

Реляционная модель представления данных

Лекция 3

Доцент каф. ВТИКГ Данилова Е.В.

План

1. Компоненты реляционной модели данных
2. Структурная часть реляционной модели данных
3. Реализация РМД в реляционных базах данных
4. Целостная часть реляционной модели данных
5. Манипуляционная часть реляционной модели данных

Компоненты реляционной модели

Реляционный: relation – отношение

Принципы реляционной модели были сформулированы в 1969-1970 годах Э. Ф. Коддом (E. F. Codd):

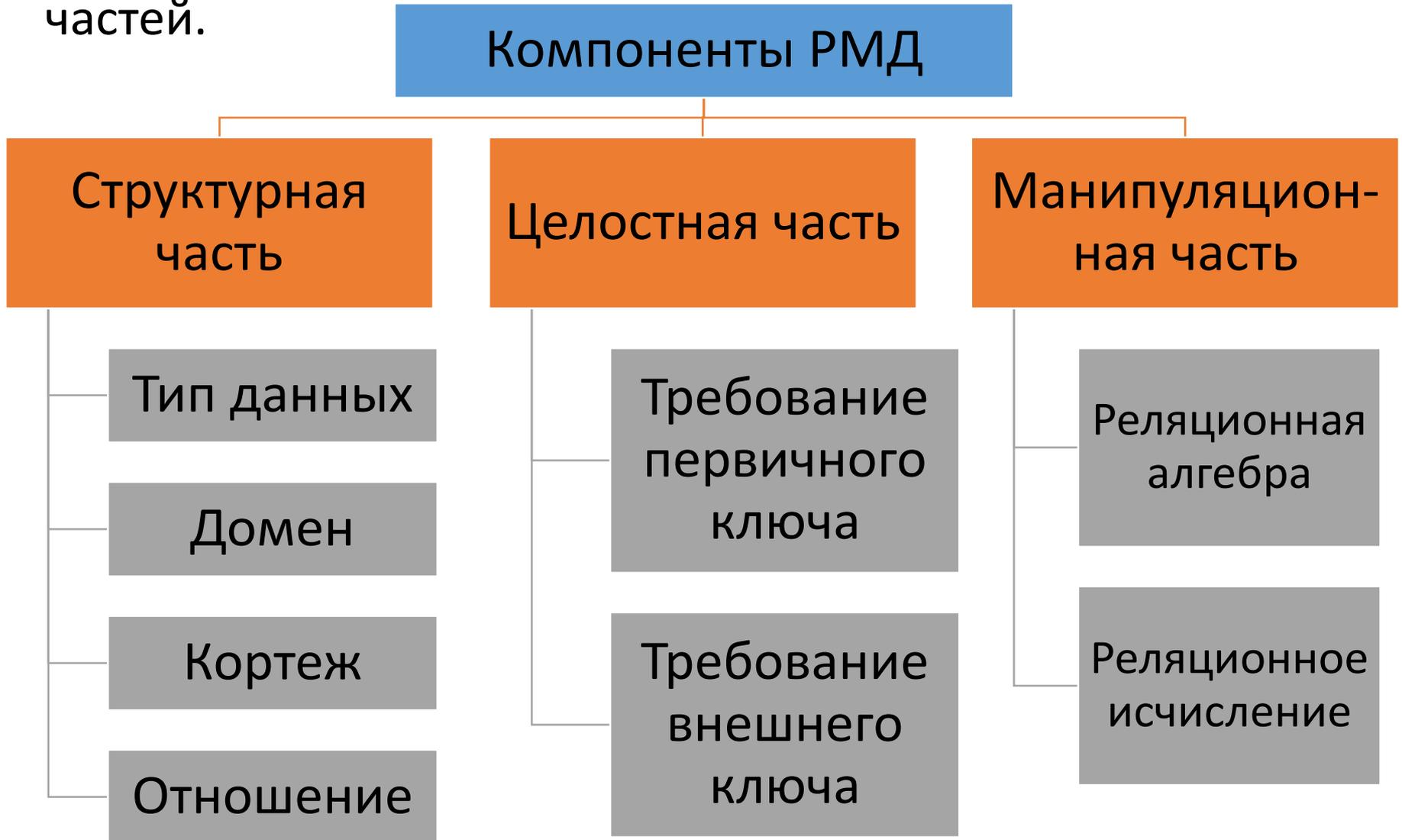
- **для внедрения в сферу управления базами данных строгих и точных принципов можно использовать математические дисциплины**

Идеи Кодда:

- впервые публично изложены в статье «A Relational Model of Data for Large Shared Data Banks», ставшей классической.
- стали общепринятыми, оказали существенное влияние на все аспекты технологии баз данных, а также на другие области: искусственный интеллект, обработка естественных языков и проектирование аппаратных средств.

Наиболее распространена трактовка реляционной модели данных, принадлежащая Крису Дейту.

Согласно Дейту, реляционная модель состоит из трех частей.



Структурная часть реляционной модели

Структурная часть.

- Описывает, какие объекты рассматриваются реляционной моделью.
- Постулируется, что единственной структурой данных, используемой в реляционной модели, являются нормализованные n -арные отношения.

1.1. Тип данных.

Понятие **тип данных** в реляционной модели данных полностью адекватно понятию типа данных в языках программирования.

Реляционная модель требует, чтобы типы используемых данных были простыми, т.е. в реляционных операциях не должна учитываться внутренняя структура данных.

Типы данных

Простые
(атомарные)

Логический

Строковый

Численный

Целый, вещественный,
дата, время,
перечислимый и др.

Структурирован-
ные

Массивы

Записи

Ссылочные

Указатели

1.2. Домен.

В реляционной модели данных с понятием тип данных тесно связано понятие домена, которое можно считать уточнением типа данных.

Домен - подмножество значений некоторого типа данных, имеющих определенный смысл.

Домен характеризуется следующими свойствами:

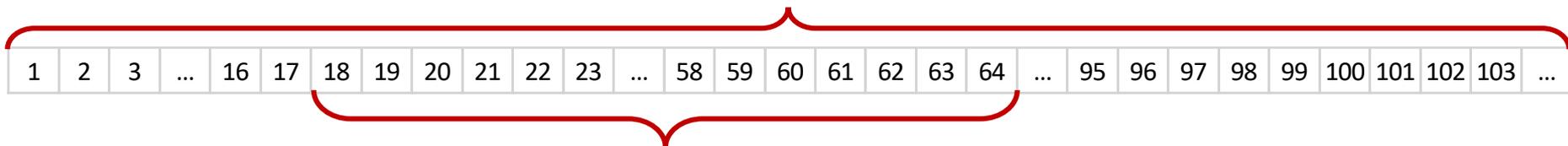
- имеет уникальное имя (в пределах базы данных);
- определен на некотором простом типе данных или на другом домене;
- может иметь некоторое логическое условие, позволяющее описать подмножество данных, допустимых для данного домена;
- несет определенную смысловую нагрузку.

Домен D , имеющий смысл «возраст сотрудника» можно описать как следующее подмножество множества натуральных чисел:

$$D = \{n \in N : n \geq 18 \text{ and } n \leq 65\}$$

Если тип данных можно считать множеством всех возможных значений данного типа, то домен напоминает подмножество в этом множестве:

множество N натуральных чисел



подмножество множества N

Отличие домена от понятия подмножества состоит в том, что домен отражает семантику, определенную предметной областью.

Может быть несколько доменов, совпадающих как подмножества, но несущих различный смысл.

Например, домены «Стоимость» и «Срок годности» можно одинаково описать как множество неотрицательных целых чисел, но смысл этих доменов будет различным, и это будут различные домены.

Препарат



Код препарата

Название препарата

Код упаковки

Стоимость

Дата выпуска

Срок годности

Основное значение доменов – ограничивать сравнения

Некорректно, с логической точки зрения, сравнивать значения из различных доменов, даже если они имеют одинаковый тип. В этом проявляется смысловое ограничение доменов.

Понятие домена помогает правильно моделировать предметную область.

1.3. Кортеж.

Кортеж – основной элемент реляционной модели.

Кортеж – это упорядоченный набор элементов, каждый из которых принадлежит определенному множеству или, иначе говоря, имеет свой тип (тип данных).

Кортеж – это набор именованных значений заданного типа.

Пусть дана коллекция типов данных T_i ($i = 1, 2, \dots, n$), которые не обязательно все должны быть разными. Значением кортежа t (или кратко – кортежем), определенным с помощью этих типов, является множество упорядоченных троек в форме:

$$\langle A_i : T_i : v_i \rangle$$

где A_i – имя атрибута, T_i – имя типа, v_i – значение типа T_i .

