

Введение в реляционные базы данных

Лекция 5: Функциональные зависимости

Артем Толканев

October 16, 2024

Хранение СУБД

На диске присутствует файл базы данных (или несколько)



Файл
БД



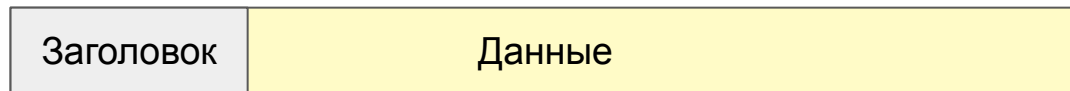
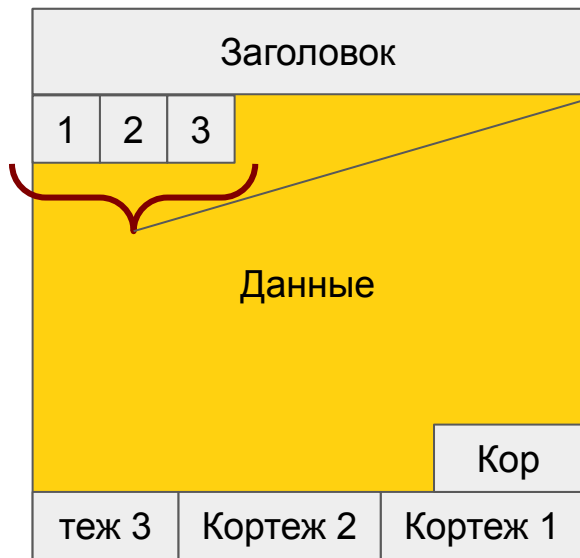
Файл базы данных состоит из нескольких частей, с которыми работает СУБД

MS SQL, Postgres
8KB

MySQL 16KB



Массив с
информацией о
расположении
кортежей



Заголовок	Данные
-----------	--------

book_id	name	theme_id	publishing_house_id
---------	------	----------	---------------------

id_данных	seq_id	data
1	0	Курс аналитической
1	1	геометрии
2	0	Война и Мир 1,2 то
2	1	м



На этой лекции

- Функциональные зависимости
- Нормализация



Отношения

Значение отношения (обозначим за r) - состоит из из заголовка и тела, где:

- **Заголовок** отношения r - представляют собой заголовок кортежа.
Отношение r имеет такие же атрибуты и такую же степень, как и заголовок.
- **Тело** отношения r представляет собой множество кортежей, имеющих один и тот же заголовок.

Кардинальность отношения r определяется как количество элементов этого множества.



Значение отношения (обозначим за r) - состоит из из заголовка и тела, где:

- **Заголовок** отношения r - представляют собой заголовок кортежа.

Отношение r имеет такие же атрибуты и такую же степень, как и заголовок.

- **Тело** отношения r представляет собой множество кортежей, имеющих один и тот же заголовок.

Кардинальность отношения r определяется как количество элементов этого множества.

В жизни определяется как количество уникальных значений в столбце к количеству значений в столбце



Значение отношения варианты:

1. Значение переменной отношения в определенный момент
2. Множество всех возможных значений, которые переменная отношения может принимать в разные моменты существования

language_id	language
0	русский
1	английский
2	французский
3	немецкий

language_id	language
0	русский
1	английский
2	французский
3	немецкий
2147483647	Язык лишайников



Функциональные зависимости

Рассмотрим отношение r и два произвольных подмножества X и Y из множества атрибутов этого отношения



Рассмотрим отношение r и два произвольных подмножества X и Y из множества атрибутов этого отношения

book_id	name	theme_id	publishing_house_id	year_of_publishing	language_id
200	Война и Мир 1,2 том	1	11	1 998	1
201	Война и Мир 3,4 том	1	11	1 998	1
202	Othello	2	12	2 005	2
203	Курс аналитической геометрии	5	13	2 005	1
204	Обломов	1	14	2 005	1
205	Капитанская Дочка	1	15	2 005	1
206	Общая физика	5	13	2 005	1
207	Дубровский	1	15	2 005	1
208	Анна Каренина	1	11	2 005	1
209	1984	2	12	2 005	2
210	Мартин Иден	2	16	2 005	1
211	Сердца Трех	2	16	2 007	1
212	Белый Клык	2	16	2 007	1
213	Три Сестры	1	14	2 007	1
214	Русские сказки	4	14	2 007	1
215	Курс аналитической геометрии	5	13	2 006	1



1. Значение переменной отношения в определенный момент

Рассмотрим отношение r и два произвольных подмножества X и Y из множества атрибутов этого отношения

$X = \{\text{publishing_house_id}\}$, $Y = \{\text{language_id}\}$

Тогда Y **функционально зависимо** от X , тогда и только тогда, когда каждое значение множества X связано точно с одним значением множества Y отношения r .

