

# Теория реляционных баз данных

# Литература

- Мейер Д. Теория реляционных баз данных. М.:«Мир», 1987
- Дейт К.Дж. Введение в системы баз данных. , 6-е изд.:Пер. с англ.. ,К.; СПб.:Издательский дом «Вильямс», 2000.
- Кодд Э.
- Джексон Г. Проектирование реляционных баз данных для использования с микроЭВМ, М.:«Мир», 1991
- Хансен Г., Хансен Дж. Базы данных. Разработка и управление. - Издательство Бином
- **Мишенин А.И. Теория экономических информационных систем. М.:ФиС, 2005.**

# Основные понятия



# Реляционная модель.

## Информационные единицы.

- База данных
- Отношение → таблица
- Запись (строка, ряд, запись, row, кортеж)
- Атрибут (поле)

Домен – множество значений атрибута

# Реляционная модель данных

- **Реляционная база данных** - совокупность взаимосвязанных плоских таблиц
- **Особенности реляционной модели:**
  - Простая линейная структура записи
  - Связи между таблицами устанавливаются динамически, в момент выполнения запроса по равенству значений полей связи
  - ЯМД – теоретико-множественный

# Влияние особенностей модели на проектирование

- Должны быть устранены все составные единицы информации
- Поля связи должны иметь соответствующие друг другу типы данных, одинаковые длины. Совпадение имен не обязательно, но желательно



# Ключи

- **Ключ** - атрибут или совокупность атрибутов однозначно идентифицирующих строку отношения
- **Ключ**, состоящий из одного атрибута, называется **простым**.
- **Ключ**, состоящий из нескольких атрибутов, называется **составным**.

# Свойства ключа

- Уникальность
- Неизбыточность
- Не может содержать пустых значений



# Ключи

- Все атрибуты, входящие в ключ, называются **ключевыми атрибутами**.
- **Атрибуты**, не являющиеся частью ключа, называются **неключевыми**.
- На роль ключа в отношении могут претендовать несколько атрибутов (совокупностей атрибутов).
- В этом случае каждый из них называется **вероятным** (альтернативным, возможным, потенциальным) **ключом**.
- Если в отношении имеется несколько потенциальных ключей, необходимо выделить один из них в качестве **первичного ключа**.

Таб_номер	фио	Дата_рождения	. . .	
<hr/>				

Код_сотрудника	Язык	Степень_владения
_____	_____	



Код_постав щика	Код_продук ции	Дата_поста ки	количество	
_____	_____	_____		

Код кафедры	Наименова ние кафедры полное	Наименова ние кафедры краткое
_____	_ _ _ _ _	_ _ _

# Расписание занятий

Код_сотрудника	Код_предмета	Группа	День недели	время
_____			_____	_____
		_____	_____	_____