Пример.

Пусть дана следующая коллекция типов данных

$$T_4: Ч, С, С, Ч$$

где **Ч** – числовой тип данных, **С** – строковый тип данных.

Тогда кортежем будет являться следующее множество:

$\langle \text{Код:Ч:1, Название:С:Здоровье, Улица:С:Ленина, Номер_дома:Ч:15} \rangle$

$$A_1 T_1 v_1 \quad A_2 T_2 v_2 \quad A_3 T_3 v_3 \quad A_4 T_4 v_4$$

Пример.

Пусть дана следующая коллекция типов данных

$T_4: Ч, С, С, Ч$

где Ч – числовой тип данных, С – строковый тип данных.

Тогда кортежем будет являться следующее множество:

<Код:Ч:1, Название:С:Здоровье, Улица:С:Ленина, Номер_дома:Ч:15>

$A_1 T_1 v_1$ $A_2 T_2 v_2$ $A_3 T_3 v_3$ $A_4 T_4 v_4$

Код : Ч	Название : С	Улица : С	Номер_дома:Ч
1	Здоровье	Ленина	15

Требования, предъявляемые к кортежу t :

1. Значение n определяет степень, или арность кортежа t .
2. Упорядоченная тройка $\langle A_i:T_i:v_i \rangle$ является компонентом кортежа t .

$\langle \text{Код:Ч:1, Название:С:Здоровье, Улица:С:Ленина, Номер_дома:Ч:15} \rangle$

компоненты кортежа

Код : Ч

1

Название : С

Здоровье

Улица : С

Ленина

Номер_дома:Ч

15

3. Упорядоченная пара $\langle A_i : T_i \rangle$ представляет собой атрибут и однозначно определяется именем атрибута A_i .

$\langle \underbrace{\text{Код:Ч:1}}, \underbrace{\text{Название:С:Здоровье}}, \underbrace{\text{Улица:С:Ленина}}, \underbrace{\text{Номер_дома:Ч:15}} \rangle$

атрибуты

Код : Ч	Название : С	Улица : С	Номер_дома:Ч
1	Здоровье	Ленина	15

В неформальном изложении принято исключать имена типов и показывать только имена атрибутов.

<Код:1, Название:Здоровье, Улица:Ленина, Номер_дома:15>

атрибуты

Код	Название	Улица	Номер_дома
1	Здоровье	Ленина	15

4. Значение v_i – это значение атрибута, соответствующее атрибуту A_i кортежа t .

Тип T_i – это соответствующий тип атрибута.

<Код:1, Название:Здоровье, Улица:Ленина, Номер_дома:15>

Значение атрибута

Код	Название	Улица	Номер_дома
1	Здоровье	Ленина	15

5. Полное множество атрибутов составляет заголовок кортежа **t**:

<Код:Ч, Название:С, Улица:С, Номер_дома:Ч>

Код : Ч	Название : С	Улица : С	Номер_дома:Ч
1	Здоровье	Ленина	15

*Заголовок
кортежа*

или в неформальном изложении:

<Код, Название, Улица, Номер_дома>

Код	Название	Улица	Номер_дома
1	Здоровье	Ленина	15

6. Тип кортежа **t** определен заголовком этого кортежа, а сам заголовок и этот тип кортежа имеют такие же атрибуты (и поэтому такие же имена и типы атрибутов) и такую же степень, как **t**.

Основные свойства кортежа.

1. Каждый кортеж содержит точно одно значение (соответствующего типа) для каждого из своих атрибутов.
2. Для компонентов кортежа не предусмотрено упорядочение (например, слева на право). Это свойство следует из того, что понятие кортежа определено на основании множества компонентов, а в математике множества не характеризуются упорядочением своих элементов.

Основные свойства кортежа.

3. Каждое подмножество кортежа представляет собой кортеж, а каждое подмножество заголовка кортежа является заголовком.
4. Кортеж степени один ($n=1$) называется одноэлементным, кортеж степени два – двухэлементным, кортеж степени три – трехэлементным (и т.д.); в общем случае кортеж степени n называют n -элементным. Кортеж нулевой степени (т.е. кортеж без компонентов) называют нуль-элементным или нуль-арным.

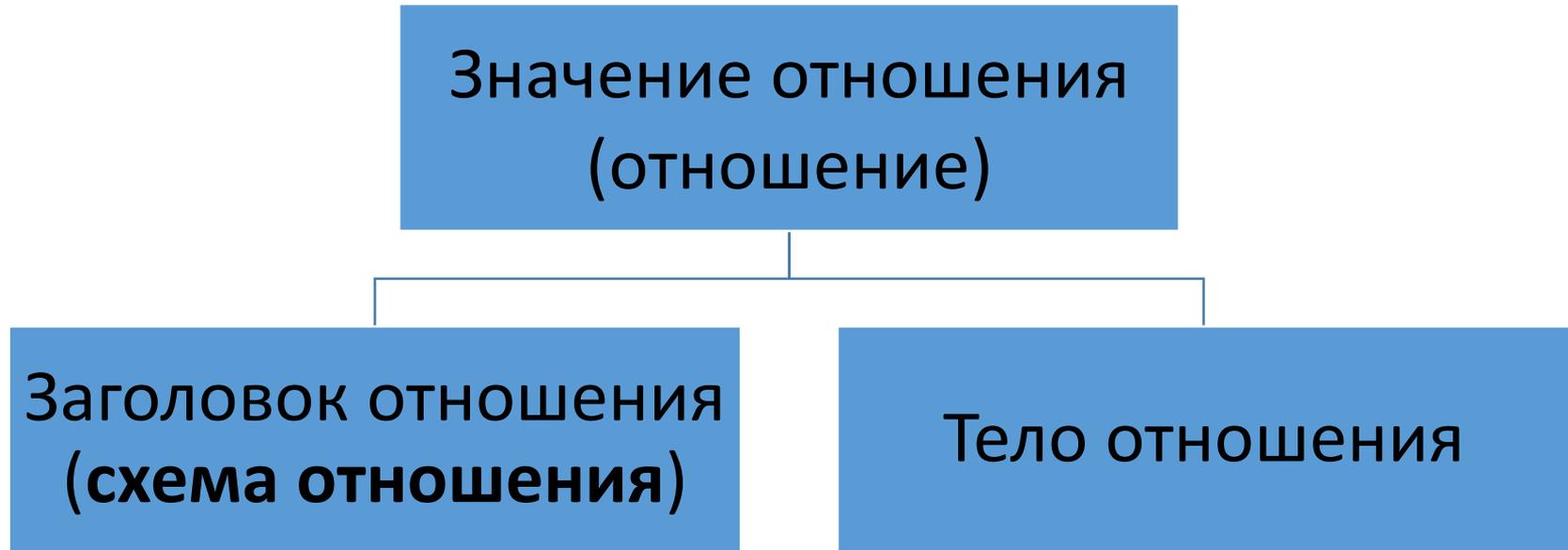
1.4. Отношение.

Фундаментальным понятием реляционной модели данных является понятие отношения.

В реляционной модели различают следующие понятия:

- значения отношений (или просто отношения);
- типы отношения;
- переменные отношения (relation variable, или сокращенно relvar).

Значение отношения r (или просто отношение) состоит из двух частей: заголовок и тела.



