

# Тема: проектирование реляционной БД на основе универсального отношения

ТабНом	Начк
125	Иванов
138	Петров
195	Петров
200	Иванов

а

ТабНом	Начк
125	Иванов
138	Петров
195	-
200	-

б

Дублирование данных, не являющихся  
избыточными

## Избыточное дублирование данных

Дублирование информации о номере телефона  
является избыточным

ТабНом	Начк	Нтел
125	Иванов	3051
138	Петров	2222
195	Петров	2222
200	Иванов	3051

а

ТабНом	Начк	Нтел
125	Иванов	3051
138	Петров	2222
195	Петров	-
200	Иванов	-

б

ТабНом	Начк		Начк	Нтел
125	Иванов		Иванов	3051
138	Петров		Петров	2222
195	Петров			
200	Иванов			

в

Исключение избыточных данных

# Универсальное отношение

Требуется разработать небольшую БД для учета успеваемости студентов, проживающих в общежитии

Первый шаг процесса проектирования состоит в определении как всех атрибутов, наличие которых обязательно в БД, так и связей между атрибутами.

**Сном:** Номер студента.

**Сфам:** Фамилия студента. *Каждый студент имеет только одну фамилию, но возможно, что одну фамилию носят несколько студентов.*

**Кном:** Номер комнаты в общежитии. *В одной комнате может проживать более одного студента.*

**Тном:** Номер телефона студента. *Каждая комната общежития имеет один телефон и им пользуются все студенты, проживающие в этой комнате.*

**Дисц:** Название дисциплины. *В БД будут храниться данные о дисциплинах, по которым проведена итоговая аттестация студента.*

**Семестр:** семестр, в котором данная дисциплина была завершена студентом.

*Возможно, что студент изучал одну и ту же дисциплину в различных семестрах.*

**Оценка:** Оценка за дисциплину. *Оценка, полученная студентом за определенную дисциплину в данном семестре.*

## УСПЕВАЕМОСТЬ

## Данные для размещения в БД

Сном	Сфам	Кном	Тном	Дисц.	Сем.	Оценка
111	Серов	120	2135	ОС	3	2
				ПО	2	4
				Физика	3	5
222	Перов	211	3257	ОС	3	4
				Химия	1	5
				АК	4	4
333	Иванов	120	2135	ПО	2	5
110	Поляков	211	3257	ОС	1	4

## Одна "строка" таблицы

111	Серов	120	2135	ОС	3	2
				ПО	2	4
				Физика	3	5

Таблица УСПЕВАЕМОСТЬ **не** является отношением

## УСПЕВАЕМОСТЬ

Сном	Сфам	Кном	Тном	Дисц.	Сем.	Оценка
111	Серов	120	2135	ОС	3	2
111	Серов	120	2135	ПО	2	4
111	Серов	120	2135	Физика	3	5
222	Перов	211	3257	ОС	3	4
222	Перов	211	3257	Химия	1	5
222	Перов	211	3257	АК	4	4
333	Иванов	120	2135	ПО	2	5
110	Поляков	211	3257	ОС	1	4

Таблица УСПЕВАЕМОСТЬ представляет собой экземпляр корректного отношения.

Его называют **универсальным отношением** проектируемой БД.

**УНИВЕРСАЛЬНОЕ ОТНОШЕНИЕ** - отношение, включающее все представляющие интерес атрибуты  
(может содержать все данные, которые предполагается размещать в БД)



Процесс разбиения **ОТНОШЕНИЯ** с целью уменьшения вероятности возникновения аномалий называется **нормализацией (декомпозицией)**

## **Функциональная зависимость (ФЗ)**

определяется следующим образом:

Если даны два атрибута **A** и **B**, то говорят, что **B** функционально зависит от **A**, если для каждого значения **A** существует ровно одно связанное с ним значение **B** (в любой момент времени). **A** и **B** могут быть составными

**$A \rightarrow B$**  - математическая форма записи

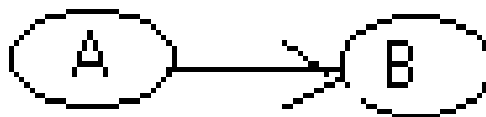


диаграмма или графическая форма записи

# УСПЕВАЕМОСТЬ

Сном	Сфам	Кном	Тном	Дисц.	Сем.	Оценка
111	Серов	120	2135	ОС	3	2
111	Серов	120	2135	ПО	2	4
111	Серов	120	2135	Физика	3	5
222	Перов	211	3257	ОС	3	4
222	Перов	211	3257	Химия	1	5
222	Перов	211	3257	АК	4	4
333	Иванов	120	2135	ПО	2	5
110	Поляков	211	3257	ОС	1	4