Вероятные составные ключи

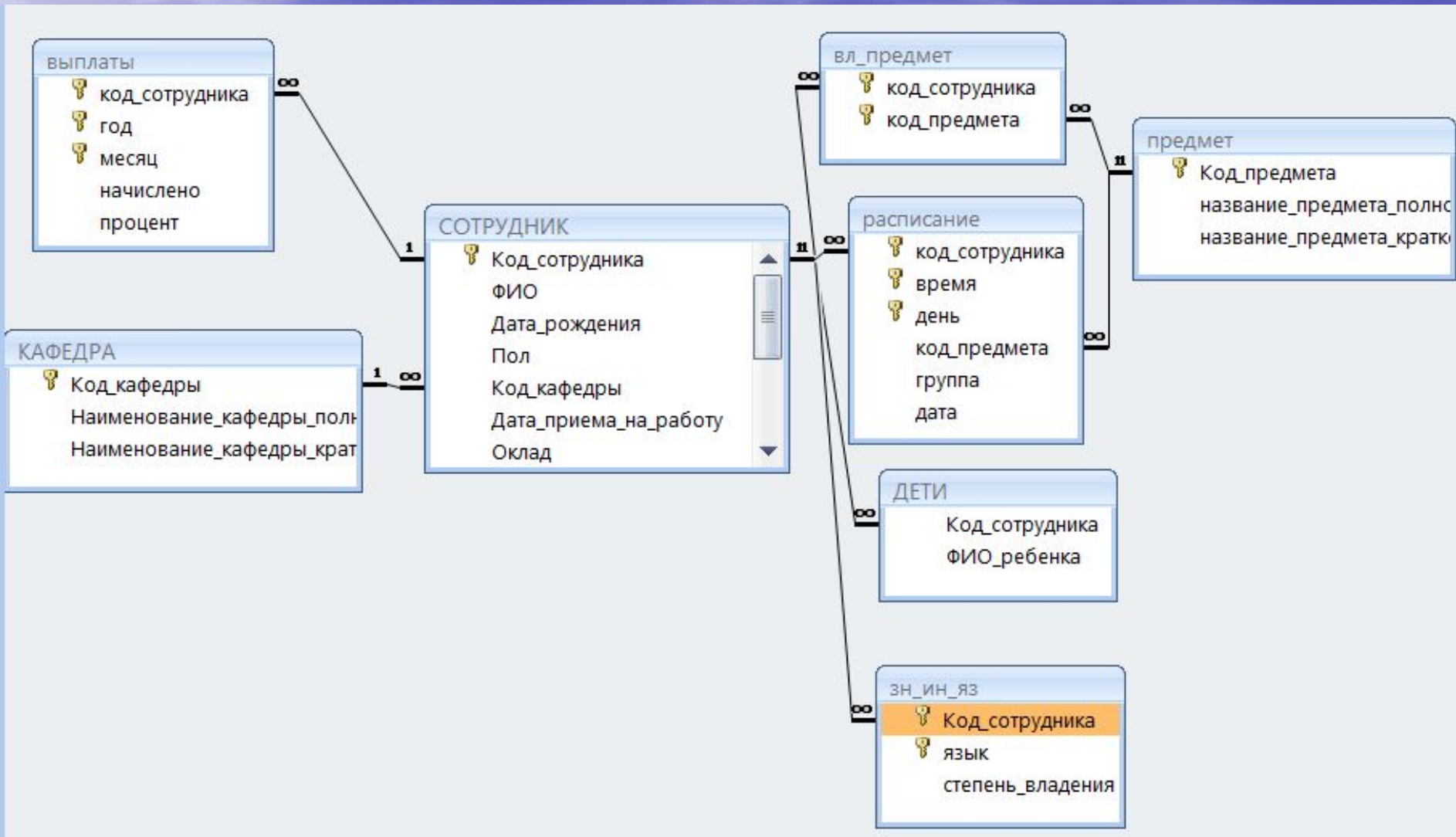


# Факторы, влияющие на выбор первичного ключа


Будут рассмотрены при изложении  
алгоритма проектирования

# Внешний ключ

- Атрибут (совокупность атрибутов), который в данном отношении ключом не является (но может входить в состав составного ключа), а в другом отношении является первичным ключом, называется **ВНЕШНИМ КЛЮЧОМ**. (\* при связи между таблицами 1:1 может являться первичным ключом)
- Связь в реляционных базах данных устанавливается от ключа к внешнему ключу







*Функциональные  
зависимости*

# Понятие функциональной зависимости

A, B – атрибут или совокупность атрибутов

- **Функциональная зависимость (functional dependency)**  
B является функционально зависимым от A тогда и только тогда, когда каждому значению A соответствует одно и только одно значение B.

Обозначается:

A → B

Или

$F(A)=B$

- **Детерминант (determinant)** — атрибут, который определяет значения других атрибутов.  
Синоним – **определитель**.

# Функциональная зависимость.

## Пример 1

Таб_ном	ФАМИЛИЯ	ГОД_Р
09	ПЕТРОВ	1970
10	СМИРНОВ	1955
11	КЛЮЕВА	1988
12	ИВАНОВ	1946

Таб\_ном --> ФАМИЛИЯ

Таб\_ном --> ГОД\_Р



# Функциональная зависимость.

## Пример 2

Код_предприятия	Код_продукции	Дата	Количество
0111	255	11.02.07	500
0111	256	11.02.07	300
0112	256	11.02.07	700
0111	256	15.02.07	400

Код\_предприятия, Код\_продукции, Дата → Количество

- **Неключевые атрибуты таблицы функционально зависят от ключа**
- **Между атрибутами первичного ключа не может быть функциональных зависимостей**