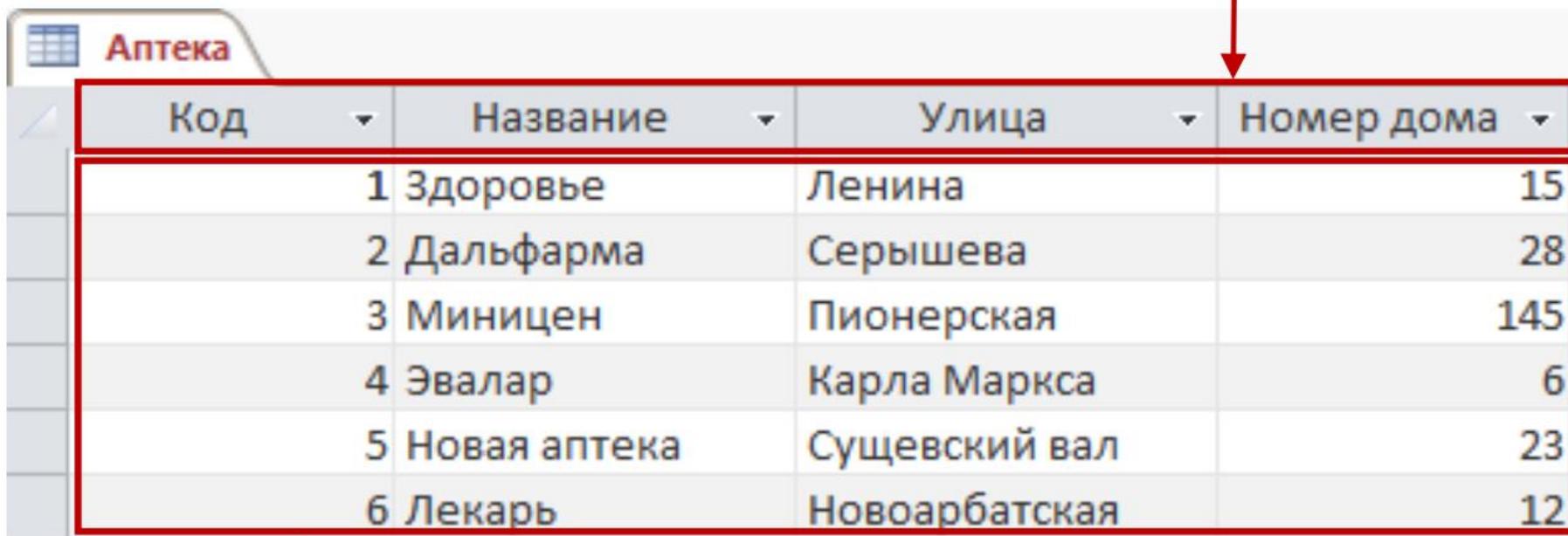
Заголовок отношения r представляет собой заголовок кортежа. Отношение r имеет такие же атрибуты (следовательно, такие же имена атрибутов и типы атрибутов) и такую же степень n , как заголовок кортежа:

$$(< A_1: T_1 >, < A_2: T_2 >, \dots, < A_n: T_n >)$$

Тело отношения r представляет собой множество кортежей, имеющих один и тот же заголовок.

$(\langle A_1: v_1 \rangle, \langle A_2: v_2 \rangle, \dots, \langle A_n: v_n \rangle)$

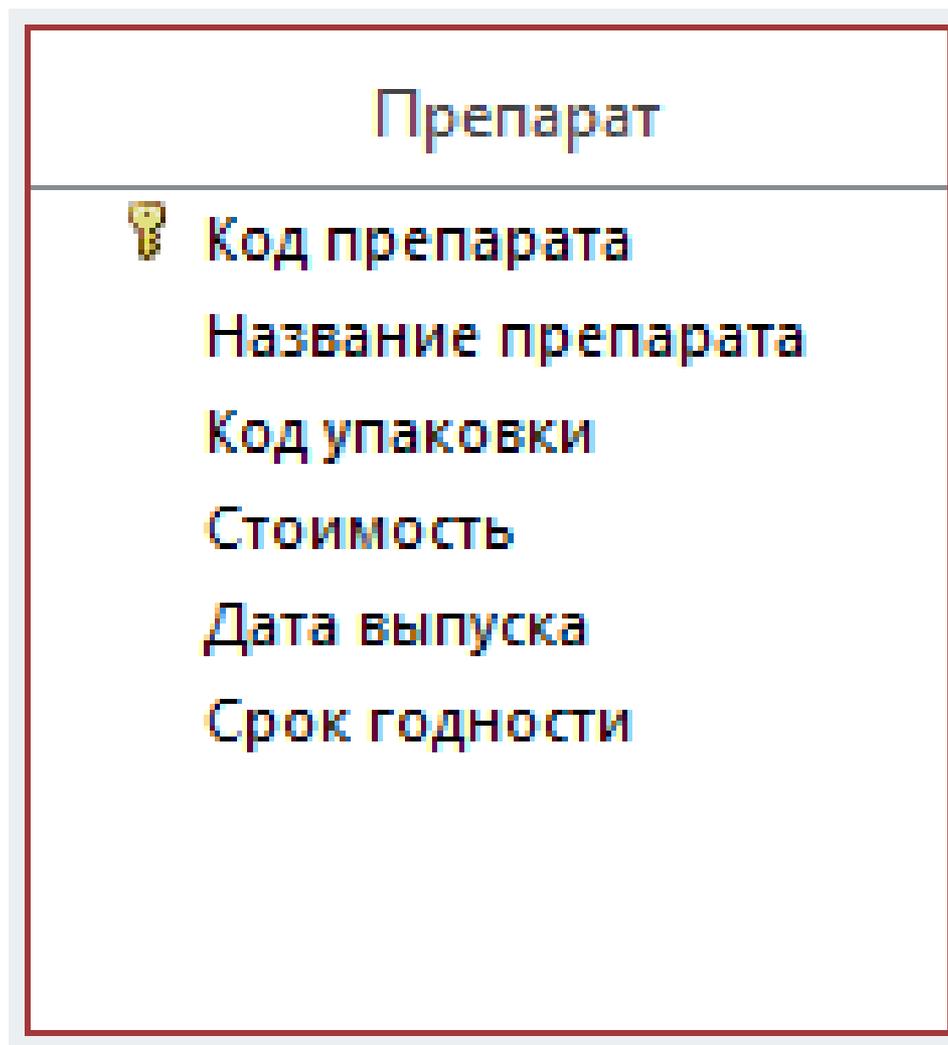
Заголовок (схема) отношения



Код	Название	Улица	Номер дома
1	Здоровье	Ленина	15
2	Дальфарма	Серышева	28
3	Миницен	Пионерская	145
4	Эвалар	Карла Маркса	6
5	Новая аптека	Сущевский вал	23
6	Лекарь	Новоарбатская	12

Тело отношения

Пример схемы отношения в СУБД MS Access



Пример записи отношения (в реляционной модели):

$$\mathbf{r} (< A_1:v_1 >, < A_2:v_2 >, \dots, < A_n:v_n >)$$

или

$$\mathbf{r} (A_1, A_2, \dots, A_n)$$

или просто \mathbf{r} .

Число атрибутов в отношении называют **степенью** (или -
арностью) **отношения**.

Тип отношения \mathbf{r} определяется заголовком \mathbf{r} и имеет такие же атрибуты (следовательно, имена и типы атрибутов) и степень, как и сам заголовок отношения.

Отношение, строго говоря, не содержит кортежей – оно содержит тело, а тело, в свою очередь, включает в себя кортежи.

Но человеку удобно формулировать свои высказывания так, как если бы отношения непосредственно включали кортежи.

Как и в случае кортежей, отношение степени один называется унарным, отношение степени два – бинарным, отношение степени три – тернарным (и т. д.).

Отношение степени n называется n -арным.

Отношение нулевой степени (т.е. отношение без атрибутов) принято называть нуль-арным.

Для отношений справедливы следующие утверждения

1. Каждое подмножество заголовка отношения является заголовком (как и в случае кортежей).
2. Каждое подмножество тела отношения является телом отношения.

Фундаментальные свойства отношений

1. Каждый кортеж содержит точно одно значение (соответствующего типа) для каждого атрибута (значения всех атрибутов являются атомарными).

То есть отношения **нормализованы**.

Таким образом, отношение всегда находится в первой нормальной форме (1НФ).

Примеры отношений

- Нормализованное отношение

Аптека				
Код	Название	Улица	Номер дома	
1	Здоровье	Ленина	15	
2	Дальфарма	Серышева	28	
3	Миницен	Пионерская	145	

- Ненормализованное «отношение»

Аптека				
Код	Название	Улица	Номер дома	
1	Здоровье, Хабаровск, 08:00 - 20:00	Ленина	15	
2	Дальфарма, Хабаровск, 09:00 - 18:00	Серышева	28	
3	Миницен, Владивосток, 10:00 - 20:00	Пионерская	145	

Фундаментальные свойства отношений

2. Атрибуты не характеризуются каким-либо упорядочением (например, слева на право).

Код	Название	Улица	Номер дома	Город
1	Здоровье	Ленина	15	Хабаровск
2	Дальфарма	Серышева	28	Хабаровск
3	Миницен	Пионерская	145	Владивосток
4	Эвалар	Карла Маркса	6	Уссурийск

Код	Город	Улица	Номер дома	Название
1	Хабаровск	Ленина	15	Здоровье
2	Хабаровск	Серышева	28	Дальфарма
3	Владивосток	Пионерская	145	Миницен
4	Уссурийск	Карла Маркса	6	Эвалар

Фундаментальные свойства отношений

3. Кортежи не характеризуются каким-либо упорядочением (например, сверху вниз).

