

*Базы данных*



# НОРМАЛИЗАЦИЯ БД

*Кафедра САПРиУ*

*Подготовил: Плonsкий В.Ю.*

## *План*

- **Функциональные зависимости**
  - Нормализация отношений
  - Первая нормальная форма
  - Вторая нормальная форма
  - Третья нормальная форма
  - Нормальная форма Бойса-Кодда
- **Многозначные зависимости**
  - Четвертая нормальная форма
- **Зависимости соединения**
  - Пятая нормальная форма

## *Проектирование при помощи нормализации*

**Нормализация отношений** — это **пошаговый обратимый процесс** декомпозиции исходных отношений в отношения, обладающие **лучшими свойствами** при включении, изменении и удалении данных.

### *«Побочный эффект»*

При нормализации выявляются все возможные функциональные зависимости и назначаются ключи отношениям.

## *Нормальные формы*

- Первая нормальная форма
- Вторая нормальная форма
- Третья нормальная форма
- Нормальная форма Бойса-Кодда

(1НФ)	
(2НФ)	<b><i>FD</i></b>
(3НФ)	
(НФБК)	

- Четвертая нормальная форма
- Пятая нормальная форма  
(НФ проекции-соединения)

(4НФ)	<b><i>MVD</i></b>
-------	-------------------

(5НФ)	<b><i>JD</i></b>
-------	------------------

***MVD*** ∈ ***JD***

- Доменно-ключевая нормальная форма (6НФ)

# Нормальные формы

**НЕТ**  
зависимостей  
соединения

**НЕТ**  
многозначных  
зависимостей

**НЕТ**  
транзитивных  
зависимостей

**НЕТ**  
частичных  
зависимостей

**НЕТ**  
повторяющихся  
групп



**1НФ**

**2НФ**

**3НФ / НФБК**

**4НФ**

**5НФ**

# *Определение 1-й нормальной формы*

Реляционная модель поддерживает только *нормализованные* отношения.



Отношение находится в *1НФ*, если:

- отсутствуют дублирующиеся строки;
- в каждом столбце данные одного типа;
- в каждой ячейке хранится атомарное значение

# Примеры



П.1

Код сотрудника	ФИО	Номер телефона
1	Иванов И. И.	283-56-82, 390-57-34
2	Петров П. П.	708-62-34



П.2

Номер	Автомобиль	Год	Стоимость	Характеристики
C065MK178 RUS	Toyota Avalon	2020	4720000	Автомат, бензин, 3.5 л, 8 ст.
B095TM177 RUS	Ford Expedition	2021	6810000	Автомат, бензин, 3.5 л, 10 ст.
T325PC197 RUS	Lada Granta	2019	595000	МКПП, бензин, 1.6 л, 5 ст.



# Примеры

## Пример №1



Код сотрудника	ФИО	Номер телефона
1	Иванов И. И.	283-56-82, 390-57-34
2	Петров П. П.	708-62-34



Код сотрудника	ФИО	Номер телефона 1	Номер телефона 2	Номер телефона 3
1	Иванов И. И.	283-56-82	390-57-34	NULL
2	Петров П. П.	708-62-34	NULL	NULL





# Примеры

## Пример №1

Код сотрудника	ФИО
1	Иванов И. И.
2	Петров П. П.

Номер телефона	Код сотрудника
283-56-82	1
390-57-34	1
708-62-34	2

ИЛИ

---

Код сотрудника	Номер телефона	ФИО
1	283-56-82	Иванов И. И.
1	390-57-34	Иванов И.И.
2	708-62-34	Петров П. П.

# Примеры

## Пример №2



Номер	Автомобиль	Год	Стоимость	Характеристики
C065MK178 RUS	Toyota Avalon	2020	4720000	Автомат, бензин, 3.5 л, 8 ст.
B095TM177 RUS	Ford Expedition	2021	6810000	Автомат, бензин, 3.5 л, 10 ст.
T325PC197 RUS	Lada Granta	2019	595000	МКПП, бензин, 1.6 л, 5 ст.



Серия	Ном	Регион	Марка	Модель	Год	Стоимость	Коробка	Топливо	Объем	Ступ
СМК	065	178	Toyota	Avalon	2020	4720000	Автомат	Бензин	3.5	8
ВТМ	095	177	Ford	Expedition	2021	6810000	Автомат	Бензин	3.5	10
ТРС	326	197	Lada	Granta	2019	595000	МКПП	Бензин	1.6	5

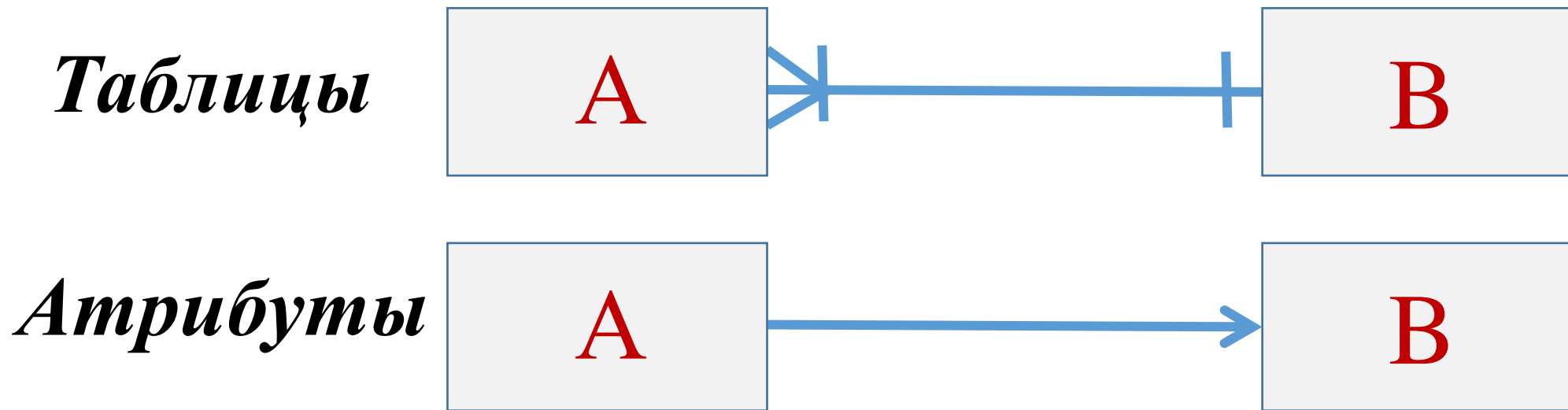
# Универсальное отношение



проект_номер	проект_наименование	сотр_номер	сотр_имя	спец	часовая_ставка	количество_часов
15	Alpha	101	Сергей Иванов	Программист	200	120
15	Alpha	102	Андрей Петров	Программист	200	100
15	Alpha	110	Антон Смирнов	Аналитик	300	40
18	Beta	103	Федор Антонов	Программист	200	250
18	Beta	102	Андрей Петров	Программист	200	280
18	Beta	111	Петр Семенов	Разработчик БД	250	80
22	Delta	104	Сергей Федотов	Программист	200	180
22	Delta	105	Иван Андреев	Программист	200	150
22	Delta	110	Антон Смирнов	Аналитик	300	60

# Функциональная зависимость

- **Функциональная зависимость** (Functional Dependency, FD) – связь многие-к-одному между множествами значений атрибутов **внутри** данного отношения.



Наличие определенных ФЗ между атрибутами отношения является **отражением семантики предметной области** и не может быть определено исходя из конкретных значений атрибутов в некоторый момент времени.

## Функциональная зависимость (ФЗ)

- **Функциональная зависимость** (functional dependency, FD) для отношения  $R$  – это ограничение вида:  
Если два кортежа отношения  $R$  совпадают на атрибутах  $A_1, A_2, \dots, A_n$ , то они должны совпадать и в другом атрибуте  $B$ .