# Взаимно-однозначное соответствие

(Пример из учебника Мишенин А.И. «Теория экономических информационных систем»)

Наименование предприятия	ИНН
ДИНАМО	77014
АТЭ	77036
МАНОМЕТР	77054

**Наименование предприятия <--> ИНН**

**\*\* утверждение было бы верно, если:**

- нет предприятий с одинаковыми названиями
- Нет одинаковых ИНН
- Ни то, ни другое утверждение не верно



# Взаимно-однозначное соответствие (пример 2)

Код_кафедр	Наименование_кафедры_полное	Наименование_кафедры_крат
+	1 Проектирования экономических информационных систем	ПЭИС
+	2 Иностранных языков	ИнЯз
+	3 Банковского дела	БД
+	4 Истории	ИСТ

Код\_кафедры ↔ Наименование\_кафедры\_полное

Код\_кафедры ↔ Наименование\_кафедры\_краткое

Наименование\_кафедры\_краткое ↔ Наименование\_кафедры\_полное

# Теория нормализации отношений

# Нормализация

**Нормализация**



**Приведение к первой  
нормальной форме  
(1 NF)**



**Приведение к более высокой  
нормальной форме  
(2,3,4,5 ... NF)**



# Первая нормальная форма (*1NF*)

Данные хранятся в плоской двухмерной таблице без:

- неповторяющихся СЕИ
- векторов
- повторяющихся групп.

Таблица находится в *первой нормальной форме (1NF)* тогда и только тогда, когда ни одна из ее строк не содержит в любом своем поле более одного значения и ни одно из ее ключевых полей не пусто.

- Приведение к 1 NF – представление данных в виде плоской двухмерной таблицы
- *Дальнейшая нормализация* – это разбиение таблицы на две или более, обладающих лучшими свойствами при включении, изменении и удалении данных.

**Приведение к 1NF.**

**«Универсальное отношение»**

Понятие **«Универсальное отношение»**

- все атрибуты записываются в одной  
таблице.

Чаще используется как теоретическая  
основа



# Пример документа

Ведомость на выплату заработной платы			
Отдел _____			
Месяц _____		год _____	
№ п/п	ФИО	Сумма на руки	Подпись получателя
Итого по отделу			

# Отношение в 1NF

№ п/п	ФИО	Отдел	Месяц	Год	Сумма на руки	Итого по отделу
1	Иванов	АСУ	1	2007	15 000	100 000
2	Сидоров	АСУ	1	2007	10 000	100 000
1	Иванов	АСУ	2	2007	15 000	100 000

# Отношение в 1NF

<b>Таб_ном</b>	<b>ФИО</b>	<b>Отдел</b>	<b>Месяц</b>	<b>Год</b>	<b>Сумма на руки</b>	<b>Итого по отделу</b>
<b>1</b>	<b>Иванов</b>	<b>АСУ</b>	<b>1</b>	<b>2007</b>	<b>15 000</b>	<b>100 000</b>
<b>2</b>	<b>Сидоров</b>	<b>АСУ</b>	<b>1</b>	<b>2007</b>	<b>10 000</b>	<b>100 000</b>
<b>1</b>	<b>Иванов</b>	<b>АСУ</b>	<b>2</b>	<b>2007</b>	<b>15 000</b>	<b>100 000</b>



# *Недостатки первой нормальной формы (1NF)*

- Аномалии по вставке
- Аномалии по корректировке
- Дублирование данных

# Вторая нормальная форма (2NF)

- Отношение находится во **второй нормальной форме**, если оно соответствует первой нормальной форме, и все неключевые атрибуты функционально полно зависят от первичного ключа.
- Атрибут **функционально полно** зависит от ключа, если он функционально зависит от всего ключа, но не зависит от любой его части

# Функциональные зависимости отношения

- Таб\_ном, месяц, год → сумма на руки
- Таб\_ном → Фамилия
- Таб\_ном → Отдел
- Отдел, месяц, год → итого по отделу



# Отношение в 2NF

Таб_ном	ФИО	Отдел
1	Иванов	АСУ
2	Сидоров	АСУ

Таб_ном	М-ц	Год	Сумма на руки
1	1	2007	15 000
2	1	2007	10 000
1	2	2007	15 000

# Отношение в 2NF (продолжение)

<b>Отдел</b> ——	<b>М-ц</b> ——	<b>Год</b> ——	<b>Итого по отделу</b>
<b>АСУ</b>	<b>1</b>	<b>2007</b>	<b>100 000</b>
<b>АСУ</b>	<b>2</b>	<b>2007</b>	<b>100 000</b>
<b>. . .</b>			