Код	Город	Улица	Номер дома	Название
20	Хабаровск	Комсомольская	53	Артлайф
19	Хабаровск	Космическая		Фарком
18	Хабаровск	Кочнева	14	Семейная аптека
17	Хабаровск	Суворова	51	Эколайф-ДВ

17	Хабаровск	Суворова	51	Эколайф-ДВ
18	Хабаровск	Кочнева	14	Семейная аптека
19	Хабаровск	Космическая		Фарком
20	Хабаровск	Комсомольская	53	Артлайф

Фундаментальные свойства отношений

4. В отношении отсутствуют дубликаты кортежей.

В классической теории множеств каждое множество состоит из различных (не одинаковых) элементов



Отношение - множество кортежей



Элементы отношения (кортежи) должны быть разными (не одинаковыми).

Из свойства 4 следует наличие у каждого отношения «**первичного ключа**».

Реализация РМД в реляционных базах данных

Таблицы и реляционная модель данных

Отношение – формальный объект, фундаментальное понятие реляционной модели.

Таблица – неформальное отображение отношения в реляционной базе данных.

1. Для каждого атрибута в заголовке отношения предусмотрено имя типа, но эти имена типов обычно не показаны на изображениях в виде таблиц.
2. Для каждого компонента каждого кортежа в теле отношения предусмотрено имя типа и имя атрибута, но эти имена типа и атрибута обычно не показаны на изображениях в виде таблицы.

3. Каждое значение атрибута в каждом кортеже в теле отношения является значением соответствующего типа, но эти значения на изображениях в виде таблицы обычно показаны в сокращенной форме.

4. Столбцы в таблице характеризуются упорядочением слева направо, а атрибуты отношения – нет. Т. о. столбцы в таблице реляционной БД могут иметь повторяющиеся имена или вообще не иметь имен.

5. Строки таблицы характеризуются упорядочением сверху вниз, а кортежи отношения – нет.

6. Таблица может содержать дубликаты строк, а отношение не содержит дубликаты кортежей.

Соглашения, принятые для интерпретации отношений в виде таблиц РБД

- Каждый столбец характеризуется соответствующим типом.
- Каждое значение атрибута представляет собой значение определенного типа.
- Упорядочение строк и столбцов не имеет значения.
- Дубликаты строк не допускаются.

SQL и реляционная модель данных

1. В SQL поддерживаются не **кортежи**, а **строки**.

Строки характеризуются упорядочением своих компонентов слева направо.

INSERT INTO WORKERS VALUES ('Иван', 22);

NAME	AGE
Иван	22
Петр	32
Павел	34

**INSERT INTO *конечный_объект* (*поле1*[, *поле2*[, ...]])
VALUES (*поле1*[, *поле2*[, ...]])**

2. В SQL поддерживаются не **отношения**, а **таблицы**.

Тело таблицы – не множество кортежей, а мультимножество строк.

Мультимножество (multiset) – представляет собой коллекцию, которая, не имеет упорядочения, но допускает наличие дубликатов элементов).

Т. о. в таблицах реляционной БД могут присутствовать дубликаты строк.

NAME	AGE
Иван	22
Иван	22
Павел	34

Целостная часть РМД

Целостная часть РМД

- Описывает ограничения специального вида, которые должны выполняться для любых отношений в любых реляционных базах данных: целостность сущностей и целостность внешних ключей.

Требование целостности сущностей

Объект
реального
мира

```
graph LR; A[Объект реального мира] --> B[Сущность]; B --> C[Набор кортежей];
```

Сущность

Набор
кортежей

Любой кортеж любого отношения должен быть отличим от любого другого кортежа этого отношения

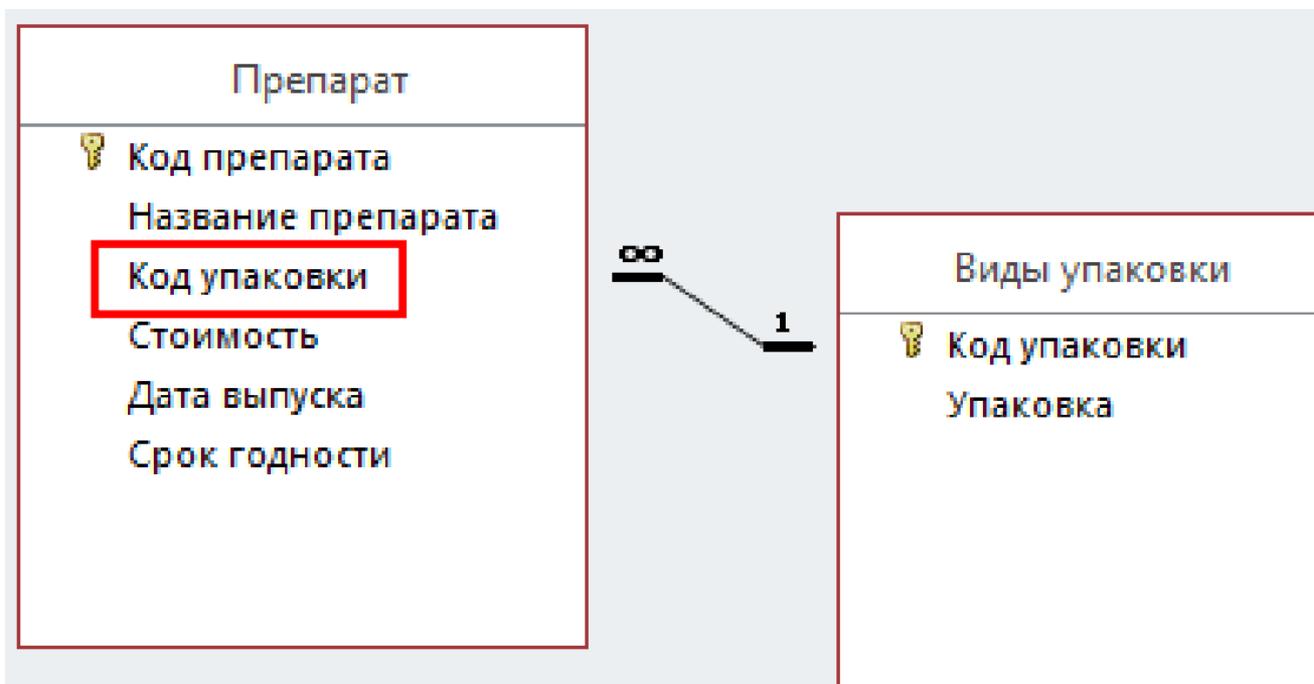
**Любое отношение должно обладать
первичным ключом**

Все элементы первичного ключа уникальны, и никакая часть первичного ключа не может быть пустой (null).

Целостность на уровне сущности гарантирует, что каждый экземпляр сущности будет иметь уникальную идентификацию, а значения внешнего ключа могут должным образом ссылаться на значения первичного ключа.

Требование целостности по ссылкам (требование внешнего ключа)

- Для каждого значения внешнего ключа, появляющегося в ссылающемся отношении, в отношении, на которое ведет ссылка, должен найтись кортеж с таким же значением первичного ключа, либо значение внешнего ключа должно быть неопределенным (т.е. ни на что не указывать)



Целостность на уровне ссылки

- База данных не должна содержать каких-либо несогласованных значений внешнего ключа.

«Если значение В ссылается на А, то А должно существовать»

- Внешний ключ может иметь пустое значение (null), если только он не является частью первичного ключа данной таблицы, или значение, совпадающее со значением первичного ключа в связанной таблице. Каждое непустое значение внешнего ключа должно ссылаться на существующее значение первичного ключа.

Манипуляционная часть РМД

Манипуляционная часть РМД

- Описывает два эквивалентных способа манипулирования реляционными данными – реляционную алгебру и реляционное исчисление.

Реляционная алгебра

Реляционная алгебра – это коллекция операций, которые принимают отношения в качестве операндов и возвращают отношение в качестве результата

Реляционное свойство замкнутости: результатом любой конкретной реляционной операции с отношениями является другое отношение.

Следствие свойства замкнутости: допустимо применение вложенных реляционных выражений, т.е. таких выражений, операнды которых, в свою очередь, представлены реляционными выражениями произвольной сложности

Реляционная алгебра, предложенная Коддом, включает следующие основные операции:

■ традиционные операции со множествами (теоретико-множественные):

- объединение;
- разность (вычитание);
- пересечение;
- декартово (прямое) произведение;

■ специальные реляционные операции:

- выборка (сокращение, селекция, ограничение);
- проекция;
- деление;
- соединение.