В символическом виде записывается как

$X \twoheadrightarrow Y$

, читается как “ X функционально определяет Y ”.



Рассмотрим отношение r и два произвольных подмножества X и Y из множества атрибутов этого отношения

book_id	name	theme_id	publishing_house_id	year_of_publishing	language_id
200	Война и Мир 1,2 том	1	11	1 998	1
201	Война и Мир 3,4 том	1	11	1 998	1
202	Othello	2	12	2 005	2
203	Курс аналитической геометрии	5	13	2 005	1
204	Обломов	1	14	2 005	1
205	Капитанская Дочка	1	15	2 005	1
206	Общая физика	5	13	2 005	1
207	Дубровский	1	15	2 005	1
208	Анна Каренина	1	11	2 005	1
209	1984	2	12	2 005	2
210	Мартин Иден	2	16	2 005	1
211	Сердца Трех	2	16	2 007	1
212	Белый Клык	2	16	2 007	1
213	Три Сестры	1	14	2 007	1
214	Русские сказки	4	14	2 007	1
215	Курс аналитической геометрии	5	13	2 006	1



1. Значение переменной отношения в определенный момент

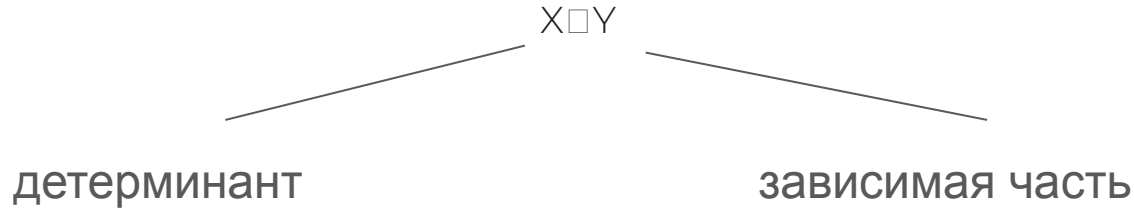
Рассмотрим отношение r и два произвольных подмножества X и Y из множества атрибутов этого отношения

Тогда Y **функционально зависимо** от X , тогда и только тогда, когда каждое значение множества X связано точно с одним значением множества Y отношения r .

Иначе говоря, если два кортежа отношения r совпадают по значению X , они совпадают и по значению Y .



1. Значение переменной отношения в
определенный момент



Тогда Y **функционально зависимо** от X , тогда и только тогда, когда каждое значение множества X связано точно с одним значением множества Y отношения r .



2. Множество всех возможных значений, которые переменная отношения может принимать в разные моменты существования

Рассмотрим отношение r и два произвольных подмножества X и Y из множества атрибутов этого отношения

Тогда Y **функционально зависимо** от X , тогда и только тогда, когда **для любого допустимого значения переменной отношения** каждое значение множества X связано точно с одним значением множества Y .



Рассмотрим отношение r и два произвольных подмножества X и Y из множества атрибутов S

book_id	name	theme_id	publishing_house_id	year_of_publishing	language_id
200	Война и Мир 1,2 том	1	11	1 998	1
201	Война и Мир 3,4 том	1	11	1 998	1
202	Othello	2	12	2 005	2
203	Курс аналитической геометрии	5	13	2 005	1
204	Обломов	1	14	2 005	1
205	Капитанская Дочка	1	15	2 005	1
206	Общая физика	5	13	2 005	1
207	Дубровский	1	15	2 005	1
208	Анна Каренина	1	11	2 005	1
209	1984	2	12	2 005	2
210	Мартин Иден	2	16	2 005	1
211	Сердца Трех	2	16	2 007	1
212	Белый Клык	2	16	2 007	1
213	Три Сестры	1	14	2 007	1
214	Русские сказки	4	14	2 007	1
215	Курс аналитической геометрии	5	13	2 006	1
216	Курс аналитической геометрии	5	13	2024	2

Потенциальный ключ

Пусть K - множество атрибутов переменной отношения r .

