



# Теория проектирования реляционных баз данных

Дж.Ульман

Основы систем баз данных

Глава 5

# Проблемы проектирования - аномалии

Поставщики(назв\_пост,адрес\_пост,товар,цена)

## АНОМАЛИИ

- ❖ **Избыточность** (повторяется АДРЕС)
- ❖ **Включение**  
(нельзя вставить ПОСТАВЩИКА, если нет *поставляемого* товара)
- ❖ **Удаление**  
(если удалить все ТОВАРЫ, то и поставщика НЕ БУДЕТ)
- ❖ **Обновление** (НУЖНО следить за тем, чтобы обновление было проведено на всех кортежах БД).  
Например, при изменении адреса у некоторого поставщика необходимо изменить значение поля «адрес\_пост» для всех кортежей этого поставщика в БД)

# Ограничения на отношения

- ❖ Отношения могут *различными способами* использоваться для моделирования «реального мира»
- ❖ Каждый кортеж отношения может, например, представлять НЕКОТОРЫЙ ОБЪЕКТ или СВЯЗЬ между объектами.
- ❖ Из фактов о реальном мире следует, что **НЕ КАЖДОЕ конечное множество кортежей** может быть **текущим значением отношения**, даже если бы они имели правильную арность и были выбраны из правильных доменов.

Персонал (фамилия, дата рождения, ном\_пасп, отдел, должность)

Иванов	13.09.1990	5001793549	1	м.н.с.
Петров	09.13.1980	500284334	3	с.н.с.

Отделы (код, название)

1	операционных систем
2	баз данных

- ❖ Существует 2 вида ограничений на отношения:
  - ✓ Ограничения, которые зависят от семантики элементов домена («дата рождения», «ном\_пасп» и т.п.)
  - ✓ Ограничения, которые накладываются в виде равенства или неравенства значений атрибутов, например, «отдел» (Персонал) «код» (Отделы) или вообще в виде некоторой функциональной зависимости

# Функциональные зависимости

Пусть  $R(A_1, A_2, \dots, A_n)$  - отношение,

$X, Y$  являются подмножествами  $\{A_1, A_2, \dots, A_n\}$

## Определение

Множество атрибутов  $Y$  функционально зависит от  $X$  на  $R$ ,  
 $(X \rightarrow Y)$

если в отношении  $R$  не могут содержаться

2 кортежа, компоненты которых

- ❖ совпадают по всем атрибутам из  $X$ ,
- ❖ но не совпадают по одному или более атрибутам из  $Y$

То есть  $\neg \exists r_1, r_2 (r_1, r_2 \in R, \pi_x(r_1) = \pi_x(r_2) \wedge \pi_y(r_1) \neq \pi_y(r_2))$  или  
 $\forall r_1, r_2 (r_1, r_2 \in R, \pi_x(r_1) = \pi_x(r_2) \Rightarrow \pi_y(r_1) = \pi_y(r_2))$

Функциональные зависимости отражают

а) зависимость неключевых атрибутов от КЛЮЧА,

б) отношение N:1 разных групп объектов

# Логические следствия функциональных зависимостей

Пусть  $R$  - схема отношения, а  $A, B, C$  – некоторые атрибуты.

Пусть известно, что ФЗ  $A \rightarrow B$  и  $B \rightarrow C$  выполняются на  $R$ .

Можно утверждать, что тогда  $A \rightarrow C$ .

Доказательство от противного.

Пусть  $(a_1, b_1, c_1)$  и  $(a_2, b_2, c_2)$  кортежи, в которых  $a_1 = a_2$ .

Предположим, что при этом  $c_1 \neq c_2$ .

Но тогда, если  $b_1 \neq b_2$ , то возникает противоречие с ФЗ  $A \rightarrow B$ .

А если  $b_1 = b_2$  и при этом  $c_1 \neq c_2$ ,  
то возникает противоречие с  $B \rightarrow C$ .

И, следовательно,  $c_1 = c_2$ .

Пусть  $F$  - множество ФЗ в  $R$ .