Бинарные и унарные реляционные операции

Аналогия из арифметики:

пример унарной операции – модуль числа:

$$|-1| = 1$$

операнд 1

результат

пример бинарной операции – сложение:

$$3 + 4 = 7$$

операнд 1

операнд 2

результат

Требование к отношениям при выполнении реляционных операторов

Некоторые реляционные операции (например, объединение) требуют, чтобы отношения имели одинаковые заголовки.

Отношение = заголовок + тело

Результат объединения отношений **A** и **B** –
отношение???

A и **B** – разнотипные кортежи (разное количество атрибутов) и/или различные наименования атрибутов.

Требование выполняется следующим образом:

- введением понятия «совместимость по типу»;
- использованием оператора переименования атрибутов.

Отношения, совместимые по типу

Отношения называют совместимыми по типу, если они имеют идентичные заголовки:

- отношения имеют одно и то же множество имен атрибутов, то есть для любого атрибута в одном отношении найдется атрибут с таким же наименованием в другом отношении;
- атрибуты с одинаковыми именами определены на одних и тех же доменах.

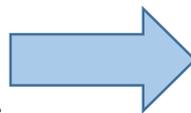
Оператор переименования атрибутов

Некоторые отношения не являются совместимыми по типу, но становятся таковыми после переименования атрибутов.

ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	2	4
Петров П.П.	1997	3	3
Сидоров С.С.	1999	1	5
Никаноров В.Е.	1998	2	3,7

← Отношение **A** – студенты, обучающиеся в ДВГУПС

Отношение **B** – студенты, обучающиеся в группе 32А



ФАМИЛИЯ ИО	ГОД	КУРС	УСПЕВАЕМОСТЬ
Завгородний А.А.	1998	2	4,5
Светличная С.В.	1998	2	5
Карандышев В.Н.	1998	2	4,3
Никаноров В.Е.	1998	2	3,7
Иванов И.И.	1998	2	4

A не совместимо по типу с **B**

R RENAME Atr_1, Atr_2, \dots AS $NewAtr_1, NewAtr_2, \dots$

где:

- **R** – отношение,
- **Atr_1, Atr_2, \dots** - исходные имена атрибутов;
- **$NewAtr_1, NewAtr_2, \dots$** - новые имена атрибутов.

Результат применения оператора переименования атрибутов – новое отношение с измененными именами атрибутов.

Традиционные операции со множествами

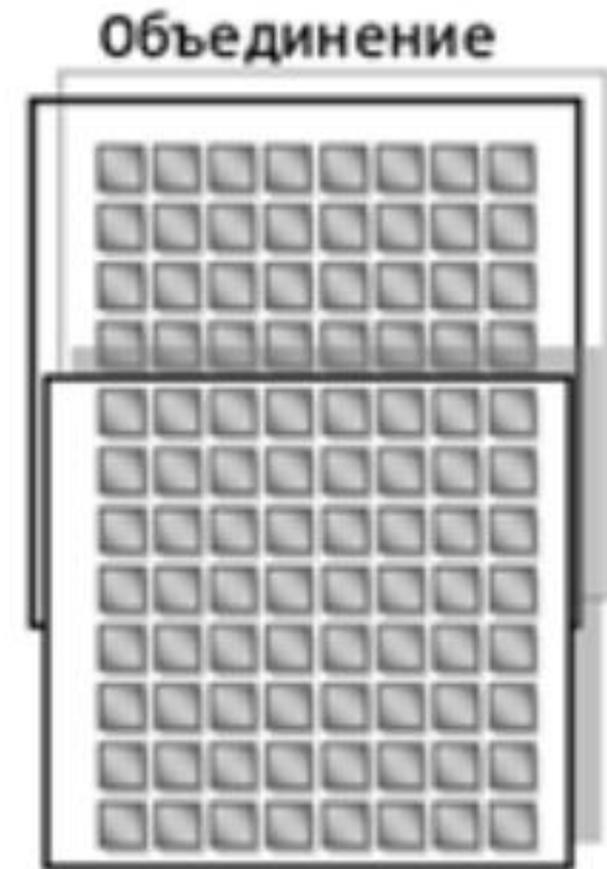
Объединение

Если даны отношения A и B одного и того же типа, то объединение этих отношений является отношением того же типа с телом, которое состоит из всех кортежей t , присутствующих в A , или B , или в обоих отношениях.

$$C = A \text{ UNION } B$$

Так как по определению отношения никогда не содержат дубликаты кортежей, то из результатов операции объединения устраняются дубликаты – одинаковые кортежи.

Операнды данной реляционной операции должны принадлежать одному и тому же типу.



ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	2	4
Петров П.П.	1997	3	3
Сидоров С.С.	1999	1	5
Никаноров В.Е.	1998	2	3,7

← Отношение **A** – студенты, обучающиеся в ДВГУПС

ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Завгородний А.А.	1998	2	4,5
Светличная С.В.	1998	2	5
Карандышев В.Н.	1998	2	4,3
Никаноров В.Е.	1998	2	3,7
Иванов И.И.	1998	2	4

Отношение **B** – студенты, обучающиеся в группе 32А →

ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	2	4
Петров П.П.	1997	3	3
Сидоров С.С.	1999	1	5
Никаноров В.Е.	1998	2	3,7
Завгородний А.А.	1998	2	4,5
Светличная С.В.	1998	2	5
Карандышев В.Н.	1998	2	4,3

← Отношение **C = A UNION B**

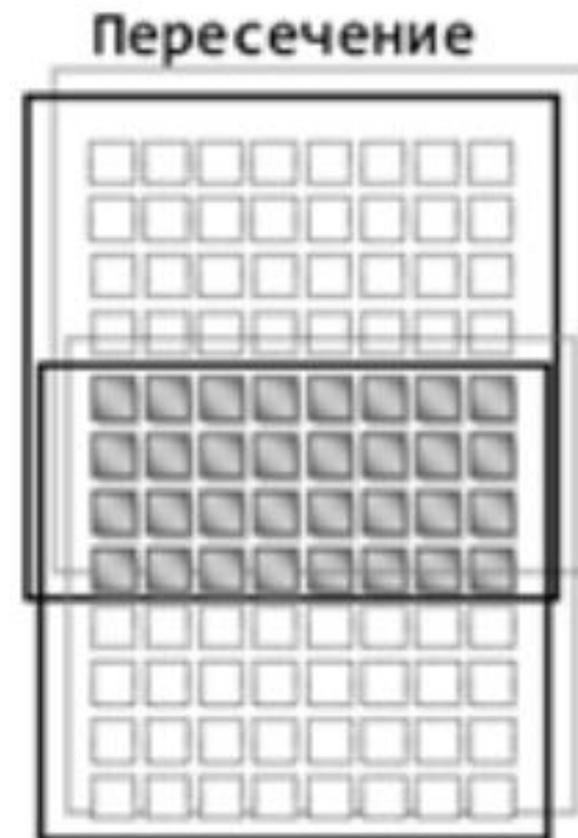
Пересечение

Если даны отношения A и B одного и того же типа, то пересечение этих отношений является отношением того же типа с телом, которое состоит из всех кортежей t , присутствующих одновременно в обоих отношениях A и B .

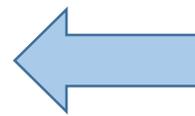
$$C = A \text{ INTERSECT } B$$

Так как по определению отношения никогда не содержат дубликаты кортежей, то из результатов операции объединения устраняются дубликаты – одинаковые кортежи.

Операнды данной реляционной операции должны принадлежать одному и тому же типу.

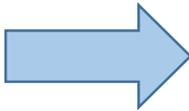


ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	2	4
Петров П.П.	1997	3	3
Сидоров С.С.	1999	1	5
Никаноров В.Е.	1998	2	3,7



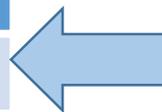
Отношение **A** – студенты, обучающиеся в ДВГУПС

Отношение **B** – студенты, обучающиеся в группе 32А



ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Завгородний А.А.	1998	2	4,5
Светличная С.В.	1998	2	5
Карандышев В.Н.	1998	2	4,3
Никаноров В.Е.	1998	2	3,7
Иванов И.И.	1998	2	4

ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	2	4
Никаноров В.Е.	1998	2	3,7



Отношение
C = A INTERSECT B

Разность

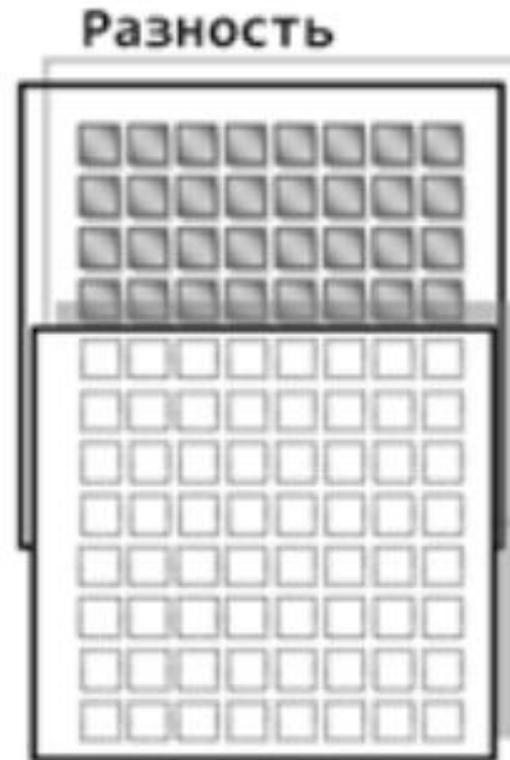
Если даны отношения **A** и **B** одного и того же типа, то разностью этих отношений **A MINUS B** (в указанном порядке), является отношение того же типа с телом, состоящим из всех кортежей **t**, таких, что **t** присутствует в **A**, но не в **B**.

$$C = A \text{ MINUS } B$$

или

$$C = A \text{ DIFFERENCE } B$$

Операнды данной реляционной операции должны принадлежать одному и тому же типу.



ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	2	4
Петров П.П.	1997	3	3
Сидоров С.С.	1999	1	5
Никаноров В.Е.	1998	2	3,7

← Отношение **A** – студенты, обучающиеся в ДВГУПС

Отношение **B** – студенты, обучающиеся в группе 32А →

ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Завгородний А.А.	1998	2	4,5
Светличная С.В.	1998	2	5
Карандышев В.Н.	1998	2	4,3
Никаноров В.Е.	1998	2	3,7
Иванов И.И.	1998	2	4

ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Петров П.П.	1997	3	3
Сидоров С.С.	1999	1	5

← Отношение **C = A MINUS B**

Отношение **C = B MINUS A** →

ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Завгородний А.А.	1998	2	4,5
Светличная С.В.	1998	2	5
Карандышев В.Н.	1998	2	4,3

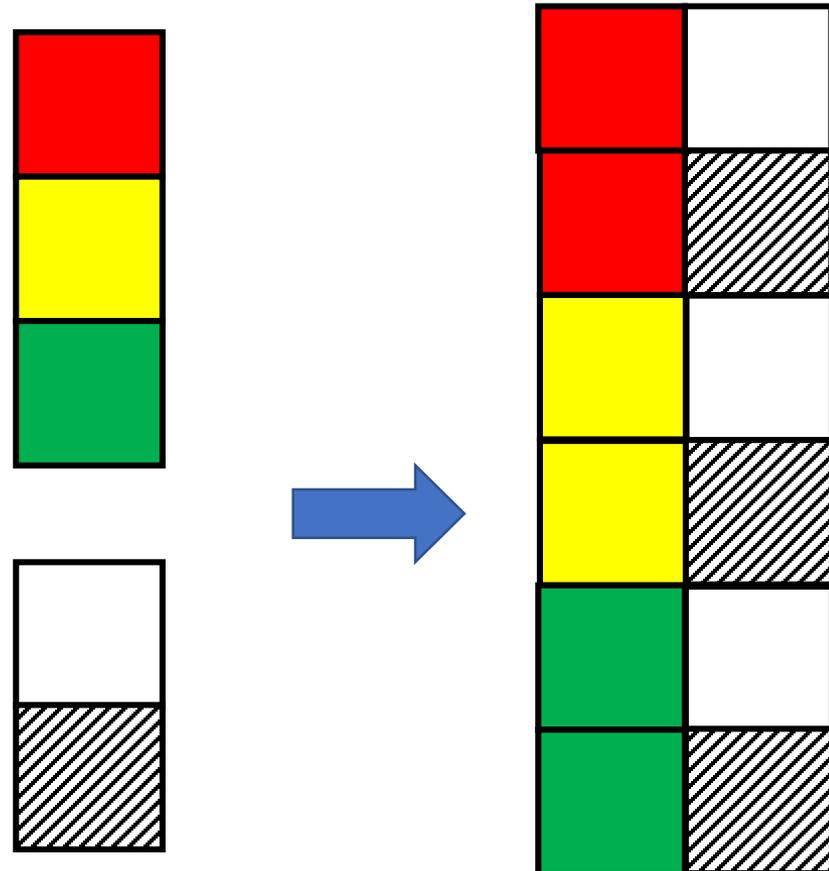
Декартово произведение

Декартово произведение отношений A и B , не имеющих общих атрибутов, – это отношение, заголовок которого представляет собой объединение заголовков отношений A и B , а тело состоит из всех всех кортежей t , таких, что t является объединением кортежа, принадлежащего к отношению A , и кортежа, принадлежащего к отношению B .

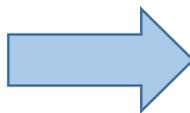
$$C = A \text{ TIMES } B$$

или

$$C = A \text{ PRODUCT } B$$



Отношение **A** – студенты,
обучающиеся в группе 32А



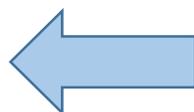
ФИО
Завгородний А.А.
Светличная С.В.
Карандышев В.Н.
Никаноров В.Е.

ФИО	ДИСЦИПЛИНА
Завгородний А.А.	Менеджмент
Завгородний А.А.	Правоведение
Завгородний А.А.	Экология
Светличная С.В.	Менеджмент
Светличная С.В.	Правоведение
Светличная С.В.	Экология
Карандышев В.Н.	Менеджмент
Карандышев В.Н.	Правоведение
Карандышев В.Н.	Экология
Никаноров В.Е.	Менеджмент
Никаноров В.Е.	Правоведение
Никаноров В.Е.	Экология

Отношение **B** –
дисциплины 3-го
семестра группы 32А



ДИСЦИПЛИНА
Менеджмент
Правоведение
Экология



Отношение
C = A TIMES B

Специальные реляционные операции

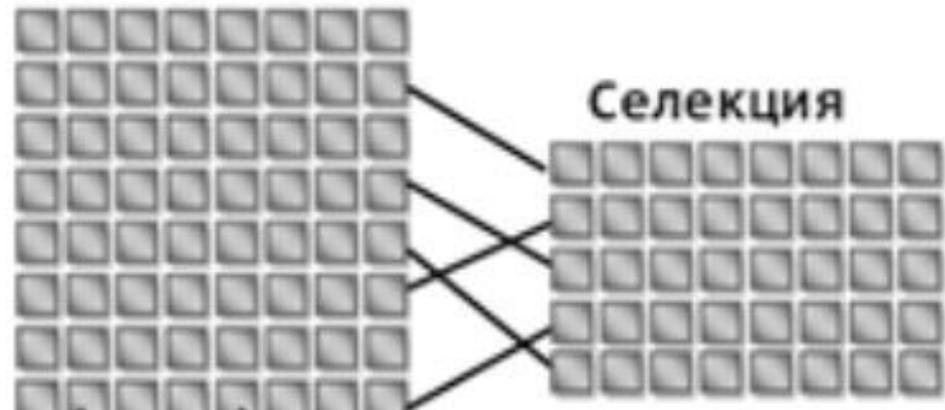
Выборка (SELECT, сокращение, селекция и др.)

Выборка отношения **A** по формуле **f** представляет собой новое отношение с таким же заголовком и телом, состоящим из кортежей отношения **A**, которые удовлетворяют истинности логического выражения, заданного формулой **f**.

A WHERE f

Оператор выборки позволяет получить «горизонтальное» подмножество заданного отношения, т.е. подмножество кортежей заданного отношения, для которых удовлетворяется некоторое указанное условие.

f – предикат,
условие сокращения.



ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	2	4
Петров П.П.	1997	3	3
Сидоров С.С.	1999	1	5
Никаноров В.Е.	1998	2	3,7
Завгородний А.А.	1998	2	4,5
Светличная С.В.	1998	2	5
Карандышев В.Н.	1998	2	4,3

Отношение **A** – студенты, обучающиеся в ДВГУПС

Отношение A WHERE ГОД РОЖДЕНИЯ = 1998

ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	2	4
Никаноров В.Е.	1998	2	3,7
Завгородний А.А.	1998	2	4,5
Светличная С.В.	1998	2	5
Карандышев В.Н.	1998	2	4,3

*условие
сокращения f*

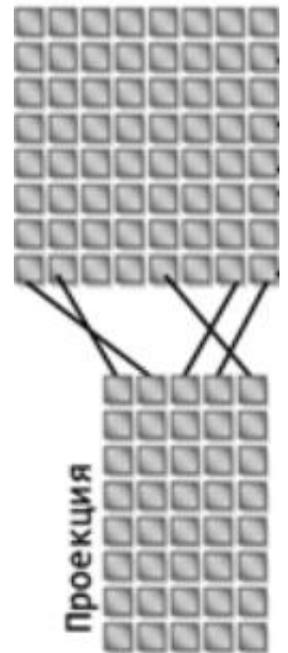
Проекция (PROJECT)

Проекцией отношения A с атрибутами X, Y, \dots, Z называется отношение $A \{X, Y, \dots, Z\}$, заголовок которого формируется из заголовка отношения A путем удаления всех атрибутов, не указанных в множестве $\{X, Y, \dots, Z\}$, а тело состоит из всех кортежей $\{X:x, Y:y, \dots, Z:z\}$, таких, что в отношении A присутствует кортеж со значением x атрибута X , y атрибута Y ... и z атрибута Z .

