



Проектирование баз данных

Цели проектирования

- ◆ Хранение всех необходимых данных
- ◆ Исключение избыточности
- ◆ Минимизация количества отношений
- ◆ Нормализация отношений для решения проблем, связанных с обновлением и удалением данных

Как определить какие данные нужны?

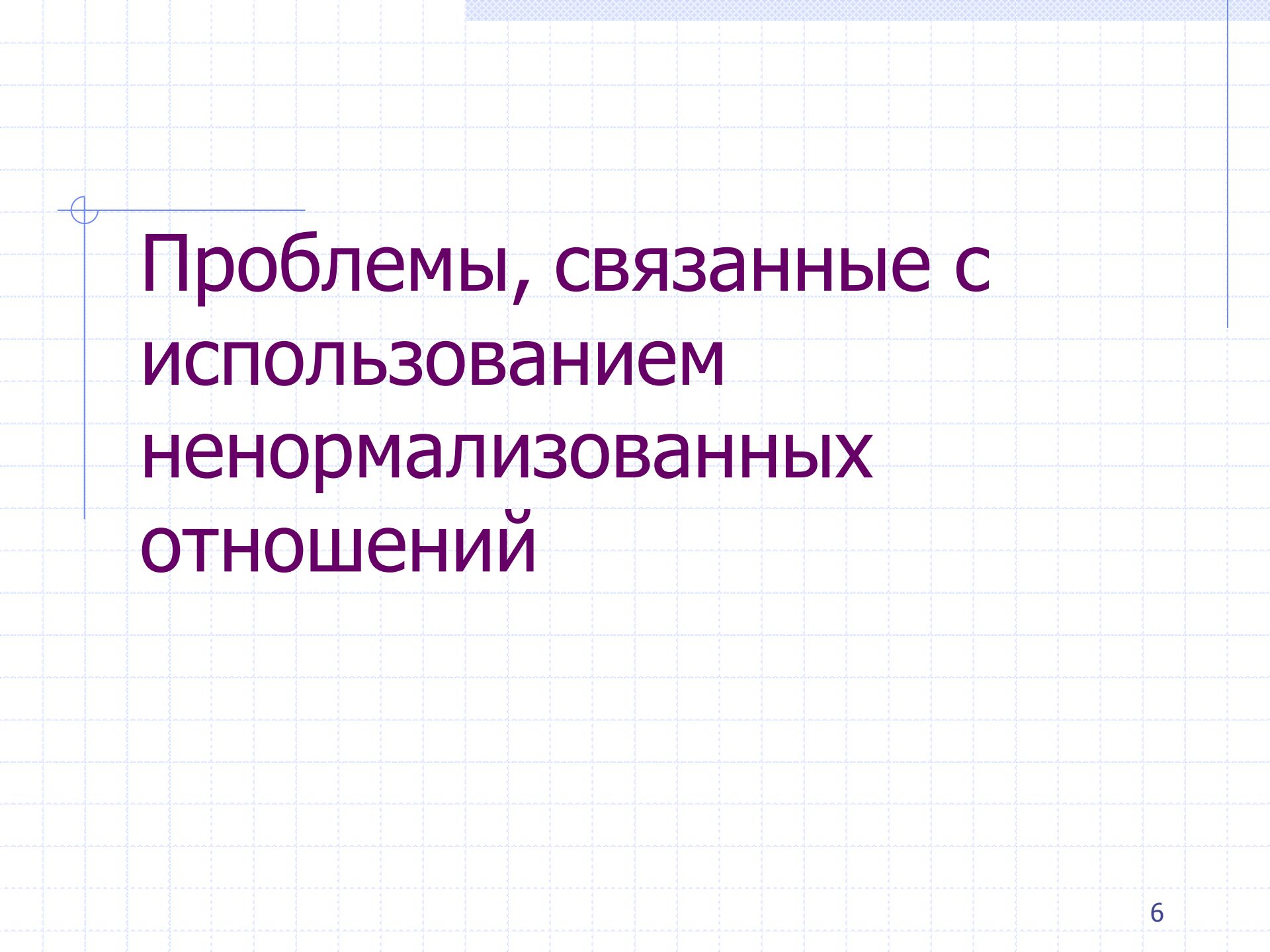
- ◆ Моделируем предметную область
- ◆ Рассматриваем известные данные, связи между ними, процессы (бизнес-процессы)
- ◆ Разделяем на подзадачи
- ◆ Используется UML и case-средства

Откуда берется избыточность?

- ◆ Одни и те же данные в разных задачах/подсистемах
- ◆ Дублирование информации в рамках одной задачи

Противоречие

- ◆ Минимизация количества отношений
- ◆ Нормализация отношений для решения проблем, связанных с обновлением и удалением данных



Проблемы, связанные с использованием ненормализованных отношений

Пример таблицы для базы данных «Тьютор»

Исходная таблица

студ#	ФИО	комн#	тел#	Курс	Предм	Балл
m3215	Иванов И.И.	120DH	2136123	PM16	MA84	12.5
					AG85	10.9
					PL83	13.7
i2345	Петров П.П.	238MP	2345671	IT211	MA84	10.7
					PL84	14.7
f1234	Сидоров С.С.	120DH	2136123	AS12	TP12	9.0

Это еще не отношение !!!