Функциональная зависимость  $X \rightarrow Y$  логически следует из  $F$ , если для каждого отношения  $r$  со схемой  $R$ , на котором выполняются зависимости  $F$  выполняется и  $X \rightarrow Y$

$F^+$  - замыкание  $F$  – множество функциональных зависимостей, которые логически следуют из  $F$ .

# Примеры

- Функциональные зависимости декларируются при проектировании БД (один раз).
- Необходим анализ семантики атрибутов (эксперт).

## База КООПЕРАТИВ

Члены (фамилия, адрес, баланс)

Заказы (фамилия, товар, количество)

Поставщики (назв\_пост, адрес\_пост, товар, цена)

### Определяем ФЗ

В «Заказы» фамилия, товар → количество

В «Поставщики» назв\_пост, товар → цена

назв\_пост → адрес\_пост

Можно вывести зависимость

назв\_пост, товар → адрес\_пост, цена

## Пример замыкания функциональных зависимостей

Пусть  $R$  - схема отношения, а  $A, B, C$  – атрибуты.

Пусть известно, что на  $R$  выполняются функциональные зависимости  $A \rightarrow B$  и  $B \rightarrow C$ .

Тогда замыкание  $F^+$  состоит из всех зависимостей  $X \rightarrow Y$ , таких что выполняется одно из следующих условий:

1.  $X$  содержит  $A$ , например,  $ABC \rightarrow AB$ ,  $AB \rightarrow BC$  или  $A \rightarrow C$
2.  $X$  содержит  $B$ , но не  $A$ , и  $Y$  не содержит  $A$ , например,  $BC \rightarrow B$ ,  $B \rightarrow C$  или  $B \rightarrow \emptyset$ .
3.  $X \rightarrow Y$  – одна из двух зависимостей:  $C \rightarrow C$  или  $C \rightarrow \emptyset$

# Ключи

Пусть  $R(A_1, A_2, \dots, A_n)$  - отношение,  $F$  – ФЗ,  
а  $X$  подмножество  $\{A_1, A_2, \dots, A_n\}$ .

$X$  называется **ключом**, если

1.  $X \rightarrow A_1, A_2, \dots, A_n$  **принадлежит**  $F^+$ ,
2. Ни для какого собственного подмножества  $Y \subset X$ ,  
 $Y \rightarrow A_1, A_2, \dots, A_n$  **не принадлежит**  $F^+$ .

Множество атрибутов, содержащее ключ в качестве подмножества, называется **суперключем** .

Иногда используется термин **ВОЗМОЖНЫЙ КЛЮЧ**



# Пример. Ключи

В примере на с. 7 только атрибут А является ключом, так как только  $A \rightarrow ABC$  и никакое подмножество А не определяет функционально ABC.

Другой пример R (Город, Адрес, Индекс), где Адрес – название улицы, номер дома, номер квартиры, Индекс – почтовый индекс отделения связи

Если имеем ФЗ:

Город, Адрес  $\rightarrow$  Индекс

Индекс  $\rightarrow$  Город

То можно проверить, что {Город, Адрес} и {Адрес, Индекс} одновременно являются ключами.

Из определения ключа

1. должны иметь место следующие ФЗ:

{Город, Адрес}  $\rightarrow$  {Город, Адрес, Индекс}

{Адрес, Индекс}  $\rightarrow$  {Город, Адрес, Индекс}

2. ни один из атрибутов {Адрес}, {Город}, {Индекс} отдельно не определяет кортеж в отношении.

Проверить это самостоятельно и уметь показать, как это получается !

# Аксиомы функциональных зависимостей

Чтобы определить ключи, нужно вычислить  $F^+$  из  $F$  или, по крайней мере, для всякой функциональной зависимости знать содержится ли она в  $F^+$ !

Для этого нужно иметь правила, которые указывают, как из одной или более зависимостей выводить другие.