**Сном -> Сфам**

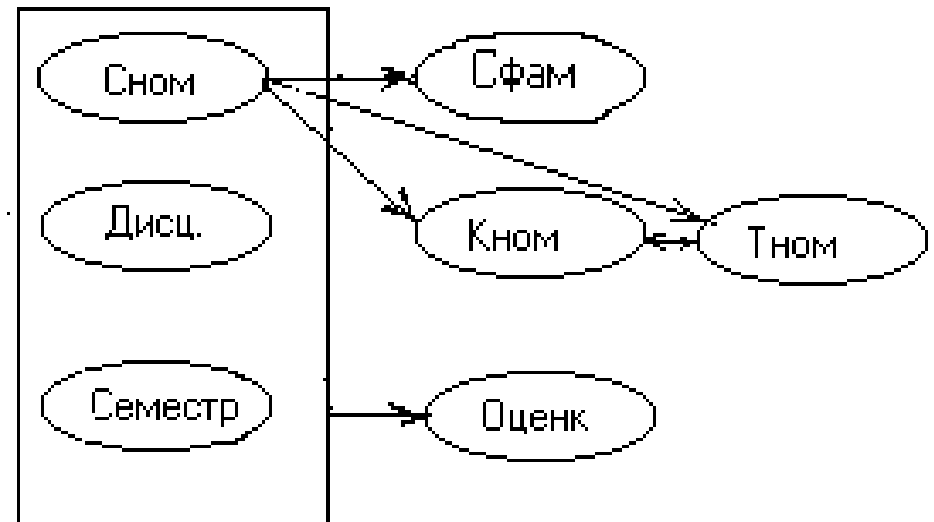
**Сном -> Кном**

**Кном -> Тном**

**Тном -> Кном**

**Сном -> Тном**

**Сном, Дисциплина, Семестр -> Оценка**





## ***Полная функциональная зависимость***

**Функциональная зависимость  $A \rightarrow B$  является полной ФЗ, если удаление какого-либо атрибута из группы атрибутов  $A$  приводит к потере этой зависимости**  
*(Неключевой атрибут функционально полно зависит от составного ключа если он функционально зависит от всего ключа в целом, но не находится в функциональной зависимости от какого-либо из входящих в него атрибутов).*

## ***Частичная ФЗ (неполная)***

**Функциональная зависимость  $A \rightarrow B$  является частичной ФЗ, если в группе атрибутов  $A$  есть один или несколько атрибутов, при удалении которых эта зависимость сохраняется**

При неполной функциональной зависимости  
возникают *аномалии*:

- . *включения*
- . *удаления*
- . *обновления*

Данные виды аномалий возникают при любой избыточной функциональной зависимости.

## **Многозначная зависимость**

**Один атрибут таблицы многозначно определяет другой атрибут той же таблицы, если для каждого значения первого атрибута существует хорошо определенное множество соответствующих значений второго атрибута.**

## Пример:

**Дано отношение**

**Книги(ISBN, Название, Автор, Область знаний).**

**Книга имеет уникальный идентификатор ISBN,  
книга может быть написана коллективом авторов,  
книга может относиться к нескольким областям знаний**

**Книга**

ISBN	Название	Автор	Область знаний
5-123-12345-1	Информатика для экономистов	Иванов А.В.	Информатика
5-123-12345-1	Информатика для экономистов	Иванов А.В.	Экономика
5-123-12345-1	Информатика для экономистов	Петров С.М.	Информатика
5-123-12345-1	Информатика для экономистов	Петров С.М.	Экономика

**Существуют следующие функциональные зависимости:**

**ISBN → Название**

**ISBN ->> Автор**

**ISBN ->> Область знаний**

## **Транзитивная функциональная зависимость.**

Пусть  $A, B, Z$  – три атрибута некоторого отношения.