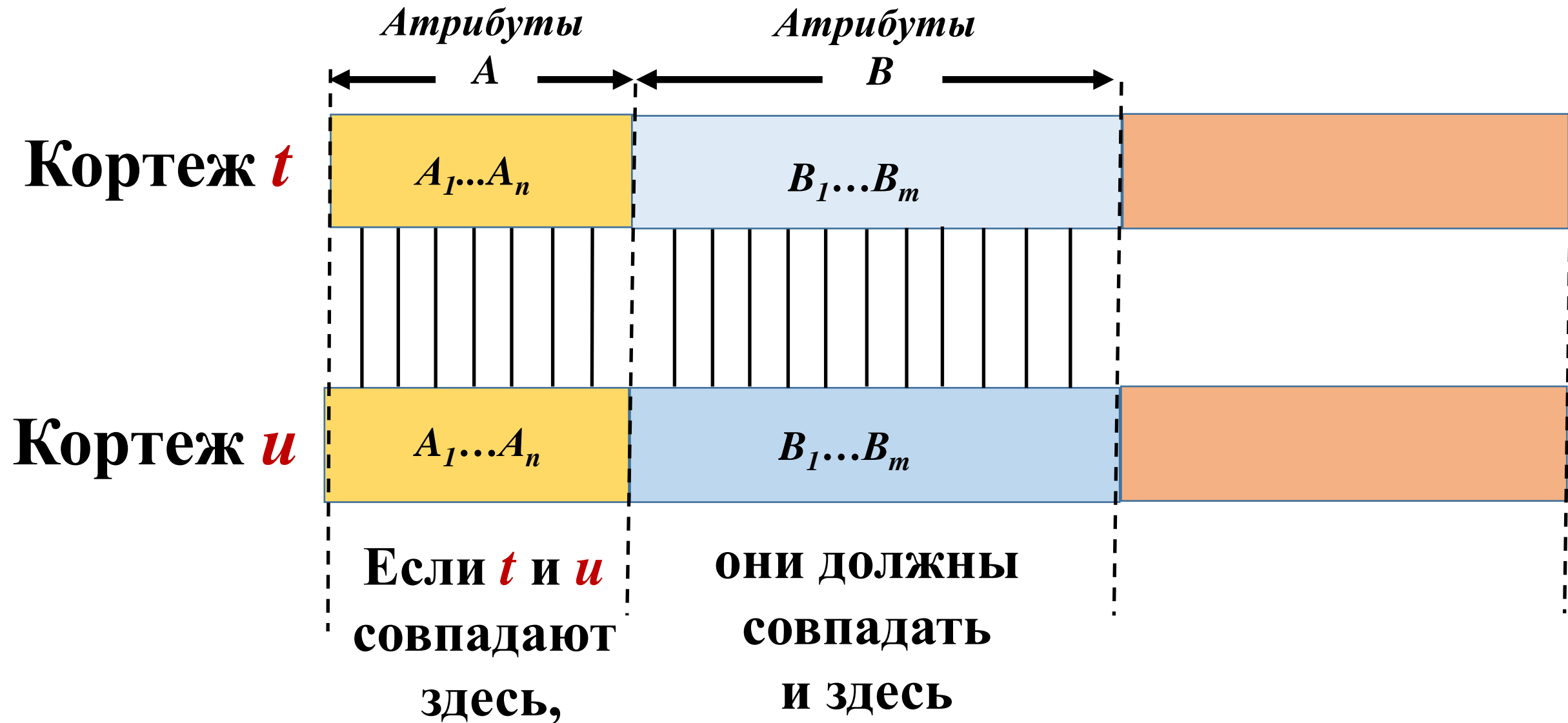
Формальная запись FD:

$$A_1 A_2 \dots A_n \rightarrow B$$

$A_1 A_2 \dots A_n$  функционально определяет  $B$ .

Левая часть – детерминант, правая часть – зависимая.

# Функциональная зависимость (графическое представление)



## Функциональная зависимость

- Выражение  $A_1 A_2 \dots A_n \rightarrow B$  называют *функциональной зависимостью*, так как оно определяет **функцию**, которая в качестве параметров принимает список значений, по одному на каждый  $A_1, A_2, \dots, A_n$  и возвращает строго определенное значение (либо не возвращает никакого значения) для атрибута  $B$ .

ВЫЧИСЛЕНИЕ



ПОИСК



ФЗ является свойством схемы, а не отдельного экземпляра отношения, и отражает семантику предметной области.



## *Ключи и функциональная зависимость*

- Множество вида  $\{A_1, A_2, \dots, A_n\}$ , состоящее из одного или нескольких атрибутов, является **потенциальным ключом отношения**, если выполняются следующие условия:
- Атрибуты  $\{A_1, A_2, \dots, A_n\}$  функционально обуславливают все остальные атрибуты (ситуация, когда два различных кортежа совпадают во всех атрибутах невозможна).
- Ни одно из допустимых подмножеств атрибутов множества  $\{A_1, A_2, \dots, A_n\}$  не является **функциональным обоснованием** всех остальных атрибутов отношения (ключ минимален).

# Примеры ФЗ

Отношение **ФИЛЬМЫ**

title	year	length	FilmType	StudioName	StarName
Звёздные войны	1977	124	color	Fox	Кэрри Фишер
Звёздные войны	1977	124	color	Fox	Марк Хэмилл
Звёздные войны	1977	124	color	Fox	Харрисон Форд
...	...	...	...	...	...

$title\ year \rightarrow length$

$title\ year \rightarrow StudioName$

$title\ year \rightarrow FilmType$

Сокращенная запись:

$title\ year \rightarrow length, FilmType, StudioName$

$title\ year \not\rightarrow StarName$

$title\ StarName \not\rightarrow year$

$year\ StarName \not\rightarrow title$

# Примеры ФЗ

ЭКЗАМЕН =

КОД_СТУД	ИМЯ_СТУД	ДИСЦИПЛИНА	ОЦЕНКА
C2	Иванов	Физика	5
C2	Иванов	Математика	4
C2	Иванов	История	4
C2	Иванов	Информатика	5
C6	Смирнов	Физика	3
C6	Смирнов	Математика	4
C6	Смирнов	Информатика	3
C9	Иванов	Физика	5
C9	Иванов	Информатика	4

**Ключ** (составной) –  
{КОД\_СТУДЕНТА,  
ДИСЦИПЛИНА}

Некоторые зависимости  
могут быть **выведены** из  
других функциональных  
зависимостей.

*Код\_студ Дисциплина → Имя\_студ*  
*Код\_студ Дисциплина → Оценка*  
*Код\_студ Дисциплина → Имя\_студ Оценка*

*Код\_студ Дисциплина Имя\_студ → Оценка*

*Код\_студ Дисциплина → Код\_студ*

*Код\_студ Дисциплина Имя\_студ → Дисциплина*

*Код\_студ Дисциплина → Имя\_студ*  
*Код\_студ Дисциплина → Оценка*



*Код\_студ Дисциплина → Имя\_студ Оценка*

## Тривиальные зависимости

- ФЗ называют **тривиальной**, когда **правая (зависимая часть)** символической записи данной зависимости является **подмножеством ее левой части** (детерминанта).

### Примеры

*Код\_студ Дисциплина → Код\_студ*

*Код\_студ Дисциплина Имя\_студ → Дисциплина*

### Не полностью нетривиальная зависимость

*title year → length year*



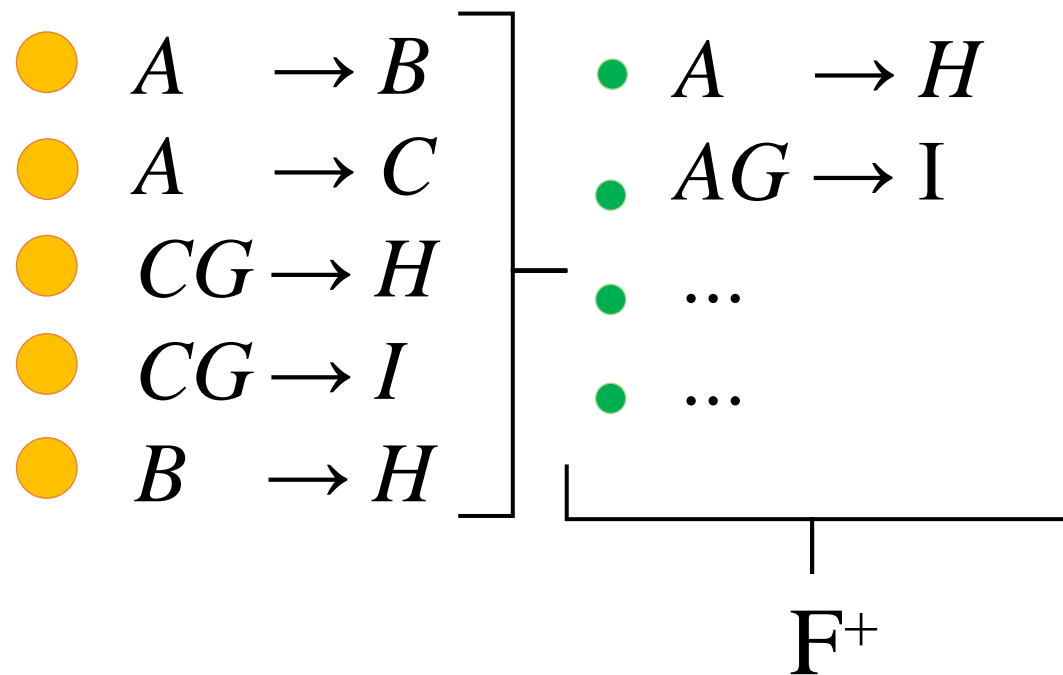
Можно изъять из **правой** части ФЗ те атрибуты, которые присутствуют в левой части.

*title year → length*

# Замыкание функциональных зависимостей

$R(A, B, C, G, H, I)$  – схема отношения.

Функциональные  
зависимости  $F$ :



Замыкание  $F^+$  – всех набор ФЗ, логически выводимых из  $F$ .

