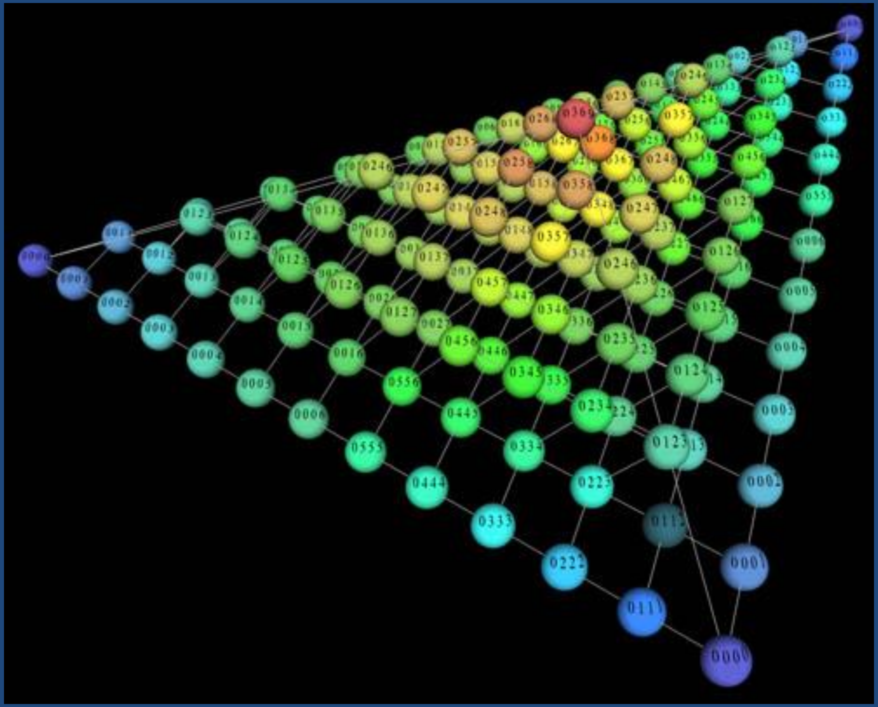


РЕЛЯЦИОННАЯ АЛГЕБРА

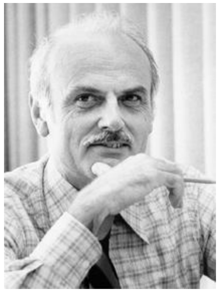
Наилучший результат дает
красивое алгебраическое тождество.
Г. Александров

Содержание

- Определение и назначение реляционной алгебры
- Традиционные операции реляционной алгебры
- Специальные операции реляционной алгебры
- Дополняющие элементы реляционной алгебры

Реляционная алгебра

- *Реляционная алгебра* – формальная система манипулирования отношениями в реляционной модели данных.
- Существует в двух несколько различающихся вариантах: классическая алгебра Т. Кодда и алгебра A 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 операторов, использующих отношения в качестве аргументов и возвращающих реляционное отношение в качестве результата.
- Реляционный оператор $\boldsymbol{f}$ выглядит как функция с реляционными отношениями в качестве аргументов: $R = f(R_1, R_2, \ldots, R_n)$.
- *Реляционная алгебра является замкнутой*, так как в качестве аргументов в реляционные операторы можно подставлять другие реляционные операторы, имеющие подходящий тип: $R = \boldsymbol{f}(f_1(R_{11}, R_{12}, \ldots), f_2(R_{21}, R_{22}, \ldots), \ldots)$
- 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 $\rho_{НовоеОтношение(НовАтр1,\ldots,\ НовАтрN)}(СтароеОтношение)$
  - $\rho_{Поставщики(КодП,\ Имя,\ Город,\ Рейтинг)}(S)$
  - $\rho_{Детали}(P)$

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.

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*.

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
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*.

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*.

