

Функциональные
зависимости
Нормализация отношений

Пример плохого отношения

Название фирмы	Адрес	Телефон	Товар	Цена (руб.)
Чижиков & Со	Уткин проезд, 5	999-99-99	Винт большой	3
Чижиков & Со	Уткин проезд, 5	999-99-99	Винт маленький	5
Винты-гайки	Ул.Ленина, 1	333-33-33	Винт	4
Чижиков & Со	Уткин проезд, 5	999-99-99	Гайка	1
Стройтовары	Стройбаза № 1	444-44-44	Саморез	2
Чижиков & Со	Уткин проезд, 5	999-99-99	Саморез	3

Фирма-товар

Недостатки

- Избыточность
- Аномалии изменения
- Аномалии удаления
- Аномалии добавления

Решение - декомпозиция

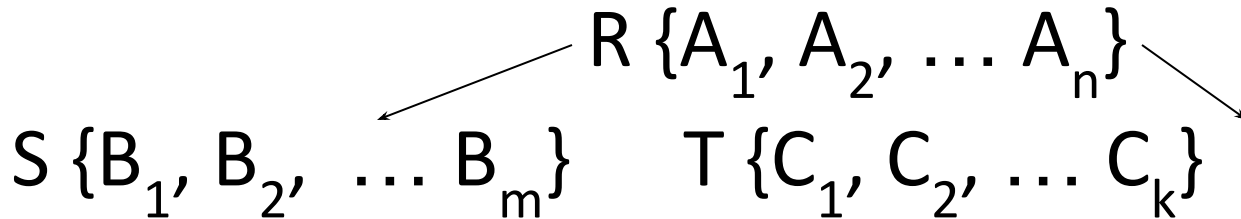
Название фирмы	Адрес	Телефон
Чижиков & Со	Уткин проезд, 5	999-99-99
Винты-гайки	Ул.Ленина, 1	333-33-33
Стройтовары	Стройбаза № 1	444-44-44

Фирма

Название фирмы	Товар	Цена (руб.)
Чижиков & Со	Винт большой	3
Чижиков & Со	Винт маленький	5
Винты-гайки	Винт	4
Чижиков & Со	Гайка	1
Стройтовары	Саморез	2
Чижиков & Со	Саморез	3

Товар

Декомпозиция



$$1) \{A_1, A_2, \dots, A_n\} = \{B_1, B_2, \dots, B_m\} \cup \{C_1, C_2, \dots, C_k\}$$

$$2) S = \pi_{B_1, B_2, \dots, B_m}(R)$$

$$3) T = \pi_{C_1, C_2, \dots, C_k}(R)$$

Ограничения на значения:

- семантические, т.е. корректность отдельных значений (год рождения больше нуля);
- ограничения на значения, которые зависят только от равенства или неравенства значений (совпадают ли компоненты двух кортежей); наиболее важные ограничения называются функциональной зависимостью.

Функциональные зависимости

- $R \{A_1, A_2, \dots, A_n\}$
- $X, Y \subset \{A_1, A_2, \dots, A_n\}$
- $X \rightarrow Y$ если любому значению X соответствует в точности одно значение Y
- $X \rightarrow Y \Leftrightarrow |\pi_Y(\sigma_{X=x}(R))| \leq 1$
- Название фирмы \rightarrow Адрес, телефон.
- Название фирмы, товар \rightarrow Цена

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

ФЗ бывают:

- Тривиальные

$$\{B_1, B_2, \dots, B_m\} \subset \{A_1, A_2, \dots, A_n\}$$

- Нетривиальные

$$\{B_1, B_2, \dots, B_m\} \not\subset \{A_1, A_2, \dots, A_n\}$$

$$\{A_1, A_2, \dots, A_n\} \cap \{B_1, B_2, \dots, B_m\} \neq \emptyset$$

- Полностью нетривиальные

$$\{A_1, A_2, \dots, A_n\} \cap \{B_1, B_2, \dots, B_m\} = \emptyset$$

Ключ

- Ключ – набор атрибутов, который функционально определяет все остальные
- F – множество функциональных зависимостей, заданных на отношении R
- $A \rightarrow C$ называется транзитивной, если существует такой атрибут B , что имеются функциональные зависимости $A \rightarrow B$ и $B \rightarrow C$ и отсутствует функциональная зависимость $C \rightarrow A$

Замыкание множества атрибутов

- $R \{A_1, A_2, \dots, A_n\}$
- $\{B_1, B_2, \dots, B_m\} \subset \{A_1, A_2, \dots, A_n\}$
- F – МН-ВО ФЗ
- $Z = \{B_1, B_2, \dots, B_m\}^+$
 - $Z_0 := \{B_1, B_2, \dots, B_m\}$
 - $B_i B_j \rightarrow C$
 - $Z_1 := Z_0 \cup C$
- $\{B_1, B_2, \dots, B_m\}^+ = \{A_1, A_2, \dots, A_n\} \Rightarrow$
 $\{B_1, B_2, \dots, B_m\}$ - КЛЮЧ

Пример

- $R \{A, B, C, D, E, F\}$
- $S = \{A \rightarrow D, AB \rightarrow E, BF \rightarrow E, CD \rightarrow F, E \rightarrow C\}$
- $\{AE\}^+ ?$

Пример

- $R \{A, B, C, D, E, F\}$
- $S = \{A \rightarrow D, AB \rightarrow E, BF \rightarrow E, CD \rightarrow F, E \rightarrow C\}$
- $\{AE\}^+ = ACDEF$

Аксиомы Армстронга

- если $B \subset A$, то $A \rightarrow B$
рефлексивность;
- если $A \rightarrow B$, то $AC \rightarrow BC$
пополнение;
- если $A \rightarrow B$ и $B \rightarrow C$, то $A \rightarrow C$
транзитивность.

