

«Базы данных»  
Лекция № 2

Начальник отдела НИЧ, к.э.н., доцент Д.Г. Корнеев

2009 год

# Реляционная алгебра

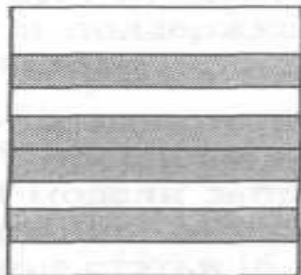
*Реляционная алгебра* — это коллекция операций, которые принимают отношения в качестве операндов и возвращают отношение в качестве результата.

Первая версия этой алгебры была определена Коддом.

Эта "оригинальная" алгебра включала **восемь операций**, которые подразделялись на описанные ниже две группы с четырьмя операциями каждая.

- Традиционные операции с множествами — *объединение, пересечение, разность и декартово произведение* (все они были немного модифицированы с учетом того факта, что их операндами являются именно отношения, а не произвольные множества).
- Специальные реляционные операции, такие как *сокращение* (известное также под названием *выборки*), *проекция, соединение и деление*.

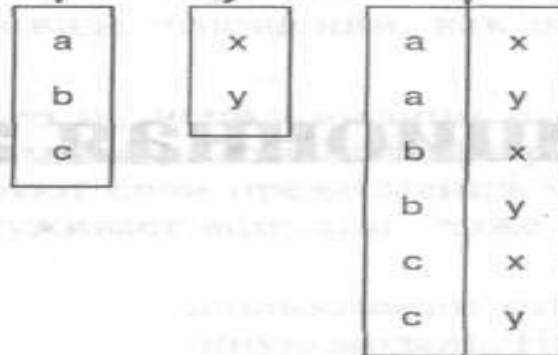
Сокращение



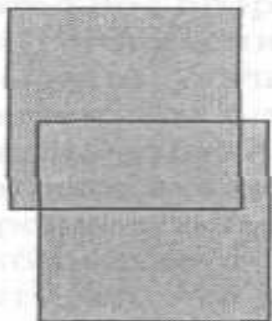
Проекция



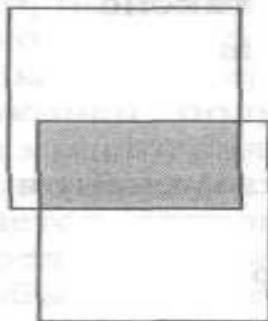
Произведение



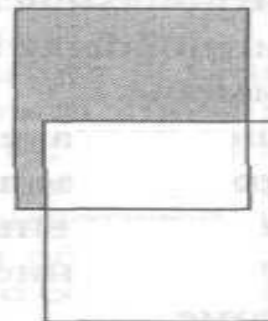
Объединение



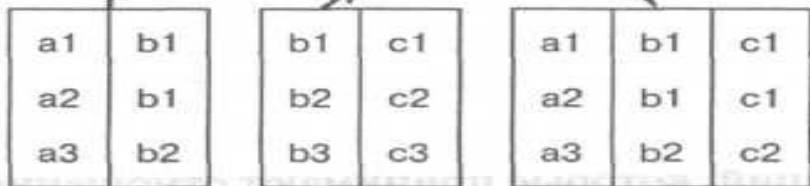
Пересечение



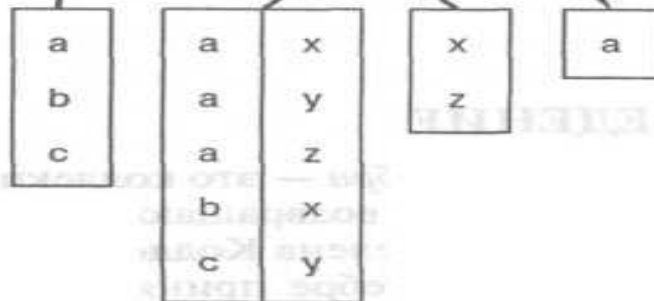
Разность



(Естественное) соединение



Деление



# Реляционная алгебра

## Объединение

*В математике объединение двух множеств представляет собой множество всех элементов, принадлежащих либо к одному из них, либо к обоим заданным множествам.*

Поскольку любое отношение представляет собой (или, скорее, содержит) множество (а именно множество кортежей), оно, безусловно, позволяет формировать объединение двух таких множеств; результатом является множество, состоящее из всех кортежей, присутствующих либо в одном, либо в обоих из заданных отношений.