R1

A	B
a1	b1
a2	b2

R2

C	D	E
c1	d1	e1
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)	$\sigma_{condition}(R)$	R **WHERE** condition
Проекция	$\pi_{Attr1,\ Attr2}(R)$	R**[**Attr1, Attr2**]**
Естественное соединение	R1 ⋈ R2	R1 **JOIN** R2
Θ-соединение	$R1 \bowtie_{\mathbf{R1.Attr1\ \Theta\ R2.Attr2}} R2$	-
Деление	R1 ÷ R2	R1 **DIVIDE** R2

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

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

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

Выборка по составному условию

- $\sigma_{C1 \text{ and } C2}(R) \equiv \sigma_{C1}(R) \cap \sigma_{C2}(R)$
- $\sigma_{C1 \text{ or } C2}(R) \equiv \sigma_{C1}(R) \cup \sigma_{C2}(R)$
- $\sigma_{\text{not } C}(R) \equiv R - \sigma_{C}(R)$

Проекция

- Результатом *проекции* отношения R по атрибутам A1, …, AN является отношение $\pi_{A1, \ldots, AN}(R)$, 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⋈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

A	B
a1	b1
a2	b2

R2

B	C
b1	c1
b2	c2
b3	c3
b1	c4

R1 ⋈ 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 $R1 \underset{A\,\Theta\,B}{\bowtie} R2 \equiv \sigma_{A\,\Theta\,B}(R1 \times R2)$.

R1

A	B
5	2
3	4

R2

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

$$\mathbf{R1} \underset{A > C}{\bowtie} \mathbf{R2}$$

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 ÷ B будет отношение с заголовком из атрибута X и телом, в которое входят кортежи <x:X> такие, что существует кортеж <x:X,y:Y>, который принадлежит отношению A для *всех* кортежей <y:Y> из отношения B.

Деление

SP ÷ **P** =

SID	PID
S1	P1
S1	P2
S1	P3
S1	P4
S1	P5
S2	P1
S2	P2
S3	P2
S4	P2
S4	P4
S5	P5

÷

PID
P1

=

SID
S1
S2

SP ÷ **P** =

SID	PID
S1	P1
S1	P2
S1	P3
S1	P4
S1	P5
S2	P1
S2	P2
S3	P2
S4	P2
S4	P4
S5	P5

÷

PID
P2
P4

=

SID
S1
S4

Деление

SP

SID	PID
S1	P1
S1	P2
S1	P3
S1	P4
S1	P5
S2	P1
S2	P2
S3	P2
S4	P2
S4	P4
S5	P5

÷

P

PID
P1
P2
P3
P4
P5

=

SID
S1

SP

SID	PID
S1	P1
S1	P2
S1	P3
S1	P4
S1	P5
S2	P1
S2	P2
S3	P2
S4	P2
S4	P4
S5	P5

÷

S

SID
S1

=

PID
P1
P2
P3
P4
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:
    $R1 \underset{C}{\bowtie} R2 \equiv \pi_L(\sigma_C(R1 \times R2))$,
    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:
    $R_n \div S_k \equiv \pi_{1,\ldots,n-k}(R) - \pi_{1,\ldots,n-k}((\pi_{1,\ldots,n-k}(R) \times S) - R)$ – *Докажите!*

Примеры

- Получить имена поставщиков, которые поставляют деталь с кодом P2.
  - R1 := SP $\bowtie$ S
  - R2 := $\sigma_{PID='P2'}$(R1)
  - Result := π_{Name}(R2)
  - $\pi_{Name}(\sigma_{PID='P2'}(SP \bowtie S))$

Примеры

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

Примеры

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

Примеры

- Получить имена поставщиков, которые поставляют все детали.
  - $R1 := \pi_{SID,PID}(SP)$
  - $R2 := \pi_{PID}(P)$
  - $R3 := R1 \div R2$
  - $R4 := R3 \bowtie S$
  - $Result := \pi_{Name}(R4)$
  - $\pi_{Name}((\pi_{SID,PID}(SP) \div \pi_{PID}(P)) \bowtie S)$

Примеры

- Получить имена поставщиков, которые не поставляют деталей.
  - R1 := SP ⋈ S
  - R2 := $\pi_{SID,Name}$ (R1)
  - R3 := $\pi_{SID,Name}$ (S)
  - R4 := R3 – R2
  - Result := π_{Name}(R4)
  - $\pi_{Name}(\pi_{SID,Name}(SP \bowtie S) - \pi_{SID,Name}(S))$

