата, льготы и др.

3. База данных «Детский лагерь отдыха», которая должна включать информацию: об отдыхающих детях: № свидетельства о рождении, ФИО, дата рождения, адрес и т.д. Информация о путевках: № путевки, дата заезда, продолжительность заезда; об отрядах: номер отряда, название, корпус, количество детей; о вожатых: код вожатого, ФИО, паспортные данные, дата рождения, стаж, в каком отряде работает.

В условиях снижения аудиторной нагрузки, а также для реализации индивидуального подхода, студентам можно предложить альтернативное задание, например, нормализовать отношения:

1. Номер гостиницы (№, тип номера, количество мест, стоимость места, место, обстановка).

Место (№ места, проживающий).

Проживающий (Ф.И.О., № удостоверения, дата заезда, дата отъезда, срок оплаты).

Обстановка (наименование, количество).

2. Дом (адрес, тип дома, квартира, кол-во квартир).

Квартира (№ кв., Ф.И.О. отв. съемщика, количество комнат, тип квартиры, общая жил площадь, комната, этаж, проживающий).

Комната (№ комнаты, площадь).

Проживающий (Ф.И.О., степень родства).

3. Школа (номер, адрес, Ф.И.О. директора, класс, кабинет).

Класс (№ класса, Ф.И.О. классного руководителя, предмет, состав).

Предмет (наименование, Ф.И.О. преподавателя, наименование кааб., количество часов по предмету).

Состав (Ф.И.О. ученика, пол)

Кабинет (Наименование кабинета, номер комнаты, пособие).

Завершает процесс разработки базы данных этап физического проектирования. Процесс создания набора реляционных таблиц и ограничений для них, а так же работа с объектами данной СУБД (таблицами, запросами, формами, отчетами) подробно описан нами в [2, 3]. Для контроля усвоения теоретических знаний по разделу «Проектирование баз данных» предлагаются тестовые задания, например:

1. Концептуальная модель ...
 - a) не учитывает методов физического хранения данных
 - b) содержит спецификации формата доступа к БД
 - c) содержит спецификации формата файла
 - d) представляет структуру объектов предметной области
2. Логическая модель - представляет собой
 - a) концептуальную модель с учетом поддержки ее средствами СУБД
 - b) концептуальную модель с учетом методов ее хранения на диске
 - c) описание способа физической реализации проекта базы данных
 - d) определение конкретных структур хранения данных и методов доступа к ним, обеспечивающих оптимальную производительность СУБД
3. Наиболее независим от средств СУБД и методов доступа к данным уровень проектирования БД:
 - a) концептуальный
 - b) логический
 - c) физический
 - d) внешний
4. Если операция над отношениями имеет цель минимизировать дублирование данных, то производится операция:
 - a) нормализации
 - b) обновления
 - c) соединения
 - d) объединения
5. Некоторая таблица содержит в одном из столбцов массив значений. Какую нормальную форму нарушает структура такой таблицы?
 - a) первую

- b) вторую
- c) третью
- d) четвертую

Отметим, что, как правило, реализация данного подхода к обучению студентов проектированию баз данных не вызывает у них затруднений, а полученные ими знания и навыки системы и осознаны.

Список литературы

1. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. / Конноли Т., Бегг Л., Страчан А. - 3-е изд. - Вильямс, 2003 - 1440 с.
2. Дробахина А.Н. Информационные системы: проектирование и реализация: Учебное пособие. / А.Н. Дробахина - Новокузнецк: РИО КузГПА, 2012. - 64 с.
3. Дробахина А.Н. Создание и ведение реляционных баз данных в СУБД Openoffice.org BASE. / А.Н. Дробахина // Информатика и образование №4(222)/2011. - С. 59-70.
4. Ледовских И.А. Примеры тем для разработки информационных систем и баз данных: Элективные курсы «Информатика для учащихся 10-11 классов» / Ледовских И.А. - Режим доступа: <http://rudocs.exdat.com>
5. Можаров М.С. Формирование профессиональной мобильности будущего учителя информатики // Педагогическая информатика. 2009. № 3. С. 31-37.
6. Можаров М.С., Коткин С.Д. О развитии содержательной линии «моделирование и формализация» в школьном курсе «информатика и икт»//Информатика и образование. 2010. № 4. С. 95-99.
7. Можарова А.Э. Основы проектирования и разработки информационных систем / А.Э. Можарова // [Информационно-коммуникационные технологии в педагогическом образовании.](#)-2012, № 6 (21).-С. 14-18.
8. Стасышин В.М. Введение в проектирование реляционных баз данных // Учебное пособие по курсу по курсу "Базы данных". - Новосибирск, НГТУ, 1999 - Режим доступа: <http://www.ict.edu.ru/ft/002359/index.html>