



tsegorah 16 июня 2012

## Основы реляционной алгебры

Реляционная алгебра базируется на теории множеств и является основой логики работы баз данных.

Когда я только изучал устройство баз данных и SQL, предварительное ознакомление с реляционной алгеброй очень помогло дальнейшим знаниям правильно уложиться в голове, и я постараюсь что бы эта статья произвела подобный эффект.

Так что если вы собираетесь начать свое обучение в этой области или вам просто стало интересно, прошу под кат.

### Реляционная база данных

Для начала введем понятие реляционной базы данных, в которой будем выполнять все действия.

Реляционной базой данных называется совокупность отношений, содержащих всю информацию, которая должна храниться в базе. В данном определении нам интересен термин отношение, но пока оставим его без строго определения.

Лучше представим себе таблицу продуктов.

таблица PRODUCTS

ID	NAME	COMPANY	PRICE
123	Печеньки	ООО "Темная сторона"	190
156	Чай	ООО "Темная сторона"	60
235	Ананасы	ОАО "Фрукты"	100
623	Томаты	ООО "Овощи"	130



+64



654



42

Таблица состоит из 4х строк, строка в таблице является кортежем в реляционной теории. Множество упорядоченных кортежей называется отношением. Перед тем как дать определение отношения, введем еще один термин — домен. Домены применительно к таблице это столбцы.

Для ясности, теперь введем строгое определение отношения.

Пусть даны  $N$  множеств  $D_1, D_2, \dots, D_n$  (домены), отношением  $R$  над этими множествами называется множество упорядоченных  $N$ -кортежей вида  $\langle d_1, d_2, \dots, d_n \rangle$ , где  $d_1$  принадлежит  $D_1$  и т.д. Множества  $D_1, D_2, \dots, D_n$  называются доменами отношения  $R$ . Каждый элемент кортежа представляет собой значение одного из атрибутов, соответствующего одному из доменов.

### Ключи в отношениях

В отношении требованием является то, что все кортежи должны различаться. Для однозначной идентификации кортежа существует первичный ключ. Первичный ключ это атрибут или набор из минимального числа атрибутов, который однозначно идентифицирует конкретный кортеж и не содержит дополнительных атрибутов. Подразумевается, что все атрибуты в первичном ключе должны быть необходимыми и достаточными для идентификации конкретного кортежа, и исключение любого из атрибутов в ключе сделает его недостаточным для идентификации. Например, в такой таблице ключом будет сочетание атрибутов из первого и второго столбца.

таблица DRIVERS

COMPANY	DRIVER
ООО "Темная сторона"	Владимир
ООО "Темная сторона"	Михаил
ОАО "Фрукты"	Руслан
ООО "Овощи"	Владимир

Видно, что в организации может быть несколько водителей, и чтобы однозначно идентифицировать водителя необходимо и значение из столбца "Название организации" и

В реляционной БД таблицы взаимосвязаны и соотносятся друг с другом как главные и подчиненные. Связь главной и подчиненной таблицы осуществляется через первичный ключ (primary key) главной таблицы и внешний ключ (foreign key) подчиненной таблицы. Внешний ключ это атрибут или набор атрибутов, который в главной таблице является первичным ключом.

Этой подготовительной теории будет достаточно для знакомства с основными операциями реляционной алгебры.

## Операции реляционной алгебры

Основные восемь операций реляционной алгебры были предложены [Э.Коддом](#).

- Объединение
- Пересечение
- Вычитание
- Декартово произведение
- Выборка
- Проекция
- Соединение
- Деление

Первая половина операций аналогична таким же операциям над множествами. Часть операций можно выразить через другие операции. Рассмотрим большую часть операций с примерами.

Для понимания важно запомнить, что результатом любой операции алгебры над отношениями является еще одно отношение, которое можно потом так же использовать в других операциях.

Создадим еще одну таблицу, которая нам пригодится в примерах.

таблица SELLERS

ID	SELLER
123	ООО "Дарт"
156	ОАО "Ведро"

Условимся, что в этой таблице ID это внешний ключ, связанный с первичным ключом таблицы PRODUCTS.

Для начала рассмотрим самую простую операцию — имя отношения. Её результатом будет такое же отношение, то есть выполнив операцию PRODUCTS, мы получим копию отношения PRODUCTS.

### Проекция

Проекция является операцией, при которой из отношения выделяются атрибуты только из указанных доменов, то есть из таблицы выбираются только нужные столбцы, при этом, если получится несколько одинаковых кортежей, то в результирующем отношении остается только по одному экземпляру подобного кортежа.

Для примера сделаем проекцию на таблице PRODUCTS выбрав из нее ID и PRICE.

Синтаксис операции:

$\pi_{(ID, PRICE)} PRODUCTS$

В результате этой операции получим отношение:

ID	PRICE
123	190
156	60
235	100
623	130

### Выборка

Выборка — это операция, которая выделяет множество строк в таблице, удовлетворяющих заданным условиям. Условием может быть любое логическое выражение.

Для примера сделаем выборку из таблицы с ценой больше 90.

Синтаксис операции:

ID	NAME	COMPANY	PRICE
123	Печеньки	ООО "Темная сторона"	190
235	Ананасы	ОАО "Фрукты"	100
623	Томаты	ООО "Овощи"	130

В условии выборки мы можем использовать любое логическое выражение. Сделаем еще одну выборку с ценой больше 90 и ID товара меньше 300:

$\sigma_{(PRICE > 90 \wedge ID < 300)}$  PRODUCTS

ID	NAME	COMPANY	PRICE
123	Печеньки	ООО "Темная сторона"	190
235	Ананасы	ОАО "Фрукты"	100

Совместим операторы проекции и выборки. Мы можем это сделать, потому что любой из операторов в результате возвращает отношение и в качестве аргументов использует также отношение.

