



РЕЛЯЦИОННА Я МОДЕЛЬ ДАННЫХ



1. Основные определения

Теоретической основой модели стала теория отношений, основу которой заложили Чарльз Содерс Пирс (1839-1914), Эрнст Шредер (1841-1902).

«... множество отношений замкнуто относительно некоторых специальных операций, то есть образует вместе с этими операциями абстрактную алгебру»



Американский математик Э. Ф. Кодд в 1970 году впервые сформулировал основные понятия и ограничения реляционной модели.

Реляционной моделью (*relational model*) данных называется совокупность реляционной структуры и реляционной алгебры.

Основной структурой данных в модели является **отношение** (от англ. relation — отношение).

N -арным отношением R называют подмножество декартова произведения $D_1 \times D_2 \times \dots \times D_n$ множеств $D_1 \times D_2 \times \dots \times D_n$ ($n \geq 1$) необязательно различных.

Исходные множества $D_1 \times D_2 \times \dots \times D_n$ называют в модели *доменами*.

$$R \subseteq D_1 \times D_2 \times \dots \times D_n ,$$

где $D_1 \times D_2 \times \dots \times D_n$ — полное декартово произведение.

Полное декартово произведение — это набор всевозможных сочетаний из n элементов каждое, где каждый элемент берется из своего домена.

Например.

$D_1 = \{\text{Иванов, Крылов, Степанов}\};$

$D_2 = \{\text{Теория автоматов, Базы данных}\};$

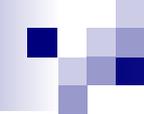
$D_3 = \{3, 4, 5\}$

<Иванов, Теория автоматов,3>; <Иванов, Теория автоматов,4>; <Иванов, Теория автоматов,5>;
<Крылов, Теория автоматов,3>; <Крылов, Теория автоматов,4>; <Крылов, Теория автоматов,5>;
<Степанов, Теория автоматов,3>; <Степанов, Теория автоматов,4>; <Степанов, Теория автоматов,5>;
<Иванов, Базы данных,3>; <Иванов, Базы данных,4>; <Иванов, Базы данных,5>;
<Крылов, Базы данных,3>; <Крылов, Базы данных,4>; <Крылов, Базы данных,5>;
<Степанов, Базы данных,3>; <Степанов, Базы данных,4>; <Степанов, Базы данных,5>;

R		
Фамилия	Дисциплина	Оценка
Иванов	Теория автоматов	4
Иванов	Базы данных	3
Крылов	Теория автоматов	5
Степанов	Теория автоматов	5
Степанов	Базы данных	4

Данная таблица обладает рядом специфических свойств:

- ✓ в таблице нет двух одинаковых строк.
- ✓ таблица имеет столбцы, соответствующие атрибутам отношения.
- ✓ каждый атрибут в отношении имеет уникальное имя.
- ✓ порядок строк в таблице произвольный.



Вхождение домена в отношение называют *атрибутом*.

Строки отношения называются *кортежами*.

Количество атрибутов в отношении – *степень* отношения.

Экземпляр отношения – отражает состояние данного объекта в текущий момент времени, и понятие схемы отношения, которая определяет структуру отношения.

Схемой отношения R называется перечень имен атрибутов данного отношения с указанием домена, к которому они относятся:

$$S_R = (A_1, A_2, A_n), A_i \subseteq D_i$$

Схемы двух отношений будут *эквивалентными*, если они имеют одинаковую степень и возможно такое упорядочение имен атрибутов в схемах, что на одинаковых местах будут находиться сравнимые атрибуты (т.е. атрибуты, принимающие значения из одного домена).