# Правила вывода зависимостей

**Аксиомы Армстронга** — позволяют из данного множества зависимостей вывести другие ФЗ.

- *Рефлексивность*: если  $B \subseteq A$ , то  $A \rightarrow B$
- *Приращение*: если  $A \rightarrow B$ , то  $A C \rightarrow B C$
- *Транзитивность*: если  $A \rightarrow B$  и  $B \rightarrow C$ , то  $A \rightarrow C$

## *Дополнительные правила (следствия из аксиом):*

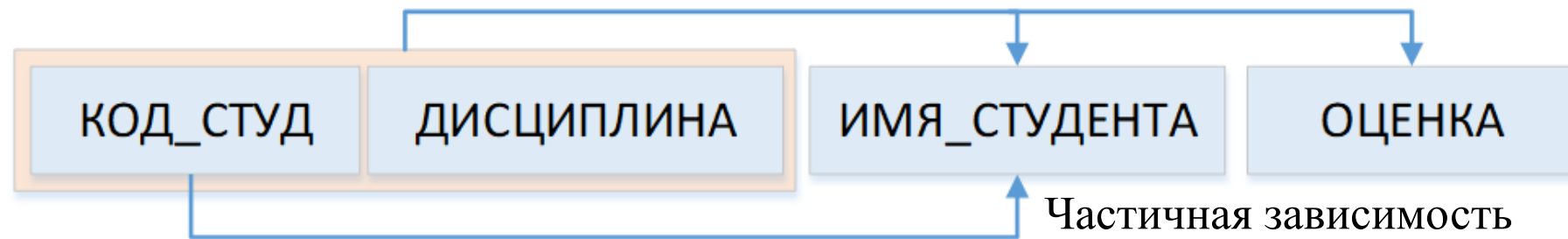
- Самоопределение:  $A \rightarrow A$
- Псевдотранзитивность: если  $A \rightarrow B$  и  $C B \rightarrow D$ , то  $A C \rightarrow D$
- Декомпозиция: если  $A \rightarrow B C$ , то  $A \rightarrow B$  и  $A \rightarrow C$
- Объединение: если  $A \rightarrow B$  и  $A \rightarrow C$ , то  $A \rightarrow B C$

# Предпосылки перехода к 2НФ

ЭКЗАМЕН =

ЭКЗАМЕН	
🔑	Код_студента
🔑	Дисциплина
	Имя_студента
	Оценка

КОД_СТУД	ДИСЦИПЛИНА	ИМЯ_СТУД	ОЦЕНКА
C6	Физика	Петров	3
C6	Математика	Петров	4
C6	Информатика	Петров	3
C2	Физика	Иванова	5
C2	Математика	Иванова	4
C2	История	Иванова	4
C2	Информатика	Иванова	5
C2	Иностр. язык	Иванова	5
C9	Иностр. язык	Попов	4
C1	История	Кузнецов	5
C1	Иностр. язык	Кузнецов	4
C8	Археология	Орлов	5







# *Аномалии обновления БД*

 **Аномалии вставки:** нельзя вставить одни данные, в связи с тем, что не определены другие.

---

 **Аномалии удаления:** при удалении одних данных можно потерять другие.

---



 **Аномалии модификации:** может возникнуть потеря целостности в связи с избыточностью данных и их частичным обновлением.



**ПРИЧИНА АНОМАЛИЙ:** в одном отношении хранится информация о нескольких сущностях.

# Аномалии при выполнении DML-операторов к 1НФ

- Операция **INSERT**: В отношение ЭКЗАМЕН нельзя вставить кортеж с данными о студенте (код и фамилия), если нет данных о сдаче студентом хотя бы одной дисциплины.
- Операция **DELETE**: При удалении информации, что конкретный студент сдавал какую-либо дисциплину, можно потерять данные о фамилии этого студента из-за того, что удаляемый кортеж был единственным, содержащим её.
- Операция **UPDATE**: Значение атрибута ИМЯ\_СТУДЕНТА повторяется в отношении многократно для каждого конкретного значения атрибута КОД\_СТУДЕНТА. Если для каких-либо кортежей такая коррекция не будет произведена, в отношении окажутся кортежи с противоречивыми данными.

ЭКЗАМЕН	
	Код_студента
	Дисциплина
	Имя_студента
	Оценка

## Неприводимая зависимость

Функциональная зависимость называется **неприводимой слева** (**функционально полной**), если ни один атрибут не может быть опущен из ее детерминанта без нарушения этой зависимости.

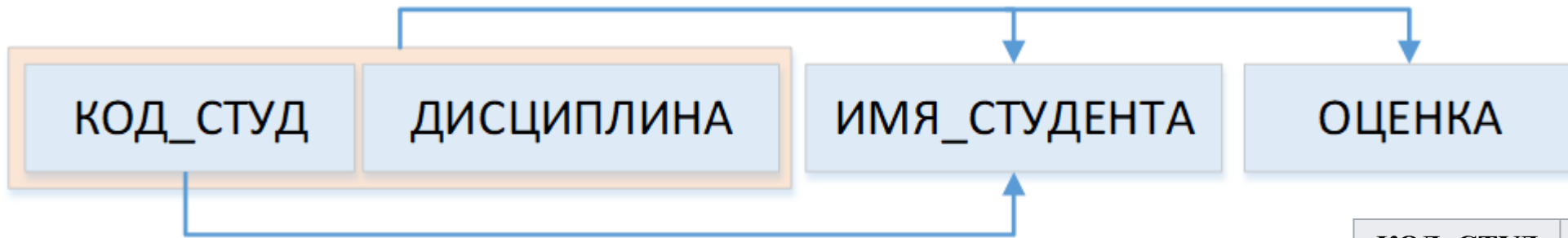
$$\text{Kod\_студ} \rightarrow \text{Имя\_студ}$$

**Приводимая** (**частичная**) функциональная зависимость, является следствием **неприводимой** ФЗ:

$$\text{Kod\_студ} \rightarrow \text{Имя\_студ} \Rightarrow \text{Kod\_студ} \text{ Дисциплина} \rightarrow \text{Имя\_студ}$$

Декомпозиция отношения ЭКЗАМЕН на два отношения СТУДЕНТ и УСПЕВАЕМОСТЬ приведет к исключению из отношения **приводимой** ФЗ атрибута от ключа отношения.