Из таблицы с продуктами выберем все компании, продающие продукты дешевле 110.

$\pi_{COMPANY} \sigma_{(PRICE < 110)}$  PRODUCTS

COMPANY
ООО "Темная сторона"
ОАО "Фрукты"

## Умножение

отношениями, в результате которой мы получаем отношение со всеми доменами из двух начальных отношений. Кортежи в этих доменах будут представлять из себя все возможные сочетания кортежей из начальных отношений. На примере будет понятнее.

Получим декартово произведения таблиц PRODUCTS и SELLERS.  
Синтаксис операции:

PRODUCTS × SELLERS

Можно заметить, что у двух этих таблиц есть одинаковый домен ID. В подобной ситуации домены с одинаковыми названиями получают префикс в виде названия соответствующего отношения, как показано ниже.

Для краткости перемножим не полные отношения, а выборки с условием ID<235

(цветом выделены одни и те же кортежи)

PRODUCTS.ID	NAME	COMPANY	PRICE	SELLERS.ID	SELLER
123	Печеньки	ООО "Темная сторона"	190	123	ООО "Дарт"
156	Чай	ООО "Темная сторона"	60	156	ОАО "Ведро"
123	Печеньки	ООО "Темная сторона"	190	156	ОАО "Ведро"
156	Чай	ООО "Темная сторона"	60	123	ООО "Дарт"

Для примера использования этой операции представим себе необходимость выбрать продавцов с ценами меньше 90. Без произведения необходимо было бы сначала получить ID продуктов из первой таблицы, потом по этим ID из второй таблицы получить нужные имена SELLER, а с использованием произведения будет такой запрос:

$\Pi_{(SELLER)} \sigma_{(PRODUCTS.ID=SELLERS.ID \wedge PRICE<90)} PRODUCTS \times SELLERS$

В результате этой операции получим отношение:

SELLER
--------

## Соединение и естественное соединение

Операция соединения обратна операции проекции и создает новое отношение из двух уже существующих. Новое отношение получается конкатенацией кортежей первого и второго отношений, при этом конкатенации подвергаются отношения, в которых совпадают значения заданных атрибутов. В частности, если соединить отношения PRODUCTS и SELLERS, этими атрибутами будут атрибуты доменов ID.

Также для понятности можно представить соединение как результат двух операций. Сначала берется произведение исходных таблиц, а потом из полученного отношения мы делаем выборку с условием равенства атрибутов из одинаковых доменов. В данном случае условием является равенство PRODUCTS.ID и SELLERS.ID.

Попробуем соединить отношения PRODUCTS и SELLERS и получим отношение.

PRODUCTS.ID	NAME	COMPANY	PRICE	SELLERS.ID	SELLER
123	Печеньки	ООО "Темная сторона"	190	123	ООО "Дарт"
156	Чай	ООО "Темная сторона"	60	156	ОАО "Ведро"
235	Ананасы	ОАО "Фрукты"	100	235	ЗАО "Овоще База"
623	Томаты	ООО "Овощи"	130	623	ОАО "Фирма"

Натуральное соединение получает схожее отношение, но в случае, если у нас корректно настроена схема в базе ( в данном случае первичный ключ таблицы PRODUCTS ID связан с внешним ключем таблицы SELLERS ID), то в результирующем отношении остается один домен ID.

Синтаксис операции:

Получится такое отношение:

PRODUCTS.ID	NAME	COMPANY	PRICE	SELLER
123	Печеньки	ООО "Темная сторона"	190	ООО "Дарт"
156	Чай	ООО "Темная сторона"	60	ОАО "Ведро"
235	Ананасы	ОАО "Фрукты"	100	ЗАО "Овоще База"
623	Томаты	ООО "Овощи"	130	ОАО "Фирма"

Пересечение и вычитание.

Результатом операции пересечения будет отношение, состоящее из кортежей, полностью входящих в состав обоих отношений.

Результатом вычитания будет отношение, состоящее из кортежей, которые являются кортежами первого отношения и не являются кортежами второго отношения.

Данные операции аналогичны таким же операциям над множествами, так что, я думаю, нет необходимости подробно их расписывать.

Источники информации

- Основы использования и проектирования баз данных — В. М. Илющечкин
- курс лекций Introduction to Databases — Jennifer Widom, Stanford University

Буду благодарен за аргументированные замечания

Теги: [реляционные базы данных](#), [база данных](#), [реляционная алгебра](#)

 +64

 165,8k

 654



33

0

Карма

Рейтинг

Юрий Соколов [@tsegorah](#)

 +64

 654

 42



Поделиться публикацией



 Комментарии 42

#### ПОХОЖИЕ ПУБЛИКАЦИИ

14 сентября 2013

**Руководство по проектированию реляционных баз данных (10-13 часть из 15) [перевод]**

11 сентября 2013

**Руководство по проектированию реляционных баз данных (7-9 часть из 15) [перевод]**

11 сентября 2013

**Руководство по проектированию реляционных баз данных (4-6 часть из 15) [перевод]**

#### ПОПУЛЯРНОЕ ЗА СУТКИ

вчера в 19:26

**Удаленное управление компьютером через браузер**

вчера в 21:19

**Андрей Терехов (продолжение): «Когда говорят, что мы отстали от американцев навсегда, я отвечаю: „Не дождетесь!“»**

 +64

 654

 42

сегодня в 12:25

**Только не говорите, что вы таким не были или мой взгляд на обучение**

сегодня в 07:35

**Деревянные игрушки, часть шестая — 1992**

сегодня в 14:20

**Инициализация в современном C++**

habr



Настройка языка

Полная версия

© 2006–2019 «ТМ»



+64



654



42