Первичный ключ отношения (PRIMARY KEY)

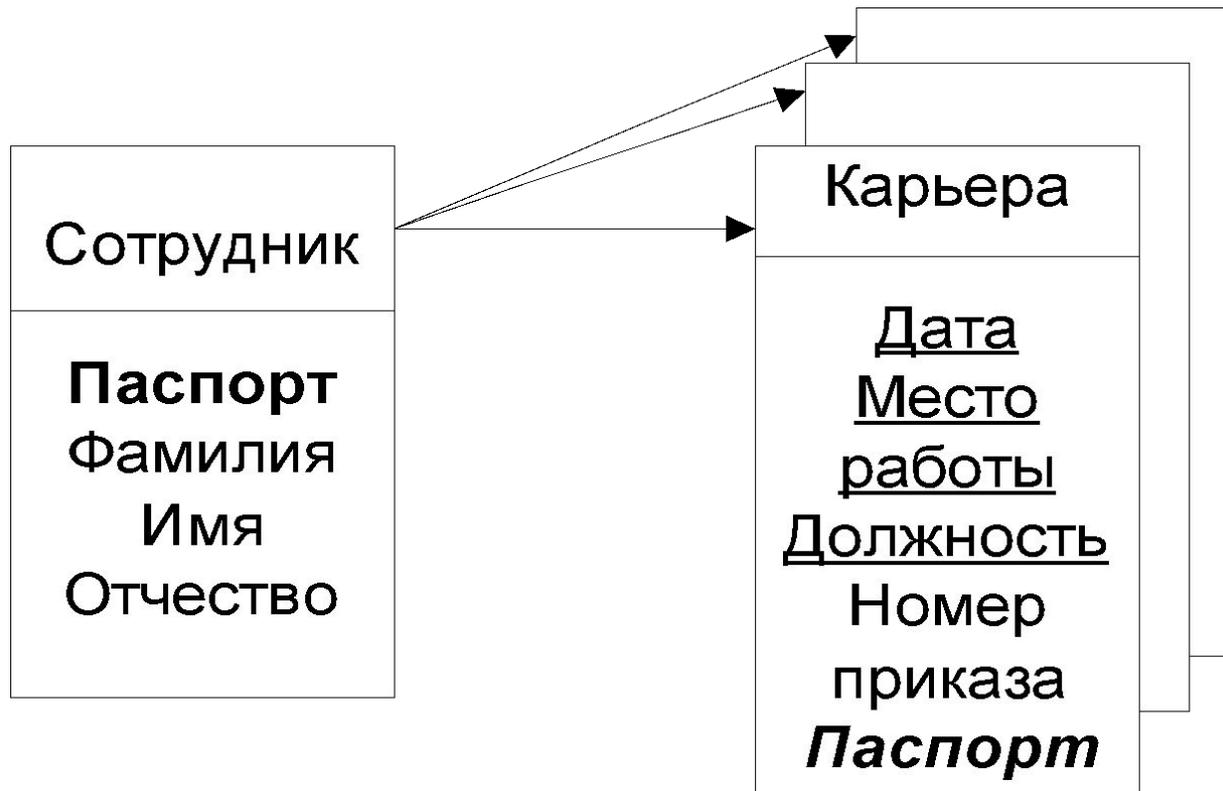
- каждое отношение имеет один и только один РК;
- значения РК не могут принимать значения null;
- значения РК не должны повторяться;

- РК не оказывает ни какого влияния на порядок атрибутов в отношении;
- РК не оказывает ни какого влияния на порядок кортежей в отношении;
- РК не оказывает ни какого влияния на доступ к данным.

Внешний ключ (FOREIGN KEY)

- значение FK должно ссылаться на существующее значение соответствующего первичного ключа;
- значения FK – могут быть null.

Например





2. ОПЕРАЦИИ НАД ОТНОШЕНИЯМИ. РЕЛЯЦИОННАЯ АЛГЕБРА

Алгеброй называется множество объектов с заданной на нем совокупностью операций, замкнутых относительно этого множества, называемого *основным множеством*.

Основным множеством в реляционной алгебре является *множество отношений*.



2.1. Теоретико-множественные операции реляционной алгебры

1). *Объединением* двух отношений называется отношение, содержащее множество кортежей, принадлежащих либо первому, либо второму исходным отношениям, либо обоим отношениям одновременно.