Например, объединение множества кортежей поставщиков, которые в настоящее время присутствуют в  $S$ , и множества кортежей деталей, присутствующих в настоящее время в отношении  $P$ , безусловно, представляет собой множество.

## Примеры отношений

- Поставщики S { S#, SNAME, STATUS, CITY }  
PRIMARY KEY { S# }
- Поставки P { P#, PNAME, COLOR, WEIGHT, CITY } PRIMARY KEY { P# }
- SP { S#, P#, QTY }  
PRIMARY KEY { S#, P# }  
FOREIGN KEY { S# } REFERENCES S  
FOREIGN KEY { P# } REFERENCES P

# Объединение отношений

- Но хотя этот результат можно назвать *множеством*, он не является *отношением*; отношения не могут содержать смесь кортежей разных типов, поскольку они должны включать однотипные кортежи.  
А нам требуется, чтобы результат представлял собой отношение, поскольку необходимо сохранить реляционное свойство замкнутости.
- Поэтому объединение в реляционной алгебре не полностью соответствует общему определению объединения в математике; скорее, оно является объединением особого рода, в котором два входных отношения должны принадлежать **к одному типу**.

# Объединение отношений

Поэтому определение операции реляционного объединения должно быть таким: если даны отношения  $A$  и  $B$  одного и того же типа, то **объединение** этих отношений  $A \text{ UNION } B$  является отношением того же типа с телом, которое состоит из всех кортежей  $t$ , присутствующих в  $A$  или  $B$  или в обоих отношениях.

A

S#	SNAME	STATUS	CITY
S1	Smith	20	London
S4	Clark	20	London

B

S#	SNAME	STATUS	CITY
S1	Smith	20	London
S2	Jones	10	Paris

1. Объединение  
(A UNION B)

S#	SNAME	STATUS	CITY
S1	Smith	20	London
S4	Clark	20	London
S2	Jones	10	Paris

2. Пересечение  
(A INTERSECT B)

S#	SNAME	STATUS	CITY
S1	Smith	20	London

3. Разность  
(A MINUS B)

S#	SNAME	STATUS	CITY
S4	Clark	20	London

4. Разность  
(B MINUS A)

S#	SNAME	STATUS	CITY
S2	Jones	10	Paris



# Объединение отношений

- **Пример.** Предположим, что отношения А и В имеют вид, показанный на рис. (оба они получены из текущего значения отношения поставщиков S;

В «А» приведены данные о поставщиках из Лондона, а в «В» — данные о поставщиках, которые поставляют деталь Р1). В таком случае в объединение  $A \cup B$  (см. рис. а) входят поставщики, которые либо находятся в Лондоне, либо поставляют деталь Р1, либо соответствуют обоим этим условиям.

*Обратите внимание на то, что в результате содержатся три кортежа, а не четыре; по определению отношения никогда не содержат дубликаты кортежей (поэтому, неформально выражаясь, из результатов операции объединения "устраняются дубликаты").*

# Пересечение отношений

- Как и для объединения, и фактически по той же причине, для реляционной операции пересечения требуется, чтобы ее операнды принадлежали к одному и тому же типу.
- Если даны отношения  $A$  и  $B$  одного и того же типа, то пересечением этих отношений  $A$  *INTERSECT*  $B$  является отношение того же типа с телом, состоящим из всех кортежей  $t$ , таких, что  $t$  присутствует одновременно в  $A$  и  $B$ .

