

Базы данных

Шехтман В. Е.

Лекция 2. Реляционная модель
данных

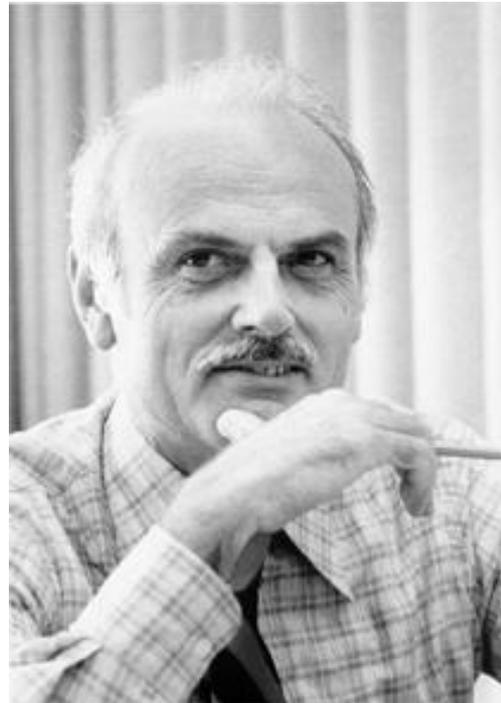
Понятие модели данных

- Структурные основы модели (каковы структуры хранения данных и как отражаются взаимосвязи между ними)
- Операции над данными (как задавать вопросы, связанные с предметной областью, моделируемой в базе данных)
- Ограничения на данные (как их выражать)

Реляционная модель данных

- Используется большинством промышленных СУБД
- Очень простая
- Запросы на высокоуровневом языке – простые но выразительные
- Эффективная реализация

Эдгар Ф. Кодд, 1970
IBM, System R, 1975



БД = множество (set) именованных **отношений** (таблиц)

Отношение имеет именованные **атрибуты** (колонки)

Кортеж (tuple) (**запись**) содержит значения атрибутов

Атрибут имеет один опр. **тип** (принадлежит **домену**)

Схема – описание структуры отношения (ий) в БД

Экземпляр – актуальное содержание отношения в опр. момент времени

NULL – специальное значение: “unknown”, “undefined”

Ключ – атрибут (или множество атрибутов), чье значение уникально для всех кортежей

Student (ID, Name, EGE, Photo)

Id	Name	EGE	Photo

Institute (Name, City, Limit)

Name	City	Limit

Создание отношений-таблиц на SQL

```
Create Table Institute  
  (name varchar(50),  
   city char(10),  
   limit integer)
```

```
Create Table Student(  
ID INT NOT NULL PRIMARY KEY,  
name  VARCHAR(50),  
EGE REAL, photo BLOB)
```

**Запрос к базе данных, состоящей из отношений,
создает отношение**

Пример: база данных об абитуриентах

Institute(cName, city, limit)

Student(sID, sName, EGE,)

Apply(sID, cName, fclt, decision)

Institute

cName	city	limit

Student

sID	sName	EGE	ol

Apply

sID	cName	fclt	decision

Простейший запрос: имя отношения

Операторы используются для фильтрации и комбинирования

Institute

cName	city	quantity

Student

sID	sName	EGE	ol

Apply

sID	cName	fct	desc

Select: извлекает требуемые строки: $Sel_{cond} Rel$

Студенты с $EGE > 70$

$Sel_{EGE > 70} Student$

Студенты с $EGE > 70$ и $ol < 1$

$Sel_{EGE > 70 \text{ and } ol < 1} Student$

Поступающие в НФИ КемГУ на ФИТ

$Sel_{cname='НФИ КемГУ' \text{ and } fclt='ФИТ'} Apply$

Institute

cName	city	quant

Student

sID	sName	EGE	ol

Apply

sID	cName	fclt	de

Проекция: извлекает требуемые колонки

$\text{Proj}_{A_1, A_2, \dots} \text{Rel}$

Выбрать sID и решения по всем абитуриентам

$\text{Proj}_{\text{sID}, \text{dec}} \text{Apply}$

Institute

cName	city	quantity

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	dec

Фильтровать и строки и колонки:

ID и имя студентов с EGE>70

Proj _{sid, sname} (**Sel** _{EGE>70} **Student**)

Institute

cName	city	quantity

Student

sID	sName	EGE	OI

Apply

sID	cName	fclt	dec

Дубликаты

Перечислить факультеты и решения о приеме

Proj fclt, dec **Apply**

SQL – мультимножества

Алгебра - множества

Institute

cName	city	quantity

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	dec

Умножение (декартово) отношений

Student * Apply

sID	sName	EGE	ol	sID	cName	fclt	dec

Institute

Student

Apply

cName	city	quant

sID	sName	EGE	ol

sID	cName	fclt	dec

Умножение отношений (декартово)

