

Тема №3 Основы реляционной теории баз данных

1. Понятие отношения
2. Формы представления отношений
3. Реляционные операции

Понятие отношения

Реляционная модель данных была предложена Е.Ф. Коддом в 1970 году и получила к настоящему времени широкое распространение и популярность.

Этому способствовали два ее существенных достоинства:

- 1) однородность представления данных в модели, которая обуславливает простоту восприятия ее конструкций пользователями базы данных,
- 2) наличие развитой математической теории реляционных баз данных, которая обуславливает корректность ее применения.

Понятие отношения

В основе реляционной модели данных лежит понятие *отношения*, которое задается списком своих элементов и перечислением их значений.

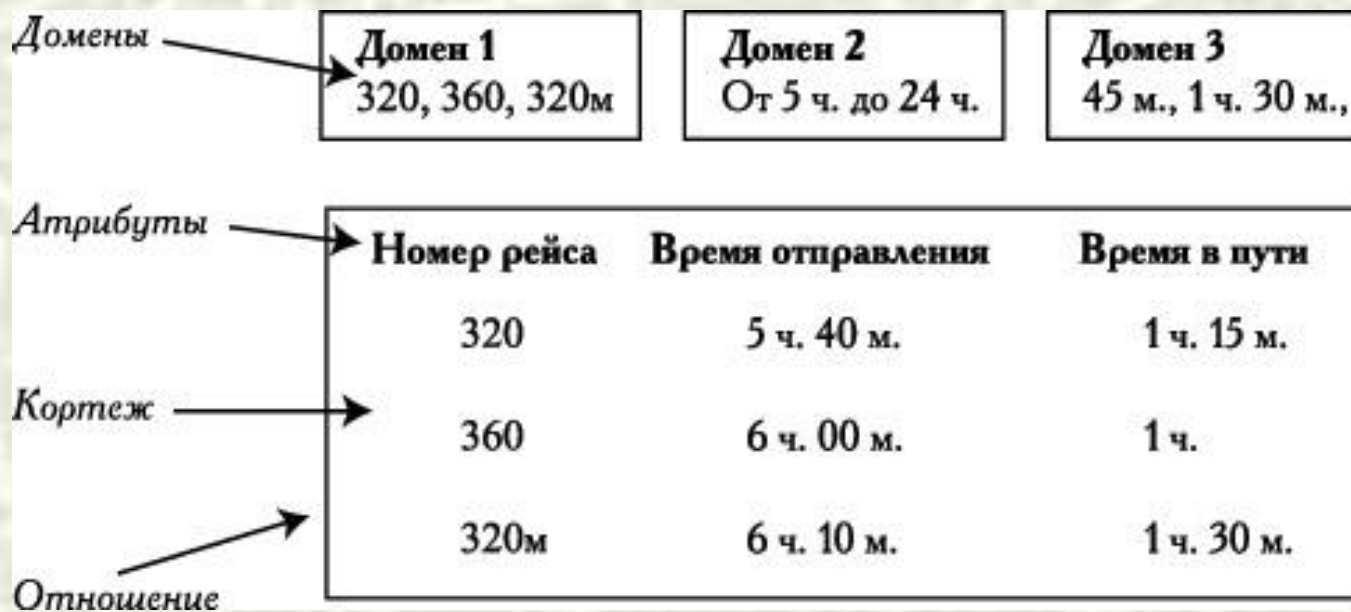
Пример:

Рассматривается расписание движения автобусов. Каждый включенный в расписание рейс имеет свой номер, время отправления и время в пути.

Расписание может быть представлено таблицей. Заголовки колонок таблицы носят название атрибутов. Список их имен носит названия схемы отношения. Каждый атрибут определяет тип представляемых им данных, который вместе с областью его значений называется доменом.

Понятие отношения

Вся таблица целиком называется отношением, а каждая строка таблицы носит название кортежа отношения. Таким образом, отношение можно представить в виде двумерной таблицы.



Понятие отношения

Математически отношение может быть определено как множество кортежей, являющееся подмножеством декартова произведения фиксированного числа областей (доменов).

В каждом кортеже должно быть одинаковое число компонент (атрибутов) и значение каждого из них выбирается из некоторого определенного домена.

Понятие отношения

Введем ряд математических определений, связанных с понятием отношения.