При этом  $A \rightarrow B$  и  $B \rightarrow Z$ , но обратное соответствие отсутствует, т.е.  $Z \not\rightarrow B$  и  $B \not\rightarrow A$ .

Тогда  $Z$  транзитивно зависит от  $A$ .

**Избыточная функциональная зависимость** – это зависимость, заключающая в себе такую информацию, которая может быть получена на основе других зависимостей, имеющихся в базе

## Пример:

Дано отношение

**Группы**(Группа, Специальность, Факультет) с  
первичным ключом **Группа**.

**Группа** однозначно определяет **специальность**, а  
**специальность** однозначно определяет **факультет**.

Т.е. существуют следующие функциональные  
зависимости:

**Группа → Специальность**

(и наоборот, Специальность  $\rightarrow$  Группа)

**Специальность → Факультет**

(Факультет  $\rightarrow$  Специальность)

**Факультет** транзитивно зависит от **Группа**

# Первая нормальная форма (1НФ)<sup>1НФ</sup>

*Таблица (отношение) находится в 1НФ, если не содержит одинаковых строк и в любом допустимом значении этой таблицы каждая ее строка содержит только одно значение для каждого атрибута. (значения всех атрибутов **атомарны**).*

*(Универсальное отношение)*

Первая нормальная форма является основой реляционной модели данных.

## Пример

Персона(Номер, Фамилия, Имя, Отчество, ДатаРождения)

Очевидно, что существует функциональные зависимости атрибутов:

Номер  $\rightarrow$  Фамилия,

Номер  $\rightarrow$  Имя,

Номер  $\rightarrow$  Отчество,

Номер  $\rightarrow$  ДатаРождения

*Полная функциональная зависимость: все атрибуты зависят от составного ключа и не зависят ни от какой его части.*



## Вторая нормальная форма (2НФ):

2НФ

***Таблица (отношение) находится во 2НФ, если она удовлетворяет определению 1НФ, и каждый неключевой атрибут функционально полно зависит от всего ключа.***

Если какой-либо атрибут зависит от части составного первичного ключа, то необходимо:

- создать новое отношение, атрибутами которого будут:
  - часть составного ключа (первичный ключ нового отношения),
  - атрибут, зависящий от нового ключа;
- из исходного отношения исключить атрибут, включенный в новое отношение.

**Пример**

Имеется отношение  **$R(\underline{k1}, \underline{k2}, a1, a2)$** , находящееся в 1НФ, где

$k1, k2$  – составной первичный ключ,

$a1$  и  $a2$  – неключевые атрибуты отношения  $R$ ,

*Функциональные зависимости:*

**$k1, k2 \rightarrow a1$**

(атрибут  $a1$  функционально полно зависит от первичного ключа  $k1, k2$ ),

**$k1 \rightarrow a2$**

(атрибут  $a2$  зависит от части первичного ключа  $k1$ , т.е. имеется неполная функциональная зависимость)

***Для приведения отношения  $R$  к 2НФ, это отношение декомпозируется на два отношения:***

***$R1(k1, a2)$  и  $R2(k1, k2, a1)$ .***

***Отношения  $R1$  и  $R2$  будут иметь связь один-ко-многим по атрибуту  $k1$ .***

## Пример:

Дано отношение

**Поставки(КодПоставщика, КодПродукта, ЕдиницаИзмерения)**

Поставщик может поставлять различные продукты, один и тот же продукт может поставляться разными поставщиками. Тогда первичным ключом отношения будут атрибуты **КодПоставщика** и **КодПродукта**.

Значит, существует функциональная зависимость:

**КодПоставщика, КодПродукта → ЕдиницаИзмерения**

С другой стороны, какой бы поставщик не поставил продукт, единица измерения от этого не изменится (например, цельное молоко измеряется литрами независимо от поставщика, а соль – килограммами).

Т.е. существует еще одна функциональная зависимость (неключевой атрибут зависит от части первичного ключа):

**КодПродукта → ЕдиницаИзмерения**

**После исключения неполной функциональной зависимости получим отношения:**

**Поставки(КодПоставщика, КодПродукта)**

**Продукты(КодПродукта, ЕдиницаИзмерения)**

## Аномалии:

- . *включения* (пока поставщиком не будет поставлен продукт, нельзя указать единицу измерения)
- . *удаления* (исключение поставщика может привести к потере единицы измерения продукта)
- . *обновления* (при изменении единицы измерения продукта, приходится менять данные везде, где встречается данный продукт)

## Пример

Кафедра(КодИнститута, КодКафедры, Название, Телефон, Адрес)

Составной ключ - «КодИнститута, КодКафедры»

(поле «КодКафедры» идентифицирует кафедру внутри института).