# Результат нормализации (2НФ)

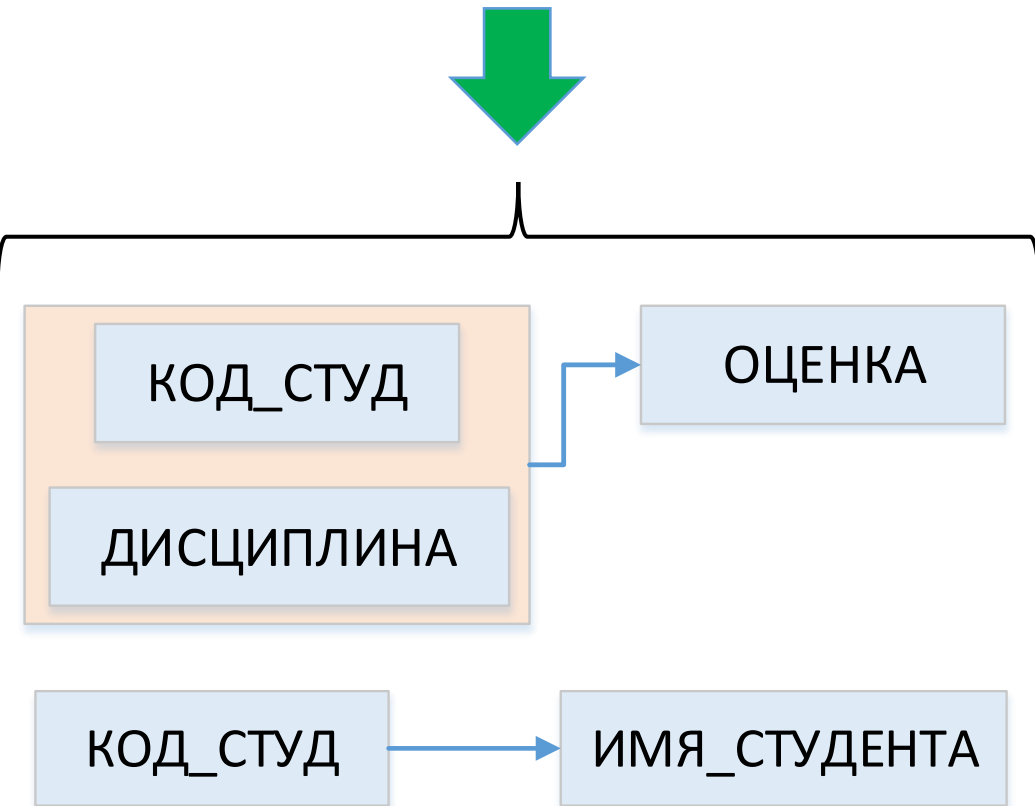


УСПЕВАЕМОСТЬ =

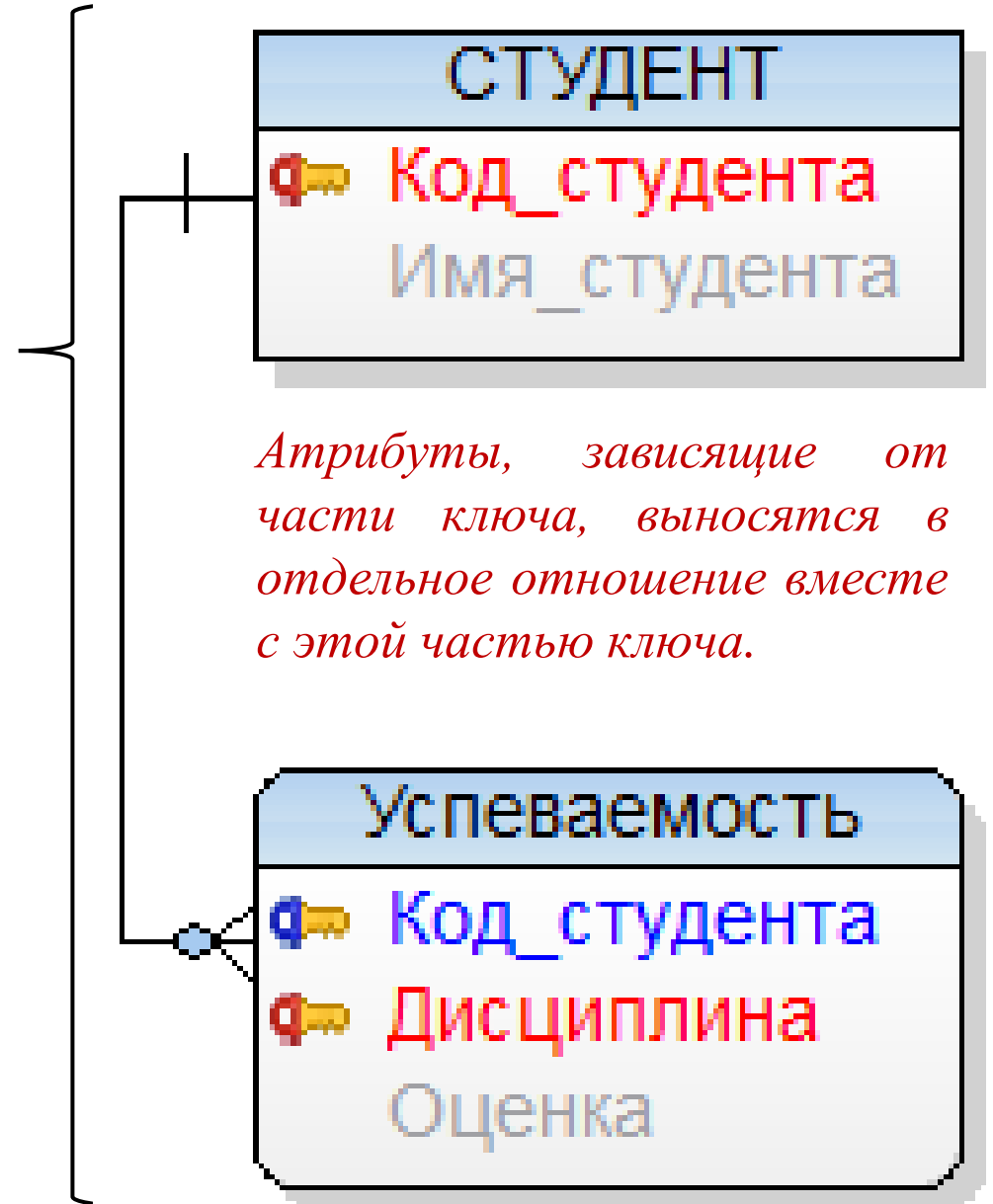
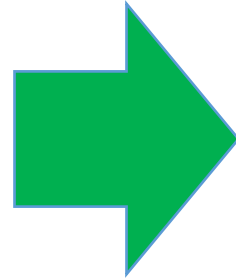
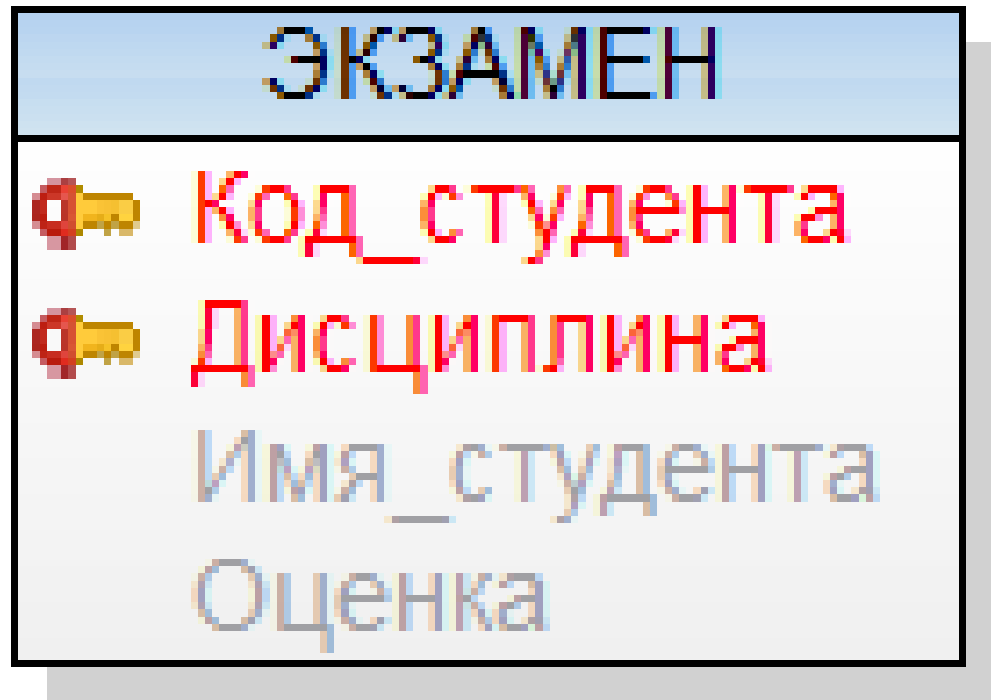
КОД_СТУД	ДИСЦИПЛИНА	ОЦЕНКА
С6	Физика	3
С6	Математика	4
С6	Информатика	3
С2	Физика	5
С2	Математика	4
С2	История	4
С2	Информатика	5
С2	Иностр. язык	5
С9	Иностр. язык	4
С1	История	5
С1	Иностр. язык	4
С8	Археология	5

СТУДЕНТ =

КОД_СТУД	ИМЯ_СТУД
С6	Петров
С2	Иванова
С9	Попов
С1	Кузнецов
С8	Орлов



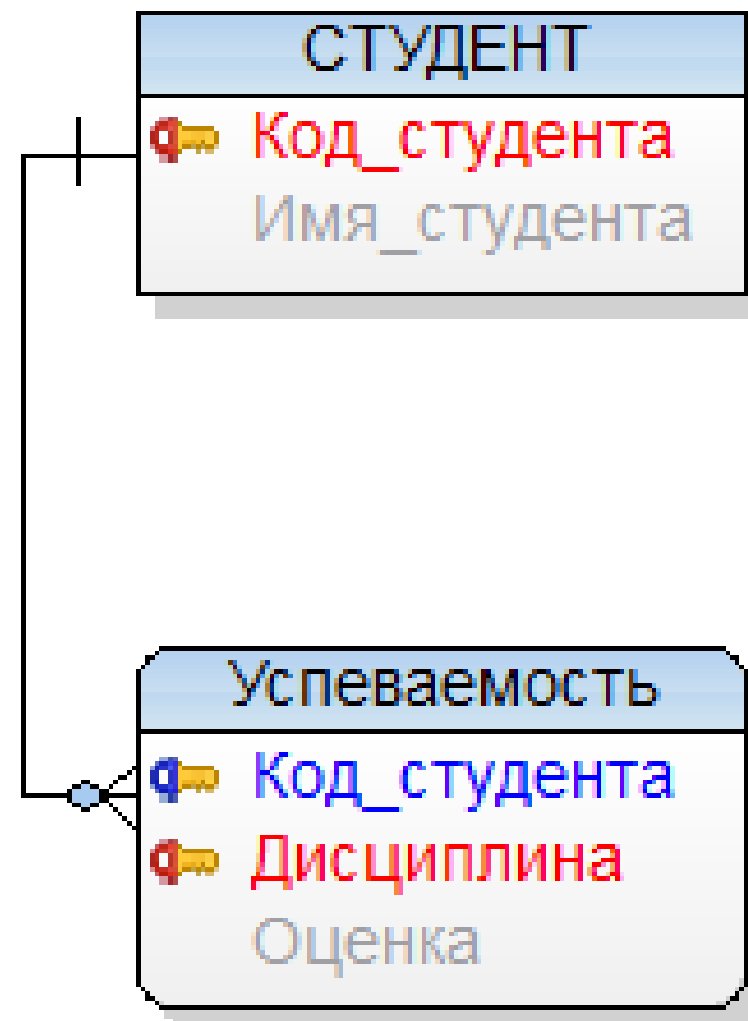
# Алгоритм нормализации (2НФ)



М – множество атрибутов отношения R.  
 $A \rightarrow V$  – частичная зависимость (A – возможный ключ, V – неключевой атрибут). R разбивается на  $R_1[A, V]$  и  $R_2[M - V]$ . Если не достигнута 2НФ, то повторно применяется данный алгоритм.

## Результат нормализации (2НФ)

- Операция **INSERT**: Информация о том, что студент с кодом С4 имеет фамилию «Смирнов», осуществляется вставкой кортежа в отношение СТУДЕНТ независимо от того, имеются ли данные о сдаче этим студентом какой-либо дисциплины.
- Операция **DELETE**: Информация о том, что студент с кодом С9 сдавал экзамен по какой-либо дисциплине (например, «Иностр. язык») может быть удалена без потери информации о фамилии этого студента.
- Операция **UPDATE**: Для изменения фамилии у студента с кодом С2 необходимо скорректировать значение соответствующего атрибута в единственной записи отношения СТУДЕНТ.



