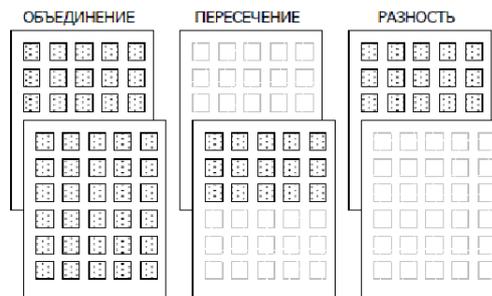


Основы реляционной алгебры.

Впервые термин "реляционная модель данных" появился в 1970 году в ИВМ. Его ввел Кодд. Он предложил использовать для обработки данных аппарат теории множеств (объединение, пересечение, разность, декартово произведение). Он показал, что любое представление данных сводится к совокупности двумерных таблиц особого вида, известного в математике как *отношение* – relation (англ.). Реляционной является БД, в которой все данные, доступные пользователю, организованы в виде набора двумерных таблиц, а все операции над данными сводятся к операциям над этими таблицами.



Предложив реляционную модель данных, Кодд создал и инструмент для удобной работы с отношениями – реляционную алгебру. Каждая операция этой алгебры использует одну или несколько таблиц (отношений) в качестве ее операндов и получает в результате новую таблицу, т.е. позволяет "разрезать" или "склеивать" таблицы.

Реляционная алгебра в явном виде представляет набор операций, которые можно использовать, чтобы сообщить системе, как в базе данных из определенных отношений реально построить необходимое отношение. Реляционные операторы обладают одним важным свойством: они замкнуты относительно понятия отношения. Это означает, что выражения реляционной алгебры определяются над отношениями реляционных БД и результатом вычисления также являются отношения.

Набор основных алгебраических операций состоит из восьми операций, которые делятся на два класса - теоретико-множественные операции и специальные реляционные операции, дополненные некоторыми специальными операциями, специфичными для баз данных.

В состав теоретико-множественных операций входят традиционные операции над множествами:

- объединение;
- пересечение;
- разность;
- декартово произведение.

Специальные реляционные операции включают:

- выборку;
- проекцию;

- естественное соединение;
- деление.

Операции объединения, пересечения и разности требуют от операндов совместимости по типу. Два отношения совместимы по типу, если:

1. каждое из них имеет одно и то же множество имен атрибутов
2. соответствующие атрибуты (с одинаковыми именами) определены в одной и той же области.

Рассмотрим примеры: есть два отношения (таблицы) А и В:

ID_NUM	NAME	CITY	AGE
1809	Иванов	Москва	45
1996	Петров	Нижний Новгород	39
1777	Сидоров	Рязань	21

ID_NUM	NAME	CITY	AGE
1809	Иванов	Москва	45
1896	Галкин	Иваново	40

Объединением двух совместимых по типу отношений А и В ($A \cup B$) называется отношение с тем же заголовком, как в отношениях А и В, и с телом, состоящим из множества строк t , принадлежащих А или В или обоим отношениям.

ID_NUM	NAME	CITY	AGE
1809	Иванов	Москва	45
1996	Петров	Нижний Новгород	39
1777	Сидоров	Рязань	21
1896	Галкин	Иваново	40

При выполнении операции объединения двух отношений создается отношение, включающее кортежи, входящие хотя бы в одно из отношений-операндов. Обратите внимание, что повторяющиеся кортежи удаляются по определению отношения.

Пересечением двух совместимых по типу отношений А и В ($A \cap B$) называется отношение с тем же заголовком, как в отношениях А и В, и с телом, состоящим из множества строк t , принадлежащих одновременно обоим отношениям А и В. Операция пересечения двух отношений создает отношение, включающее все кортежи, входящие в оба отношения-операнда.

ID_NUM	NAME	CITY	AGE
1809	Иванов	Москва	45

Разностью двух совместимых по типу отношений А и В ($A - B$) называется отношение с тем же заголовком, как в отношениях А и В, и с телом, состоящим из множества кортежей t , принадлежащих отношению А и

не принадлежащих отношению В. Отношение, являющееся разностью двух отношений включает все кортежи, входящие в первое отношение, такие, что ни один из них не входит во второе отношение.