На самом деле можно задать **полное множество правил** для вывода из  $F$  множество  $F^+$

## Аксиомы АРМСТРОНГА

Пусть дана некоторая схема отношения с универсальным множеством атрибутов  $U$  и множеством функциональных зависимостей  $F$

- А1. Рефлексивность.

Если  $Y \subseteq X \subseteq U$ , то  $X \rightarrow Y$ ,

- А2. Пополнение.

Если  $X \rightarrow Y$  и  $Z \subseteq U$ , то  $XZ \rightarrow YZ$

- А3. Транзитивность.

Если  $X \rightarrow Y$  и  $Y \rightarrow Z$ , то  $X \rightarrow Z$

# Надежность и полнота аксиом (1)

Мы определили аксиомы (правила) вывода из одних функциональных зависимостей - другие. Эти аксиомы (правила) являются надежными и полными.

**Надежными** они являются потому, что все выводимые с помощью этих правил функциональные зависимости принадлежат  $F^+$ .

**Полными** они являются потому, что любая функциональная зависимость из  $F^+$  может выведена из  $F$  с помощью этих аксиом.

Лемма 1. Аксиомы АРМСТРОНГА являются надежными.

Если зависимость  $X \rightarrow Y$  выведена из  $F$  с помощью этих аксиом, то она **справедлива в любом отношении**, в котором выполняются  $F$ .

То есть мы не можем, используя аксиомы, вывести ФЗ, не принадлежащую  $F^+$

Доказательство.

**Аксиома А1** (Если  $Y \subseteq X \subseteq U$ , то  $X \rightarrow Y$ ) очевидно является надежной. Мы не можем иметь отношение  $R$ , в котором 2 кортежа совпадают по  $X$ , но не совпадают по некоторому подмножеству  $X$ .

Для доказательства надежности **аксиомы А2** (Если  $X \rightarrow Y$  и  $Z \subseteq U$ , то  $XZ \rightarrow YZ$ ) предположим, что существуют 2 кортежа  $t$  и  $u$ , которые совпадают по  $XZ$  и не совпадают по  $YZ$ . Но так как они не могут не совпадать по  $Z$ , то значит они не совпадают по  $Y$ , что противоречит тому факту, что  $X \rightarrow Y$  справедлива для  $R$ .

Доказательство надежности **аксиомы А3** (Если  $X \rightarrow Y$  и  $Y \rightarrow Z$ , то  $X \rightarrow Z$ ) можно воспользоваться методом от противного, как мы это сделали раньше.

## Надежность и полнота аксиом (2)

**Лемма 2.** Справедливы следующие правила.

1. Правило **объединения**. Если  $X \rightarrow Y$  и  $X \rightarrow Z$ , то  $X \rightarrow YZ$ .

2. Правило **псевдотранзитивности**.

Если  $X \rightarrow Y$  и  $WY \rightarrow Z$ , то  $WX \rightarrow Z$

3. Правило **декомпозиции**. Если  $X \rightarrow Y$  и  $Z \subseteq Y$ , то  $X \rightarrow Z$

Правила объединения и декомпозиции порождают **важное следствие**:  
если  $A_1, A_2, \dots, A_n$  – атрибуты, то зависимость  $X \rightarrow A_1, A_2, \dots, A_n$  справедлива,  
если и только если зависимость  $X \rightarrow A_i$  справедлива для каждого  $i$ .

Перед обсуждением полноты аксиом определим замыкание множества атрибутов относительно множества функциональных зависимостей.

Пусть  $F$  – множество ФЗ на множестве атрибутов  $U$  и пусть  $X \subseteq U$ .

Тогда **множество атрибутов  $A$** , таких, что зависимость  $X \rightarrow A$  может быть выведена из  $F$  по аксиомам Армстронга называется **замыканием  $X$  относительно  $F$**  и обозначается  $X^+$ .

# Надежность и полнота аксиом (3)

## Лемма 3

Функциональная зависимость  $X \rightarrow Y$  следует из аксиом Армстронга, если и только если  $Y \subseteq X^+$ .

Доказательство.

Пусть  $Y \subseteq A_1, A_2, \dots, A_n$ . И кроме того  $Y \subseteq X^+$ .