## Определение 2-й нормальной формы

Отношение находится в **2НФ** тогда и только тогда, когда оно находится в **1НФ**, и каждый неключевой атрибут **функционально полно зависит** от первичного ключа.



Отношение находится в **2НФ** тогда и только тогда, когда оно находится в **1НФ** и не содержит **частичных зависимостей** (нет атрибутов, зависящих от части первичного ключа).



## *Условие обратимости*

**Процедура нормализации отношений должна быть  
обратимой.**




- Преобразование отношений в НФ более высокой степени должно осуществляться **без потерь информации**. Исходное отношение можно восстановить с помощью операции естественного соединения (**декомпозиция без потерь**).

```
select *  
from (select R1 from r)  
natural join  
(select R2 from r)
```

- ФЗ исходного и результирующих отношений должны быть эквиваленты (**сохранение зависимостей**).

# *Примеры*

## Отношение СТУДЕНТ



КОД_СТУДЕНТА	ИМЯ_СТУДЕНТА	ФАКУЛЬТЕТ
C2	Иванов	Физический
C5	Петров	Химический
C7	Иванов	Исторический

# Варианты декомпозиции

Отношение A1

КОД_СТУДЕНТА	ИМЯ_СТУДЕНТА
C2	Иванов
C5	Петров
C7	Иванов

Отношение A2

КОД_СТУДЕНТА	ФАКУЛЬТЕТ
C2	Физический
C5	Физический
C7	Исторический

Отношение Б1

КОД_СТУДЕНТА	ИМЯ_СТУДЕНТА
C2	Иванов
C5	Петров
C7	Иванов

Отношение Б2

ИМЯ_СТУДЕНТА	ФАКУЛЬТЕТ
Иванов	Физический
Петров	Физический
Иванов	Исторический

Отношение В1

КОД_СТУДЕНТА	ФАКУЛЬТЕТ
C2	Физический
C5	Химический
C7	Исторический

Отношение В2

ИМЯ_СТУДЕНТА	ФАКУЛЬТЕТ
Иванов	Физический
Петров	Физический
Иванов	Исторический



# Анализ вариантов декомпозиции

При проведении декомпозиции по варианту «б» потеряна функциональная зависимость атрибута ФАКУЛЬТЕТ от атрибута КОД\_СТУДЕНТА:

Б1	КОД_СТУДЕНТА	ИМЯ_СТУДЕНТА
	C2	Иванов
	C5	Петров
	C7	Иванов

Б2	ИМЯ_СТУДЕНТА	ФАКУЛЬТЕТ
	Иванов	Физический
	Петров	Физический
	Иванов	Исторический

Отношение Б1 JOIN Отношение Б2

КОД_СТУДЕНТА	ИМЯ_СТУДЕНТА	ФАКУЛЬТЕТ
C2	Иванов	Физический
C2	Иванов	Исторический
C5	Петров	Физический
C7	Иванов	Исторический
C7	Иванов	Физический

## *Теорема Хита (Heath's theorem)*

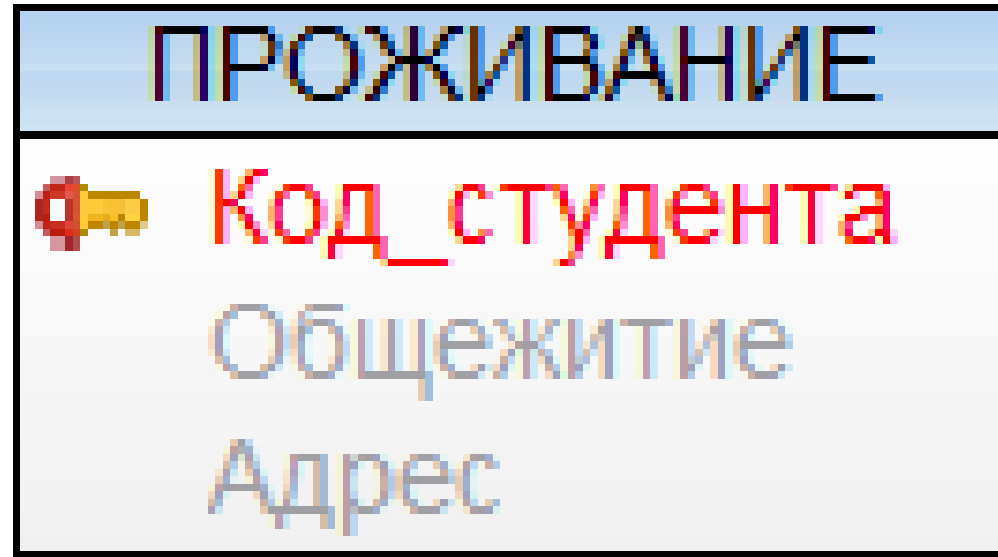
Отношение  $R$ , включающее атрибуты  $A$ ,  $B$  и  $C$ , декомпозируется без потерь на проекции  $\{A, B\}$  и  $\{A, C\}$ , если имеется функциональная зависимость:  $A \rightarrow B$

$$R1 \{A, B\} \text{ JOIN } R2 \{A, C\} = R$$



*Любую ФЗ отношения можно вынести в отдельное отношение, оставив ее детерминант в исходном отношении. При этом никакая информация не будет утеряна.*


## Предпосылки перехода к 3НФ



КОД_СТУДЕНТА	ОБЩЕЖИТИЕ	АДРЕС
C2	1	ул. Здоровцева, 14
C6	1	ул. Здоровцева, 14
C9	2	ул. Пионерстроя, 14
C1	3	ул. Здоровцева, 12
C7	3	ул. Здоровцева, 12

# Аномалии при выполнении DML-операторов к 2НФ

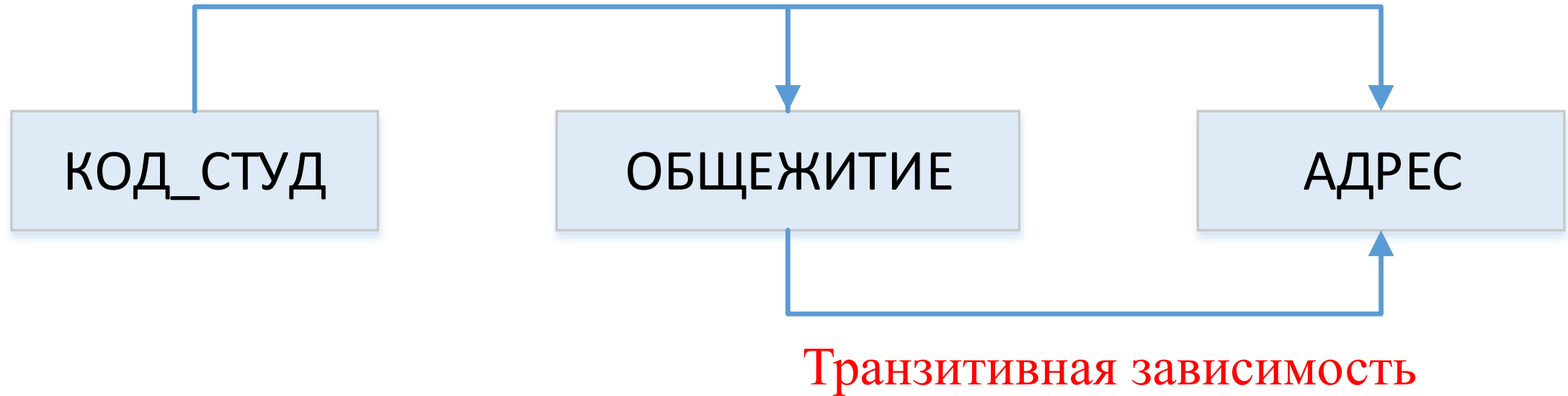
- Операция **INSERT**: Нельзя включить в БД информацию об адресе нового общежития до тех пор, пока нет данных хотя бы об одном студенте, проживающем в этом общежитии.
- Операция **DELETE**: При удалении информации о проживающих в общежитии студентов можно потерять информацию об адресе общежития. Например, при удалении кортежа для студента с кодом С9 теряется информация об адресе общежития №2.
- Операция **UPDATE**: При смене адреса общежития, необходимо внести изменения для всех студентов, проживающих в нём. При некорректном завершении операции в отношении могут оказаться кортежи с противоречивой информацией.

