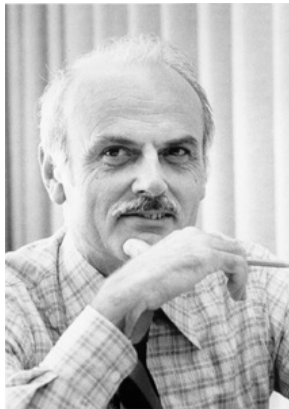




# Реляционная модель данных

Дж. Ульман,  
Основы систем баз данных, 1983.  
Глава 3

# История



## Эдгар Франк Кодд Edgar Frank Codd

1923 – 2003

- Родился в Портланде (Дорсет) в Англии.
- Обучался математике и химии в Оксфордском университете.
- Работал в Альмаденском Исследовательском Центре IBM, Сан-Хосе (Калифорния),

**A Relational Model of Data for Large Shared Data Banks, 1970,**  
*Communications of the ACM* **13** (6): 377–387

- В **1976** Кодд получил почетное звание **IBM Fellow**.
- В **1981** он получил престижную премию **Тьюринга**.
- В **2002** журнал **Forbes** поместил реляционную модель данных в список важнейших инноваций последних 85 лет.

# Аспекты модели

- ❖ Структура данных
- ❖ Манипулирование данными
  - Реляционная алгебра
  - Реляционное исчисление
    - на кортежах
    - на доменах
- ❖ Целостность данных

# Структура данных (1)

Основой структур данных в модели является теоретико-множественное понятие **ОТНОШЕНИЯ**

## Определение 1

- Пусть даны  $D_1, D_2, \dots, D_n$  – множества (домены).
- Декартово произведение  $(D_1 \times D_2 \times \dots \times D_n)$  – множество кортежей  $(v_1, v_2, \dots, v_n)$ , длины  $n$ , таких что  $v_1 \in D_1, v_2 \in D_2, \dots, v_n \in D_n$
- Отношением называется **КОНЕЧНОЕ** подмножество декартового произведения доменов  $D_1 \times D_2 \times \dots \times D_n$
- Отношение  $R (A_1:D_1, A_2:D_2, \dots, A_n:D_n)$  или  $R (A_1, A_2, \dots, A_n)$ 
  - $R$  – имя отношения
  - $A_1, A_2, \dots, A_n$  – имена атрибутов отношения
  - Элемент отношения  $r_1 r_2 \dots r_n$  – кортеж отношения

# Структура данных (2)

## Определение 2

Пусть  $A_1, A_2, \dots, A_k$  – множество имен атрибутов,  
 $D_1, D_2, \dots, D_n$  – множества значений атрибутов

Множество пар  $\langle A_i, D_j \rangle$  называется **схемой отношения**,  
если  $A_t \neq A_s$ , при  $t \neq s$ .

Множество пар  $\langle A_i, d_j \rangle$  называется **кортежем отношения**.  
где  $A_i$  – имя атрибута из  $\{A_1, A_2, \dots, A_k\}$ ,  
а  $d_j$  – значение из  $D_j$

Множество кортежей, соответствующих **одной схеме отношения**  
называется **отношением**

# Неформальные определения

- Неформально, **отношение** - это *таблица* данных, состоящая из **строк, столбцов**.
- Каждая строка содержит **набор элементов** данных, который представляет конкретные факты соответствующие **сущностям или отношениям**
  - В формальной модели *строки* называются **кортежами**
- Каждый столбец имеет **заголовок**, что позволяет судить о значении элементов данных в этом столбце
  - В формальной модели заголовок столбца, называется **именем атрибута** (или просто **атрибут**)
- Схема отношения – имя отношения с набором атрибутов

# Пример отношения

Имя отношения  
↓  
Студенты

Атрибуты

кортежи

Имя	<u>№ студ</u>	тел	адрес	возраст	балл
Иванов	090041	330-15-34	Терешковой, 40	19	3.26
Петров	090042	330-87-38	Ученых, 7	19	4.05
Сидоров	090140	230-15-32	Морской, 2	18	4.87
Голубев	090141	335-42-43	Правды, 5	20	3.87
Федоров	090142	231-48-32	Пирогова, 22	20	4.50
Костин	090086	330-45-44	Мальцева, 4	18	4.90

**Схема** - Студенты (имя, № студ, тел, адрес, возраст, балл<sub>7</sub>)

# Основные понятия в РМД

- Основные понятия в РМД:
  - ❖ отношение,
  - ❖ кортеж,
  - ❖ атрибут,
  - ❖ домен,
  - ❖ тип данных,
  - ❖ ключ
- **Тип данных БД полностью адекватен типу данных в языках программирования**  
Обычно допускается хранение
  - ❖ символьных данных (строки),
  - ❖ числовых данных (целые и вещественные),
  - ❖ битовых строк,
  - ❖ даты, времени
  - ❖ специальные числовые данные (деньги)



# Домен

- Домен – понятие специфичное для баз данных
- Домен задается **базовым типом** и логическим выражением, определяющим элементы домена