$R_1 = \{r_1\}$, $R_2 = \{r_2\}$, где r_1 и r_2 — соответственно кортежи отношений R_1 и R_2

$$R_1 \cup R_2 = \{r \mid r \in R_1 \vee r \in R_2\}$$

R₁	
Шифр детали	Наименование детали
73	Гайка М1
75	Гайка М2
76	Гайка М3
03	Болт М1
06	Болт М3
66	Шайба М1

R₂	
Шифр детали	Наименование детали
73	Гайка М1
76	Гайка М3
77	Гайка М4
04	Болт М2
06	Болт М3

Результат объединения:

R₃	
Шифр детали	Название детали
73	Гайка М1
75	Гайка М2
76	Гайка М3
03	Болт М1
06	Болт М3
66	Шайба М1
77	Гайка М4
04	Болт М2

2). *Пересечением* отношений называется отношение, которое содержит множество кортежей, принадлежащих одновременно и первому и второму отношениям \mathbf{R}_1 и \mathbf{R}_2

$$\mathbf{R}_4 = \mathbf{R}_1 \cap \mathbf{R}_2 = \{ \mathbf{r} \mid \mathbf{r} \in \mathbf{R}_1 \wedge \mathbf{r} \in \mathbf{R}_2 \}$$

Результат пересечения отношений **R1** и **R2** :

R₄	
Шифр детали	Название детали
73	Гайка М1
76	Гайка М3
06	Болт М3

3). *Разностью* отношений \mathbf{R}_1 и \mathbf{R}_2 называется отношение, содержащее множество кортежей, принадлежащих \mathbf{R}_1 и не принадлежащих \mathbf{R}_2 :

$$\mathbf{R}_5 = \mathbf{R}_1 \setminus \mathbf{R}_2 = \{ \mathbf{r} \mid \mathbf{r} \in \mathbf{R}_1 \wedge \mathbf{r} \notin \mathbf{R}_2 \}$$

$$\mathbf{R}_6 = \mathbf{R}_2 \setminus \mathbf{R}_1 = \{ \mathbf{r} \mid \mathbf{r} \in \mathbf{R}_2 \wedge \mathbf{r} \notin \mathbf{R}_1 \}$$

Результат разности отношений **R1** и **R2** :

R₅	
Шифр детали	Наименование детали
75	Гайка M2
03	Болт M1
66	Шайба M1

R₆	
Шифр детали	Наименование детали
77	Гайка M4
04	Болт M2

ЗАДАНИЕ

$R_1 = (\text{ФИО}, \text{Паспорт}, \text{Школа})$

- список абитуриентов, сдававших
подготовительные экзамены;

$R_2 = (\text{ФИО}, \text{Паспорт}, \text{Школа})$

- список абитуриентов, сдававших экзамены на
общих условиях;

$R_3 = (\text{ФИО}, \text{Паспорт}, \text{Школа})$

- список абитуриентов, принятых в институт.

Вопросы:

-  список абитуриентов, которые поступали два раза и не поступили;
-  список абитуриентов, которые поступили с первого раза;
-  список абитуриентов, которые поступили только со второго раза;
-  список абитуриентов, которые поступали только один раз и не поступили.

4). Расширенное декартово произведение.

Сцеплением (конкатенацией) кортежей

$$r = \langle r_1, r_2, \dots, r_n \rangle \text{ и } q = \langle q_1, q_2, \dots, q_m \rangle$$

называется кортеж, полученный

добавлением значений второго в конец

первого:

$$(r, q) = \langle r_1, r_2, \dots, r_n, q_1, q_2, \dots, q_m \rangle$$

где n - число элементов в первом кортеже r ;

m - число элементов во втором кортеже q .