K является **потенциальным ключом** тогда и только тогда, когда это множество обладает следующими свойствами:

- **Уникальность.** Для любого допустимого значения r все кортежи имеют уникальное значение K
- **Несократимость.** Никакое подмножество K не обладает свойством уникальности



book_id	name	theme_id	publishing_house_id	year_of_publishing	language_id
200	Война и Мир 1,2 том	1	11	1 998	1
201	Война и Мир 3,4 том	1	11	1 998	1
202	Othello	2	12	2 005	2
203	Курс аналитической геометрии	5	13	2 005	1
204	Обломов	1	14	2 005	1
205	Капитанская Дочка	1	15	2 005	1
206	Общая физика	5	13	2 005	1
207	Дубровский	1	15	2 005	1
208	Анна Каренина	1	11	2 005	1
209	1984	2	12	2 005	2
210	Мартин Иден	2	16	2 005	1
211	Сердца Трех	2	16	2 007	1
212	Белый Клык	2	16	2 007	1
213	Три Сестры	1	14	2 007	1
214	Русские сказки	4	14	2 007	1
215	Курс аналитической геометрии	5	13	2 006	1
216	Курс аналитической геометрии	5	13	2024	2



book_id	name	theme_id	publishing_house_id	year_of_publishing	language_id
200	Война и Мир 1,2 том	1	11	1 998	1
201	Война и Мир 3,4 том	1	11	1 998	1
202	Othello	2	12	2 005	2
203	Курс аналитической геометрии	5	13	2 005	1
204	Обломов	1	14	2 005	1
205	Капитанская Дочка	1	15	2 005	1
206	Общая физика	5	13	2 005	1
207	Дубровский	1	15	2 005	1
208	Анна Каренина	1	11	2 005	1
209	1984	2	12	2 005	2
210	Мартин Иден	2	16	2 005	1
211	Сердца Трех	2	16	2 007	1
212	Белый Клык	2	16	2 007	1
213	Три Сестры	1	14	2 007	1
214	Русские сказки	4	14	2 007	1
215	Курс аналитической геометрии	5	13	2 006	1
216	Курс аналитической геометрии	5	13	2024	2



Пусть K - множество атрибутов переменной отношения r .

K является **потенциальным ключом** тогда и только тогда, когда это множество обладает следующими свойствами:

- **Уникальность.** Для любого допустимого значения r все кортежи имеют уникальное значение K
- **Несократимость.** Никакое подмножество K не обладает свойством уникальности

Надмножество потенциального ключа - суперключ (superset\superkey)



Аксиомы Армстронга

Пусть A , B и C - произвольные подмножества множества атрибутов заданной переменной отношени.

Тогда правила вывода определяются следующим:

- Правило **рефлексивности**. Если множество B является подмножеством множества A , то $A \sqsubseteq B$.
- Правило **дополнения**. Если $A \sqsubseteq B$, то $A \cup C \sqsubseteq B \cup C$.
- Правило **транзитивности**. Если $A \sqsubseteq B$ и $B \sqsubseteq C$, то $A \sqsubseteq C$.



Пусть A , B и C - произвольные подмножества множества атрибутов заданной переменной отношени.

тривиальная зависимость

Тогда правила вывода определяются следующим:

- Правило **рефлексивности**. Если множество B является подмножеством множества A , то $A \sqsubseteq B$.
- Правило **дополнения**. Если $A \sqsubseteq B$, то $A \cup C \sqsubseteq B \cup C$.
- Правило **транзитивности**. Если $A \sqsubseteq B$ и $B \sqsubseteq C$, то $A \sqsubseteq C$.



- Правило **рефлексивности**. Если множество В является подмножеством множества А, то $A \sqsubseteq B$.