ПРОЖИВАНИЕ	
	Код_студента
	Общежитие
	Адрес



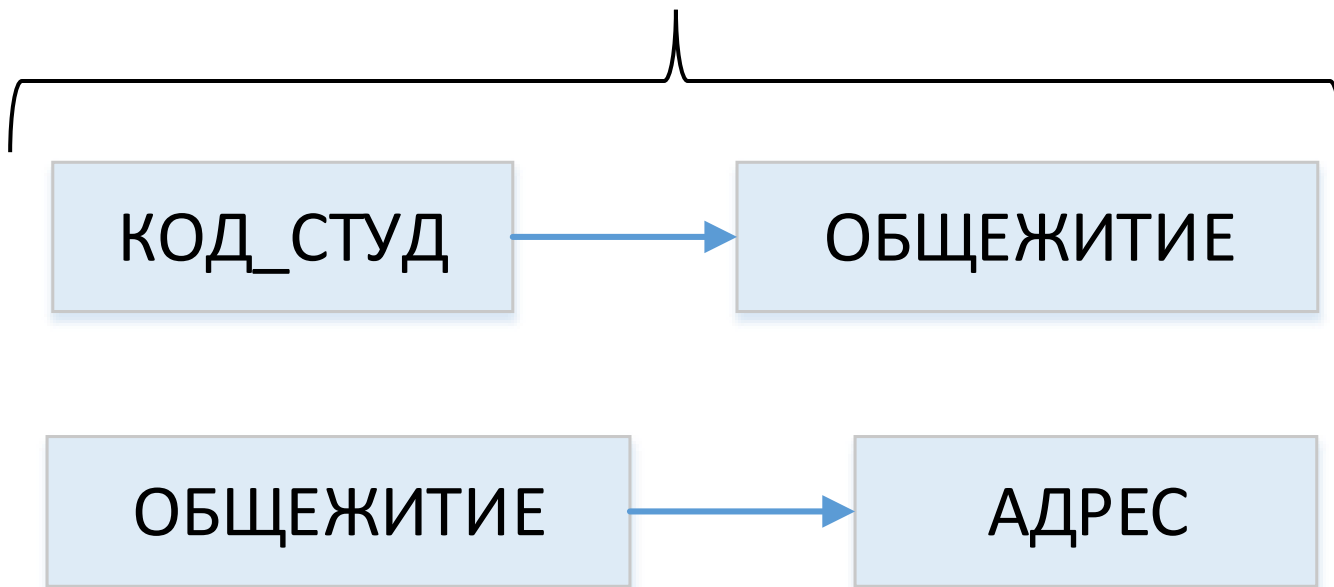
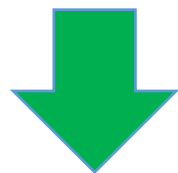
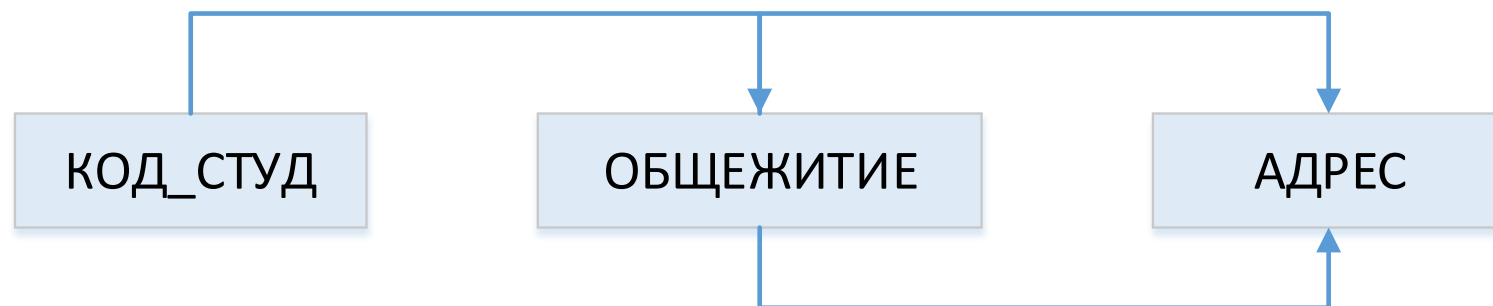
## Предпосылки перехода к 3НФ

Причина аномалий — наличие **транзитивной** функциональной зависимости атрибута АДРЕС от атрибута ОБЩЕЖИТИЕ:



*Наличие транзитивной зависимости — признак, что отношение содержит информацию более, чем одной сущности.*

## Результат нормализации (3НФ)



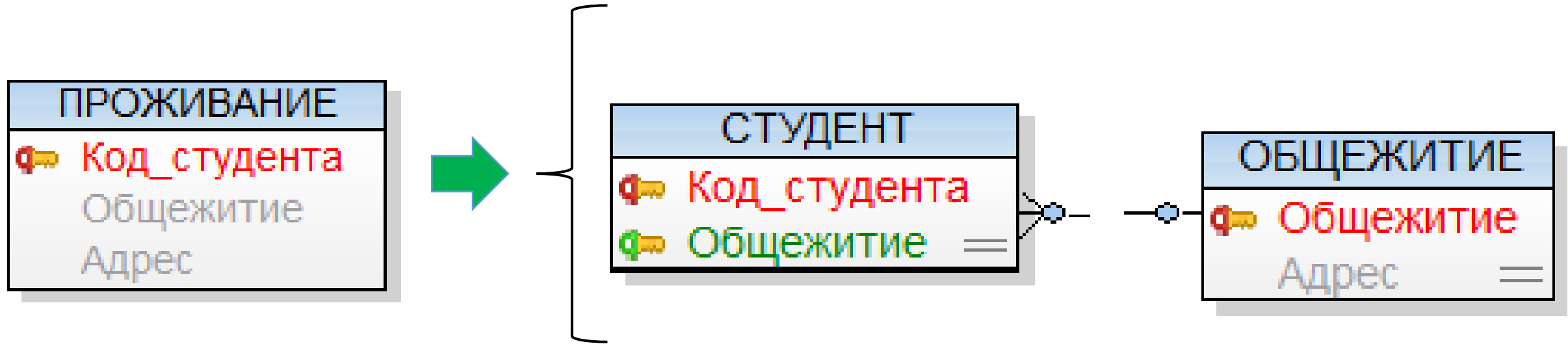
СТУДЕНТ =

КОД_СТУДЕНТА	ОБЩЕЖИТИЕ
C2	1
C6	1
C9	2
C1	3
C7	3

ОБЩЕЖИТИЕ =

ОБЩЕЖИТИЕ	АДРЕС
1	ул. Здоровцева, 14
2	ул. Пионерстроя, 14
3	ул. Здоровцева, 12

## Результат нормализации (3НФ)



$M$  – множество атрибутов отношения  $R$ .

$A \rightarrow B$  и  $B \rightarrow C$  – частичная зависимость ( $A$  – возможный ключ,  $C$  – неключевой атрибут).

$R$  разбивается на  $R1[B, C]$  и  $R2[M - C]$ .

Если не достигнута 3НФ, то повторно применяется данный алгоритм.

*Неключевые атрибуты, зависящие от других неключевых атрибутов, выносятся в отдельное отношение вместе с детерминантом.*

## Определение 3-й нормальной формы

Отношение находится в **3НФ**, когда оно находится в **2НФ**, и его каждый **неключевой** атрибут **нетранзитивно** зависит от первичного ключа.





### Предпосылки перехода к НФ Бойса-Кодда (усиленная 3НФ, $3\frac{1}{2}$ НФ)

- Отношение имеет **два** или **более** потенциальных ключа.
- Потенциальные ключи — **составные** (включают в себя несколько атрибутов).
- Составные ключи **перекрываются** (один или несколько общих атрибутов).

Для отношений, в которых имеется **ОДИН** потенциальный ключ, третья нормальная форма отношения эквивалентна нормальной форме Бойса-Кодда.