Имена и ЕГЕ студентов, подавших документы на ФИТ, но не принятых

Proj sName, EGE (**Sel** Student.Sid = Apply.sId (Student * Apply))
and fclt='ФИТ'
and dec = 'NO'

Institute

cName	city	quantity

Student

sID	sName	EGE	ol

Apply

sID	cName	fclt	dec

Естественное соединение

- Умножение двух отношений с оставлением только тех пар кортежей в результате, в которых значения одноименных атрибутов совпадают
- Устранение дубликатных атрибутов

Institute

cName	city	quantity

Student

sID	sName	EGE	ol

Apply

sID	cName	flc	dec

Естественное соединение

Имена и ЕГЭ студентов, участвовавших более чем в одной олимпиаде, подавших документы на ФИТ, в институт с общим числом учащихся более 3000 и не принятых

Proj sName, EGE (Sel ol>1 and fclt='ФИТ' (Student join Apply join Institute))
and dec= 'NO'
and quant > 3000

Institute

cName	city	quant

Student

sID	sName	EGE	ol

Apply

sID	cName	fclt	dec

Естественное соединение

Exp1 join Exp2 \equiv

Proj $\text{Scheme}(\text{Exp1}) \cup \text{Scheme}(\text{Exp2})$ (**Sel** $\text{Exp1.A1} = \text{Exp2.A1}$ (**Exp1 * Exp2**))
and $\text{Exp2.A2} = \text{Exp2.A2}$
and ...

Institute

cName	city	quantity

Student

sID	sName	EGE	ol

Apply

sID	cName	flc	dec

Theta соединение

- Основная операция в РСУБД
- Слово “соединение” обычно означает theta - соединение

$\text{Exp1 join}_{\text{cond}} \text{Exp2} \equiv \text{Sel}_{\text{cond}} (\text{Exp1} * \text{Exp2})$

Institute

cName	city	quant

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	de

Разность

имена выпускников школ, которые никуда не подали документы

$R = \text{Proj}_{sID} \text{ Student} - \text{Proj}_{sID} \text{ Apply}$

$\text{Proj}_{sName} (R \text{ join Student})$

Institute

cName	city	quant

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	de

Переименование

$E(B1, B2, \dots)$

1. Rename_R E

2. Rename_{A1, A2, \dots} E

3. Rename_{R(A1, A2, \dots)} E

Institute

cName	city	quantity

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	dec

Объединение

Перечислить имена студентов и институтов

Proj_{cName} **Institute** **U** **Proj**_{sName} **Student**

Institute

cName	city	quantity

Student

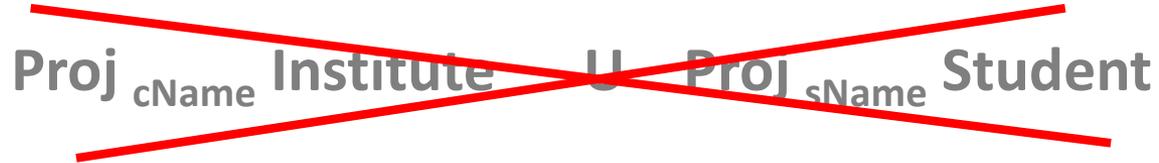
sID	sName	EG	ol

Apply

sID	cName	fclt	desc

Объединение

Перечислить имена студентов и институтов



Institute

cName	city	quantity

Student

sID	sName	EGE	ol

Apply

sID	cName	fclt	de

Переименование – для унификации схем при работе с множествами

Перечислить имена студентов и институтов

~~Proj_{cName} Institute U Proj_{sName} Student~~

Rename_{C1(NAME)} (Proj_{cName} Institute) U
Rename_{C2(NAME)} (Proj_{sName} Student)

Institute

Student

Apply

cName	city	quantity

sID	sName	EG	ol

sID	cName	flc	de

Пересечение

Имена, встречающиеся как среди абитуриентов так и среди ВУЗов

$\text{Proj}_{\text{cName}} \text{Institute} \cap \text{Proj}_{\text{sName}} \text{Student}$

Institute

cName	city	quant

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	de

Пересечение

Имена, встречающиеся как среди абитуриентов так и среди ВУЗов

~~Proj_{cName} Institute \cap Proj_{sName} Student~~

Rename_{C1(NAME)} (Proj_{cName} Institute) \cap
Rename_{C2(NAME)} (Proj_{sName} Student)