# Пересечение отношений

*Пример.*

Снова предположим, что отношения A и B показаны на рис. Тогда пересечение  $A \text{ INTERSECT } B$  (рис. б) включает всех поставщиков, которые находятся в Лондоне и поставляют деталь P1.

# Разность отношений

Как и для объединения и пересечения, для *реляционной операции разности* требуется, чтобы ее операнды принадлежали к одному и тому же типу.

- Если даны отношения  $A$  и  $B$  одного и того же типа, то **разностью** этих отношений  $A \text{ MINUS } B$  (в указанном порядке), является отношение того же типа с телом, состоящим из всех кортежей  $t$ , таких, что  $t$  присутствует в  $A$ , но не в  $B$ .

# Разность отношений

*Пример.* Снова предположим, что отношения  $A$  и  $B$  показаны на рис.

- Тогда результат *операции разности  $A \text{ MINUS } B$*  (рис. *в*) включает поставщиков, которые находятся в Лондоне и не поставляют деталь  $P1$ ;
- Результат *операции разности  $B \text{ MINUS } A$*  (рис. *г*) включает поставщиков, которые поставляют деталь  $P1$  и не находятся в Лондоне.
- Обратите внимание на то, что оператор  $\text{MINUS}$  характеризуется направленностью (некоммутативностью), так же, как вычитание в обычной арифметике (например, " $5 - 2$ " и " $2 - 5$ " не являются одним и тем же).

# Произведение отношений

- В математике *декартовым произведением* (или сокращенно *произведением*) двух множеств является множество всех таких упорядоченных пар, что в каждой паре первый элемент берется из первого множества, а второй — из второго множества.
- Поэтому декартово произведение двух отношений, неформально выражаясь, представляет собой множество упорядоченных пар кортежей. Но мы снова должны сохранить свойство замкнутости;
- Иными словами, необходимо, чтобы результат содержал кортежи как таковые, а не упорядоченные пары кортежей.

# Произведение отношений

Поэтому реляционной версией *декартова произведения* служит расширенная форма этой операции, в которой каждая упорядоченная пара кортежей заменяется одним кортежем, являющимся *объединением* двух рассматриваемых кортежей.

Это означает, что если даны следующие кортежи:

$\{ A_1, A_2, \dots, A_m \}$

и

$\{ B_1, B_2, \dots, B_n \}$

то теоретико-множественное объединение этих двух кортежей представляет собой приведенный ниже единственный кортеж:

$\{ A_1, A_2, \dots, A_m, B_1, B_2, \dots, B_n \}$

# Произведение отношений

Итак, определим (реляционное) **декартово произведение**  $A \text{ TIMES } B$  отношений  $A$  и  $B$ , не имеющих общих атрибутов, как отношение, заголовок которого представляет собой (теоретико-множественное) объединение заголовков отношений  $A$  и  $B$ , а тело состоит из всех кортежей  $t$ , таких, что  $t$  является (теоретико-множественным) объединением кортежа, принадлежащего к отношению  $A$ , и кортежа, принадлежащего к отношению  $B$ .



A

S #
S1
S2
S3
S4
S5

B

P #
P1
P2
P3
P4
P5
P6

Декартово произведение (A TIMES B)

S#	P#
S1	P1
S1	P2
S1	P3
S1	P4
S1	P5
S1	P6
..	..

..	..
S2	P1
S2	P2
S2	P3
S2	P4
S2	P5
S2	P6
..	..

..	..
S3	P1
S3	P2
S3	P3
S3	P4
S3	P5
S3	P6
..	..

..	..
S4	P1
S4	P2
S4	P3
S4	P4
S4	P5
S4	P6
..	..

..	..
S5	P1
S5	P2
S5	P3
S5	P4
S5	P5
S5	P6

# Оператор сокращения

- Оператор сокращения по сути позволяет получить "горизонтальное" подмножество заданного отношения, т.е. подмножество кортежей заданного отношения, для которых удовлетворяется некоторое указанное условие.