# Отношение в 2NF (пример 2 – расширена – не является в 3NF )

Таб_ном —	ФИО	Отдел	Руководитель отдела
1	Иванов	АСУ	Петров
2	Сидоров	АСУ	Петров



# *Недостатки отношений 2NF*

# Третья нормальная форма (3NF)

Отношение находится в **третьей нормальной форме**, если оно соответствует второй нормальной форме, и в нем не существует транзитивных зависимостей.

( $A \rightarrow B$  и  $B \rightarrow C$ , поэтому  $A \rightarrow C$ )

# Отношение в 3NF (пример)

<u>Таб_ном</u>	ФИО	Отдел
1	Иванов	АСУ
2	Сидоров	АСУ

<u>Отдел</u>	Руководитель отдела
АСУ	Петров
...	



# Нормальная форма Бойса-Кодда

Отношение соответствует *нормальной форме Бойса-Кодда*, если оно соответствует третьей нормальной форме, и все определители являются кандидатами на использование в качестве ключа.

# Четвертая нормальная форма

- **Отношение находится в четвертой нормальной форме**, если оно соответствует нормальной форме Бойса-Кодда, и в ней нет **многозначных зависимостей**.
- Атрибут А **многозначно определяет** атрибут В, если для каждого значения атрибута А существует хорошо определенное множество соответствующих значений В.

# Многозначные зависимости (multivalued dependency)

- **Многозначная зависимости** существует, если каждому значению атрибута  $A$  соответствует конечное множество значений атрибута  $B$ , связанных с  $A$ , и конечное ; множество значений атрибута  $C$ , также связанных с  $A$ . Атрибуты  $B$  и  $C$  друг от друга не зависят.
- $A \twoheadrightarrow B, A \twoheadrightarrow C$



## иллюстрация многозначных зависимостей

<b>Дисциплина</b>	<b>Преподаватель</b>	<b>Учебник</b>
Информатика	Шипилов П.А.	Форсайт Р. Паскаль для всех
Информатика	Шипилов П.А.	Уэйт М. и др. Язык Си
Информатика	Голованевский Г.Л.	Форсайт Р. Паскаль для всех
Информатика	Голованевский Г.Л.	Уэйт М. и др. Язык Си
...	...	...

Дисциплина -> Преподаватель

Дисциплина -> Учебник

# Отношения в 4 NF

Дисциплина	Преподаватель
Информатика	Шипилов П.А.
Информатика	Голованевский Г.Л.

Дисциплина	Учебник
Информатика	Форсайт Р. Паскаль для всех
Информатика	Уэйт М. и др. Язык Си
...	...

# Правила вывода



# Правила вывода

**Аксиомы** (правила, теоремы) **вывода** – правила, устанавливающие, что если некоторое отношение удовлетворяет некоторым  $F$ -зависимостям, то оно должно удовлетворять и некоторым другим  $F$ -зависимостям.

# Правила вывода

1.  $A, B \rightarrow A$  и  $A, B \rightarrow B$
2. Если  $A \rightarrow B$  и  $A \rightarrow C$  то  $A \rightarrow BC$
3. Если  $A \rightarrow B$  и  $B \rightarrow C$  то  $A \rightarrow C$
4. Если  $A \rightarrow B$  то  $AC \rightarrow B$
5. Если  $A \rightarrow B$  и  $BC \rightarrow D$  то  $AC \rightarrow D$

# Алгоритм нормализации



# Алгоритм нормализации

## Шаг 1.

Получение исходного множества функциональных зависимостей.

Рассматриваются все сочетания атрибутов (1,2,3, .....n).

Не рассматриваются варианты, которые являются следствием теорем о функциональных зависимостях.

Шаг 2. Поиск *минимального покрытия* функциональных зависимостей: множество, из которого удалены зависимости, являющиеся следствием оставшихся зависимостей.