Institute

cName	city	quant

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	dec

Пересечение

имена выпускников школ, подавших заявления в какой-либо ВУЗ

$$W = \text{Proj}_{sID} \text{Apply} \cap \text{Proj}_{sID} \text{Student}$$

$$\text{Proj}_{sName} (W \text{ join } \text{Student})$$

Institute

cName	city	quant

Student

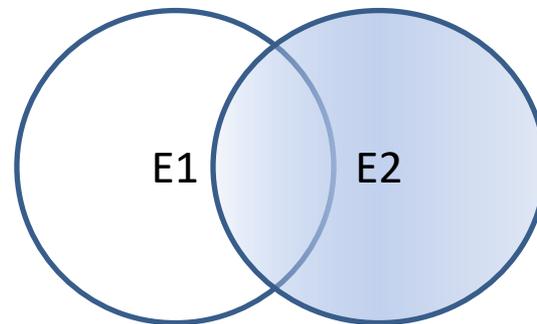
sID	sName	EG	ol

Apply

sID	cName	fclt	de

Пересечение не добавляет ничего нового (1)

$$E1 \cap E2 \equiv E1 - (E1 - E2)$$



Institute

cName	city	quantity

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	dec

Пересечение не добавляет ничего нового (2)

$E1 \cap E2 \equiv \text{Proj}_{\text{атр. в условии соединения}} (E1 \text{ join } E2)$

Institute

cName	city	quantity

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	desc

Переименование

Для самосоединений:

Пары институтов в одном и том же городе

C1 = Rename n_1, s, q_1 **Institute**

C2 = Rename n_2, s, q_2 **Institute**

C = C1 join C2

R = sel $n_1 \langle \rangle n_2$ **C**

Institute

cName	city	quantity

Student

sID	sName	EGE	ol

Apply

sID	cName	flc	de

Пары институтов в одном городе одним выражением

Sel $n1 \leftrightarrow n2$ (**Rename** $c1(n1, c, e1)$ **Institute** **join**
Rename $c2(n2, c, e2)$ **Institute**)

Institute

cName	city	quantity

Student

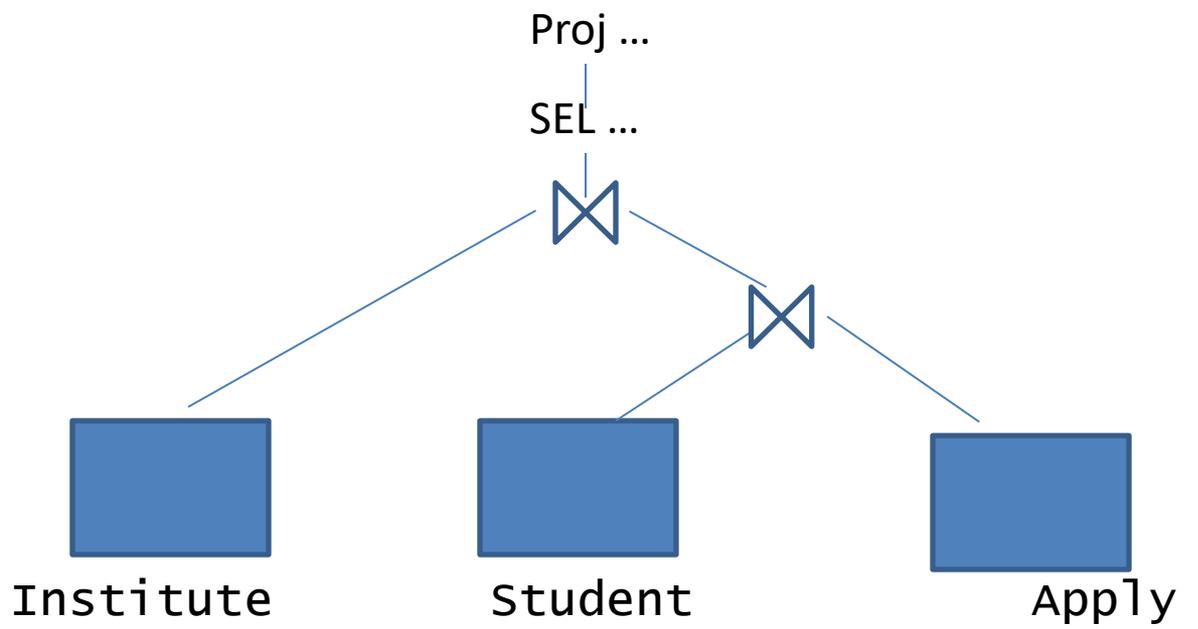
sID	sName	EG	ol

Apply

sID	cName	flc	dec

Дерево выражения – EGE абитуриентов подавших документы на ФИТ НФИ КемГУ

Proj_{EGE} (Sel_{fclt='ФИТ' and cName='НФИ КемГУ'} (Institute Join Student Join Apply))