Возвращает кортежи с сокращенным набором атрибутов.

$A \{X, Y, \dots, Z\}$

Применение операции проекции приводит к получению «вертикального» подмножества, т.е. подмножества, полученного путем удаления всех атрибутов, не указанных в разделенном запятыми списке имен атрибутов, и последующего устранения дубликатов кортежей из множества оставшихся кортежей.



Отношение **A** – студенты, обучающиеся в ДВГУПС

ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	2	4
Петров П.П.	1997	3	3
Сидоров С.С.	1999	1	5
Никаноров В.Е.	1998	2	3,7
Завгородний А.А.	1998	2	4,5
Светличная С.В.	1998	2	5
Липатова У.В.	1998	3	4,9

Отношение **A** {КУРС, ГОД РОЖДЕНИЯ}

КУРС	ГОД РОЖДЕНИЯ
2	1998
3	1997
1	1999
3	1998

Соединение

Наиболее важная реляционная операция

Виды операции «соединение»:

- естественное соединение
- экви-соединение
- \bowtie -соединение (тэта-соединение)

*Частные
случаи общей
операции
соединения*

Соединение (естественное, natural JOIN)

Пусть отношение **A** имеет следующие атрибуты:

$\{ X_1, X_2, \dots, X_m, Y_1, Y_2, \dots, Y_n \}$

отношение **B** имеет атрибуты:

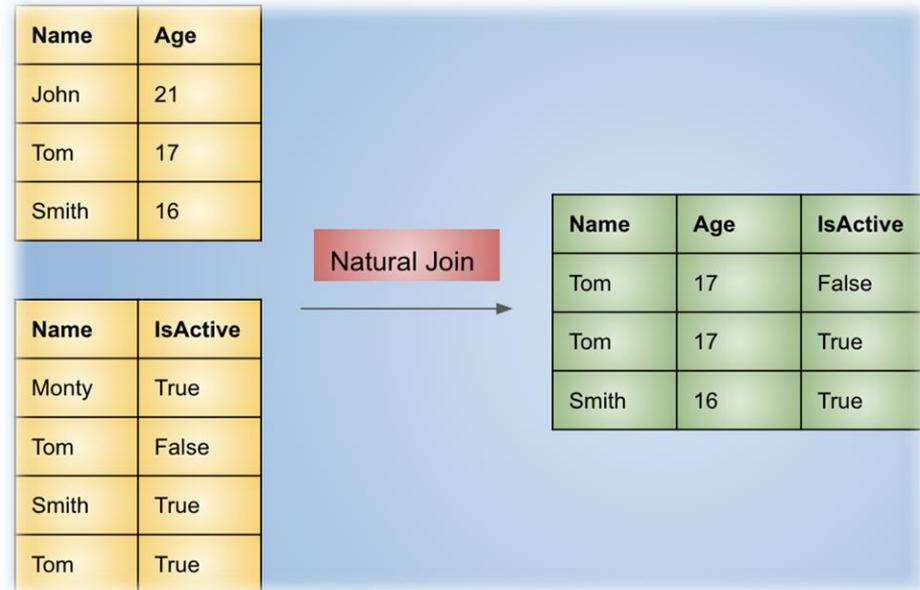
$\{ Y_1, Y_2, \dots, Y_n, Z_1, Z_2, \dots, Z_p \}$

То есть два рассматриваемых отношения имеют общее множество атрибутов $Y: \{Y_1, Y_2, \dots, Y_n\}$, а остальные атрибуты образуют множества:

$X: \{ X_1, X_2, \dots, X_m \}$ и $Z: \{ Z_1, Z_2, \dots, Z_p \}$

Естественное соединение отношений **A** и **B** представляет собой отношение с заголовком $\{X, Y, Z\}$ и телом, состоящим из всех таких кортежей $\{X:x, Y:y, Z:z\}$, что любой из этих кортежей присутствует и в отношении **A** со значением x атрибута **X** и значением y атрибута **Y**, и в отношении **B** со значением y атрибута **Y** и значением z атрибута **Z**.

A JOIN B



Соединение (естественное, natural JOIN)

Если отношения **A** и **B** имеют общий атрибут, то результат естественного соединения содержит только строки, в которых значения общего атрибута совпадают.

Связывает таблицы, выбирая только строки с общими значениями их общих атрибутов [Роб, Ролланд].

КУРС	ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ
2	Иванов И.И.	1998	4
3	Петров П.П.	1997	3
1	Сидоров С.С.	1999	5

Отношение А

КУРС	ГРУППА	СПЕЦИАЛЬНОСТЬ
1	34И	ИВТ
2	32А	БИ
1	921	ФОИ
4	921	УК

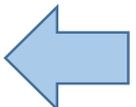
Отношение В

Y: {КУРС}

X: {ФИО, ГОД РОЖДЕНИЯ, СРЕДНИЙ БАЛЛ}

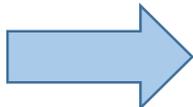
Z: { ГРУППА, СПЕЦИАЛЬНОСТЬ }

КУРС	ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ
2	Иванов И.И.	1998	4
3	Петров П.П.	1997	3
1	Сидоров С.С.	1999	5



Отношение **A**

Отношение **B**



КУРС	ГРУППА	СПЕЦИАЛЬНОСТЬ
1	34И	ИВТ
2	32А	БИ
1	921	ФОИ
4	921	УК

Отношение **C = A JOIN B**

КУРС	ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ	ГРУППА	СПЕЦИАЛЬНОСТЬ
2	Иванов И.И.	1998	4	32А	БИ
1	Сидоров С.С.	1999	5	34И	ИВТ
1	Сидоров С.С.	1999	5	921	ФОИ

Естественное соединение можно представить как последовательность операций:

1. Переименовать одинаковые атрибуты в отношениях

Отношение А

А.КУРС	ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ
2	Иванов И.И.	1998	4
3	Петров П.П.	1997	3
1	Сидоров С.С.	1999	5

Отношение В

В.КУРС	ГРУППА	СПЕЦИАЛЬНОСТЬ
1	34И	ИВТ
2	32А	БИ
1	921	ФОИ
4	921	УК

Естественное соединение можно представить как последовательность операций:

2. Получить декартово произведение отношений **A** и **B**

А.КУРС	ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ	В.КУРС	ГРУППА	СПЕЦИАЛЬНОСТЬ
2	Иванов И.И.	1998	4	1	34И	ИВТ
2	Иванов И.И.	1998	4	2	32А	БИ
2	Иванов И.И.	1998	4	1	921	ФОИ
2	Иванов И.И.	1998	4	4	921	УК
3	Петров П.П.	1997	3	1	34И	ИВТ
3	Петров П.П.	1997	3	2	32А	БИ
3	Петров П.П.	1997	3	1	921	ФОИ
3	Петров П.П.	1997	3	4	921	УК
1	Сидоров С.С.	1999	5	1	34И	ИВТ
1	Сидоров С.С.	1999	5	2	32А	БИ
1	Сидоров С.С.	1999	5	1	921	ФОИ
1	Сидоров С.С.	1999	5	4	921	УК

Естественное соединение можно представить как последовательность операций:

3. Выполнить выборку по условию **А.КУРС = В.КУРС** (по совпадающим атрибутам, столбцам соединения)

А.КУРС	ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ	В.КУРС	ГРУППА	СПЕЦИАЛЬНОСТЬ
2	Иванов И.И.	1998	4	1	34И	ИВТ
2	Иванов И.И.	1998	4	2	32А	БИ
2	Иванов И.И.	1998	4	1	921	ФОИ
2	Иванов И.И.	1998	4	4	921	УК
3	Петров П.П.	1997	3	1	34И	ИВТ
3	Петров П.П.	1997	3	2	32А	БИ
3	Петров П.П.	1997	3	1	921	ФОИ
3	Петров П.П.	1997	3	4	921	УК
1	Сидоров С.С.	1999	5	1	34И	ИВТ
1	Сидоров С.С.	1999	5	2	32А	БИ
1	Сидоров С.С.	1999	5	1	921	ФОИ
1	Сидоров С.С.	1999	5	4	921	УК

