



Базы данных

Реляционная модель
данных

Основные определения

- **Домен** – *множество* возможных значений некоторой величины из предметной области.

Примеры доменов

Фамилия = {Иванов, Петров, Сидоров}

Дисциплина = {БД, СПО, ПЯВУ}

$D1 = \{2,4\}$ $D2 = \{1,3,5\}$

- **Декартово произведение множеств** – множество всевозможных пар элементов из $D1$ и $D2$

$D1 \times D2 = \{(2,1), (2,3), (2,5), (4,1), (4,3), (4,5)\}$

Основные определения

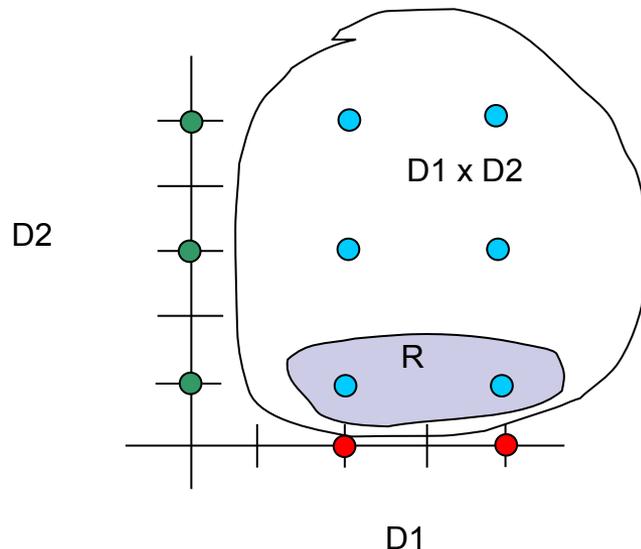
Отношение – любое подмножество из декартова произведения доменов.

Не формально: отношение (relationship) – зависимость одних данных от других

Например,

$$R = \{(2, 1), (4, 1)\}$$

$$D1 \times D2 = \{(2, 1), (2, 3), (2, 5), (4, 1), (4, 3), (4, 5)\}$$



Характеристики отношения

Отношение моделирует реальную ситуацию, т.е. для каждого элемента из R можно утверждать, что он соответствует действительности

Фамилия студента : фамилия	Учебная дисциплина : дисциплина	Экзаменационная оценка : оценка
Иванов	СПО	2
Петров	БД	4
Петров	СПО	5

- **Кортеж** = Строка = n-ка
- **Атрибут** - входение домена в отношение
- **Степень отношения** – количество атрибутов
- **Кардинальность** – количество кортежей
- **Схема отношения** – перечень имен атрибутов с указанием соответствующих доменов

Свойства отношения

- В отношении нет двух одинаковых кортежей
- Порядок следования кортежей – произвольный
- Атрибуты имеют уникальные имена

Пример атрибутов из одного домена

R = <Фамилия студента : Фамилия, Год рождения : Год, Год поступления : Год>

Свойства отношения

- Если атрибуты из одного домена, то они называются **θ -сравнимыми**, где θ - множество операций сравнения для заданного домена. Например, место рождения и место жительства – сравнимы, место рождения и год рождения не сравнимы (разные домены).
 - Для домена «Год» $\theta = \{=, <>, >, <, >=, <= \}$
 - Для домена «Место» $\theta = \{=, <> \}$
- **Эквивалентные схемы** – одинаковая степень и одинаковый порядок следования атрибутов

Реляционные ключи

Реляционная модель данных – совокупность взаимосвязанных отношений. Для поддержки иерархических связей предназначены **ключи**.

- **Суперключ** – атрибут или множество атрибутов, однозначно определяющие кортеж данного отношения.
- **Потенциальный ключ** – суперключ, который не содержит подмножества, также являющегося суперключем данного отношения. Т.о. потенциальный ключ обладает свойствами уникальности и неприводимости.
- **Первичный ключ** – это потенциальный ключ, который выбран для уникальной идентификации кортежей внутри отношения
- **Внешний ключ** – это атрибут или множество атрибутов одного отношения, которые принимают значения потенциального ключа другого отношения (может быть и того же)

На схемах названия атрибутов первичного ключа выделяют подчеркиванием

Реляционные ключи

Студенты

ФИО (1)	№ паспорта (2)	№ зачетной книжки (3)	Год рождения (4)
Петров	1234567	321	2000
Сидоров	4321233	123	2000
Петров	7777888	234	2000