Примеры

- Получить имена поставщиков, которые не поставляют деталь с кодом P2.
  - $R1 := \pi_{SID}(S)$
  - $R2 := \sigma_{PID='P2'}(SP)$
  - $R3 := \pi_{SID}(R2)$
  - $R4 := R1 - R3$
  - $R5 := R4 \bowtie S$
  - $Result := \pi_{Name}(R5)$
  - $\pi_{Name}((\pi_{SID}(S) - \pi_{SID}(\sigma_{PID='P2'}(SP))) \bowtie S))$

Примеры

- Получить имена поставщиков, которые поставляют все детали, поставляемые поставщиком с кодом S1.
  - R1 := $\pi_{SID,PID}$ (SP)
  - R2 := $\sigma_{SID='S1'}$(SP)
  - R3 := π_{PID}(R2)
  - R4 := R1 ÷ R3
  - R5 := R4 ⋈ S
  - Result := π_{Name}(R5)
  - $\pi_{Name}((\pi_{SID,PID}(SP) \div \pi_{PID}(\sigma_{SID='S1'}(SP))) \bowtie S)$

Ассоциативность и коммутативность

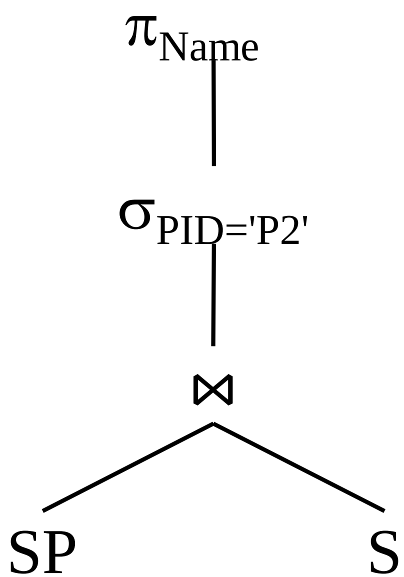
- Ассоциативные и коммутативные операции
  - $\cup$
  - $\cap$
  - $\times$
  - $\bowtie$

Назначение реляционной алгебры

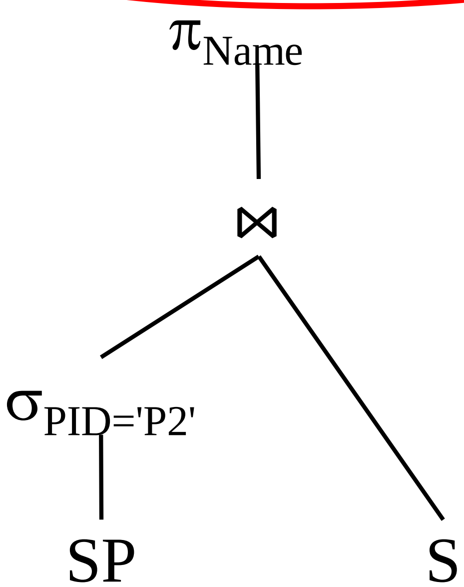
- Выборка данных из отношений
- Запись реляционных выражений
  - Определение именованных отношений
  - Определение правил целостности данных
  - Базис для оптимизации запросов

Реляционная алгебра как базис оптимизации запросов

- Получить имена поставщиков, которые поставляют деталь с кодом P2.
- $\pi_{Name}(\sigma_{PID='P2'}(SP \bowtie S))$



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



Реляционная алгебра как средство записи правил целостности

- Ссылочная целостность
  - $\pi_{SID}(SP) \subseteq \pi_{SID}(S)$
  - $\pi_{PID}(SP) \subseteq \pi_{PID}(P)$
- Целостность первичного ключа
  - $\rho_{SP1}(SP)$, $\rho_{SP2}(SP)$, $\sigma_{SP1.SID=SP2.SID \text{ and } SP1.PID=SP2.PID \text{ and } SP1.Qty \neq SP2.Qty}(SP1 \times SP2)=\varnothing$
- Ограничения домена
  - $\sigma_{Color \neq \text{'красный'} \text{ and } Color \neq \text{'синий'} \text{ and } Color \neq \text{'белый'}}(P)=\varnothing$
- Ограничения базы данных
  - $\sigma_{Rating<5}(SP \bowtie S)=\varnothing$

