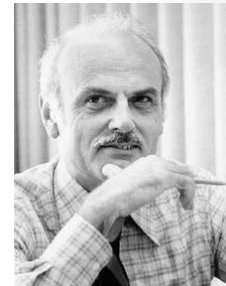


Реляционная алгебра

Технологии баз данных. Лекция 7

Реляционная алгебра

- **Реляционная алгебра** – формальная система манипулирования отношениями в реляционной модели данных.
- Существует в двух несколько различающихся вариантах: классическая алгебра Т. Кодда и алгебра А К. Дейта и Х. Дарвена.



Эдгар Кодд
1923-2003



Крис Дейт
(р. 1941)



Хью Дарвен
(р. 19??)

Замкнутость реляционной алгебры

- Реляционная алгебра представляет собой набор таких операций над отношениями, что результат каждой из операций также является отношением. Это свойство алгебры называется **замкнутостью**.

- Реляционный оператор f выглядит как функция с реляционными отношениями в качестве аргументов:

$$R = f(R_1, R_2, \dots, R_n).$$

- В реляционных выражениях *можно использовать вложенные выражения* сколь угодно сложной структуры.

Традиционные операции над множествами

Операция	Обозначение матем.	Обозначение лат. алф.
Объединение	$R1 \cup R2$	R1 UNION R2
Пересечение	$R1 \cap R2$	R1 INTERSECT R2
Разность	$R1 - R2$	R1 MINUS R2
Декартово произведение	$R1 \times R2$	R1 TIMES R2

Совместимость отношений по типу

- Два отношения являются **совместимыми по типу**, если они имеют идентичные заголовки:
 - множества имен атрибутов этих отношений совпадают
 - атрибуты с одинаковыми именами определены на одном и том же домене.
- Для приведения отношений к одному типу следует использовать операцию переименования

$\rho_{\text{НовоеОтношение}(\text{НовАтриб1}, \dots, \text{НовАтрибN})}(\text{СтароеОтношение})$

1) $\rho_{\text{Поставщики}(\text{КодП}, \text{Имя}, \text{Город}, \text{Рейтинг})}(S)$

2) $\rho_{\text{Детали}}(P)$

Объединение

- Результатом **операции объединения** двух совместимых по типу отношений $R1$ и $R2$ является отношение с тем же заголовком, что и в $R1$ и $R2$, и телом, состоящим из кортежей, принадлежащих $R1$ или $R2$ или обоим отношениям.
- В SQL это операция **union**:

`R1 union R2`

R1

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40
P2	Гайка	Челябинск	20	24

R2

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40
P3	Шуруп	Одесса	14	33

$R1 \cup R2$

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40
P2	Гайка	Челябинск	20	24
P3	Шуруп	Одесса	14	33

Пересечение

- Результатом **операции пересечения** двух совместимых по типу отношений $R1$ и $R2$ является отношение с тем же заголовком, что и в $R1$ и $R2$, и телом, состоящим из кортежей, принадлежащих обоим отношениям $R1$ и $R2$.

- В SQL это операция **intersect**:

`R1 intersect R2`

R1

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40
P2	Гайка	Челябинск	20	24

R2

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40
P3	Шуруп	Одесса	14	33

$R1 \cap R2$

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40

Вычитание

- Результатом **операции вычитания** двух совместимых по типу отношений $R1$ и $R2$ является отношение с тем же заголовком, что и в $R1$ и $R2$, и телом, состоящим из кортежей, принадлежащих отношению $R1$ и не принадлежащим отношению $R2$.
- В SQL это операция **except**:
 $R1 \text{ except } R2$

R1

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40
P2	Гайка	Челябинск	20	24

R2

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40
P3	Шуруп	Одесса	14	33

R1 - R2

PID	Name	City	Weight	Price
P2	Гайка	Челябинск	20	24

R2 - R1

PID	Name	City	Weight	Price
P3	Шуруп	Одесса	14	33

Декартово произведение

- Результатом *операции декартова произведения* двух отношений $R1$ и $R2$, не имеющих общих имен атрибутов, является отношение с заголовком, который представляет собой сцепление заголовков $R1$ и $R2$, и телом, состоящим из всех возможных кортежей, каждый из которых представляет собой сцепление кортежа из $R1$ и кортежа из $R2$.
- В SQL это операция, которая производится над таблицами из секции **from** оператора **select**:

```
select *  
from R1, R2;
```