1+2+3+4 – суперключ

1+2+3 – суперключ

2 – потенциальный ключ

3 – потенциальный ключ

2 – первичный ключ (например)

Реляционные ключи

первичный

Отделы

<u>№ отдела</u>	Название	Телефон
1	Отдел кадров	004
2	Бухгалтерия	123

Сотрудники

внешний

<u>ФИО</u>	№ отдела	Должность
Иванов	2	Начальник отдела
Сидоров	2	Кассир

Помещения

внешний

<u>№ комнаты</u>	№ отдела	Площадь
1	2	40
5	2	15

Реляционные ключи

Составные ключи

~~Жильцы~~ Жильцы Ф.И.О. № дома № квартиры

<u>Улица</u>	<u>№ дома</u>	<u>№ квартиры</u>	ФИО	№ паспорта
Горького	202	4	Петров	1234567
Горького	321	4	Иванов	4321234

Первичный ключ

Ремонт

Внешний ключ

Вид ремонта	Дата	Улица	№ дома	№ квартиры
Капитальный	02.02.2002	Горького	321	4
Крыша	03.03.2003	Горького	202	4

Реляционная алгебра

Алгебра – множество элементов с заданной на нем совокупностью операций, замкнутых относительно этого множества

Реляционная алгебра – множество отношений и совокупность операций над отношениями

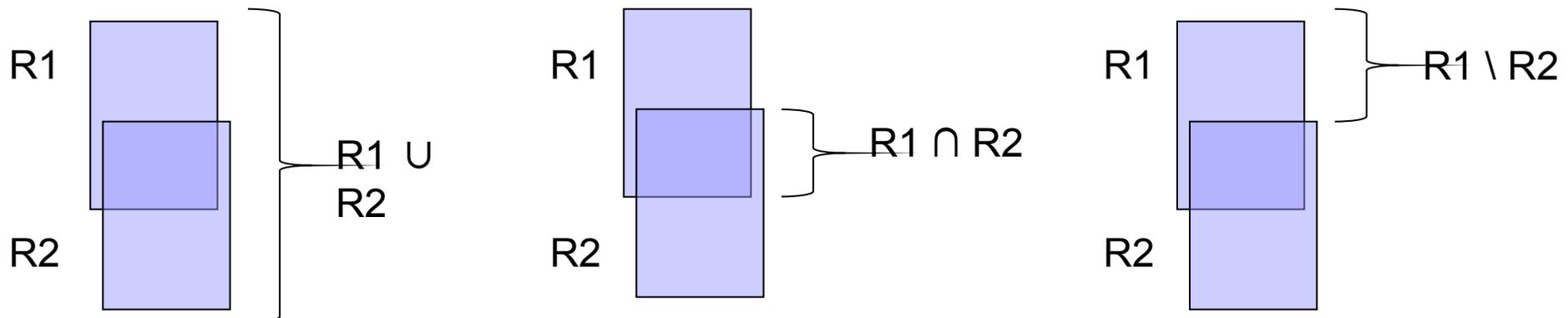
Реляционная база данных – совокупность некоторого числа отношений

Концептуальная модель базы данных (концептуальная схема базы данных) – множество всех реляционных схем отношений

Теоретико-множественные операции

- **Объединение отношений:** $R1 \cup R2 = \{r \mid r \in R1 \vee r \in R2\}$
- **Пересечение отношений:** $R1 \cap R2 = \{r \mid r \in R1 \wedge r \in R2\}$
- **Разность отношений:** $R1 \setminus R2 = \{r \mid r \in R1 \wedge r \notin R2\}$
- **Декартово произведение отношений** (моделирует ситуацию всех возможных исходов некоторого события):
 $R1 \times R2 = \{(r1, r2) \mid r1 \in R1, r2 \in R2\}$

Операции объединения, пересечения и вычитания определены только для отношений с эквивалентными схемами



$$R1 = \{1, 3, 5, 6\} \quad R2 = \{1, 6, 9\} \quad R1 \cup R2 = \{1, 3, 5, 6, 9\} \quad R1 \cap R2 = \{1, 6\} \quad R1 \setminus R2 = \{3, 5\}$$

Теоретико-множественные операции

Примеры

$R1 = \langle \text{ФИО, № зач.кн} \rangle$ - студенты, сдававшие экзамен в первый день