## Пример.

- Домен ИМЕНА может базироваться на строковом типе, но содержать только те строки, которые НЕ начинаются на Ъ или Ь знак

# Ключ

- Атрибут или множество атрибутов, *однозначно определяющих кортеж в отношении*, называется **ключом отношения**.
  - В таблице Студент ключ - № студ. Билета
- Иногда *идентификатор строки* или *порядковое число* используются как **ключи** для идентификации строки в таблице
  - Такой ключ называют **искусственным или суррогатным** ключ

# Типы



# Фундаментальные свойства отношений

- Отсутствие кортежей дубликатов
- Отсутствие упорядоченности кортежей
- Отсутствие упорядоченности атрибутов
- Атомарность значений атрибутов

# Атомарность атрибутов

Номер отдела	Отдел		
	Сотр_номер	Сотр_имя	Сотр_зарп
310	2934	Иванов	112,00
	2935	Петров	112,50
313	2937	Федоров	110
315	2938	Голубев	112

Сотр_номер	Сотр_имя	Сотр_зарп	Сотр_отд_номер
2934	Иванов	112,00	310
2935	Петров	112,50	310
2937	Федоров	110	313
2938	Голубев	112	315

# Реляционная алгебра

## ■ Теоретико-множественные операции

Условия выполнения операций

- **Объединение, пересечение, разность –**  
**схемы отношений – одинаковы (кол. атрибутов,**  
**домены соответствующих атрибутов )**

1.  $R_1(A:\text{цел}, V:\text{строk}(10), C:\text{вещ})$        $S_1(A:\text{цел}, V:\text{строk}(10))$   
**6 , Петров , 12.83**      **7 , Голубев**
2.  $R_2(A:\text{вещ}, C:\text{вещ})$        $S_2(A:\text{цел}, C:\text{строk}(10))$   
**2.87 , 3.12**      **10 , Сидоров**
3.  $R_3(A:\text{цел}, C:\text{строk}(10))$        $S_3(V:\text{цел}, E:\text{строk}(10))$   
**1010 , директор**      **2011, инженер**

- **Декартово произведение – нет ограничений**

- $R(A,B,C) \times S(E,D) = T(A,B,C,E,D)$
- $R(A,B,C) \times S(A,D) = T(A_1,B,C,A_2,D)$

# Реляционная алгебра (продолжение)

## ■ Примеры

$R (A:\text{цел}, V:\text{строка}(10)) \cup S (A:\text{цел}, V:\text{строка}(10)) =$

6	, Петров	7	, Голубев
7	, Голубев	9	, Соколов

$T (A:\text{цел}, V:\text{строка}(10))$

6	, Петров
7	, Голубев
9	, Соколов

$R (A,B,C) \times S(A,D) = T(A_1,B,C,A_2,D)$

1	3	5	2	4	1	3	5	2	4
7	9	3	6	8	7	9	3	2	4
			3	7	1	3	5	6	8
					7	9	3	6	8
					1	3	5	3	7
					7	9	3	3	7

# Реляционная алгебра (продолжение)

- Специальные операции
  - Селекция
  - Проекция
  - Соединения (JOIN)
    - Эквисоединение
    - Натуральное соединение (Natural JOIN)
    - Left Inner JOIN
    - Right Inner JOIN
    - OUTER JOIN
  - Деление



# Селекция

Операция  $\sigma_f(R(A_1, A_2, \dots, A_k))$  порождает  
отношение  $R_1(A_1, A_2, \dots, A_k)$  кортежи которого

- 1) принадлежат исходному отношению  $R(A_1, A_2, \dots, A_k)$
- 2) удовлетворяют логической функции  $f$ , которая строится из

а) операндов

- атрибутов отношения  $R$
- констант

б) операторов сравнения:  $>, <, =, \neq, \leq, \geq$ ,

в) логических операторов: *and*, *or*, *not*

г) скобок:  $(, )$

SELECT  
FROM  $R$   
WHERE  $f$

$R(A, B, C, D)$

1	3	5	7
2	8	5	9
3	7	6	9

$\sigma_{A < 3 \text{ and } C = 5}(R)$

$R_1(A, B, C, D)$

1	3	5	7
2	8	5	9

# Проекция

Пусть  $R(A_1, A_2, \dots, A_k)$  – отношение.

Операция  $\pi_{A_{i1}, A_{i2}, \dots, A_{is}}(R)$  порождает отношение  $R_1$  с

атрибутами  $A_{i1}, A_{i2}, \dots, A_{is}$  из отношения  $R$ , где  $i_s \leq k$ , следующим образом:

1. Каждый кортеж в  $R_1$  формируется из кортежа отношения  $R$  путем удаления не отмеченных в операции  $\pi_{A_{i1}, A_{i2}, \dots, A_{is}}(R)$  значений атрибутов

2. В получившемся отношении одинаковые кортежи удаляются.