**Покажем**, что в этом случае  $X \rightarrow Y$  выводится с помощью аксиом Армстронга.

По определению  $X^+$  функциональные зависимости  $X \rightarrow A_i$  выводятся для каждого  $i$  с помощью аксиом Армстронга.

Из правила объединения (п.1 лемма 2) следует, что  $X \rightarrow Y$  выводится с помощью аксиом Армстронга.

**Наоборот**, предположим, что  $X \rightarrow Y$  следует из аксиом Армстронга.

Тогда для каждого  $A_i$  из  $Y$  по правилу декомпозиции (п.3 Лемма 2) имеет место зависимость  $X \rightarrow A_i$ .

Поэтому  $Y \subseteq X^+$ .

# Теорема надежности и полноты

**Теорема.** Аксиомы АРМСТРОНГА являются надежными и полными.

В лемме 1 показано, что аксиомы АРМСТРОНГА являются надежными.

Доказательство полноты ведется от противного. Самостоятельно разобрать доказательство. (Ульман. Стр. 159)

Следствия:

1.  $X^+$  - множество атрибутов  $A$ , таких, что  $X \rightarrow A$  логически следует из  $F$  (а также по аксиомам Армстронга)
2.  $F^+$  - множество зависимостей, следующих из  $F$  по аксиомам Армстронга (вместо логически следующих из  $F$ )

У нас появляется конструктивный алгоритм, позволяющий проверять является ли ФЗ  $X \rightarrow Y$  логически выводимой ( выводимой по аксиомам Армстронга ) из множества функциональных зависимостей  $F$  уже определенных на отношении

Таким образом, **можно проверить является ли некоторый набор атрибутов ключом отношения, или находится ли отношение в нормальной форме.**

# Декомпозиция отношений

Декомпозиция отношений – разбиение отношения на несколько других.

$R(A_1, A_2, \dots, A_n)$  разбивается на  $S(B_1, B_2, \dots, B_m)$  и  $T(C_1, C_2, \dots, C_k)$  если:

1.  $\{A_1, A_2, \dots, A_n\} = \{B_1, B_2, \dots, B_m\} \cup \{C_1, C_2, \dots, C_k\}$
2. Кортежи  $S$  являются проекцией кортежей  $R$  на атрибуты  $B_1, B_2, \dots, B_m$
3. Кортежи  $T$  являются проекцией кортежей  $R$  на атрибуты  $C_1, C_2, \dots, C_k$

# Отношение проекты

N_проекта	N_сотр	N_эт	Дата_нач	Дата_кон	Назв_проекта	N_отдела	Назв_отдела
12-08	1104	1	1.01.2012	31.03.2012	Лексический анализ	5	Разработки компиляторов
12-08	1105	1	1.01.2012	31.03.2012	Лексический анализ	5	Разработки компиляторов
12-09	1106	2	1.04.2012	30.06.2012	Синтаксический анализ	5	Разработки компиляторов
12-09	1107	2	1.04.2012	30.06.2012	Синтаксический анализ	5	Разработки компиляторов

N_проекта	N_сотр	N_эт	Дата_нач	Дата_кон	N_отдела	Назв_отдела
12-08	1104	1	1.01.2012	31.03.2012	5	Разработки компиляторов
12-08	1105	1	1.01.2012	31.03.2012	5	Разработки компиляторов
12-09	1106	2	1.04.2012	30.06.2012	5	Разработки компиляторов
12-09	1107	2	1.04.2012	30.06.2012	5	Разработки компиляторов

N_проекта	Назв_проекта
12-08	Лексический анализ
12-09	Синтаксический анализ



# Нормальные формы отношений (1)

- Отношение R находится в I НФ, если все атрибуты R определены на простых доменах.

Выполняется по определению.

- Отношение R находится в II НФ, если оно находится в I НФ и все неключевые атрибуты R **функционально полно зависят** от ключа.