S WHERE CITY = 'London'	<table border="1"><thead><tr><th>S#</th><th>SNAME</th><th>STATUS</th><th>CITY</th></tr></thead><tbody><tr><td>S1</td><td>Smith</td><td>20</td><td>London</td></tr><tr><td>S4</td><td>Clark</td><td>20</td><td>London</td></tr></tbody></table>	S#	SNAME	STATUS	CITY	S1	Smith	20	London	S4	Clark	20	London			
S#	SNAME	STATUS	CITY													
S1	Smith	20	London													
S4	Clark	20	London													
P WHERE WEIGHT < WEIGHT (14.0)	<table border="1"><thead><tr><th>P#</th><th>PNAME</th><th>COLOR</th><th>WEIGHT</th><th>CITY</th></tr></thead><tbody><tr><td>P1</td><td>Nut</td><td>Red</td><td>12.0</td><td>London</td></tr><tr><td>P5</td><td>Cam</td><td>Blue</td><td>12.0</td><td>Paris</td></tr></tbody></table>	P#	PNAME	COLOR	WEIGHT	CITY	P1	Nut	Red	12.0	London	P5	Cam	Blue	12.0	Paris
P#	PNAME	COLOR	WEIGHT	CITY												
P1	Nut	Red	12.0	London												
P5	Cam	Blue	12.0	Paris												
SP WHERE S# = S# ('S6') OR P# = P# ('P7')	<table border="1"><thead><tr><th>S#</th><th>P#</th><th>QTY</th></tr></thead><tbody><tr><td></td><td></td><td></td></tr></tbody></table>	S#	P#	QTY												
S#	P#	QTY														

# Операция проекции

Предположим, что отношение  $A$  имеет атрибуты  $X, Y, \dots, Z$  (и, возможно, другие атрибуты). В таком случае **проекция** отношения  $A$  по атрибутам  $X, Y, \dots, Z$ , которая определяется с помощью следующего выражения:

$$A \{ X, Y, \dots, Z \}$$

является отношением, соответствующим описанным ниже требованиям:

- Его заголовок формируется из заголовка отношения  $a$  путем удаления всех атрибутов, не указанных в множестве  $\{ X, Y, \dots, Z \}$ .
- Тело состоит из всех кортежей  $\{ x, y, \dots, z \}$ , таких что в отношении  $A$  присутствует кортеж со значением  $x$  атрибута  $X$ ,  $y$  атрибута  $Y$ ... и  $z$  атрибута  $Z$ .

# Операция проекции

S { CITY }

CITY
London
Paris
Athens

P { COLOR, CITY }

COLOR	CITY
Red	London
Green	Paris
Blue	Oslo
Blue	Paris

( S WHERE CITY = 'Paris' ) { S# }

S#
S2
S3

# Операция соединения

Предположим, что отношения  $A$  и  $B$ , соответственно, имеют следующие атрибуты.

$X_1, X_2, \dots, X_m, Y_1, Y_2, \dots, Y_n$  ( $A$ )

$Y_1, Y_2, \dots, Y_n, Z_1, Z_2, \dots, Z_p$  ( $B$ )

- Это означает, что два рассматриваемых отношения имеют **общее множество атрибутов**  $Y$ , состоящее из атрибутов  $Y_1, Y_2, \dots, Y_n$  (и только из этих атрибутов),
- Другие атрибуты отношения  $A$  образуют множество  $X$ , состоящее из атрибутов  $X_1, X_2, X_m$ ,
- Другие атрибуты отношения  $B$  образуют множество  $Z$ , состоящее из атрибутов  $Z_1, Z_2, \dots, Z_p$ .