Естественное соединение можно представить как последовательность операций:

4. Получить проекцию (для вывода единственной копии каждого атрибута, т. е. удаляются дублирующие столбцы)

А.КУРС	ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ	ГРУППА	СПЕЦИАЛЬНОСТЬ
2	Иванов И.И.	1998	4	32А	БИ
1	Сидоров С.С.	1999	5	34И	ИВТ
1	Сидоров С.С.	1999	5	921	ФОИ

Естественное соединение можно представить как последовательность операций:

5. Переименовать атрибуты, вернув им первоначальные имена

КУРС	ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ	ГРУППА	СПЕЦИАЛЬНОСТЬ
2	Иванов И.И.	1998	4	32А	БИ
1	Сидоров С.С.	1999	5	34И	ИВТ
1	Сидоров С.С.	1999	5	921	ФОИ

Замечания

- Если $n = 0$ (а это означает, что отношения **A** и **B** не имеют общих атрибутов), то операция **A JOIN B** вырождается в операцию декартова произведения **A TIMES B**
- Если $m = p = 0$ (это означает, что отношения **A** и **B** относятся к одинаковому типу), то операция **A JOIN B** вырождается в операцию пересечения **A INTERSECT B**
- В синтаксисе естественного соединения не указываются, по каким атрибутам производится соединение. Естественное соединение производится по всем одинаковым атрибутам.

⊕-соединение (тета-соединение)

Пусть отношения **A** и **B** удовлетворяют требованиям для декартового произведения (то есть, не имеют общих имен атрибутов). Пусть отношение **A** имеет атрибут **x**, отношение **B** – атрибут **y**.

Результат тета-соединения – отношение **C** с тем же заголовком, как и у декартова произведения **A** и **B**, и с телом, состоящим из множества всех кортежей **t**, таких, что **t** присутствует в этом декартовом произведении, и выражение **x ⊕ y** принимает значение **TRUE** для данного кортежа **t**

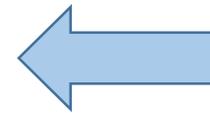
$$C = (A \text{ TIMES } B) \text{ WHERE } x \oplus y$$

$$C = A [x \oplus y] B$$

где **⊕** – один из операторов сравнения:

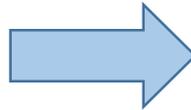
> >= < <= **=** <>

ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	4
Петров П.П.	1997	3
Сидоров С.С.	1999	2



Отношение А

Отношение В



ОЦЕНКА	ГРУППА	СПЕЦИАЛЬНОСТЬ
1	34И	ИВТ
2	32А	БИ
3	921	ФОИ
4	921	УК

Отношение С = (А TIMES В) WHERE СРЕДНИЙ БАЛЛ < ОЦЕНКА

ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ	ОЦЕНКА	ГРУППА	СПЕЦИАЛЬНОСТЬ
Петров П.П.	1997	3	4	921	УК
Сидоров С.С.	1999	2	3	921	ФОИ
Сидоров С.С.	1999	2	4	921	УК

Эквисоединение (equiJOIN, соединение по эквивалентности) – частный случай тета-соединения

Пусть отношения **A** и **B** удовлетворяют требованиям для декартового произведения (то есть, не имеют общих имен атрибутов). Пусть отношение **A** имеет атрибут **x**, отношение **B** – атрибут **y**.

Результат эквисоединения – отношение **C** с тем же заголовком, как и у декартова произведения **A** и **B**, и с телом, состоящим из множества всех кортежей **t**, таких, что **t** присутствует в этом декартовом произведении, и выражение **x = y** принимает значение **TRUE** для данного кортежа **t**

$$C = (A \text{ TIMES } B) \text{ WHERE } x = y$$

$$C = A [x = y] B$$

Эквисоединение (equiJOIN, соединение по эквивалентности)

Связывает таблицы по условию равенства, при котором сравниваются указанные столбцы каждой таблицы.

В результирующем выводе дублирующиеся столбцы не удаляются, условие (критерий) соединения должно быть явно определено.

Экви-соединение (соединение по равенству, или соединение по эквивалентности)

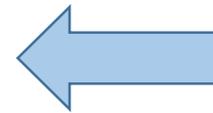
Пусть отношения **A** и **B** удовлетворяют требованиям для декартового произведения (то есть, не имеют общих имен атрибутов). Пусть отношение **A** имеет атрибут **x**, отношение **B** – атрибут **y**.

Если Θ представляет собой операцию проверки на равенство ($=$), то Θ -соединение называется экви-соединением.

Результат экви-соединения – отношение **C** с тем же заголовком, как и у декартова произведения **A** и **B**, и с телом, состоящим из множества всех кортежей **t**, таких, что **t** присутствует в этом декартовом произведении, и выражение $x = y$ принимает значение **TRUE** для данного кортежа **t**

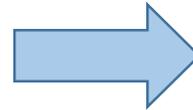
$$C = (A \text{ TIMES } B) \text{ WHERE } x = y$$

ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	4
Петров П.П.	1997	3
Сидоров С.С.	1999	5



Отношение А

Отношение В



ОЦЕНКА	ГРУППА	ДИСЦИПЛИНА
3	34И	История
5	32А	Алгебра
3	921	Геометрия
4	921	Физика

Отношение С = (А TIMES В) WHERE СРЕДНИЙ БАЛЛ = ОЦЕНКА

ФИО	ГОД РОЖДЕНИЯ	СРЕДНИЙ БАЛЛ	ОЦЕНКА	ГРУППА	ДИСЦИПЛИНА
Иванов И.И.	1998	4	4	921	Физика
Петров П.П.	1997	3	3	34И	История
Петров П.П.	1997	3	3	921	Геометрия
Сидоров С.С.	1999	5	5	32А	Алгебра

Ⓜ-соединение (тета-соединение)

В языке SQL: **INNER JOIN** (внутреннее соединение)

Используется, когда нужно включить все строки из обеих таблиц, удовлетворяющие условию объединения

Операция «соединение» в SQL

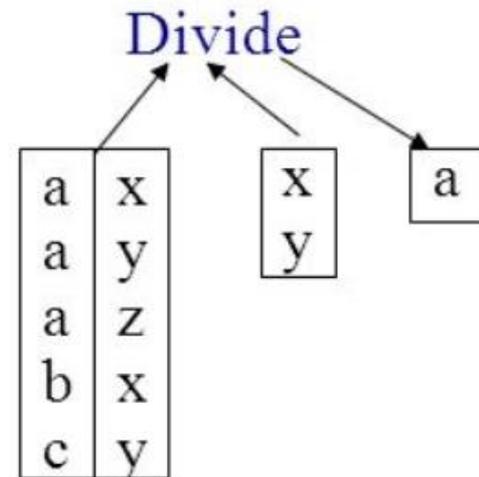
- INNER JOIN
- OUTER JOIN
 - LEFT OUTER JOIN
 - RIGHT OUTER JOIN
 - FULL OUTER JOIN
- CROSS JOIN

Деление

Применяется к отношениям, множества атрибутов которых перекрываются, причем атрибуты второго отношения **B** являются подмножеством атрибутов первого отношения **A**.

Оператор возвращает «укороченные» кортежи первого отношения **A** (без перекрывающихся атрибутов), которые образуют кортежи со всеми значениями атрибутов второго (меньшего) отношения **B**.

A DIVIDE B



Отношение А

ОЦЕНКА	ДИСЦИПЛИНА
4	Физика
3	История
3	Геометрия
5	Алгебра
4	История
5	Геометрия
3	Алгебра
4	Алгебра
4	Геометрия
3	Физика

Отношение В

ДИСЦИПЛИНА
История
Алгебра
Геометрия

Отношение В1

ДИСЦИПЛИНА
Физика
История

Отношение
C = A DIVIDE B

ОЦЕНКА
3
4
5

Отношение
C = A DIVIDE B1

ОЦЕНКА
4
3

Список литературы

1. Коннолли, Томас; Бегг, Каролин. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. 3-е издание. : Пер. с англ. — М. : Издательский дом "Вильямс", 2017. — 1440 с. : ил.
2. Дейт, К. Д. Введение в системы баз данных : Пер. с англ. / К. Д. Дейт. - 8-е изд. - Москва : Диалектика, 2019. - 1328 с. : ил.