Определение 1:

Декартово произведение Пусть D_1, D_2, \dots, D_n - произвольные конечные множества, не обязательно различные. Декартовым произведением этих множеств

$$D_1 \times D_2 \times \dots \times D_n$$

называется множество вида

$$\{(d_1, d_2, \dots, d_n), d_1 \in D_1, d_2 \in D_2, \dots, d_n \in D_n\}$$

Пример:

$$D_1 = \{a, b\}, D_2 = \{d, c\} \quad D_1 \times D_2 = \{ad, ac, bd, bc\}$$

Понятие отношения

Определение 2:

Пусть $A_1(D_1), A_2(D_2), \dots, A_n(D_n)$ - имена атрибутов. Схемой r отношения R называется конечное множество имен атрибутов

$$r = \{A_1(D_1), A_2(D_2), \dots, A_n(D_n)\}$$

Определение 3:

Отношением со схемой r на конечных множествах D_1, D_2, \dots, D_n называется подмножество R декартового произведения

$$D_1 \times D_2 \times \dots \times D_n$$

Понятие отношения

Элементы отношения (d_1, d_2, \dots, d_n) называются *кортежами*.

О каждом отношении, являющемся подмножеством декартового произведения $D_1 \times D_2 \times \dots \times D_n$, можно сказать, что оно имеет арность n .

Кортеж (d_1, d_2, \dots, d_n) имеет n компонентов.

Использование понятия декартового произведения для определения отношения в реляционной модели данных делает модель *конструктивной*. На математическом языке это означает, что все остальные понятия модели определяются в рамках строго математического построения на базе декартового произведения.

Понятие отношения

Важные моменты:

1. Атрибуты разных отношений могут быть определены на одном домене, так же как и атрибуты одного отношения. Это очень важное обстоятельство, позволяющее *устанавливать связи по значению между отношениями*.
2. Множество математически по своему определению не может иметь совпадающих элементов, и, следовательно, кортежи в отношении можно различить лишь по значению их компонент. Таким образом, в реляционной модели *полностью исключается дублирование данных о сущностях реального мира!*

Понятие отношения

3. Схема отношения также есть множество, что позволяет работать с ними с помощью теоретико-множественных операций. Это является важным моментом для построения теории проектирования реляционных схем баз данных.

Понятие отношения

Существует различие между математическим определением отношения и действительным хранением отношений в памяти компьютера.

По определению, отношение не может иметь два идентичных кортежа. Однако СУБД, поддерживающие реляционную модель данных, хранят отношения в файлах операционной системы компьютера.

Размещение отношений в файлах операционной системы допускает хранение идентичных кортежей. Большинство промышленных СУБД допускают хранение двух идентичных кортежей в базе данных.

Формы представления отношений

Отношения можно представлять в виде таблиц. Но в табличном представлении сложно показывать некоторые свойства отношений.

Предпринимаются попытки строить более четкие схемы описания отношений в реляционных базах данных.

Пример:

СХЕМА «ОТДЕЛ_КАДРОВ»

ДОМЕН T_Табельный_номер ТИП целое

ДОМЕН T_ФИО ТИП символьное

ДОМЕН T_Зарплата ТИП десятичное с фиксированной точкой

.....

ОТНОШЕНИЕ Служащий (Табельный-номер / КЛЮЧ / ДОМЕН = T_Табельный-номер, ФИО / ДОМЕН = T_ФИО,)

....

КОНЕЦ ОПИСАНИЯ СХЕМЫ.

Формы представления отношений

Ключом или ключевым полем называется уникальное значение, которое позволяет тем или иным способом идентифицировать сущность или часть сущности предметной области.

Другими словами, ключ - это значение некоторого атрибута или атрибутов в кортеже отношения, который представляет экземпляр сущности в реляционной модели данных.

Формы представления отношений

Принято различать *первичные ключи* и *частичные ключи*.

Математически первичным ключом отношения R со схемой r является подмножество сужения декартового произведения, которое позволяет однозначно идентифицировать кортеж.

Если первичный ключ содержит несколько атрибутов, то он называется *составным ключом*, в противном случае - *атомарным*.

Частичным ключом называется атрибут составного ключа, если он однозначно определяет совокупность неключевых атрибутов отношения. Атрибут кортежа, который является первичным ключом другого отношения, называется *внешним (иногда посторонним) ключом*.

Формы представления отношений

Важное свойство реляционной модели данных: каждое отношение должно иметь первичный ключ.

Отсутствие первичного ключа в отношении может привести к приобретению кортежей, которые не определены текущим состоянием предметной области, или к потере уже существующих кортежей при выполнении теоретико-множественных операций.