Множество атрибутов X **функционально полно зависит** от ключа K, если

- ❖ имеется зависимость  $K \rightarrow X$

- ❖  $\neg \exists Z \subset K$ , что  $Z \rightarrow X$

Проект(N\_проекта, N\_сотр, N\_эт,

Дата\_нач, Дата\_кон, Назв\_проекта, N\_отдела, Назв\_отдела)

ФЗ:

1.  $N\_проекта, N\_сотр, N\_эт \rightarrow Дата\_нач, Дата\_кон, Назв\_проекта, N\_отдела, Назв\_отдела$

2. **N\_проекта** **Назв\_проекта** %нарушает II НФ%

Нарушение II НФ **НЕ ВОЗМОЖНО**, если **ключ** состоит из **одного атрибута**

## Нормальные формы отношений (2)

- Отношение R находится в III НФ, если оно находится в II НФ и все неключевые атрибуты R **нетранзитивно зависят** от ключа.

Проект(N\_проекта, N\_сотр, N\_эт,  
Дата\_нач, Дата\_кон, Назв\_проекта, N\_отдела, Назв\_отдела)

**Ф3:**

$N\_проекта, N\_сотр, N\_эт \rightarrow Дата\_нач, Дата\_кон, Назв\_проекта, N\_отдела, Назв\_отдела$

$N\_проекта \rightarrow Назв\_проекта$

$N\_отдела \rightarrow Назв\_отдела$

Транзитивность

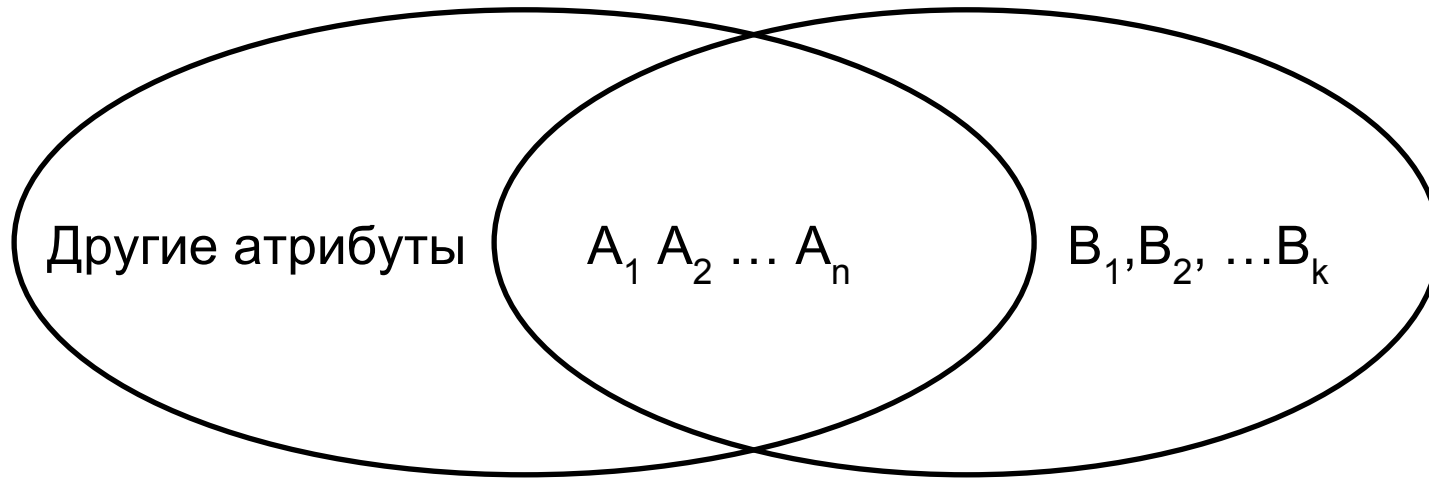
$N\_проекта, N\_сотр, N\_эт \rightarrow N\_отдела \rightarrow Назв\_отдела$