Расширенным декартовым произведением отношения R_1 степени n и отношения R_2 степени m называется отношение R_3 степени $n+m$ содержащее кортежи, полученные сцеплением каждого кортежа r отношения R_1 с каждым кортежем q отношения R_2 . То есть $R_1 = \{r\}$, $R_2 = \{q\}$, тогда

$$R_1 \otimes R_2 = \{ (r, q) \mid r \in R_1 \wedge q \in R_2 \}$$

R₇	
Шифр детали	Название детали
73	Гайка М1
75	Гайка М2
03	Болт М1
04	Болт М2
66	Шайба М3

R₈
Цех
Цех1
Цех2
Цех3

R₉		
Шифр детали	Название детали	Цех
73	Гайка М1	Цех1
75	Гайка М2	Цех1
03	Болт М1	Цех1
04	Болт М2	Цех1
66	Шайба М3	Цех1
73	Гайка М1	Цех2
75	Гайка М2	Цех2
03	Болт М1	Цех2
04	Болт М2	Цех2
66	Шайба М3	Цех2
73	Гайка М1	Цех3
75	Гайка М2	Цех3
03	Болт М1	Цех3
04	Болт М2	Цех3
66	Шайба М3	Цех3

R₁₀ <i>(реальный выпуск)</i>		
Шифр детали	Название детали	Цех
73	Гайка М1	Цех1
75	Гайка М2	Цех1
03	Болт М1	Цех2
75	Гайка М2	Цех3



2.2. Специальные операции реляционной алгебры

1). *Горизонтальный выбор* или *операция фильтрации*.

Результатом операции выбора, заданной на отношении **R** в виде булевского выражения, определенного на атрибутах отношения **R**, называется отношение **R** [α], включающее те кортежи из исходного отношения, для которых истинно условие выбора или фильтрации:

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

Например:

$$R_{12} = R_{10} [\text{Шифр детали} = \langle\langle 75 \rangle\rangle]$$

R_{12}		
Шифр детали	Название детали	Цех
75	Гайка M2	Цех1
75	Гайка M2	Цех3

2). *Вертикальный выбор* или *операция проектирования*.

Проекцией отношения **R** на набор атрибутов **V**, обозначаемой **R[V]**, называется отношение со схемой соответствующей набору атрибутов **V** $S_{R[V]} = V$, содержащему кортежи, получаемые из кортежей исходного отношения **R** путем удаления из них значений, не принадлежащих атрибутам из набора **V**

$$R[V] = \{ r[V] \}$$

Например:

R13 = R10 [Название детали = «Гайка М2»]

R14 = R13 [Цех]

R₁₃		
Шифр детали	Название детали	Цех
75	Гайка М2	Цех1
75	Гайка М2	Цех3

R₁₄
Цех
Цех1
Цех3

3). Операция условного соединения.

Операция условного соединения является бинарной, то есть исходными для нее являются два отношения, а результатом – одно.

Пусть схема отношения R содержит набор атрибутов N , а схема отношения S - набор атрибутов M . Пусть N и M - θ -сравнимы.

Тогда *соединением* отношений R и S по условию $M \theta N$ называется такое отношение U , обозначаемое $R[M \theta N]S$, которое имеет схему $U(R,S)$ и кортежи которого получаются сцеплением тех кортежей отношений R и S , для которых на наборах атрибутов N и M условие сравнения дает значение истина:

$$U = R[M \theta N]S = \{r, s \mid r \in R \wedge s \in S \wedge r[M] \theta s[N]\}$$

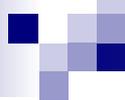
Иногда обозначается:

$$R \stackrel{F}{\sim} S,$$

где F – условие соединения.



Соединение по условию равенства (=)
называется *экви-соединением*



Соединение по условию равенства, когда один из сравниваемых атрибутов (или группа сравниваемых атрибутов) удаляется из результирующего отношения, называется *естественным соединением*

Например:

$R(A, B, C, D)$ и $S(C, D, E)$,

в результате выполнения операции $Q = R * S$
получится отношение со схемой

$Q(A, B, C, D, E)$

Операция *полусоединение* - при которой из результата исключаются все атрибуты одного из соединяемых отношений.

Обозначается как $R \ltimes_M \theta N \bowtie S$

$$R \ltimes_M \theta N \bowtie S = \{r \mid r \in R \wedge s \in S \wedge r[M] \theta s[N]\}$$

выражение через соединение и проекцию

$$R \llbracket M \theta N \rrbracket S = (R \llbracket M \theta N \rrbracket S) \llbracket A \rrbracket,$$

где A – множество всех атрибутов отношения R .