Ключ всегда отражает ту или иную степень связи между атрибутами сущностей предметной области, т.е. семантически ключ есть средство моделирования связей в модели.

Формы представления отношений

Наибольшее распространение получило представление отношений в виде графических диаграмм, например ER-диаграмм.

Преимуществами такого представления являются наглядность диаграмм и возможность их построения в ряде CASE-средств проектирования баз данных.

Обычно CASE-средства позволяют поддерживать несколько уровней представления отношений. Так, например, ErWin поддерживает уровень логической и физической моделей базы данных.

Формы представления отношений

Представление фрагментов реального мира через отношения даже в рамках одной модели данных не характеризуется единственностью.

Например: "Что есть цвет автомобиля? Связь, объект или атрибут?"

Если за объект принять автомобиль, то цвет может выступать в качестве атрибута автомобиля. Если рассматривать зависимость отражательной способности покрытия автомобиля от его цвета, то цвет можно считать объектом. Если рассматривать взаимосвязь между цветом модели автомобиля и ее номером, то цвет можно считать связью.

Формы представления отношений

При представлении какого-либо качества реального мира в модели следует четко понимать, какие запросы в рамках создаваемой модели данных должны быть разрешимыми.

Пример:

Рассмотрим отношение КРАСНЫЙ (модель). При использовании такого отношения на вопрос: "Является ли модель X красного цвета?" может быть получен ответ: "Да" или "Нет". Вопрос: "Какой цвет у модели X?" ответа не имеет, так как в отношении отсутствует атрибут "цвет".

Формы представления отношений

Основные свойства реляционной модели данных:

1. Все кортежи одного отношения должны иметь одно и то же количество атрибутов.
2. Значение каждого из атрибутов должно принадлежать некоторому определенному домену.
3. Каждое отношение должно иметь первичный ключ.
4. Никакие два кортежа не могут иметь полностью совпадающих наборов значений.
5. Каждое значение атрибутов должно быть атомарными, т.е. не должно иметь внутренней структуры и содержать в качестве компонента другое отношение.

Формы представления отношений

6. Реляционная модель данных должна быть непротиворечивой, в частности должен выполняться 1) принцип ссылочной целостности - связи между отношениями должны быть замкнутыми, 2) значения колонок должны принадлежать одному и тому же определенному для них домену.
7. Порядок следования кортежей в отношении не имеет значения. Порядок есть в большей степени свойство хранения данных, чем свойство непосредственно самой реляционной модели данных.

Реляционные операции

Классическая *реляционная модель данных* предусматривает использование восьми реляционных операций:

- объединение,
- пересечение,
- разность,
- декартово произведение,
- деление,
- проекция,
- соединение и
- выбор.

Реляционные операции

Важно: Операции выполняются над отношением в целом, а не над отдельным *кортежем отношения!*

Определение 4: *Степень отношения есть число входящих в него атрибутов или мощность схемы отношения (как множества).*

Определение 5: *Мощность отношения есть число входящих кортежей или кардинальное число отношения (как множества).*

Определение 6: *Два отношения называются совместными, если они имеют совместные схемы (совпадают схемы отношений и домены соответствующих атрибутов).*

Реляционные операции

Объединение отношений

Пусть Q_a , Q_b , Q_c - множество кортежей отношений A , B , C соответственно.

Операция объединения выполняется над двумя совместными отношениями A и B . Результатом операции объединения является отношение C , которое включает в себя все кортежи отношения A и кортежи отношения B , отличные от кортежей отношения A .

Таким образом, объединение отношений можно представить с помощью теоретико-множественной операции объединения:

$$Q_c = Q_a \cup Q_b = \{Q_a + Q_b, \text{отличные от } Q_a\}.$$

Реляционные операции

Пример: Объединение отношений. Выполним операцию объединения отношений КЛИЕНТ_1 и КЛИЕНТ_2.

Исходные отношения:

КЛИЕНТ_1 (#, Фамилия, Возраст) и КЛИЕНТ_2 (#, Фамилия, Возраст)

КЛИЕНТ_1			КЛИЕНТ_2		
1	Иванов	20	1	Иванов	20
3	Петров	23	2	Исаев	30
4	Фролов	49			

Результирующее отношение:

1	Иванов	20
3	Петров	23
4	Фролов	49
2	Исаев	30

Реляционные операции

Пересечение отношений

Операция пересечения выполняется над двумя совместными отношениями A и B.