$F = \{f_1, f_2, \dots, f_n\}$

Шаг 3. Для каждого  $f_i$  создать отношение

Шаг 4. Если первичный ключ исходного отношения не вошел ни в одну проекцию, то создать дополнительное отношение, содержащее этот ключ

Примечание:

Для взаимно однозначных зависимостей принято выделять «старший» атрибут, который затем представляет все атрибуты взаимно однозначного соответствия.

# Рекомендация

При проведении нормализации таблиц, в которые введены заменители составных первичных ключей (искусственные идентификаторы), следует хотя бы мысленно поменять их на исходные ключи, а после окончания нормализации снова восстанавливать.

# Недостатки нормализации

- Совместная обработка связанных таблиц может существенно замедлить обработку.
- Понятие «денормализация»



# Реляционная алгебра

# Реляционная алгебра

- Язык процедурного типа
- Операндами являются отношения
- Результатом является отношение

# Операция Проекция

- Унарная операция
- $T = R[X]$ ,

Где  $R$  – исходное отношение

$T$  – результирующее отношение

$X$  – список атрибутов, входящих в результирующее отношение. Является подмножеством атрибутов исходного отношения.



# Операция Проекция. Пример (абстрактный).

R

A	B	C
$a_1$	$b_1$	$c_1$
$a_1$	$b_1$	$c_2$
$a_2$	$b_2$	$c_2$
$a_2$	$b_3$	$c_3$

T=R[A,B]

A	B
$a_1$	$b_1$
$a_2$	$b_2$
$a_2$	$b_3$

# Операция Проекция. Пример 2.

Поставщик	Продукция	Дата	Количество
<b>31</b>	<b>П1</b>	<b>21.010.07</b>	<b>100</b>
<b>32</b>	<b>П1</b>	<b>21.010.07</b>	<b>120</b>
<b>31</b>	<b>П2</b>	<b>22.010.07</b>	<b>200</b>
<b>32</b>	<b>П2</b>	<b>21.010.07</b>	<b>150</b>

Поставщик
<b>31</b>
<b>32</b>

# Операция Проекция.

## Пример 3.

Поставщик	Продукция	Дата	Количество
31	П1	21.010.07	100
32	П1	21.010.07	120
31	П2	22.010.07	200
32	П2	21.010.07	150

Продукция	Количество
П1	100
П1	120
П2	200
П2	150

Операция нежелательна



# Операция Выборка

- $T = R[p]$ ,

Где R – исходное отношение

T – результирующее отношение

p – Условие выборки

Условие выборки:

- ИМЯ\_АТТРИБУТА <знак сравнения> ЗНАЧЕНИЕ
- ИМЯ\_АТТРИБУТА <знак сравнения> ИМЯ\_АТТРИБУТА

Условия выборки могут быть сложными

# Операция Выборка. Пример (абстрактный).

R

A	B	C
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>2</sub>
a <sub>2</sub>	b <sub>2</sub>	c <sub>2</sub>
a <sub>2</sub>	b <sub>3</sub>	c <sub>3</sub>

T=R[C= c<sub>1</sub>]

A	B	C
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>

# Операция Выборка.

## Пример 2

Предмет	Ном_зачетки	Дата	Оценка
<b>БД</b>	<b>07321</b>	<b>09.01.07</b>	<b>5</b>
<b>БД</b>	<b>07322</b>	<b>09.01.07</b>	<b>4</b>
<b>МО</b>	<b>07321</b>	<b>19.01.07</b>	<b>5</b>
<b>МО</b>	<b>07322</b>	<b>19.01.07</b>	<b>3</b>

**T=R[Предмет = БД]**



# Операция объединения

$$T = R1 \cup R2$$

R1

<b>A</b>	<b>B</b>
a <sub>1</sub>	b <sub>1</sub>
a <sub>2</sub>	b <sub>2</sub>
a <sub>2</sub>	b <sub>3</sub>

R2

<b>A</b>	<b>B</b>
a <sub>1</sub>	b <sub>1</sub>
a <sub>1</sub>	b <sub>2</sub>
a <sub>2</sub>	b <sub>3</sub>

T

<b>A</b>	<b>B</b>
a <sub>1</sub>	b <sub>1</sub>
a <sub>1</sub>	b <sub>2</sub>
a <sub>2</sub>	b <sub>2</sub>
a <sub>2</sub>	b <sub>3</sub>

# Операция объединения. Пример

**Сотрудники**

Таб_ном	ФАМИЛИЯ
09	ПЕТРОВ
10	СМИРНОВ
11	КЛЮЕВА

**Студенты**

Таб_ном	ФАМИЛИЯ
11	КЛЮЕВА
12	ИВАНОВ
....	