$R(A, B, C, D)$

1 3 5 7

2 8 5 9

3 7 6 9

1 8 5 4

$\pi_{A,C}(R)$

1)  $R_1(A, C)$

1 5

2 5

3 6

1 5

2)  $R_1'(A, C)$

1 5

2 5

3 6

# Соединения (JOIN)

- **Эквисоединение.**

$$R \bowtie_f S = \sigma_f (R \times S)$$

- **Натуральное соединение** (Natural JOIN, Inner JOIN). Для R и S имеющих одинаковые атрибуты

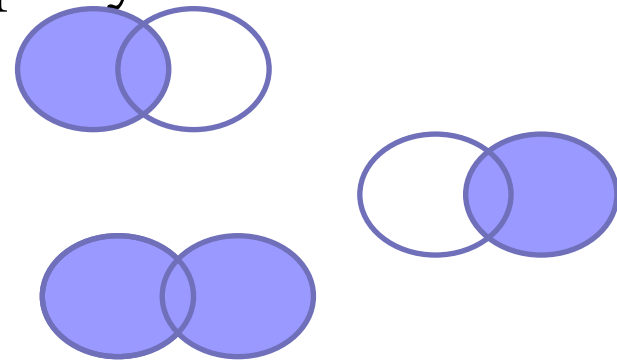
$$R \bowtie S = \pi_{\text{атр } R \cup \text{атр } S} (\sigma_f (R \times S)), \text{ где } f \text{ функция вида}$$
$$A_{r1} = A_{s1} \& A_{r2} = A_{s2} \& \dots \& A_{rl} = A_{sl},$$

а  $A_{ri} = A_{si}$  — одинаковые атрибуты из R и S

- **Left OUTER JOIN**

- **Right OUTER JOIN**


- **Full OUTER JOIN**




# Примеры соединений

1) R  S = T(A, B, C, B<sub>1</sub>, D)


R(A, B, C)	S(B, D)				
1 3 4	3 1	1	3	4	3 1
2 3 5	4 7				1 3 4 4 7
4 2 2					2 3 5 3 1

2) R  S = T(A, B, C, D)

4	2 <sup>1</sup>	3	4	1	1
2	3	5	1	3	
4	2	2	4	7	

3) R  S = T(A, B, C, B<sub>1</sub>, D)

1	3	4	3	1	
2	3	5	3	1	
		4	2	2	Null Null

4) R  S = T(A, B, C, B<sub>1</sub>, D)

1	3	4	3	1	
2	3	5	3	1	
		4	2	2	Null Null
		Null Null	Null	4	7

# Деление

Пусть  $R(A:\text{цел}, B:\text{цел}, C:\text{цел})$ ,  $S(A:\text{цел}, B:\text{цел})$

$$R(A, B, C) \div S(A, B) = T(C)$$

Отношение  $T$  есть множество кортежей  $t$  длины  $|R|-|S|$ , таких что

для всех кортежей  $u$  длины  $|S|$ , принадлежащих  $S$  кортеж  $tu$  принадлежит  $R$ .

$R(A, B, C)$

1	2	3
2	4	3
1	2	1
2	4	3

$S(A, B)$

1	2
2	4

$T(C)$

3
---

# Целостность

- Целостность сущностей
  - ❖ Ключи
- Ссылочная целостность
  - ❖ Только сообщение о нарушении целостности
  - ❖ Метод значений NULL
  - ❖ Каскадный метод
  - ❖ Метод значений по умолчанию

Пример (возможно ли такое состояние таблиц в БД?)

Сотрудник (ТН, ФИО, НомПасп, НомОтдела)

Отдел (НомОтдела, НазОтдела)

Сотрудник.НомОтдела → Отдел.НомОтдела

Сотрудник

Отдел

ТН	ФИО	НомПасп	НомОтдела	НомОтдела	НазвОтдела
1002	Иванов	5001593549	18	01	Отдел 1

# Целостность

## Пример

Сотрудник(ТН, ФИО, НомПасп, НомОтдела)

Отдел(НомОтдела, НазОтдела)

Сотрудник

ТАБНомер	ФИО	НомПасп	НомОтдела
1002	Иванов	5001593549	01
1020	Голубев	5007123321	03

Отдел

НомОтдела	НазвОтдела
01	Отдел 1
03	Отдел ИС

## Что будет ?

1. Вставить в Сотрудник строку (1005, 'Петров', 5003374727, 03)
2. Вставить в Сотрудник строку (1010, 'Сидоров', 5005374727, 05)
3. Изменить в Сотрудник строку где ТН=1005 НомОтдела на 01
4. Вставить в Отдел строку (02, 'Второй отдел')
5. И ...

## Достоинства модели

- Небольшой набор абстракций  
(отношение, кортеж, атрибут, домен, тип данных, ключ)
- Наличие простого и мощного математического аппарата  
(теория множеств, математическая логика)
- Возможность ненавигационного манипулирования данными (реляционное исчисление)

## Недостатки модели

- Небольшой набор типов данных
- Небольшой набор средств манипулирования
- Отсутствие возможности отразить семантику объектов