cName	city	quantity

sID	sName	EGE	ol

sID	cName	fclt	desc

ИТОГИ

R

Sel _{cond} E

$\sigma_{\text{cond}} E$

Proj _{A1, A2, ...} E

$\pi_{A1, A2, \dots} E$

E1 * E2

E1 U E2

E1 – E2

E1 \cap E2

E1 join E2

$E1 \bowtie E2$

E1 join_{cond} E2

$E1 \bowtie_{\theta} E2$

Rename _a R

$\rho_a R$

Пусть отношение **МагО** содержит список обязательного минимума товаров для каждого магазина. Схема отношения **МагО** совпадает со схемами **Маг1** и **Маг2**.

Какие продукты есть и в первом и во втором магазинах, но нет в **МагО**?

Пусть отношение **МагО** содержит список обязательного минимума товаров для каждого магазина. Схема отношения **МагО** совпадает со схемами **Маг1** и **Маг2**.

Какие продукты есть и в первом и во втором магазинах, но нет в **МагО**?

$$(\text{Маг1} \cap \text{Маг2}) - \text{МагО}$$

Какие продукты есть в первом магазине, отсутствуют во втором магазине, но входят в обязательный список?

Какие продукты есть в первом магазине, отсутствуют во втором магазине, но входят в обязательный список?

$(\text{Mag1} - \text{Mag2}) \cap \text{Mag0}$

Какие продукты из обязательных
есть и в первом и во втором
магазинах?

Какие продукты из обязательных
есть и в первом и во втором
магазинах?

$(\text{Mag1} \cap \text{MagO}) \cap (\text{Mag2} \cap \text{MagO})$

Какие продукты из обязательных
есть или в первом или во втором
магазинах?

Какие продукты из обязательных
есть или в первом или во втором
магазинах?

$$(\text{Mag1} \cap \text{MagO}) \cup (\text{Mag2} \cap \text{MagO})$$

Каких продуктов из числа
обязательных нет ни в первом, ни
во втором магазинах?

Каких продуктов из числа
обязательных нет ни в первом, ни
во втором магазинах?

$\text{MagO} - (\text{Mag1} \cup \text{Mag2})$

МагО(Код, Название):

(Ф1, Бананы)

(Ф9, Апельсины)

(Ф11, Ананасы)

Магазины(Наим.магазина):

Привет

Ласточка

Рябинка

Перечень магазинов с

обязательными товарами:

МагО(Код, Название):

(Ф1, Бананы)

(Ф9, Апельсины)

(Ф11, Ананасы)

Магазины(Наим.магазина):

Привет

Ласточка

Рябинка

Перечень магазинов с

обязательными товарами:

МагО ⊗ Магазины

МагО ⊗ Магазины:

Шифр	Наименование	Магазин
Ф1	Бананы	Привет
Ф1	Бананы	Ласточка
Ф1	Бананы	Рябинка
Ф9	Апельсины	Привет
Ф9	Апельсины	Ласточка
Ф9	Апельсины	Рябинка
Ф11	Ананасы	Привет
Ф11	Ананасы	Ласточка
Ф11	Ананасы	Рябинка

МагР:

Шифр	Наименование	Магазин
Ф1	Бананы	Привет
Ф1	Бананы	Ласточка
Ф9	Апельсины	Ласточка
Ф9	Апельсины	Рябинка
Ф11	Ананасы	Привет

Какие продукты из обязательных отсутствуют в каких магазинах?

Какие продукты из обязательных
отсутствуют в каких магазинах?

Какие продукты из обязательных отсутствуют в каких магазинах?

МагО ⊗ Магазины - МагР

Шифр	Наименование	Магазин
Ф1	Бананы	Рябинка
Ф9	Апельсины	Привет
Ф11	Ананасы	Ласточка
Ф11	Ананасы	Рябинка

R1 (ФИО, Дисциплина, Оценка)

R2 (ФИО, Группа)

R3 (Группа, Дисциплина)

Кто сдал экзамен по базам
данных?

R1 (ФИО, Дисциплина, Оценка)

R2 (ФИО, Группа)

R3 (Группа, Дисциплина)

Кто сдал экзамен по базам
данных?

Proj_{ФИО} (sel_{Дисциплина = "базы данных"} R1)

Список ФИО, сдавших на 5 экзамен по базам данных?

Кто должен сдавать экзамен по БД?