Пример отношения для базы данных «Тьютор»

студ#	ФИО	комн#	тел#	Курс	Предм	Балл
m3215	Иванов И.И.	120DH	2136123	PM16	MA84	12.5
m3215	Иванов И.И.	120DH	2136123	PM16	AG85	10.9
m3215	Иванов И.И.	120DH	2136123	PM16	PL83	13.7
i2345	Петров П.П.	238MP	2345671	IT211	MA84	10.7
i2345	Петров П.П.	238MP	2345671	IT211	PL84	14.7
f1234	Сидоров С.С.	120DH	2136123	AS12	TP12	9.0

Проблема вставки

студ#	ФИО	КОМН#	тел#	Курс	Предм	Балл
m3215	Иванов И.И.	120DH	2136123	PM16	MA84	12.5
m3215	Иванов И.И.	120DH	2136123	PM16	AG85	10.9
m3215	Иванов И.И.	120DH	2136123	PM16	PL83	13.7
i2345	Петров П.П.	238MP	2345671	IT211	MA84	10.7
i2345	Петров П.П.	238MP	2345671	IT211	PL84	14.7
f1234	Сидоров С.С.	120DH	2136123	AS12	TP12	9.0
m3456	Егоров Е.Е.	121MP	2123341	PM11		0.0

```
insert into TUTOR values('m3456','Егоров Е.Е.',  
                          '121MP','2123341','PM11','',0)
```

```
select ФИО from TUTOR where Балл<5
```

Проблема обновления

студ#	ФИО	комн#	тел#	Курс	Предм	Балл
m3215	Иванов И.И.	120DH	2136123	PM16	MA84	12.5
m3215	Иванов И.И.	120DH	2136123	PM16	AG85	10.9
m3215	Иванов И.И.	120DH	2136123	PM16	PL83	13.7
i2345	Петров П.П.	238MP	2345671	IT211	MA84	10.7
i2345	Петров П.П.	238MP	2345671	IT211	PL84	14.7
f1234	Сидоров С.С.	120DH	2177777	AS12	TP12	9.0

**update TUTOR set тел#='2177777' where
ФИО='Сидоров С.С.'**

select тел# from TUTOR where комн# = '120DH'

Проблема удаления

студ#	ФИО	комн#	тел#	Курс	Предм	Балл
m3215	Иванов И.И.	120DH	2136123	PM16	MA84	12.5
m3215	Иванов И.И.	120DH	2136123	PM16	AG85	10.9
m3215	Иванов И.И.	120DH	2136123	PM16	PL83	13.7
i2345	Петров П.П.	238MP	2345671	IT211	MA84	10.7
i2345	Петров П.П.	238MP	2345671	IT211	PL84	14.7
f1234	Сидоров С.С.	120DH	2177777	AS12	TP12	9.0

delete from TUTOR

where ФИО='Сидоров С.С.' and Предм='TP12'

select * from TUTOR where ФИО='Сидоров С.С.'

Как решить проблемы?

- ◆ Может быть помогут NULL
- ◆ Можно более жестко контролировать операции в приложении
- ◆ Выделить часть информации в отдельную таблицу
 - Какую информацию?
 - Достаточно ли будет еще одной таблицы или нужно больше?
 - Возможно ли будет восстановить исходную таблицу после этого?

Нормализация

- ◆ Декомпозиция отношения на два или более отношений с целью разрешения проблем обновления и удаления данных
- ◆ Декомпозиция без потерь – операция join всегда позволит восстановить исходную таблицу
- ◆ Нормальная форма – способ классификации таблиц на основе содержащихся в них функциональных зависимостей (functional dependencies, FD)

Функциональные зависимости

- ◆ ФЗ устанавливает связь между двумя непересекающимися подмножествами атрибутов внутри одной переменной-отношения
 - ФЗ могут быть определены всегда, такие ФЗ можно считать ограничениями целостности
 - ФЗ могут определяться для конкретного момента времени (для конкретных кортежей)

Функциональные зависимости

$A \rightarrow B$



А и В могут быть множествами из атрибутов

Если имеется ФЗ $\{X\} \rightarrow \{Y\}$, то X – детерминант, Y – зависимая часть

Функциональные зависимости

- ◆ Если X является потенциальным (возможным) ключом, то все не входящие в X атрибуты должны быть функционально зависимы от X
- ◆ Если переменная-отношение R удовлетворяет ФЗ: $\{A\} \rightarrow \{B\}$ и A не является потенциальным ключом, то R будет характеризоваться некоторой избыточностью

Функциональные зависимости

- ◆ Количество определенных для переменной-отношения R ФЗ F может быть достаточно велико
- ◆ Контроль такого большого количества ограничений целостности может быть затруднен
- ◆ Следует минимизировать количество ФЗ S , но так, чтобы любая ФЗ из F могла быть заменена на ФЗ из S