A - B

ID_NUM	NAME	CITY	AGE
1996	Петров	Нижний Новгород	39
1777	Сидоров	Рязань	21

Декартово произведение двух отношений А и В ($A \times B$), где А и В не имеют общих имен атрибутов, определяется как отношение с заголовком, представляющим собой сцепление двух заголовков исходных отношений А и В, и телом, состоящим из множества кортежей t таких что первым является любой кортеж отношения А, а вторым – любой кортеж, принадлежащий отношению В. Кардинальное число результирующего отношения равно произведению кардинальных чисел исходных отношений, а степень равняется сумме степеней.

Пусть отношение А – содержит имена всех текущих поставщиков, а отношение В – номера всех текущих деталей. Тогда $A \times B$ – это все текущие пары поставщик – деталь и деталь – поставщик.

S1
S2
S3

P1
P2
P3
P4

S1	P1
S1	P2
S1	P3
S1	P4

S2	P1
S2	P2
S2	P3
S2	P4

S3	P1
S3	P2
S3	P3
S3	P4

Операция выборка (или операция ограничение отношения) - создает новое отношение, содержащее только те строки отношения – операнда, которые удовлетворяют некоторому условию ограничения.

A where CITY = 'Москва' and AGE < 40

Операция выборки

Операция выборки работает с одним отношением R_1 и определяет результирующее отношение R, которое содержит только те кортежи (или строки, или записи), отношения R_1 , которые удовлетворяют заданному условию (предикату P).

Таким образом, операция выборки - унарная операция - и записывается следующим образом:

$$R = \sigma_P(R_1)$$

где P - предикат (логическое условие).

В таблице ниже дано одно отношение, с которым работает эта операция.

R3

A1	A2	A3	A4
3	hh	yl	ms
4	pp	a1	sr
1	rr	yl	ms

Просматриваем столбец A3 и устанавливаем, что предикату $A3 > 'd0'$ удовлетворяют записи в первой и третьей строках исходного отношения (так как номер буквы y в алфавите больше номера буквы d). В результате получаем следующее новое отношение, в котором две строки:

R			
A1	A2	A3	A4
3	hh	yl	ms
1	rr	yl	ms

Операция проекции

Операция проекции ($R = \pi_{a_1, \dots, a_n}(R_1)$) работает, как и операция выборки, только с одним отношением R_1 и определяет новое отношение R , в котором есть лишь те атрибуты (столбцы), которые заданы в операции, и их значения.

Пусть вновь дано то же отношение $R3$:

R3			
A1	A2	A3	A4
3	hh	yl	ms
4	pp	a1	sr
1	rr	yl	ms

Из исходного отношения выбираем только столбцы A4 и A3 и видим, что строки со значениями - первая и третья - идентичны. Исключаем дубликат и получаем следующее новое отношение, в котором два атрибута и две строки (записи):

R	
A4	A3
ms	yl
sr	a1

Операция деления

Результатом операции деления ($R = R_1 \div R_2$) является набор кортежей (строк) отношения R_1 , которые соответствуют комбинации всех кортежей отношения R_2 . Для этого нужно, чтобы в отношении R_2 была часть атрибутов (можно и один), которые есть в отношении R_1 . В результирующем

отношении присутствуют только те атрибуты отношения R1, которых нет в отношении R2.

Даны два отношения R5 и R6:

R5				R6	
A1	A2	A3	A4	A2	A3
2	S3	4	sun	R4	8
3	X8	7	kab	X8	7
3	R4	8	kab		

Комбинации всех кортежей отношения R6 соответствуют вторая и третья строки отношения R5. Но после исключения атрибутов (столбцов) A2 и A3 эти строки становятся идентичными. Поэтому в новом отношении присутствует эта строка один раз. Новое отношение:

R	
A1	A4
3	kab

Операция тета-соединения

В результате этой операции получается отношение, которое содержит кортежи из декартова произведения отношений R1 и R2 удовлетворяющие предикату P. Значением предиката P может быть один из операторов сравнения (<, <=, >, >=, = или !=).

Даны два отношения R3 и R4:

R3				R4	
A1	A2	A3	A4	A5	A6
3	hh	yl	ms	3	hh
4	pp	a1	sr	4	pp
1	rr	yl	ms		

Это таблицы (отношения) из главы о декартовом произведении. Выполняем операцию декартового произведения. Видим, что условию предиката P удовлетворяет один кортеж декартового произведения - первый (так как $3 \geq 3$). Получаем следующее новое отношение:

R					
A1	A2	A3	A4	A5	A6
3	hh	yl	ms	3	hh