**Кадры**

Таб_ном	ФАМИЛИЯ
09	ПЕТРОВ
10	СМИРНОВ
11	КЛЮЕВА
12	ИВАНОВ

# Операция Пересечения

$$T = R1 \wedge R2$$

R1

<b>A</b>	<b>B</b>
a <sub>1</sub>	b <sub>1</sub>
a <sub>2</sub>	b <sub>2</sub>
a <sub>2</sub>	b <sub>3</sub>

R2

<b>A</b>	<b>B</b>
a <sub>1</sub>	b <sub>1</sub>
a <sub>1</sub>	b <sub>2</sub>
a <sub>2</sub>	b <sub>3</sub>

T

<b>A</b>	<b>B</b>
a <sub>1</sub>	b <sub>1</sub>
a <sub>2</sub>	b <sub>3</sub>



# Операция Пересечения. Пример

**Сотрудники**

Таб_ном	ФАМИЛИЯ
09	ПЕТРОВ
10	СМИРНОВ
11	КЛЮЕВА

**Студенты**

Таб_ном	ФАМИЛИЯ
11	КЛЮЕВА
12	ИВАНОВ
....	

**Студенты-Сотрудники**

Таб_ном	ФАМИЛИЯ
11	КЛЮЕВА

# Операция Вычитания

$$T = R1 \setminus R2$$

R1

<b>A</b>	<b>B</b>
a <sub>1</sub>	b <sub>1</sub>
a <sub>2</sub>	b <sub>2</sub>
a <sub>2</sub>	b <sub>3</sub>

R2

<b>A</b>	<b>B</b>
a <sub>1</sub>	b <sub>1</sub>
a <sub>1</sub>	b <sub>2</sub>
a <sub>2</sub>	b <sub>3</sub>

T

<b>A</b>	<b>B</b>
a <sub>2</sub>	b <sub>2</sub>

# Операция Вычитания . Пример

**Сотрудники**

Таб_ном	ФАМИЛИЯ
09	ПЕТРОВ
10	СМИРНОВ
11	КЛЮЕВА

**Студенты**

Таб_ном	ФАМИЛИЯ
11	КЛЮЕВА
12	ИВАНОВ
....	

**Сотрудники «не студенты»**

Таб_ном	ФАМИЛИЯ
09	ПЕТРОВ
10	СМИРНОВ



# Операция Соединения

$$T=R1[p]R2,$$

где  $p$  – условие соединения

R1

A	B
a <sub>1</sub>	b <sub>1</sub>
a <sub>2</sub>	b <sub>2</sub>
a <sub>2</sub>	b <sub>3</sub>

R2

A	C
a <sub>1</sub>	c <sub>1</sub>
a <sub>1</sub>	c <sub>2</sub>
a <sub>2</sub>	c <sub>3</sub>

T

A	B	C
a <sub>1</sub>	b <sub>1</sub>	c <sub>1</sub>
a <sub>1</sub>	b <sub>1</sub>	c <sub>2</sub>
a <sub>2</sub>	b <sub>2</sub>	c <sub>3</sub>
a <sub>2</sub>	b <sub>3</sub>	c <sub>3</sub>

# Операция Соединения. Пример

Сотрудники

Зн\_ин\_яз

Таб_ном	ФАМИЛИЯ
09	ПЕТРОВ
10	СМИРНОВ
11	КЛЮЕВА

Таб_ном	Язык
11	английский
10	английский
10	немецкий

Таб_ном	ФАМИЛИЯ	Язык
10	СМИРНОВ	английский
10	СМИРНОВ	немецкий
11	КЛЮЕВА	английский

# Операция Соединения

- В «условии соединения» может использоваться любой знак сравнения
- Чаще всего используется знак «=».
- Такое соединение называется натуральным.
- В ЯМД реляционных СУБД включены разновидности Соединения : внутреннее, левое, правое и др.



# Операции реляционной алгебры (сводная диаграмма)