$R2 = \langle \text{ФИО, № зач.кн} \rangle$ - студенты, сдававшие экзамен во второй день

$R3 = \langle \text{ФИО, № зач.кн} \rangle$ - студенты, перешедшие на следующий курс

- Студенты, сдававшие экзамен 2 раза, но отчисленные
 $R = (R1 \cap R2) \setminus R3$
- Студенты, сдававшие экзамен 1 раз, и сдавшие его
 $R = ((R1 \setminus R2) \cap R3) \cup ((R2 \setminus R1) \cap R3)$

Декартово произведение

R1 (фамилии)

Иванов
Петров
Сидоров

R2 (номера зач.кн.)

001
002

$R1 \times R2$

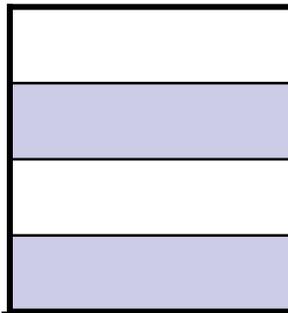
Иванов	001
Иванов	002
Петров	001
Петров	002
Сидоров	001
Сидоров	002

Специальные операции реляционной алгебры

Горизонтальная выборка (фильтрация, выборка)

$$R[\alpha(r)] = \{r \mid r \in R \wedge \alpha(r) = \text{истина}\}$$

где $\alpha(r)$ – предикат от атрибутов отношения



$R1 = \langle \text{ФИО, № зач.кн., стипендия} \rangle$

Иванов	001	500
Петров	011	300
Иванов	005	200
Сидоров	003	600

$R2 = R1[\text{ФИО} = \text{'Иванов'}]$

Иванов	001	500
Иванов	005	200

Одноместная (унарная) операция

Специальные операции реляционной алгебры

Проекция

$R = \langle a_1, a_2, \dots, a_m \rangle$, $V \subseteq \{a_i\}$ – множество атрибутов

$R[V] = \{r[V]\}$ – отношение с атрибутами V

Проекция – отношение со схемой V , содержащее кортежи из исходного отношения, после удаления атрибутов, не входящих в V . Дубликаты кортежей из результата удаляются

Одноместная (унарная) операция

Специальные операции реляционной алгебры

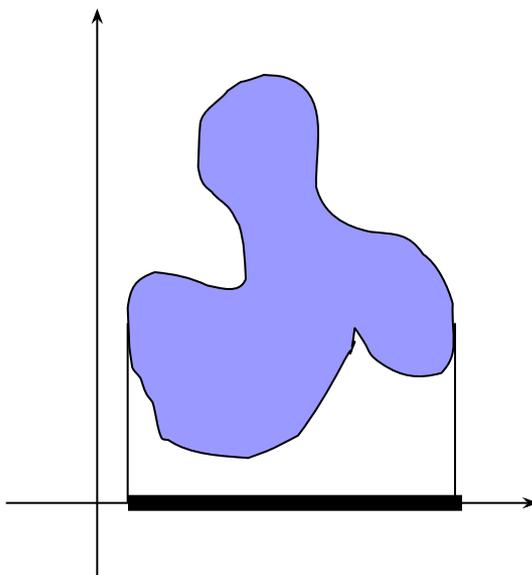
$R1 = \langle \text{ФИО}, \text{№ зач.кн.}, \text{стипендия} \rangle$

Иванов	001	500
Петров	011	300
Иванов	005	200
Сидоров	003	600

$R2 = R1[\text{ФИО}]$

Иванов
Петров
Сидоров

Аналогия



Специальные операции реляционной алгебры

Условное соединение

Двуместная (бинарная) операция

$$R = \langle a_1, a_2, \dots \rangle \quad T = \langle b_1, b_2, \dots \rangle$$

$\theta_k \in \{<, >, <=, >=, =, <>\}$ - операции сравнения

$\beta = \{R.a_i \theta_k T.b_j\}$, $k=1, k_m$ – предикат сравнения, определенный для атрибутов отношений

Тогда

$$R[\beta] T = \{(r,t) \mid r \in R, t \in T \text{ и } \beta(r.a_i \theta_k t.b_j) = \text{истина}, k=1, k_m\}$$

Условное соединение – конкатенация кортежей по заданному **условию** (поэтому **условное** соединение)

