

Базы данных

Шехтман В. Е.

Лекция 5. Нормализация
реляционных баз данных I

Степени нормализации

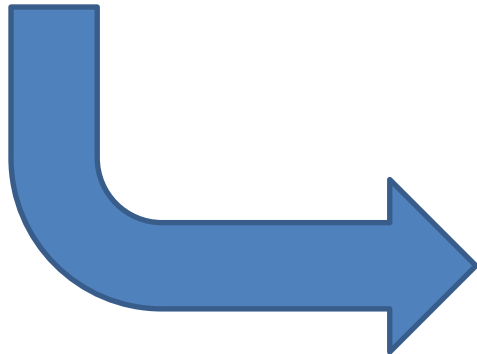
- 1 НФ – все реляционные таблицы (атрибуты таблицы не содержат таблиц (“плоские”))
- 2 НФ
- 3 НФ
- НФБК
- 4НФ
- 5НФ

1 НФ?

Студент	Средн Балл	Дисциплины
Иван	3.7	Математика
		Английский
		Физкультура
Кирилл	4.9	Физкультура
		Пение

1 НФ

Студент	Средн Балл	Дисциплины
Иван	3.7	Математика
		Английский
		Физкультура
Кирилл	4.9	Физкультура
		Пение



Студент	СреднБалл
Иван	3.7
Кирилл	4.9

Дисциплина
Математика
Физика
Английский
Пение
Физкультура
Культурология

Студент	Дисциплина
Иван	Математика
Иван	Английский
Иван	Физкультура
Кирилл	Пение

Конструирование реляционных схем

- Концептуальная модель
- Реляционная модель
- Нормализация (исключение *аномалий*)





Без нормализации в реляционной схеме возможны проблемы

- Избыточность (повторение одних и тех же сведений)
- Аномалия вставки (нельзя ввести данные об одном объекте без того, чтобы не вводить данные о другом, хотя они независимы)
- Аномалия изменения (приходится изменить несколько записей, а не одну)
- Аномалия удаления (можно потерять данные не желая того)

Пример плохой схемы

<u>Имя</u>	Таб №	<u>Тел</u>	Предприятие
Иван	1001	530909	ЗСМК
Иван	1001	340087	ЗСМК
Анна	7318	743535	НКМК

Избыточность – повторение имени и Таб№

Аномалия вставки – нельзя ввести сведения о сотруднике, если у него нет телефона (или он не известен)

Аномалия изменения - Иван переходит с ЗСМК на КФЗ, что требует изменения нескольких записей

Аномалия удаления – если Анна отказывается от телефона, то где она работает?

Декомпозиция плохой схемы

Имя	Таб №	Предприятие
Иван	1001	ЗСМК
Анна	7318	НКМК

Таб №	Тел
1001	530909
1001	340087
7318	743535

Данные не повторяются
Легко перевести Ивана на КЗФ
Легко удалить все телефоны Анны

Аномалий больше нет

Функциональные зависимости

ФЗ $A \rightarrow B$ означает, что

если два кортежа совпадают в атрибуте A , то они совпадают и в атрибуте B

$A_1, A_2, \dots, A_n \rightarrow B_1, B_2, \dots, B_m$

Пример:

ИНН предприятия \rightarrow Предприятие

Пример ФЗ

<u>Таб№</u>	Имя	Тел	Должность
12345	Иван	543262	Бухгалтер
34587	Петр	987201	Продавец
83746	Антон	987201	Продавец
98271	Лариса	543262	Юрист

Таб№ -> Имя, Тел, Должность

Должность -> Тел

~~Тел -> Должность~~

Даны 4 карты – с лицевой стороны число, на рубашке – один из двух цветов.

Гипотеза ФЗ: если на лицевой стороне чётное число, то её рубашка – голубая.

Какие карты минимально необходимо перевернуть для проверки гипотезы?