Дополняющие элементы реляционной алгебры

- Операторы обновления
  - targetR **:=** sourceR
  - **INSERT** sourceR **INTO** targetR,
    где sourceR и targetR совместимы по типу
  - **UPDATE** targetR attribute **:=** value
  - **DELETE** targetR
- Реляционные сравнения
  - $=$
  - $\neq$
  - $\subseteq$
- Прочее
  - **IS_EMPTY**(R)
  - t **IN** R

Дополняющие элементы реляционной алгебры

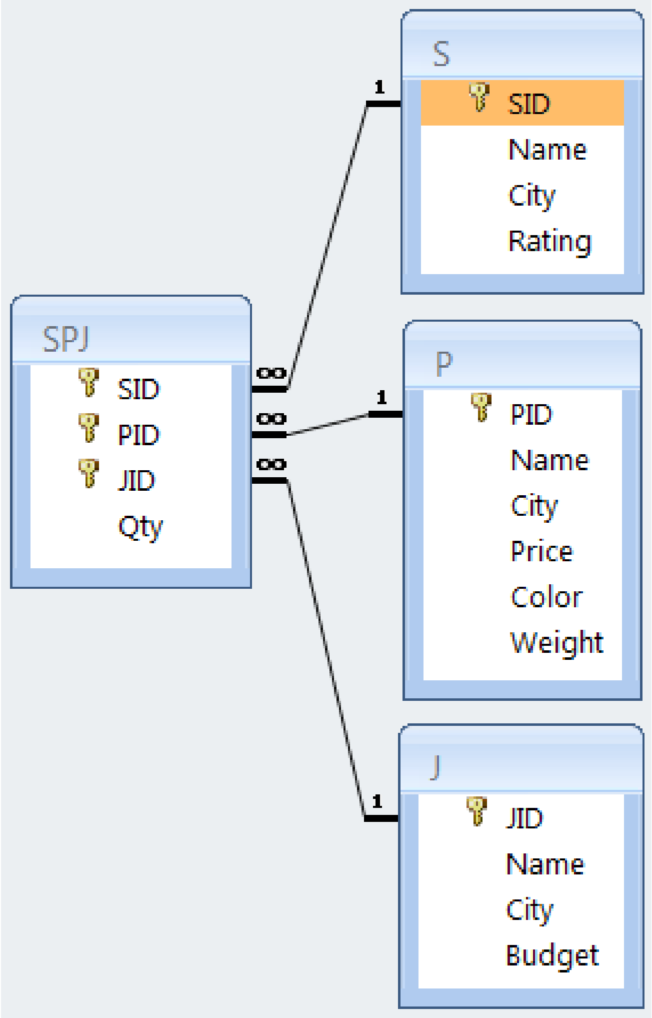
- Операция расширения
  - **EXTEND** targetR **ADD** scalarExpr **AS** newAttribute
- Операция подведения итогов
  - **SUMMARIZE** targetR **BY** (attribute)
    ADD aggrExpr **AS** newAttribute

Полнота языка баз данных

- Язык баз данных является *реляционно полным*, если выражения этого языка позволяют определить каждое отношение, которое определяется с помощью реляционной алгебры.

Практикум по составлению запросов реляционной алгебры

- S – Поставщики
- P – Детали
- J – Проекты
- SPJ – Поставки



Заключение

- Реляционная алгебра – средство записи реляционных выражений для определения данных (области выборки, обновления, представлений, снимков, правил безопасности и целостности и др.).
- Состав реляционной алгебры
  - Традиционные операции: объединение, пересечение, разность, декартово произведение.
  - Специальные операции: выборка, Θ-выборка, естественное соединение, Θ-соединение, деление.
  - Операции выборки, проекции, произведения, объединения и вычитания являются примитивными.
  - Дополняющие элементы: операции расширения и подведения итогов, операторы обновления, реляционные сравнения.
- Реляционная алгебра – базис для оптимизации запросов и средство записи ограничений целостности.
- Язык баз данных является *реляционно полным*, если выражения этого языка позволяют определить каждое отношение, которое определяется с помощью реляционной алгебры.