Правила вывода (из аксиом Армстронга)

- **1. Объединение**

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

$X \rightarrow Y + A2 = X \rightarrow XY$, $X \rightarrow Z + A2 = YX \rightarrow YZ + A3 = X \rightarrow YZ$

- **2. Псевдотранзитивность**

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

$X \rightarrow Y + A2 = WX \rightarrow WY$. $WY \rightarrow Z + A3 = WX \rightarrow Z$.

- **3. Декомпозиция**

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

$A1 + A3$.

Замыкание множества функциональных зависимостей

F^+ - множество всех зависимостей, которые можно вывести из F , называют замыканием множества ФЗ F

Любое множество функциональных зависимостей, из которого можно вывести все остальные ФЗ, называется базисом

Если ни одно из подмножеств базиса базисом не является, то такой базис минимален

Замыкание множества функциональных зависимостей

- $R \{A_1, A_2, \dots, A_n\}$
- F – МН-ВО ФЗ
- $B_1, B_2, \dots, B_m \rightarrow C$
- $(B_1, B_2, \dots, B_m \rightarrow C) \in F^+$, if
 $C \in \{B_1, B_2, \dots, B_m\}^+$

Пример:

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

$AB \rightarrow C, C \rightarrow D, D \rightarrow A$

- Найти все нетривиальные ФЗ, которые следуют из заданных
- Возможные ключи

Покрытие множества функциональных зависимостей

- Множество ФЗ F_2 называется покрытием множества ФЗ F_1 , если любая ФЗ, выводимая из F_1 , выводится также из F_2 .
- $F_1^+ \subseteq F_2^+$
- F_1 и F_2 называются эквивалентными, если $F_1^+ = F_2^+$.

Минимальное покрытие множества функциональных зависимостей

- правая часть любой ФЗ из F является множеством из одного атрибута (простым атрибутом);
- удаление любого атрибута из левой части любой ФЗ приводит к изменению замыкания F^+ ;
- удаление любой ФЗ из F приводит к изменению F^+ .

ДЕКОМПОЗИЦИЯ

- Декомпозиция – это разбиение на множества, может быть пересекающиеся, такие, что их объединение – это исходное отношение.
- Восстановить исходное отношение можно только естественным соединением.
- Говорят, что декомпозиция обладает свойством соединения без потерь, если для любого отношения

$$r = \pi_{R_1}(r) \bowtie \pi_{R_2}(r) \bowtie \dots \bowtie \pi_{R_n}(r).$$

А что происходит с зависимостями при декомпозиции?

- Можно определить $\pi_Z(F): X \rightarrow Y \text{ } XY \subseteq Z$
- Декомпозиция сохраняет множество зависимостей, если из объединения всех проекций зависимостей логически следует F .

Проектирование реляционных отношений

- 1 нормальная форма (НФ)– значения не являются множествами и кортежами.
- Атрибут называется первичным, если входит в состав любого возможного ключа.
- 2 нормальная форма – 1 НФ + любой атрибут, не являющийся первичным, полностью зависит от любого его ключа, но не от подмножества ключа.
- Фирма, Адрес, Телефон, Товар, Цена

3 НФ

- Транзитивная зависимость: пусть A, B, C – атрибуты, $A \rightarrow B, B \rightarrow C$, A не зависит от B и B не зависит от C . Тогда говорят, что C транзитивно зависит от A .
- 3 нормальная форма – если отношение находится во 2 нормальной форме и любой атрибут, не являющийся первичным, нетранзитивно зависит от любого возможного ключа.

Примеры:

- Универмаг, Товар, Номер отдела, Заведующий
- Город, Индекс, Адрес

Примеры:

- 3 нормальная форма – (Город, Индекс, Адрес)
- 2 нормальная форма, но не 3 нормальная форма – (Универмаг, Товар, Номер отдела, Заведующий)
- $UT \rightarrow H$, $UH \rightarrow Z$, ключ – UT .

НФ Бойса-Кодда

- Нормальная форма Бойса–Кодда – если $X \rightarrow A$, $A \notin X$, то $X \supseteq \text{ключ } R$.
- (Город, Индекс, Адрес) – 3 нормальная форма, но не форма Бойса–Кодда.
Если разобьем на две (Город, Индекс), (Индекс, Адрес), пропадает зависимость Город, Адрес \rightarrow Индекс.

НФ Бойса-Кодда

- (Город, Индекс, Адрес) – 3 нормальная форма, но не форма Бойса–Кодда.
Если разобьем на две (Город, Индекс),
(Индекс, Адрес), пропадает зависимость
Город, Адрес → Индекс.

Вывод:

- Каждая схема отношений может быть приведена к форме Бойса–Кодда, так что декомпозиция обладает свойством соединения без потерь.
- Любая схема может быть приведена к 3 нормальной форме с соединением без потерь и с сохранением функциональной зависимости.
- Но не всегда можно привести к форме Бойса–Кодда с сохранением функциональных зависимостей.

Шаги при декомпозиции

1. Находим минимальное покрытие множества функциональных зависимостей
2. Выделяем зависимость, нарушающую НФ
 $X \rightarrow Y$ (и нет атрибутов, зависящих от Y).
3. Находим зависимости с такой же левой частью.
 $X \rightarrow W, X \rightarrow Z$
4. Выделяем в отдельное отношение $XYWZ$

Пример

- S Студент
- G Группа
- H Время
- R Аудитория
- C Предмет
- T Преподаватель