# Нормальные формы Бойса-Кодда (BCNF)

- Тривиальные ФЗ. Говорят, что функциональная зависимость  $A_1 A_2 \dots A_n \rightarrow B$  является тривиальной, если атрибут  $B$  совпадает с любым из атрибутов  $A_i, i = 1, \dots, n$
- Нетривиальные ФЗ.  $A_1 A_2 \dots A_n \rightarrow B_1, B_2, \dots, B_k$  называется нетривиальной, если по меньшей мере один из атрибутов  $B_i$  не является элементом множества  $\{A_1, A_2, \dots, A_n\}$
- Отношение  $R$  удовлетворяет **BCNF**, т. и т. т., когда для любой нетривиальной зависимости  $R$  вида  $A_1 A_2 \dots A_n \rightarrow B$  множество  $\{A_1, A_2, \dots, A_n\}$  образует суперключ для  $R$ .

# Приведение к нормальным формам

- Поиск нетривиальных функциональных зависимостей  $A_1 A_2 \dots A_n \rightarrow V_1, V_2, \dots V_k$  нарушающих НФ.
- Разбиение на 2 отдельных отношения



- Исходное отношение должно восстанавливаться из 2-ух получившихся отношений с помощью **НАТУРАЛЬНОГО СОЕДИНЕНИЯ**

# Отношение проекты

Проекты(N\_проекта, N\_сотр, N\_эт, Дата\_нач, Дата\_кон, Назв\_проекта, N\_отдела, Назв\_отдела)

Ф3:

1. N\_проекта, N\_сотр, N\_эт → Дата\_нач, Дата\_кон, Назв\_проекта, N\_отдела, Назв\_отдела

2. N\_проекта → Назв\_проекта, Дата\_нач, Дата\_кон

3. N\_отдела → Назв\_отдела



Проекты(N\_проекта, N\_сотр, N\_эт, Дата\_нач, Дата\_кон, Назв\_проекта, N\_отдела, Назв\_отдела)

1. По Ф3 2 разбиваем отношение ПРОЕКТЫ на 2 отношения (приводим к II НФ):

Проекты (N\_проекта, Назв\_проекта, Дата\_нач, Дата\_кон )

Участие (N\_проекта, N\_сотр, N\_эт, N\_отдела, Назв\_отдела )



2. По Ф3 3 разбиваем отношение УЧАСТИЕ на 2 отношения (приводим к III НФ):

Проекты (N\_проекта, Назв\_проекта , Дата\_нач, Дата\_кон )

Отделы (N\_отдела, Назв\_отдела )

Сотрудники-проекты (N\_проекта, N\_сотр, N\_эт, N\_отдела)

ВСЕ ?

# Схема БД в нормальной форме

N_проекта	Назв_проекта	Дата_нач	Дата_кон
12-08	Лексический анализ	1.01.2012	31.03.2012
12-09	Синтаксический анализ	1.04.2012	30.06.2012

N_проекта	N_сотр	N_эт	N_отдела
12-08	1104	1	5
12-08	1105	1	5
12-09	1106	2	5
12-09	1107	2	5

N_отдела	Назв_отдела
5	Разработки компиляторов

# Отношение Сессия

Сессия (N\_зач, N\_пред, Фам, Имя, Отч, Назв\_предмета, Дата, Оценка)

ФЗ:

N\_зач, N\_пред → Фам, Имя, Отч, Назв\_предмета, Дата, Оценка

N\_зач → Фам, Имя, Отч

N\_пред → Назв\_предмета

По ФЗ N\_зач → Фам, Имя, Отч

разбиваем отношение Сессия на 2 отношения ( **первый шаг к II НФ** ):

Студент (N\_зач, Фам, Имя, Отч )

Сессия (N\_зач, N\_пред, Назв\_предмета, Дата, Оценка)

По ФЗ N\_пред → Назв\_предмета

разбиваем отношение Сессия на 2 отношения ( **второй шаг к II НФ** ):

Итоговая схема ( в III НФ ) почему ?

Студент (N\_зач, Фам, Имя, Отч )

Предмет (N\_пред, Назв\_предмета)

Сессия (N\_зач, N\_пред, Дата, Оценка)