A		B	
year_of_publishing	language_id	year_of_publishing	
2 007	1	2 007	
2 005	2	2 005	
1 998	1	1 998	
2 006	1	2 006	
2 005	1	2 005	



- Правило **дополнения**. Если $A \sqsubseteq B$, то $A \cup C \sqsubseteq B \cup C$.

book_id	name
207	Дубровский
205	Капитанская Дочка
203	Курс аналитической геометрии
215	Курс аналитической геометрии
210	Мартин Иден
204	Обломов

A

B

book_id	theme_id	name	theme_id
207	1	Дубровский	1
205	1	Капитанская Дочка	1
203	5	Курс аналитической геометрии	5
215	5	Курс аналитической геометрии	5
210	2	Мартин Иден	2
204	1	Обломов	1

A ∪ C

B ∪ C



Правило **транзитивности**. Если $A \sqsubseteq B$ и $B \sqsubseteq C$, то $A \sqsubseteq C$.

book_id	name
207	Дубровский
205	Капитанская Дочка
203	Курс аналитической геометрии
215	Курс аналитической геометрии
210	Мартин Иден
204	Обломов

A

name	theme_id
Дубровский	1
Капитанская Дочка	1
Курс аналитической геометрии	5
Мартин Иден	2
Обломов	1

B

book_id	theme_id
207	1
205	1
203	5
215	5
210	2
204	1

A

Аксиомы Армстронга



Пусть A , B и C - произвольные подмножества множества атрибутов заданной переменной отношени.

Дополнительно:

- Правило **самоопределения**. $A \sqsubseteq A$.
- Правило **декомпозиции**. Если $A \sqsubseteq B \cup C$, то $A \sqsubseteq B$ и $A \sqsubseteq C$.
- Правило **объединения**. Если $A \sqsubseteq B$ и $A \sqsubseteq C$, то $A \sqsubseteq B \cup C$.
- Правило **композиции**. Если $A \sqsubseteq B$ и $C \sqsubseteq D$, то $A \cup C \sqsubseteq B \cup D$.



Пусть A , B и C - произвольные подмножества множества атрибутов заданной переменной отношени.

Дополнительно

Теорема Дарвена (**общая теорема объединения**):

- Если $A \sqsubseteq B$ и $C \sqsubseteq D$, то $A \cup (C \setminus B) \rightarrow B \cup D$.



4. THE GENERAL UNIFICATION THEOREM

Theorem: $A \rightarrow B \ \& \ C \rightarrow D \implies A \cup (C - B) \rightarrow B \cup D$

Proof:

- | | | |
|-----|---|-----------------------------|
| <1> | $A \rightarrow B$ | (given) |
| <2> | $C \rightarrow D$ | (given) |
| <3> | $A \rightarrow B \cap C$ | (joint dependence, <1>) |
| <4> | $C - B \rightarrow C - B$ | (self-determination) |
| <5> | $A \cup (C - B) \rightarrow$
$(B \cap C) \cup (C - B)$ | (composition, <3> and <4>) |
| <6> | $A \cup (C - B) \rightarrow C$ | (simplifying <5>) |
| <7> | $A \cup (C - B) \rightarrow D$ | (transitivity, <6> and <2>) |
| <8> | $A \cup (C - B) \rightarrow B \cup D$ | (composition, <1> and <7>) |

This completes the proof. ■



Теорема Дарвена (общая теорема объединения):

- Если $A \sqsubseteq B$ и $C \sqsubseteq D$, то $A \cup (C \setminus B) \rightarrow B \cup D$.

name	theme_id
1984	2
Othello	2
Анна Каренина	1
Белый Клык	2
Война и Мир 1,2 том	1
Война и Мир 3,4 том	1
Дубровский	1
Капитанская Дочка	1
Курс аналитической геометрии	5

theme_id	publishing_house_id	year_of_publishing	language_id
2	12	2 005	2
2	12	2 005	2
1	11	2 005	1
2	16	2 007	1
1	11	1 998	1
1	11	1 998	1
1	15	2 005	1
1	15	2 005	1
5	13	2 005	1
5	13	2 006	1



Теорема Дарвена (**общая теорема объединения**):

- Если $A \sqsubseteq B$ и $C \sqsubseteq D$, то $A \cup (C \setminus B) \rightarrow B \cup D$.