# Операция соединения

- Теперь множества  $\{ X_1, X_2, \dots, X_m \}, \{ Y_1, Y_2, \dots, Y_n \}$  и  $\{ Z_1, Z_2, \dots, Z_p \}$  могут рассматриваться, соответственно, как три составных атрибута  $X, Y$  и  $Z$ .
- В таком случае (естественное) **соединение**  $A$  и  $B$  выражается следующим образом.

## ***A JOIN B***

Оно представляет собой отношение с заголовком  $\{X, Y, Z\}$  и телом, состоящим из всех таких кортежей  $\{X_x, Y_y, Z_z\}$ , что любой из этих кортежей присутствует и в отношении  $A$ , со значением  $x$  атрибута  $X$  и значением  $y$  атрибута  $Y$ , и в отношении  $B$ , со значением  $y$  атрибута  $Y$  и значением  $z$  атрибута  $Z$ .

S#	SNAME	STATUS	CITY	P#	PNAME	COLOR	WEIGHT
S1	Smith	20	London	P1	Nut	Red	12.0
S1	Smith	20	London	P4	Screw	Red	14.0
S1	Smith	20	London	P6	Cog	Red	19.0
S2	Jones	10	Paris	P2	Bolt	Green	17.0
S2	Jones	10	Paris	P5	Cam	Blue	12.0
S3	Blake	30	Paris	P2	Bolt	Green	17.0
S3	Blake	30	Paris	P5	Cam	Blue	12.0
S4	Clark	20	London	P1	Nut	Red	12.0
S4	Clark	20	London	P4	Screw	Red	14.0
S4	Clark	20	London	P6	Cog	Red	19.0

# Операция деления

Предположим, что отношения **A** и **B**, соответственно, имеют следующие атрибуты:

- $X_1, X_2, \dots, X_m$  (**A**);  $Y_1, Y_2, \dots, Y_n$  (**B**)

Здесь ни один из атрибутов  $X_i$  ( $i = 1, 2, \dots, m$ ) не имеет одинакового имени с любым из атрибутов  $Y_j$  ( $j = 1, 2, \dots, n$ ).

Пусть отношение **C** имеет следующие атрибуты:

- $X_1, X_2, \dots, X_m, Y_1, Y_2, \dots, Y_n$  (**C**)

ЭТО означает, что **C** имеет заголовок, представляющий собой (теоретико-множественное) объединение заголовков **A** и **B**.

Будем рассматривать множества  $\{X_1, X_2, \dots, X_m\}$

$\{Y_1, Y_2, \dots, Y_n\}$ , соответственно, как составные атрибуты **X** и

**Y**. В таком случае операция *деления A на B по C* (где *a* — делимое, *B* — делитель, а *C* — посредник) может быть

представлена с помощью следующего выражения.

- ***A DIVIDE BY B PER C***



# Операция деления

- **Деление** представляет собой отношение с заголовком  $\{X\}$  и телом, состоящим из всех кортежей  $\{X x\}$ , присутствующих в  $A$ , причем таких, что кортеж  $\{X x, Y y\}$  присутствует в  $C$  для всех кортежей  $\{Y y\}$ , присутствующих в  $B$ .
- Иными словами, данный результат состоит из тех значений  $x$ , присутствующих в  $A$ , для которых соответствующие значения  $y$  в  $C$  включают все значения  $y$  из  $B$ .

# Пример деления

На рис. приведены некоторые примеры деления. В каждом случае делимое (DEND) представляет собой проекцию текущего значения отношения  $S$  по атрибуту  $S\#$ ; посредник (MED) в каждом случае является проекцией текущего значения отношения  $SP$  по атрибутам  $S\#$  и  $P\#$ ; а три делителя (DOR) являются такими, как указано на этом рисунке.