Имеет место функциональная зависимость:

**КодИнститута, КодКафедры → Название, Телефон, Адрес**

Для каждой кафедры указывается адрес. Однако, адрес на самом деле зависит только от кода института (т.е. от части ключа).

*Для исправления ситуации адрес должен стать атрибутом института.*

## Третья нормальная форма (3НФ):

3НФ

*Таблица (отношение) находится в 3НФ, если она находится во 2НФ и ни один из её неключевых атрибутов не связан ФЗ с любым другим неключевым атрибутом.*

*(Каждый неключевой атрибут нетранзитивно зависит от первичного ключа).*

## Пример

Имеется отношение  $R(\underline{k1}, a1, a2)$ , находящееся в 2НФ, где  $k1$  – первичный ключ, а

$a1$  и  $a2$  – неключевые атрибуты отношения  $R$ , имеются функциональные зависимости:

$k1 \rightarrow a1$

$a1 \rightarrow a2$

тогда атрибут  $a2$  транзитивно зависит от  $k1$ .

*Для приведения отношения  $R$  к 3НФ, это отношение декомпозируется на два отношения:*

*$R1(k1, a1)$  и  $R2(a1, a2)$ .*

*Отношения  $R1$  и  $R2$  будут иметь связь многие-к-одному по атрибуту  $a1$ .*

## Пример:

Дано отношение

Группы(Группа, Специальность, Факультет)

Группа однозначно определяет специальность, а специальность однозначно определяет факультет:

Группа → Специальность

(и наоборот, Специальность  $\nrightarrow$  Группа)

Специальность → Факультет

(Факультет  $\nrightarrow$  Специальность)

*После исключения транзитивной функциональной зависимости получим отношения:*

Группы(Группа, Специальность)

Специальности(Специальность, Факультет)



## Пример

Кафедра(КодКафедры, Название, Телефон, Корпус, Адрес)

Первичный ключ состоит только из одного атрибута. Следовательно, все неключевые атрибуты связаны с ним полной функциональной зависимостью (условие **2НФ** соблюдено).

Можно выделить ФЗ:

КодКафедры → Название, Телефон, Корпус

Корпус → Адрес

Адрес определяется корпусом, в котором размещается кафедра. Следовательно, имеет место *транзитивная зависимость*:

КодКафедры → Корпус → Адрес

*После исключения транзитивной ФЗ получим отношения:*

Кафедра(КодКафедры, Название, Телефон, Корпус)

Корпус(Корпус, Адрес)

## **Нормальная форма Бойса-Кодда**

**НФБК**

**Таблица (отношение) находится в нормальной форме Бойса-Кодда (НФБК), если она удовлетворяет определению 3НФ и любая ФЗ между ее атрибутами сводится к полной ФЗ от возможного первичного ключа.**

Ситуация, когда отношение будет находиться в 3НФ, но не в нормальной форме Бойса-Кодда (НФБК), возникает при условии, что отношение имеет два (или более) возможных ключа, которые являются составными и имеют общий атрибут.

## Пример

Имеется отношение  $R(a1, a2, a3, a4)$ ,  
находящееся в 3НФ,  
где  $a1, a2$  – возможный ключ,  
 $a2, a3$  – возможный ключ,  
 $a4$  – неключевой атрибут отношения  $R$ ,  
Имеются функциональные зависимости:

$a1 \rightarrow a3$

$a3 \rightarrow a1$

$a1, a2 \rightarrow a4$

$a2, a3 \rightarrow a4$

Для приведения отношения  $R$  к НФБК, это отношение  
декомпозируется на два отношения:

$R1(\underline{a1}, a3)$  и  $R2(\underline{a1}, \underline{a2}, a4)$

или  $R1(\underline{a3}, a1)$  и  $R2(\underline{a2}, \underline{a3}, a4)$

## Пример:

Дано отношение

**Экзамен(№ зачетки, № паспорта, Дисциплина, Дата, Оценка)**

Возможными ключами будут атрибуты:

№ зачетки, Дисциплина, Дата и № паспорта, Дисциплина, Дата.