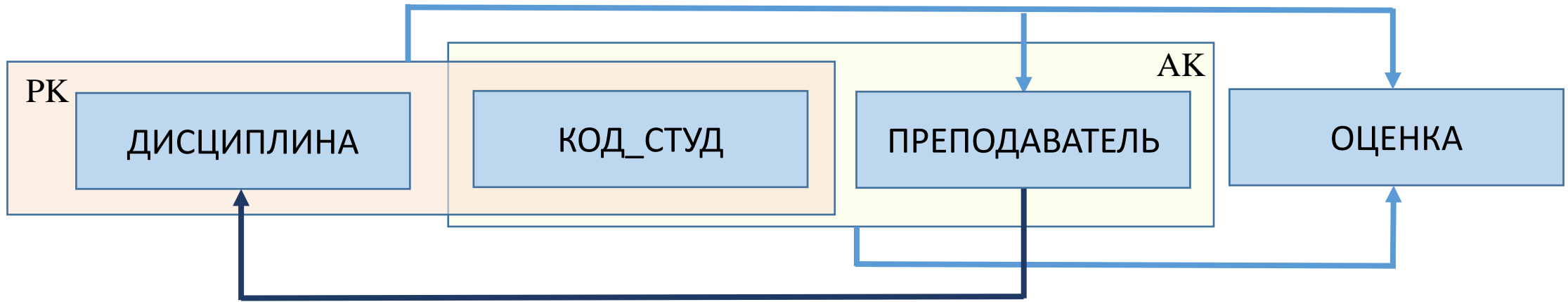
# Примеры

ЭКЗАМЕН =

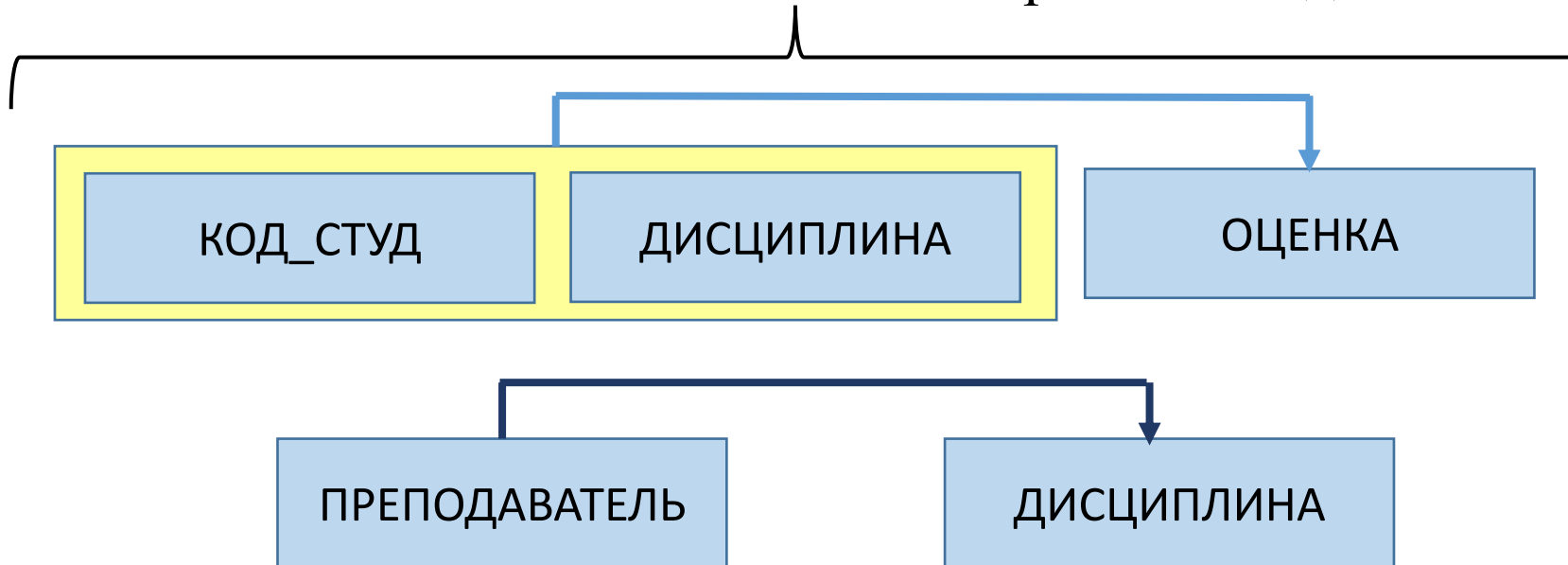
ЭКЗАМЕН	
	Код_студ
	Дисциплина
	Преподаватель
	Оценка

КОД_СТУД	ПРЕДМЕТ	ПРЕПОДАВАТЕЛЬ	ОЦЕНКА
C6	Математика	← Иванов И.И.	3
C6	Программирование	← Петров П.П.	4
C6	Базы данных	Алексеев А.А.	3
C2	Математика	← Смирнов С.С.	5
C2	Базы данных	Алексеев А.А.	4
C2	Сети	Яковлев Я.Я.	4
C2	Физика	Фёдоров Ф.Ф.	5
C2	Программирование	← Петров П.П.	5
C9	Физика	Фёдоров Ф.Ф.	4
C1	Сети	Яковлев Я.Я.	5
C1	Программирование	← Сидоров С.С.	4
C8	Математика	← Иванов И.И.	5

# Результат нормализации (НФБК)





Теряем ФЗ, но сохраняем  
обратимость декомпозиции



# Алгоритм нормализации (НФБК)

ПРЕПОДАВАТЕЛЬ =

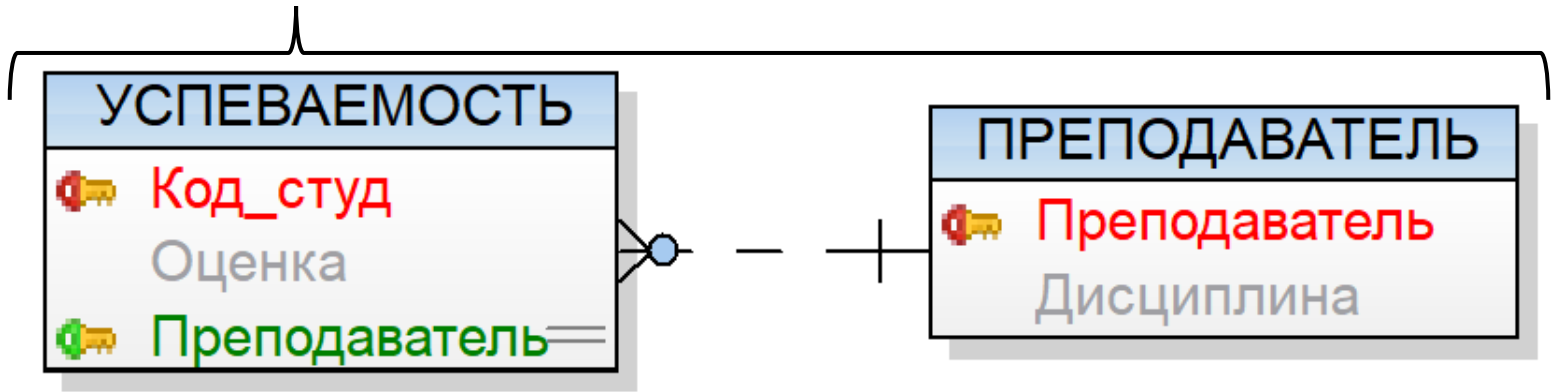
УСПЕВАЕМОСТЬ =

ЭКЗАМЕН
 Код_студ
 Дисциплина
Преподаватель
Оценка



ПРЕПОДАВАТЕЛЬ	ДИСЦИПЛИНА
Иванов И.И.	Математика
Петров П.П.	Программирование
Смирнов С.С.	Математика
Алексеев А.А.	Базы данных
Яковлев Я.Я.	Сети
Фёдоров Ф.Ф.	Физика
Сидоров С.С.	Программирование

КОД_СТУД	ПРЕПОДАВАТЕЛЬ	ОЦЕНКА
С6	Иванов И.И.	3
С6	Петров П.П.	4
С6	Алексеев А.А.	3
С2	Смирнов С.С.	5
С2	Алексеев А.А.	4
С2	Яковлев Я.Я.	4
С2	Фёдоров Ф.Ф.	5
С2	Петров П.П.	5
С9	Фёдоров Ф.Ф.	4
С1	Яковлев Я.Я.	5
С1	Сидоров С.С.	4
С8	Иванов И.И.	5



АЛГОРИТМ аналогичен 3НФ. Атрибуты, зависящие от детерминантов, не являющихся потенциальными ключами выносятся в отдельное отношение вместе с детерминантами.

## *Определение НФБК*

Отношение находится в нормальной форме **Бойса-Кодда**, когда детерминанты каждой нетривиальной и полной функциональной зависимости отношения являются **потенциальными ключами**.



Отношение находится в **НФБК**, если на диаграммах функциональных зависимостей стрелки исходят **только от потенциальных ключей**.



# Предпосылки перехода к 4НФ

Атрибуты  
(\*) и (\*\*) в  
предметной  
области **НЕ  
СВЯЗАНЫ**

Сотрудник	Навык *	Язык **
Иванов	Программист	
Иванов	Администратор	
Иванов		Английский
Иванов		Немецкий
Иванов		Французский

– Непересекающаяся  
форма

с повторениями ← случайная форма → с NULL-значениями

Сотруд ник	Навык	Язык
Иванов	Программист	Английский
Иванов	Администратор	Немецкий
Иванов	Администратор	Французский

Сотруд ник	Навык	Язык
Иванов	Программист	Английский
Иванов	Администратор	Немецкий
Иванов	NULL	Французский

# Предпосылки перехода к 4НФ

Сотрудник	Навык *	Язык **
Иванов	Программист	
Иванов	Администратор	
Иванов		Английский
Иванов		Немецкий
Иванов		Французский

без ограничений

случайная форма

Сотрудник	Навык	Язык
Иванов	Программист	Английский
Иванов	Администратор	
Иванов		Немецкий
Иванов	Администратор	Французский



перекрестная форма

Сотрудник	Навык	Язык
Иванов	Программист	Английский
Иванов	Программист	Немецкий
Иванов	Программист	Французский
Иванов	Администратор	Английский
Иванов	Администратор	Немецкий
Иванов	Администратор	Французский

ВАРИАНТ РЕШЕНИЯ:  
Все значения атрибутов (\*) связаны со всеми значениями атрибута (\*\*)

НФБК

Ключ (составной) –  
{Сотрудник, Навык, Язык}

# Предпосылки перехода к 4НФ

Сотрудник	Навык	Язык
Иванов	Программист	Английский
Иванов	Программист	Немецкий
Иванов	Программист	Французский
Иванов	Администратор	Английский
Иванов	Администратор	Немецкий
Иванов	Администратор	Французский

Сотрудник : Навык (M:N)



Сотрудник : Язык (M:N)



Сотрудник	Навык	Язык
Иванов	Аналитик	????????????



Сотрудник	Навык	Язык
Иванов	Аналитик	????????????



Сотрудник	Навык	Язык
Иванов	Аналитик	????????????