2 НФ

Отношение в 2НФ если неключевые атрибуты зависят от всех атрибутов ключа (а не от части его атрибутов)

<u>Магазин</u>	<u>Товар</u>	Цена	НДС	Продавец	Директор
Лента	Гречка	50	0	Иванов	Васюткин
Метро	Гречка	47	0	Петров	Сергуткин
Мосмарт	Арбуз	20	10	Сидоров	Кузькин
Лента	Лимон	75	10	Иванцов	Васюткин

Какая НФ?

2 НФ

<u>Товар</u>	<u>НДС</u>
Гречка	0
Арбуз	10
Лимон	10

<u>Магазин</u>	<u>Товар</u>	<u>Цена</u>
Лента	Гречка	50
Метро	Гречка	47
Мосмарт	Арбуз	20
Лента	Лимон	75

...

Еще пример ФЗ

Продукт(Название, Категория, Отдел, Цвет, Цена)

Название -> Цвет

Категория -> Отдел

Цвет, Категория -> Цена

следовательно

Название, Категория -> Цена

Почему?

Правила Армстронга 1

- Разбиение и объединение

$A_1, A_2, \dots, A_n \rightarrow B_1, B_2, \dots, B_m$

\Leftrightarrow

$A_1, A_2, \dots, A_n \rightarrow B_1$

$A_1, A_2, \dots, A_n \rightarrow B_2$

$A_1, A_2, \dots, A_n \rightarrow B_m$

Правила Армстронга 2

- Тривиальная ФЗ

$A_1, A_2, \dots, A_n \rightarrow A_i$

Правила Армстронга 3

- Транзитивность

$A_1, A_2, \dots, A_n \rightarrow B_1, B_2, \dots, B_m$

$B_1, B_2, \dots, B_m \rightarrow C_1, C_2, \dots, C_k$

\Rightarrow

$A_1, A_2, \dots, A_n \rightarrow C_1, C_2, \dots, C_k$

Пример

Продукт(Название, Категория, Отдел, Цвет, Цена)

1. Название -> Цвет

2. Категория -> Отдел

3. Цвет, Категория -> Цена

=>

4. Название, Категория -> Название *Тривиальная ФЗ*

5. Название, Категория -> Цвет *Транзит ФЗ 4, ФЗ 1*

6. Название, Категория -> Категория *Тривиальная ФЗ*

7. Название, Категория -> Цвет, Категория *Комбинация ФЗ 5, ФЗ 6*

8. Название, Категория -> Цена *Транзит ФЗ 7, ФЗ 3*

Замыкание множества атрибутов

A^+ = множество атрибутов B , для которых $A \rightarrow B$

Пример:

Продукт(Название, Категория, Отдел, Цвет, Цена)

Название \rightarrow Цвет

Категория \rightarrow Отдел

Цвет, Категория \rightarrow Цена

$\{\text{Название}\}^+ = \{\text{Название, Цвет}\}$

$\{\text{Название, Категория}\}^+ = \{\text{Название, Категория, Отдел, Цвет, Цена}\}$

$\{\text{Цвет}\}^+ = \{\text{Цвет}\}$

Алгоритм вычисления замыкания

$X = \{A_1, A_2, \dots, A_n\}$

Пока X^+ не перестанет изменяться:

Если $B_1, B_2, \dots, B_m \rightarrow C,$

и B_1, B_2, \dots, B_m входят в $X,$

то добавить C к X^+

Применение замыканий

Для определения всех ФЗ!

Чтобы проверить имеется ли $X \rightarrow A$, следует

1. Вычислить X^+
2. Проверить входит ли A в X^+

Пример вычисления замыкания

$R(A, B, C, D, E, F)$

$A, B \rightarrow C$ $A, D \rightarrow E$ $B \rightarrow D$ $A, F \rightarrow B$

$\{A, B\}^+ = \{A, B, C, D, E\}$

$\{A, F\}^+ = \{A, F, B, C, D, E\}$