Результатом операции пересечения является отношение C, которое включает в себя кортежи отношения A, полностью совпадающие с кортежами отношения B.

Таким образом, *пересечение* отношений можно представить с помощью теоретико-множественной операции пересечения:

$$Q_c = Q_a \cap Q_b = \{Q_a, \text{совпадающие с } Q_b\}$$

Реляционные операции

Пример: Пересечение отношений. Выполним операцию пересечения отношений КЛИЕНТ_1 и КЛИЕНТ_2.

Исходные отношения:

КЛИЕНТ_1 (#, Фамилия, Возраст) и КЛИЕНТ_2 (#, Фамилия, Возраст)

КЛИЕНТ_1			КЛИЕНТ_2		
1	Иванов	20	1	Иванов	20
3	Петров	23	2	Исаев	30
4	Фролов	49			

Результирующее отношение:

1	Иванов	20
---	--------	----

Реляционные операции

Разность отношений

Операция *разности* выполняется над двумя совместными отношениями A и B.

Результатом операции разности является отношение C, которое включает в себя кортежи отношения A, отличные от кортежей отношения B.

Таким образом, разность отношений можно представить с помощью теоретико-множественной операции разности:

$$Q_c = Q_a - Q_b = \{Q_a, \text{отличные от } Q_b\}.$$

Реляционные операции

Пример: Разность отношений. Выполним операцию разности отношений КЛИЕНТ_1 и КЛИЕНТ_2.

Исходные отношения:

КЛИЕНТ_1 (#, Фамилия, Возраст) и КЛИЕНТ_2 (#, Фамилия, Возраст)

КЛИЕНТ_1			КЛИЕНТ_2		
1	Иванов	20	1	Иванов	20
3	Петров	23	2	Исаев	30
4	Фролов	49			

Результирующее отношение:

3	Петров	23
4	Фролов	49

Реляционные операции

Декартово произведение отношений

Операция декартова произведения выполняется над двумя произвольными отношениями A и B.

Результатом операции декартова произведения является отношение C, степень которого равна сумме степеней исходных отношений, а мощность - произведению мощностей исходных отношений.

Таким образом, *декартово произведение* отношений можно представить с помощью декартова произведения множеств:

$$Q_c = Q_a \times Q_b.$$

Реляционные операции

Пример: Операция "декартово произведение".

Выполним операцию декартова произведения отношений СЛУЖАЩИЕ и МЕДОСМОТР.

Исходные отношения:

СЛУЖАЩИЕ (#, Фамилия, Пол)

МЕДОСМОТР (Процедура, Дата)

СЛУЖАЩИЙ			МЕДОСМОТР	
1	Иванов	М	ЭКГ	17.08
5	Антонова	Ж	Анализ крови	20.08
			Терапевт	23.08

Реляционные операции

Результирующее отношение:

РЕЗУЛЬТАТЫ_МЕДОСМОТРА (#, Фамилия, Пол, Процедура, Дата) =
СЛУЖАЩИЙ X МЕДОСМОТР

1	Иванов	м	ЭКГ	17.08
1	Иванов	м	Анализ крови	20.08
1	Иванов	м	Терапевт	23.08
5	Антонова	ж	ЭКГ	17.08
5	Антонова	ж	Анализ крови	20.08
5	Антонова	ж	Терапевт	23.08

Реляционные операции

Проекция отношения

Операция проекции выполняется над одним отношением A .

Результатом выполнения операции проекции над отношением A является отношение C , которое включает в себя все *кортежи* отношения A , но только с теми атрибутами, на которые выполняется *проекция*.

Операцию проекции отношения можно представить следующим образом:

$$Q_C = \pi_{i_1, i_2, \dots, i_N}(Q_A)$$

Реляционные операции

Пример: Проекция отношения.

Выполним операцию проекции отношения СОТРУДНИК на атрибуты ОТДЕЛ и ДОЛЖНОСТЬ.

Исходное отношение:

СОТРУДНИК (#, Фамилия, Отдел, Должность)

1	Иванов	12	Инженер
2	Исаев	11	Гл.специалист
3	Петров	11	Инженер
4	Фролов	11	Инженер
5	Антонова	12	Конструктор

Реляционные операции

Результирующее отношение:

ДОЛЖНОСТЬ (Отдел, Должность) = р (Отдел, Должность)

12	Инженер
11	Гл.специалист
11	Инженер
12	Конструктор

Реляционные операции

Деление отношений

Операция *деления* выполняется над двумя отношениями A и B , где A - отношение-делимое, а B - отношение-делитель.