R1		R2		
A	B	C	D	E
a1	b1	c1	d1	e1
a2	b2	c2	d2	e2
		c3	d3	e3

R1 × R2				
A	B	C	D	E
a1	b1	c1	d1	e1
a1	b1	c2	d2	e2
a1	b1	c3	d3	e3
a2	b2	c1	d1	e1
a2	b2	c2	d2	e2
a2	b2	c3	d3	e3

Специальные реляционные операции

Операция	Обозначение греч. алф.	Обозначение лат. алф.
Выборка (ограничение)	$\sigma_{\text{condition}}(R)$	R WHERE condition
Проекция	$\pi_{\text{Attr1}, \dots, \text{AttrN}}(R)$	R[Attr1, ..., AttrN]
Естественное соединение	$R1 \bowtie R2$	R1 JOIN R2
Θ-соединение	$R1 \bowtie_{R1.\text{Attr1} \Theta R2.\text{Attr2}} R2$	-
Деление	$R1 \div R2$	R1 DIVIDE R2
Группировка	$\gamma_{\text{Attr1}, \dots, \text{AttrN}}(R)$	
Сортировка	$\tau_{\text{Attr1}, \dots, \text{AttrN}}(R)$	
Удаление дубликатов	$\delta_{\text{Attr1}, \dots, \text{AttrN}}(R)$	

Выборка

- Результатом **Θ -выборки** из отношения R с помощью операции сравнения Θ над атрибутами (или литералами) A1 и A2 является отношение $\sigma_{A1\Theta A2}(R)$, имеющее тот же заголовок, что и R, и тело, состоящее из тех кортежей R, для которых вычисление выражения A1 Θ A2 дает истину.

R

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40
P2	Гайка	Челябинск	20	24
P3	Шуруп	Одесса	14	33

$\sigma_{Price>30}(R)$

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40
P3	Шуруп	Одесса	14	33

Выборка по составному условию

- $\sigma_{C_1 \text{ and } C_2}(R) \equiv \sigma_{C_1}(R) \cap \sigma_{C_2}(R)$
- $\sigma_{C_1 \text{ or } C_2}(R) \equiv \sigma_{C_1}(R) \cup \sigma_{C_2}(R)$
- $\sigma_{\text{not } C}(R) \equiv R - \sigma_C(R)$

Проекция

- Результатом **проекции** отношения R по атрибутам A1, ..., AN является отношение $\pi_{A1, \dots, AN}(R)$, имеющее заголовок, состоящий из атрибутов A1, ..., AN, и тело, которое состоит из всех кортежей R, получаемых отбрасыванием атрибутов, не входящих в список A1, ..., AN.

R

PID	Name	City	Weight	Price
P1	Болт	Париж	15	40
P2	Гайка	Челябинск	20	24
P3	Болт	Одесса	15	33

$\pi_{\text{Name, Weight}}(R)$

Name	Weight
Гайка	20
Болт	15

Естественное соединение

- Результатом *естественного соединения* отношений $R1(A,B)$ и $R2(B,C)$ по общему атрибуту B является отношение $R1 \bowtie R2$ с заголовком из атрибутов A, B, C , и телом, состоящим из соединенных кортежей, которые имеют совпадающие значения в общем атрибуте.
- $R1 \bowtie R2 \equiv \pi_{A,B,C} (\sigma_{R1.B=R2.B} (R1 \times R2))$

R1		R2	
A	B	B	C
a1	b1	b1	c1
a2	b2	b2	c2
		b3	c3
		b1	c4

R1 \bowtie R2		
A	B	C
a1	b1	c1
a1	b1	c4
a2	b2	c2

Θ -соединение

- Результатом операции **Θ -соединения** отношения R1 по атрибуту A и отношения R2 по атрибуту B является отношение

$$R1 \bowtie_{A \Theta B} R2 \equiv \sigma_{A \Theta B}(R1 \times R2)$$

R1

A	B
5	2
3	4

R2

B	C	D
2	3	1
4	6	5
2	1	7

R1 $\bowtie_{A > C}$ R2

A	R1.B	R2.B	C	D
5	2	2	3	1
5	2	2	1	7
3	4	2	1	7

Деление

- Пусть имеются отношения $A(X, Y)$ и $B(Y)$, где атрибуты Y определены на одном и том же домене.

Тогда результатом **деления** $A \div B$ будет отношение с заголовком из атрибута X и телом, в которое входят кортежи $\langle x:X \rangle$ такие, что существует кортеж $\langle x:X, y:Y \rangle$, который принадлежит отношению A для всех кортежей $\langle y:Y \rangle$ из отношения B .