Операция соединения эквивалентна операции декартова произведения и последующей выборке в соответствии с предикатом соединения

Специальные операции реляционной алгебры

Частные случаи условного соединения:

- **Соединение по эквивалентности.** Это соединение в котором все θ_k – операции сравнения на равенство
- **Естественное соединение.** Это соединение по эквивалентности двух отношений R и T , выполняемое по общим атрибутам X , из результатов которого исключаются по одному экземпляру каждого общего атрибута
- **Внешние соединения.** (Будут рассмотрены позже)

R1

x	y
a	1
b	2

R2

y	z
1	H
1	C
3	N

R1[R1.y=R2.y]R2

x	y	y	z
a	1	1	H
a	1	1	C

Соединение
по эквивалентности

x	y	z
a	1	H
a	1	C

Естественное
соединение

Примеры

Концептуальная схема базы данных

$E = \langle \text{ФИО}, \text{Дисц}, \text{Оценка} \rangle$ - результаты сдачи экзаменов

$G = \langle \text{ФИО}, \text{Группа} \rangle$ - состав группы

$P = \langle \text{Группа}, \text{Дисц} \rangle$ - набор дисциплин, по которым надо сдавать экзамены группам

1. Получить список студентов, сдавших на отлично БД

$R = (E[\text{Оценка}=5 \text{ и } \text{Дисц}=\text{'БД'}])[\text{ФИО}]$

Примеры

2. Получить список тех, кто должен был сдавать экзамен по БД, но пока еще не сдавал

а) Соединить G и P, чтобы получить студентов и дисциплины, которые они должны сдавать

$R1 = (G[G.Группа = P.Группа \text{ и } P.Дисц = 'БД']P) [ФИО]$

б) Получить студентов, сдавших экзамен по БД

$R2 = (E[E.Дисц='БД'])[ФИО]$

в) Вычесть из всех, кто должен сдавать тех, кто уже сдал

$R = R1 \setminus R2$

Примеры

3. Получить список студентов, имеющих несколько двоек (более одной)

Введем E' – ссылка на то же самое отношение E (переименование отношения).

$R = (E[E.ФИО=E'.ФИО \text{ и } E.Оц=E'.Оц \text{ и } E.Оц=2 \text{ и } E.Дисц \neq E'.Дисц]E')[E.ФИО]$

E			E'		
ФИО	Дисц	Оценка	ФИО	Дисц	Оценка
Иванов	БД	2	Иванов	БД	2
Петров	БД	4	Петров	БД	4
Иванов	ОС	2	Иванов	ОС	2

Примеры

4. Получить список круглых отличников

а) Получить список студентов, которые должны что-либо сдавать с соответствующими дисциплинами

$R1 = (G[G.Группа = P.Группа]P) [ФИО, Дисц]$

б) Получить список студентов и дисциплин, уже сданных на отлично. Но среди студентов будут еще те, которые не все сдали на отлично или что-то еще не сдали.

$R2 = (E[Оценка = 5])[ФИО, Дисц]$

в) Список студентов, что-либо не сдавших на отлично, или не сдавших какие-то экзамены

$R3 = (R1 \setminus R2)[ФИО]$

г) Из всех студентов, которые должны сдавать экзамены, вычесть не сдавших что-либо на отлично и не сдававших какие-то экзамены

$R = R1[ФИО] \setminus R3$

(здесь нельзя делать $G[ФИО] \setminus R3$, т.к. в результат попадут студенты, которые не должны сдавать экзамены вообще)

Примеры

Концептуальная модель производства деталей

$P = \langle \text{Шифр}, \text{Название} \rangle$ - номенклатура деталей

$D = \langle \text{Цех} \rangle$ - все цеха завода

$F = \langle \text{Шифр}, \text{Цех} \rangle$ - детали, выпускаемые цехами

5. Определить перечень цехов, в которых выпускаются все детали (вся номенклатура)

$R1 = P[\text{Шифр}]$ получить шифры всех деталей

$R2 = R1 \times D$ моделируется ситуация, когда во всех цехах выпускаются все детали

$R3 = R2 \setminus F$ остаются цеха и детали, не выпускаемые в этих цехах

$R4 = R3[\text{Цех}]$ цеха, в которых выпускаются не все детали

$R5 = D \setminus R4$ цеха, в которых выпускаются все детали