При этом атрибуты B должны являться подмножеством атрибутов A .

Результатом выполнения операции деления является отношение C , которое включает в себя атрибуты отношения A , отличные от атрибутов отношения B , и только те кортежи, декартовы произведения которых с отношением B дают отношение A :

$$Q_c = Q_a \div Q_b = \pi_{1,2,\dots,(n-m)}(Q_a) - \pi_{1,2,\dots,(n-m)}(\pi_{1,2,\dots,(n-m)}(Q_a) \times Q_b) - Q_a.$$

Реляционные операции

Пример: Деление отношений.

Выполним операцию деления отношения РЕЗУЛЬТАТЫ_МЕДОСМОТРА на отношение МЕДОСМОТР.

Исходные отношения:

РЕЗУЛЬТАТЫ_МЕДОСМОТРА (#, Фамилия, Пол, Процедура, Дата)
МЕДОСМОТР (Процедура, Дата)

1	Иванов	М	ЭКГ	17.08
1	Иванов	М	Анализ крови	20.08
1	Иванов	М	Терапевт	23.08
5	Антонова	Ж	ЭКГ	17.08
5	Антонова	Ж	Анализ крови	20.08
5	Антонова	Ж	Терапевт	23.08

ЭКГ	17.08
Анализ крови	20.08
Терапевт	23.08

Реляционные операции

Результирующее отношение:

СЛУЖАЩИЙ (#, Фамилия, Пол) = РЕЗУЛЬТАТЫ_МЕДОСМОТРА / МЕДОСМОТР

1	Иванов	М
5	Антонова	Ж

Реляционные операции

Выбор из отношения

Операция выбора (*селекции*) выполняется над одним отношением A .

Результатом выполнения операции выбора является отношение C , которое включает в себя кортежи отношения A , удовлетворяющие заданному условию (критерию выбора).

Операция выбора из отношения может быть представлена следующим образом:

$$Q_c = \sigma_F(Q_a),$$

Реляционные операции

где σ - обозначает операцию выбора, F - критерий выбора на множестве атрибутов в форме логического выражения, образованного с помощью определенных операндов (константы, имена атрибутов, арифметические операции сравнения, логические операции).

Реляционные операции

Пример: Селекция отношения.

Произведем выбор из отношения СЛУЖАЩИЕ по критерию "Возраст \geq 30".

Исходное отношение:
СЛУЖАЩИЕ (#, Фамилия, Возраст)

1	Иванов	20
2	Исаев	30
3	Петров	23
4	Фролов	49
5	Антонова	25

Реляционные операции

Критерий выбора: Возраст \geq 30

Результирующее отношение:
СЛУЖАЩИЕ (#, Фамилия, Возраст)

2	Исаев	30
4	Фролов	49

Реляционные операции

Соединение отношений

Операция *q-соединения* выполняется над двумя отношениями А и В.

Результатом выполнения операции *-соединения* является отношение С, которое включает в себя все кортежи со всеми атрибутами исходных отношений А и В, удовлетворяющими заданному условию.

В каждом отношении выделяется *атрибут*, по которому выполняется соединение.

Реляционные операции

Операция соединения отношений может быть представлена следующим образом:

$$Q_c = Q_a \bowtie Q_b = \sigma_{i\theta(n+j)}(Q_a \times Q_b),$$

где n - степень отношения Q_a ; θ - арифметический оператор сравнения; i, j - номера атрибутов в Q_a и Q_b соответственно, по которым выполняется *соединение*.

Реляционные операции

Пример: Соединение отношений.

Выполним операцию естественного соединения отношений ЭКЗАМЕН_ВЕДОМОСТЬ и ГРУППА по атрибуту "Группа".

Исходные отношения:

ЭКЗАМЕН_ВЕДОМОСТЬ (Студент, Дисциплина, Оценка, Группа)

ГРУППА (Курс, Группа, Наименование)

Иванов	Математика	5	12	5	10	АСУ
Петров	Математика	3	10	5	11	Прикладная математика
Исаев	Математика	4	11			
Антонова	Математика	3	12			

Реляционные операции

Результирующее отношение:

РЕЗУЛЬТАТ (Студент, Дисциплина, Оценка, Группа, Курс, Наименование)

Петров	Математика	3	10	5	АСУ
Исаев	Математика	4	11	5	Прикладная математика