Функциональные зависимости

◆ Сокращение множества ФЗ

- Функциональная зависимость *тривиальна* тогда и только тогда, когда ее правая (зависимая) часть является подмножеством ее левой части (детерминанта)
- Будем рассматривать только нетривиальные ФЗ

◆ Замыкание множества ФЗ

- Если S – множество ФЗ переменной R , то все ФЗ, которые следуют из S , образуют замыкание S^+

Функциональные зависимости

◆ Аксиомы Армстронга

R – кортежная переменная

$\{A\}, \{B\}, \{C\}, \{D\}$ – подмножества атрибутов в R

A1. *Рефлексивность* если $B \subset A$, то $A \rightarrow B$

A2. *Дополнение* если $A \rightarrow B$, то $\{A \cup C\} \rightarrow \{B \cup C\}$

A3. *Транзитивность* если $A \rightarrow B$ и $B \rightarrow C$, то $A \rightarrow C$

Данный набор правил является *полным* и *непротиворечивым*

Функциональные зависимости

◆ Дополнительные правила

Д1. Самоопределение $A \rightarrow A$

Д2. Декомпозиция если $A \rightarrow \{B \cup C\}$, то $A \rightarrow B$ и $A \rightarrow C$

Д3. Объединение если $A \rightarrow B$ и $A \rightarrow C$, то $A \rightarrow \{B \cup C\}$

Д4. Композиция

если $A \rightarrow B$ и $C \rightarrow D$, то $\{A \cup C\} \rightarrow \{B \cup D\}$

◆ Общая теорема объединения

T0. Если $A \rightarrow B$ и $C \rightarrow D$, то $\{A \cup C \setminus B\} \rightarrow \{B \cup D\}$

Пример

$R(a,b,c,d,e,f)$

ФЗ:

$\{a\} \rightarrow \{b,c\}$

$\{b\} \rightarrow \{e\}$

$\{c,d\} \rightarrow \{e,f\}$

есть ли ФЗ $\{a,d\} \rightarrow \{f\}$?

$\{a\} \rightarrow \{b,c\}$

Д2

$\{a\} \rightarrow \{c\}$ $\{a\} \rightarrow \{b\}$

А2

$\{a,d\} \rightarrow \{c,d\}$

А3

$\{a,d\} \rightarrow \{e,f\}$

Д2

$\{a,d\} \rightarrow \{f\}$

Какие ФЗ важны?

- ◆ Является ли детерминант потенциальным ключом?

Пусть Z – интересующее нас множество атрибутов

Обозначим $S^+ \subset S$, которое состоит из всех функциональных зависимостей с множеством атрибутов Z

Обозначим через Z^+ множество атрибутов, которые функционально зависимы от Z (замыкание множества атрибутов)

Замыкание множества атрибутов

- ◆ ФЗ $X \rightarrow Y$ будет принадлежать замыканию S^+ , если $Y \subset X^+$
- ◆ Z является возможным (потенциальным) ключом R , если Z^+ содержит все атрибуты R

Замыкание множества атрибутов

Алгоритм

$$Z^+ = Z$$

повторять пока Z^+ изменяется

для каждой ФЗ $X \rightarrow Y$

если $X \in Z^+$, то добавить в Y

$$Z^+ = Z^+ \cup Y$$

Пример

$R(a,b,c,d,e,f)$

$\{a\} \rightarrow \{b,c\}$

$\{e\} \rightarrow \{c,f\}$

$\{b\} \rightarrow \{e\}$

$\{c,d\} \rightarrow \{e,f\}$

Найти замыкание для множества атрибутов $\{a,b\}$

Пример

$$Z+ = Z = \{a,b\}$$

$$\{a\} \rightarrow \{b,c\}$$

$$\{e\} \rightarrow \{c,f\}$$

$$\{b\} \rightarrow \{e\}$$

$$\{c,d\} \rightarrow \{e,f\}$$

$$Z+ = Z+ \cup \{b,c\} = \{a,b,c\}$$

$$Z+ = Z+ \cup \{e\} = \{a,b,c,e\}$$

Пример

$$Z+ = \{a, b, c, e\}$$

$$\{a\} \rightarrow \{b, c\}$$

$$\{e\} \rightarrow \{c, f\}$$

$$\{b\} \rightarrow \{e\}$$

$$\{c, d\} \rightarrow \{e, f\}$$

$$Z+ = Z+ \cup \{c, f\} = \{a, b, c, e, f\}$$

Пример

$Z_+ = \{a, b, c, e, f\}$

$\{a\} \rightarrow \{b, c\}$

$\{e\} \rightarrow \{c, f\}$

$\{b\} \rightarrow \{e\}$

$\{c, d\} \rightarrow \{e, f\}$

замыкание $Z_+ = \{a, b, c, e, f\}$

Пример

$$\mathbb{Z}^+ \neq \mathbb{R}$$

значит, \mathbb{Z} не является возможным ключом

Показать, что возможным ключом является
 $\{a, d\}$