Деление

SP		÷	P	=	SID
SID	PID		PID		SID
S1	P1		P1		S1
S1	P2				S2
S1	P3				
S1	P4				
S1	P5				
S2	P1				
S2	P2				
S3	P2				
S4	P2				
S4	P4				
S5	P5				

SP		÷	P	=	SID
SID	PID		PID		SID
S1	P1		P2		S1
S1	P2		P4		S4
S1	P3				
S1	P4				
S1	P5				
S2	P1				
S2	P2				
S3	P2				
S4	P2				
S4	P4				
S5	P5				

Деление

SP			P	=	SID
SID	PID	÷	PID	=	SID
S1	P1		P1		S1
S1	P2		P2		
S1	P3		P3		
S1	P4		P4		
S1	P5		P5		
S2	P1				
S2	P2				
S3	P2				
S4	P2				
S4	P4				
S5	P5				

SP			S	=	PID
SID	PID	÷	SID	=	PID
S1	P1		S1		P1
S1	P2				P2
S1	P3				P3
S1	P4				P4
S1	P5				P5
S2	P1				
S2	P2				
S3	P2				
S4	P2				
S4	P4				
S5	P5				

Примитивные и выражаемые операции реляционной алгебры

- **Примитивные** операции не могут быть выражены через другие: *выборка, проекция, произведение, объединение, вычитание, переименование.*

- **Выражаемые** операции

- пересечение:

$$R1 \cap R2 \equiv R1 - (R1 - R2)$$

- Θ -соединение:

$$R1 \bowtie R2 \equiv \sigma_C(R1 \times R2)$$

- естественное соединение:

$$R1 \bowtie R2 \equiv \pi_L(\sigma_C(R1 \times R2)),$$

где L – список атрибутов из R1 и атрибутов из R2, отсутствующих в R1; C – условие вида $R1.Attr1=R2.Attr1$ and ... and $R1.AttrN=R2.AttrN$ (Attr1, ..., AttrN – атрибуты соединения)

- деление:

$$R_n \div S_k \equiv \pi_{1,\dots,n-k}(R) - \pi_{1,\dots,n-k}((\pi_{1,\dots,n-k}(R) \times S) - R)$$

Примеры

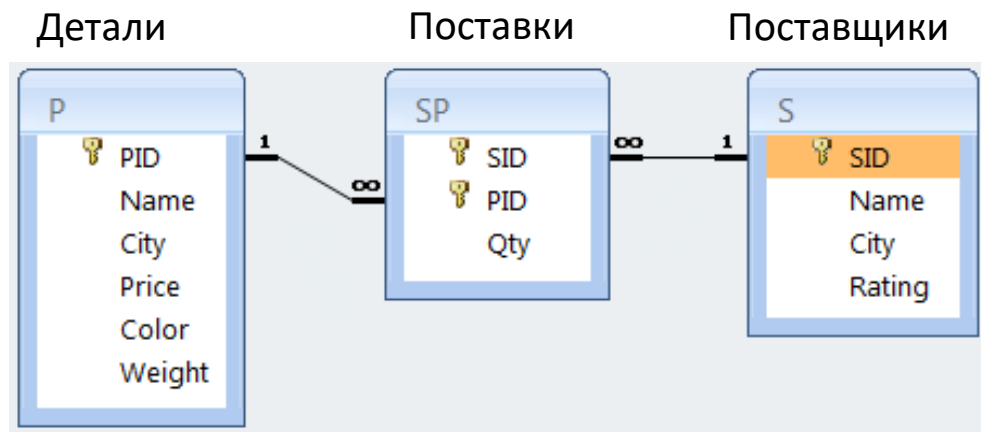
- Получить имена поставщиков, которые поставляют деталь с кодом P2.
- SQL:

```
select Name
from SP, S
where SP.SID=S.SID and
      PID='P2';
```

- Операции реляционной алгебры:

- 1) $R1 := SP \bowtie S$
- 2) $R2 := \sigma_{PID='P2'}(R1)$
- 3) $Result := \pi_{Name}(R2)$

$$\pi_{Name}(\sigma_{PID='P2'}(SP \bowtie S))$$



Примеры

- Получить имена поставщиков, которые поставляют детали красного цвета.
 - $R1 := \sigma_{\text{Color}='красный'}(P)$
 - $R2 := R1 \bowtie SP$
 - $R3 := \pi_{SID}(R2)$
 - $R4 := R3 \bowtie S$
 - $\text{Result} := \pi_{\text{Name}}(R4)$
 - $\pi_{\text{Имя_п}}(\pi_{\text{Код_п}}(\sigma_{\text{Цвет}='красный'}(P) \bowtie SP) \bowtie S)$