R				S			?					?						
A	B	C	D	C	D	E	A	B	C	D	C	D	E	A	B	C	D	E
a ₁	b ₁	c ₁	d ₁	c ₁	d ₁	e ₂	a ₁	b ₁	c ₁	d ₁	c ₁	d ₁	e ₂	a ₁	b ₁	c ₁	d ₁	e ₂
a ₁	b ₁	c ₂	d ₁	c ₂	d ₁	e ₃	a ₁	b ₂	c ₁	d ₁	c ₁	d ₁	e ₂	a ₁	b ₂	c ₁	d ₁	e ₂
a ₁	b ₂	c ₁	d ₁	c ₂	d ₁	e ₁	a ₂	b ₃	c ₁	d ₁	c ₁	d ₁	e ₂	a ₂	b ₃	c ₁	d ₁	e ₂
a ₁	b ₂	c ₂	d ₁	c ₂	d ₂	e ₅	a ₂	b ₃	c ₁	d ₁	c ₁	d ₁	e ₂	a ₂	b ₃	c ₁	d ₁	e ₂
a ₂	b ₃	c ₁	d ₁				a ₁	b ₁	c ₂	d ₁	c ₂	d ₁	e ₃	a ₁	b ₁	c ₂	d ₁	e ₃
a ₂	b ₄	c ₂	d ₃				a ₁	b ₂	c ₂	d ₁	c ₂	d ₁	e ₃	a ₁	b ₂	c ₂	d ₁	e ₃
							a ₁	b ₁	c ₂	d ₁	c ₂	d ₁	e ₁	a ₁	b ₁	c ₂	d ₁	e ₁
							a ₁	b ₂	c ₂	d ₁	c ₂	d ₁	e ₁	a ₁	b ₂	c ₂	d ₁	e ₁



4). Операція делення.

ЗАДАНИЕ

$R_1 = \langle \text{ФИО, Дисциплина, Оценка} \rangle$

– информация о попытках (как успешных, так и неуспешных) сдачи экзаменов студентами ;

$R_2 = \langle \text{ФИО, Группа} \rangle$

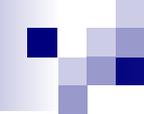
– состав групп ;

$R_3 = \langle \text{Группа, Дисциплина} \rangle$

– список дисциплин, которые надо сдавать каждой группе.

Вопросы:

-  список студентов, которые сдали экзамен по БД на «отлично»;
-  список студентов, которые должны сдавать экзамен по БД, но пока не сдавали;
-  список студентов, которые сдавали экзамен по БД;
-  список студентов, которые имеют несколько двоек;
-  список студентов, сдавших всё на «отлично».



Задание для самостоятельной работы

Даны отношения, моделирующие работу банка и его филиалов.

Клиент может иметь несколько счетов, при этом они могут быть размещены как в одном, так и в разных филиалах банка. В отношении **R1** содержится информация обо всех клиентах и их счетах в филиалах нашего банка. Каждый клиент, в соответствии со своим счетом, может рассчитывать на кредит от банка (сумма допустимого кредита зафиксирована).

R₁				
ФИО клиента	№ филиала	№ счета	Остаток	Кредит

R₂	
№ филиала	Район

С использованием языка реляционной алгебры составить запросы, позволяющие выбрать:

- a) филиалы, клиенты которых имеют счета с остатком, превышающим 5000грн.
- b) клиентов, которые имеют счета во всех филиалах данного банка.
- c) клиентов, которые имеют только по одному счету в разных филиалах (т.е. в общем у этих клиентов может быть несколько счетов, но в одном филиале не более одного счета).

- d) клиенты, которые имеют счета в нескольких филиалах банка, расположенных только в одном районе.
- e) филиалы, которые не имеют ни одного клиента.
- f) филиалы, которые имеют клиентов с остатком на счету ноль.
- g) филиалы, у которых есть клиенты с кредитом, превышающим остаток на счету в 2 раза.