Имеются следующие функциональные зависимости:

**№ зачетки, Дисциплина, Дата → Оценка**

**№ паспорта, Дисциплина, Дата → Оценка**

**№ зачетки → № паспорта**

**№ паспорта → № зачетки**

**После приведения отношения к НФБК могут быть получены отношения:**

**Студент(№ зачетки, № паспорта),**

**Экзамен(№ зачетки, Дисциплина, Дата, Оценка)**

**или**

**Студент(№ паспорта, № зачетки),**

**Экзамен(№ паспорта, Дисциплина, Дата, Оценка)**

# Высшие нормальные формы

Четвертая нормальная форма (4НФ)

*Пятая нормальная форма (5НФ)*  
*(нормальная форма проекции-соединения)*

## Алгоритм декомпозиционного проектирования БД

- 1) Построение универсального отношения для БД.
- 2) Определение всех ФЗ, существующих между атрибутами универсального отношения.
- 3) Удаление всех избыточных ФЗ из исходного набора ФЗ в соответствии с процедурой нормализации:

1НФ

2НФ

3НФ

НФ Бойса-Кодда

*Коддом доказано, что большинство потенциальных аномалий в БД будет устранено в случае должной декомпозиции каждого отношения в нормальную форму Бойса-Кодда (НФБК)*

## 1 этап

Универсальное отношение

$R1(\underline{Сном}, \underline{Дисц}, \underline{Семестр}, \text{Оценка}, \text{Сфам}, \text{Кном}, \text{Тном})$

## 2 этап

Определение всех ФЗ:

**Сном -> Сфам**

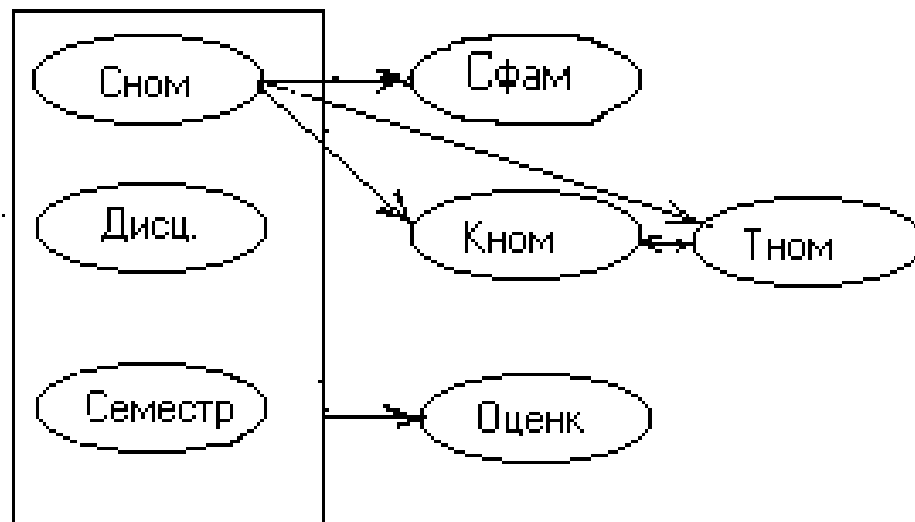
**Сном -> Кном**

**Кном -> Тном**

**Тном -> Кном**

**Сном -> Тном**

**Сном, Дисциплина, Семестр -> Оценка**



### 3 этап - процедура нормализации

1НФ

2НФ

R2(Сном, Дисц, Семестр, Оценка)

R3(Сном, Сфам, Кном, Тном)

3НФ

R2(Сном, Дисц, Семестр, Оценка)

R4(Кном, Тном)

R5(Сном, Сфам, Кном)



**R2**

<u>Сном</u>	<u>Дисц</u>	<u>Сем</u>	<u>Оценка</u>
111	ОС	3	2
111	ПО	2	4
222	АК	4	4
333	ПО	2	5
110	ОС	1	4

**R4**

<u>Кном</u>	<u>Тном</u>
120	2135
211	3257

**R5**

<u>Сном</u>	<u>Сфам</u>	<u>Кном</u>
111	Серов	120
222	Перов	211
333	Иванов	120
110	Поляков	211
444	Белов	401