<p>DEND</p> <table border="1"> <thead> <tr><th>S#</th></tr> </thead> <tbody> <tr><td>S1</td></tr> <tr><td>S2</td></tr> <tr><td>S3</td></tr> <tr><td>S4</td></tr> <tr><td>S5</td></tr> </tbody> </table>	S#	S1	S2	S3	S4	S5	<p>MED</p> <table border="1"> <thead> <tr><th>S#</th><th>P#</th></tr> </thead> <tbody> <tr><td>S1</td><td>P1</td></tr> <tr><td>S1</td><td>P2</td></tr> <tr><td>S1</td><td>P3</td></tr> <tr><td>S1</td><td>P4</td></tr> <tr><td>S1</td><td>P5</td></tr> <tr><td>S1</td><td>P6</td></tr> <tr><td>..</td><td>..</td></tr> </tbody> </table>	S#	P#	S1	P1	S1	P2	S1	P3	S1	P4	S1	P5	S1	P6	..	..	<table border="1"> <tbody> <tr><td>..</td><td>..</td></tr> <tr><td>S2</td><td>P1</td></tr> <tr><td>S2</td><td>P2</td></tr> <tr><td>S3</td><td>P2</td></tr> <tr><td>S4</td><td>P2</td></tr> <tr><td>S4</td><td>P4</td></tr> <tr><td>S4</td><td>P5</td></tr> </tbody> </table>	..	..	S2	P1	S2	P2	S3	P2	S4	P2	S4	P4	S4	P5
S#																																						
S1																																						
S2																																						
S3																																						
S4																																						
S5																																						
S#	P#																																					
S1	P1																																					
S1	P2																																					
S1	P3																																					
S1	P4																																					
S1	P5																																					
S1	P6																																					
..	..																																					
..	..																																					
S2	P1																																					
S2	P2																																					
S3	P2																																					
S4	P2																																					
S4	P4																																					
S4	P5																																					
<p>DOR</p> <table border="1"> <thead> <tr><th>P#</th></tr> </thead> <tbody> <tr><td>P1</td></tr> </tbody> </table>	P#	P1	<p>DOR</p> <table border="1"> <thead> <tr><th>P#</th></tr> </thead> <tbody> <tr><td>P2</td></tr> <tr><td>P4</td></tr> </tbody> </table>	P#	P2	P4	<p>DOR</p> <table border="1"> <thead> <tr><th>P#</th></tr> </thead> <tbody> <tr><td>P1</td></tr> <tr><td>P2</td></tr> <tr><td>P3</td></tr> <tr><td>P4</td></tr> <tr><td>P5</td></tr> <tr><td>P6</td></tr> </tbody> </table>	P#	P1	P2	P3	P4	P5	P6																								
P#																																						
P1																																						
P#																																						
P2																																						
P4																																						
P#																																						
P1																																						
P2																																						
P3																																						
P4																																						
P5																																						
P6																																						
<p>DEND DIVIDEBY DOR PER MED</p>																																						
<table border="1"> <thead> <tr><th>S#</th></tr> </thead> <tbody> <tr><td>S1</td></tr> <tr><td>S2</td></tr> </tbody> </table>	S#	S1	S2	<table border="1"> <thead> <tr><th>S#</th></tr> </thead> <tbody> <tr><td>S1</td></tr> <tr><td>S4</td></tr> </tbody> </table>	S#	S1	S4	<table border="1"> <thead> <tr><th>S#</th></tr> </thead> <tbody> <tr><td>S1</td></tr> </tbody> </table>	S#	S1																												
S#																																						
S1																																						
S2																																						
S#																																						
S1																																						
S4																																						
S#																																						
S1																																						

# Нормализация. 1НФ

FIRST	S#	STATUS	CITY	P#	QTY
	S1	20	London	P1	300
	S1	20	London	P2	200
	S1	20	London	P3	400
	S1	20	London	P4	200
	S1	20	London	P5	100
	S1	20	London	P6	100
	S2	10	Paris	P1	300
	S2	10	Paris	P2	400
	S3	10	Paris	P2	200
	S4	20	London	P2	200
	S4	20	London	P4	300
	S4	20	London	P5	400

# Нормализация. 2НФ