Список ФИО, сдавших на 5 экзамен по базам данных?

Proj_фио (Sel_Оценка=5 AND Дисциплина = "базы данных" R1)

Кто должен сдавать экзамен по БД?

Список ФИО, сдавших на 5 экзамен по базам данных?

Proj ФИО (Sel Оценка=5 AND Дисциплина = "базы данных" R1)

Кто должен сдавать экзамен по БД?

Proj ФИО (R2 join R3.Группа = R2.Группа AND R3.Дисциплина = "базы данных" R3)

Кто имеет двойки более чем по одной дисциплине?

Кто имеет двойки более чем по одной дисциплине?

$R1' = \text{RENAME}(R1)$

$R1'' = \text{RENAME}(R1)$

$\text{Proj}_{\text{ФИО}}(R1' \text{ join } R1''.\text{ФИО}=R1''.\text{ФИО AND}$
 $R1'.\text{Дисциплина} \neq R1''.\text{Дисциплина}$
 $\text{AND } R1'.\text{Оценка} \leq 2$
 $\text{AND } R1''.\text{Оценка} \leq 2 \text{ } R1'')$

Кто является круглым отличником?

Кто является круглым отличником?

Список всех пар (студент, дисциплина):

$R4 = \text{Proj}_{\text{ФИО}, \text{Дисц}}$
 $(R2 \text{ join}_{R2. \text{Группа} = R3. \text{Группа}} R3)$

Кто является круглым отличником?

Список всех пар (студент, дисциплина):

$$R4 = \text{Proj}_{\text{ФИО}, \text{Дисц}} (\text{R2} \text{ join}_{\text{R2.Группа}=\text{R3.Группа}} \text{R3})$$

Список всех пар (студент, дисциплина), для которых получены отличные оценки:

$$R5 = \text{Proj}_{\text{ФИО}, \text{Дисц}} (\text{sel}_{\text{оценка} = 5} \text{R1})$$

Кто является круглым отличником?

Список всех пар (студент, дисциплина):

$$R4 = \text{Proj}_{\text{ФИО}, \text{Дисц}} (\text{R2} \text{ join}_{\text{R2.Группа}=\text{R3.Группа}} \text{R3})$$

Список всех пар (студент, дисциплина), для которых получены отличные оценки:

$$R5 = \text{Proj}_{\text{ФИО}, \text{Дисц}} (\text{sel}_{\text{Оценка} = 5} \text{R1})$$

Список студентов, не сдавших что-либо на 5:

$$R6 = \text{Proj}_{\text{ФИО}} (\text{R4} - \text{R5})$$

Кто является круглым отличником?

Список всех пар (студент, дисциплина):

$$R4 = \text{Proj}_{\text{ФИО}, \text{Дисц}} (\text{R2} \text{ join}_{\text{R2.Группа}=\text{R3.Группа}} \text{R3})$$

Список всех пар (студент, дисциплина), для которых получены отличные оценки:

$$R5 = \text{Proj}_{\text{ФИО}, \text{Дисц}} (\text{sel}_{\text{Оценка} = 5} \text{R1})$$

Список студентов, не сдавших что-либо на 5:

$$R6 = \text{Proj}_{\text{ФИО}} (\text{R4} - \text{R5})$$

Исключаем R6 из общего списка студентов:

$$(\text{Proj}_{\text{ФИО}} \text{R2}) - \text{R6}$$

Ограничения целостности

Среди студентов в Apply могут быть только такие, которые есть в Student

Institute

cName	city	quantity

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	de

Ограничения целостности

Среди студентов в Apply могут быть только такие, которые есть в Student

$$|\pi_{sID} \text{ Apply}| \leq |\pi_{sID} \text{ Student}|$$

ИЛИ

$$\pi_{sID} \text{ Apply} - \pi_{sID} \text{ Student} = \emptyset$$

Institute

cName	city	quantity

Student

sID	sName	EGE	ol

Apply

sID	cName	fclt	desc

Ограничения целостности

В Apply может быть не более одного заявления от студента на каждый из факультетов института

Institute

cName	city	quantity

Student

sID	sName	EGE	ol

Apply

sID	cName	fclt	desc

Ограничения целостности

В Apply может быть не более одного заявления от студента на каждый из факультетов института

A1 = ρ Apply

A2 = ρ Apply

A1 join $A1.cName=A2.cName$ **A2 = ∅**

and $A1.fclt=A2.fclt$

and $A1.sId=A2.sId$

and $A1.dec \neq A2.dec$

Institute

cName	city	quant

Student

sID	sName	EG	ol

Apply

sID	cName	fclt	dec