## Определение многозначной зависимости

**Многозначная зависимость** (multivalued dependency, **MVD**) – для отношения  $R$ , включающего непересекающиеся подмножества атрибутов  $A$ ,  $B$  и  $C$ , выполняется утверждение:

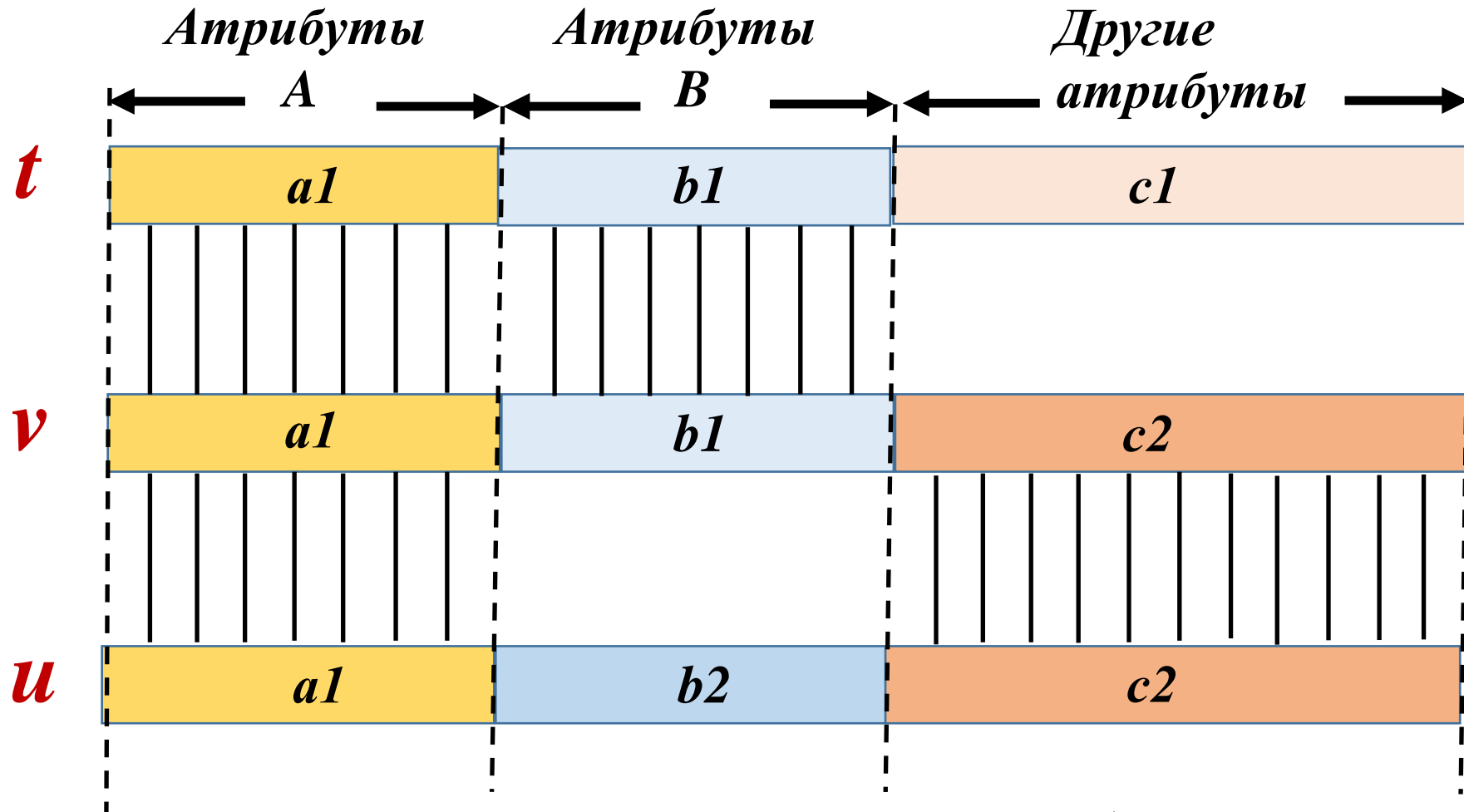
При заданных значениях атрибутов из  $A$  существует множество связанных значений атрибутов из  $B$ , и это множество не зависит от значений атрибутов из  $C$ .

$$\underbrace{A \twoheadrightarrow B \qquad \qquad \qquad A \twoheadrightarrow C}_{A \twoheadrightarrow B / C}$$



Для отношения  $R(A, B, C)$  существует зависимость  $A \twoheadrightarrow B$  тогда и только тогда, когда  $R = R[A, B] \text{ JOIN } R[A, C]$ .

# Графическое представление МЗФ



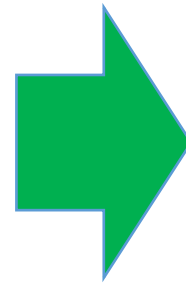
Для каждой пары кортежей  $t$  и  $u$ , которые совпадают во всех атрибутах  $A_1, A_2, \dots, A_n$  можно найти некоторый кортеж  $v$ , совпадающий: с кортежами  $t$  и  $u$  в атрибутах  $A_1, A_2, \dots, A_n$ ; с кортежем  $t$  в атрибутах  $B_1, B_2, \dots, B_m$ ; с кортежем  $u$  во всех атрибутах, которые не принадлежат ни множеству  $\{A_1, A_2, \dots, A_n\}$ , ни множеству  $\{B_1, B_2, \dots, B_m\}$ .

## Теорема Фейджина. Результат нормализации (4НФ)

Отношение  $R$ , включающее атрибуты  $A$ ,  $B$  и  $C$ , декомпозируется без потерь на проекции  $R[A, B]$  и  $R[A, C]$ , если имеется нетривиальная зависимость:  $A \twoheadrightarrow B / C$

*Сотрудник*  $\twoheadrightarrow$  *Навык / Язык*

Сотрудник	Навык	Язык
Иванов	Программист	Английский
Иванов	Программист	Немецкий
Иванов	Программист	Французский
Иванов	Администратор	Английский
Иванов	Администратор	Немецкий
Иванов	Администратор	Французский



Сотрудник	Навык
Иванов	Программист
Иванов	Администратор

Сотрудник	Язык
Иванов	Английский
Иванов	Немецкий
Иванов	Французский

## *Определение 4-й нормальной формы*

***MVD является обобщением FD (аксиома репликации):***

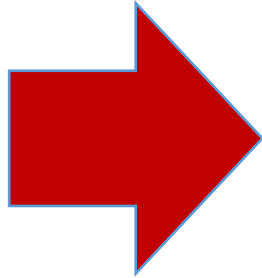
*Если  $A \longrightarrow B$ , то  $A \twoheadrightarrow B$*

Многозначная зависимость  $A \twoheadrightarrow B$  называется **тривиальной**, если  $B \subseteq A$  или  $B \cup A = R$ .

Отношение находится в **4НФ**, когда оно находится в **НФБК** и не содержит нетривиальных многозначных зависимостей.

## Предпосылки перехода к 5НФ

Агент	Поставщик	Деталь
Иванов	Север	Балка
Иванов	Север	Опора
Иванов	Юг	Балка
Иванов	Юг	Опора
Петров	Север	Балка



Агент	Поставщик
Иванов	Север
Иванов	Юг
Петров	Север

Поставщик	Деталь
Север	Балка
Север	Опора
Юг	Балка
Юг	Опора

$R1 [A, B] \text{ JOIN } R2 [B, C] \Leftrightarrow R$

$R2 [B, C] \text{ JOIN } R3 [A, C] \Leftrightarrow R$

$R3 [A, C] \text{ JOIN } R1 [A, B] \Leftrightarrow R$

Если при естественном соединении отношений, полученных при декомпозиции, **создаются ложные кортежи**, то исходное отношение не удовлетворяет **зависимости соединения (JD)** относительно атрибутов, по которым была выполнена проекция исходного отношения.

Агент	Поставщик	Деталь
Иванов	Север	Балка
Иванов	Север	Опора
Иванов	Юг	Балка
Иванов	Юг	Опора
Петров	Север	Балка
<b>Петров</b>	<b>Север</b>	<b>Опора</b>

**Невозможно** выполнить декомпозицию на две проекции без потерь.



## Результат нормализации (5НФ)

Агент	Поставщик
Иванов	Север
Иванов	Юг
Петров	Север

Поставщик	Деталь
Север	Балка
Север	Опора
Юг	Балка
Юг	Опора

Агент	Деталь
Иванов	Балка
Иванов	Опора
Петров	Балка

$$R1[A, B] \text{ JOIN } R2[B, C] \text{ JOIN } R3[A, C] = R$$

*Зависимость соединения называется **тривиальной**, если одно из подмножеств атрибутов  $A, B, \dots Z$  совпадает с множеством всех атрибутов отношения  $R$ .*

Отношение находится в **5НФ** (НФ проекции-соединения), если оно находится в **4НФ**, и для каждой его нетривиальной зависимости соединения все множества атрибутов являются суперключами.

## Пример 5НФ

РЕЙТИНГ =

Наименование	Код	Обучение	Исследование
Политехнический	P_01	21	4
Технологический	T_01	19	6

**Потенциальные ключи** – Наименование, Код

**JD1\*** ([Наименование, Код, Обучение], [Наименование, Исследование])

**JD2\*** ([Наименование, Код, Исследование], [Наименование, Обучение])

**JD3\*** ([Наименование, Код, Обучение], [Код, Исследование])

**JD4\*** ([Наименование, Код, Исследование], [Код, Обучение])

**JD5\*** ([Наименование, Код], [Наименование, Обучение], [Код, Исследование])

**ВСЕ** проекции – суперключи, отношение РЕЙТИНГ находится в 5НФ.