

Реляционная модель данных

Концепция реляционной модели была предложена в 1969 году Эдгаром Коддом, известным специалистом в области баз данных, а в 1970 году она была им опубликована.

Реляционная модель является удобной и наиболее привычной формой представления данных.

Реляционная модель представляет собой совокупность данных, состоящую из набора двумерных таблиц, в этой модели вся информация хранится в виде **таблиц (отношений)**, состоящих из строк и столбцов.

С помощью одной таблицы удобно описывать сведения о группах однородных (имеющих одинаковые свойства) объектов, явлений или процессов реального мира. Каждая строка таблицы содержит сведения о конкретном объекте, явлении или процессе. Таблица состоит из совокупности столбцов, каждый из которых имеет уникальное имя. Строка (запись) имеет одинаковую структуру и описывает с помощью полей свойства объектов, то есть имеют одно и тоже количество реквизитов и имеют одинаковую длину. Количество строк в таблице практически неограниченно.

Например, таблица может содержать сведения о группе обучаемых, о каждом из которых известны следующие характеристики: фамилия, имя и отчество, пол, дата рождения, адрес проживания. Поскольку в рамках одной таблицы не удастся описать все данные из предметной области, то создается несколько таблиц, между которыми устанавливаются связи. Данные двух таблиц связываются общими столбцами, а не физическими ссылками или указателями.

При табличной организации данных отсутствует иерархия элементов, строки и столбцы могут быть просмотрены в любом порядке.

Основные элементы реляционной модели

Элемент реляционной модели	Форма представления
Отношение	Двумерная таблица
Схема отношения	Строка заголовков таблицы
Кортеж	Строка таблицы
Кардинальность	Количество кортежей, которое содержит отношение
Сущность	Описание свойств объекта
Атрибут	Заголовок столбца таблицы
Домен	Множество допустимых значений атрибута
Значение атрибута	Значение поля в записи
Первичный ключ	Один или несколько атрибутов отношения, однозначно идентифицирующий каждый из кортежей
Тип данных	Тип значений элементов таблицы

Сущность – это объект любой природы, данные о котором хранятся в базе данных. Данные о сущности хранятся в отношении. Атрибуты представляют собой свойства, характеризующие сущность. Схема отношения представляет собой список имен атрибутов. Отношение должно обладать свойствами:

- не может быть одинаковых строк (кортежей);
- кортежи не упорядочены (порядок строк произвольный);
- все столбцы однородны;
- все элементы в одном столбце имеют один тип данных;
- каждый атрибут в отношении имеет уникальное имя;
- все значения атрибутов атомарные. В теории баз данных это атрибуты, которые хранят единственное значение и не являются ни списком, ни множеством значений.

Иными словами, это такие данные, разделение которых на составляющие приводит к потере их смысла с точки зрения решаемой задачи. Например, если атрибут «Цена» содержит значение 15, то попытка разделить его на 1 и 5 приведет к полной бессмыслице.

Данные, не являющиеся атомарными, называются составными. Например, в базе, содержащей сведения о сотрудниках организации, скорее всего, имеется поле «Адрес», которое может быть представлено в виде «почтовый индекс, город, улица, дом, корпус, квартира». С точки зрения кадрового учета это атомарный атрибут, поскольку маловероятно, что потребуются отдельное использование его частей, например, сортировка сотрудников по номеру квартиры. Поэтому его разбиение на несколько полей бессмысленно и только усложняет структуру базы данных.

В то же время, если производится анализ издательского рынка, где требуется определить число подписчиков на определенные издания в тех или иных городах и регионах, то атрибуты «Почтовый индекс» и «Город» могут иметь самостоятельную ценность. Тогда адрес перестает быть атомарным атрибутом. Следовательно, одни и те же данные в зависимости от ситуации могут рассматриваться и как атомарные, и как составные.

Знание, являются ли атрибуты в базе данных атомарными с точки зрения решения конкретной задачи, очень важно при организации их хранения, поскольку при нормализации реляционной СУБД все отношения должны быть приведены в 1-ю нормальную форму, которая может содержать только атомарные атрибуты.

Пример представления отношения СОТРУДНИК

Атрибут Отдел		Схема отношения (строка заголовков)	
кортеж	ФИО	ОТДЕЛ	ДОЛЖНОСТЬ
	Иванов И.И.	002	Начальник
	Матвеев Т.Р.	001	Заместитель
	Кузьмин Л.Д.	002	Инженер

Отношение СОТРУДНИК содержит 3 кортежа, включает 4 домена. Схема отношения имеет вид СОТРУДНИК (ФИО, Отдел, Должность, Дата Рождения).

Достоинства реляционной модели данных:

- простота представления данных, изложение информации в простой и понятной для пользователя форме (таблица);
- реляционная модель данных основана на строгом математическом аппарате, что позволяет лаконично описывать необходимые операции над данными;
- минимальная избыточность данных (целью организации базы данных должно быть уничтожение избыточных данных там, где это выгодно, и контроль над теми противоречиями, которые вызываются наличием избыточных данных. Избегается дублирование информации);
- независимость приложения пользователя от данных;
- для работы с моделью данных нет необходимости полностью знать организацию базы данных

Недостатки реляционной модели данных:

- относительно медленный доступ к данным;
- требуется относительно большой объем памяти;
- трудность в переводе в таблицу сложных отношений

Развитием способов организации данных является постреляционная модель данных.

Классическая реляционная модель предполагает неделимость данных, хранящихся в полях записей таблиц, это ограничение мешает эффективной реализации приложений.

Постреляционная модель данных является расширением реляционной модели.

Она снимает ограничение неделимости данных, допуская многозначные поля, значения которых состоят из подзначений. Набор значений многозначных полей считается самостоятельной таблицей, встроенной в главную таблицу.

Рассмотрим ту и другую модель на примере информации о накладных.

Таблица Накладные содержит данные о номерах накладных и номерах покупателей. В таблице Накладные Товары содержатся данные о каждой из накладных: номер накладной, название товара и количество товара. Таблица Накладные связана с таблицей Накладные Товары по полю Номер накладной.

Реляционная модель данных

Таблица Накладные

Номер накладной	Номер покупателя
0373	8723
8374	8232
7364	8723

Таблица Накладные товары

Номер накладной	Товар	Количество
0373	Сыр	3
0373	Рыба	2
8374	Шоколад	2
8374	Сок	6
8374	Печенье	2
7364	Йогурт	1

Постреляционная модель данных

Таблица Накладные

Номер накладной	Номер покупателя	Товар	Количество
0373	8723	Сыр	3
		Рыба	2
8374	8232	Шоколад	2
		Сок	6
		Печенье	2
7364	8723	Йогурт	1

Как видно из рисунка, по сравнению с реляционной моделью в постреляционной модели данные хранятся более эффективно, а при обработке не требуется выполнять операцию соединения данных из двух таблиц.

Помимо обеспечения вложенности полей постреляционная модель поддерживает ассоциированные многозначные поля (множественные группы). Совокупность ассоциированных полей называется ассоциацией. При этом в строке первое значение одного столбца ассоциации соответствует первым значениям всех других столбцов ассоциации. Аналогичным образом связаны все вторые значения столбцов и т.д.

На длину полей и количество полей в записях таблицы не накладывается требование постоянства. Это означает, что структура данных и таблиц имеет большую гибкость.

Преимуществом постреляционной модели является возможность представления совокупности связанных реляционных таблиц одной постреляционной таблицей. Это обеспечивает высокую наглядность представления информации и повышение эффективности ее обработки.

Недостатком постреляционной модели является сложность решения проблемы обеспечения целостности и непротиворечивости хранимых данных.