name	publishing_house_id	year_of_publishing	language_id	theme_id
1984	12	2 005	2	2
Othello	12	2 005	2	2
Анна Каренина	11	2 005	1	1
Белый Клык	16	2 007	1	2
Война и Мир 1,2 том	11	1 998	1	1
Война и Мир 3,4 том	11	1 998	1	1
Дубровский	15	2 005	1	1
Капитанская Дочка	15	2 005	1	1
Курс аналитической геометрии	13	2 005	1	5
Курс аналитической геометрии	13	2 006	1	5



Множество всех функциональных зависимостей, которые следуют из данного множества функциональных зависимостей S , называют **замыканием** множества S , будем обозначать его, как S^+



Неприводимые множества зависимостей

Пусть S_1 и S_2 - два множества функциональных зависимостей.



Пусть $S1$ и $S2$ - два множества функциональных зависимостей.

Если любая функциональная зависимость, которая следует из множества зависимостей $S1$, следует также из множества зависимостей $S2$, то множество $S2$ называется **покрытием** для множества $S1$.



Пусть $S1$ и $S2$ - два множества функциональных зависимостей.

Если любая функциональная зависимость, которая следует из множества зависимостей $S1$, следует также из множества зависимостей $S2$, то множество $S2$ называется **покрытием** для множества $S1$.

Если множество $S1$ является покрытием $S2$ и множество $S2$ является покрытием $S1$, то множества $S1$ и $S2$ **эквивалентны**.



Множество функциональных зависимостей называется

неприводимым или минимальным

тогда и только тогда, когда оно обладает следующими свойствами:

- Правая часть каждой функциональной зависимости из множества S содержит только один атрибут.
- Детерминант каждой функциональной зависимости из множества S , в свою очередь, является неприводимым, т.е. ни один атрибут из детерминанта не может быть убран.
- Ни одна функциональная зависимость из множества S не может быть удалена из множества S без изменения его замыкания S^+ .



Нормальные формы

Принцип хорошего тона - по одному факту в одном месте. Предметом нормализации становится формализации подобных простых идей.

© Джейсон Стэтхэм



Теорема Хита

Пусть r - переменная отношения, а A, B, C - непересекающиеся множества атрибутов этой переменной отношения.

Если r удовлетворяет функциональной зависимости $A \twoheadrightarrow B$, то r равна соединению её проекций по атрибутам $r_1(A, B)$ и $r_2(A, C)$.



Теорема Хита

Пусть r - переменная отношения, а A, B, C - непересекающиеся множества атрибутов этой переменной отношения.

Если r удовлетворяет функциональной зависимости $A \twoheadrightarrow B$,
то r равна соединению её проекций по атрибутам $r_1(A,B)$ и $r_2(A,C)$.

```
WITH
a AS (
SELECT DISTINCT name, theme_id FROM book
),
b AS (
SELECT DISTINCT name, language_id FROM book
)
SELECT
a.name, theme_id, language_id
FROM
a INNER JOIN b ON a.name = b.name
```

=

```
SELECT DISTINCT name, theme_id, language_id FROM book
```



Теорема Хита

Пусть r - переменная отношения, а A, B, C - непересекающиеся множества атрибутов этой переменной отношения.

Если r удовлетворяет функциональной зависимости $A \twoheadrightarrow B$,
то r равна соединению её проекций по атрибутам $r_1(A,B)$ и $r_2(A,C)$.

```
WITH a AS (  
  SELECT DISTINCT name, book_id FROM book  
) , b AS (  
  SELECT DISTINCT name, year_of_publishing, language_id FROM book)  
SELECT a.name, book_id, year_of_publishing, language_id FROM a INNER JOIN b ON a.name =  
b.name
```



1НФ

Переменная отношения находится в 1НФ тогда и только тогда, когда в любом допустимом значении этой переменной отношения каждый ее кортеж содержит только одно значение для каждого из атрибутов.

name	year_of_publishing
Капитанская Дочка	{2005}
Дубровский	{2005}
Война и Мир 3,4 том	{1998}
Мартин Иден	{2005}
Война и Мир 1,2 том	{1998}
Курс аналитической геометрии	{2005,2006}

name	year_of_publishing
Капитанская Дочка	{2005}
Дубровский	{2005}
Война и Мир 3,4 том	{1998}
Мартин Иден	{2005}
Война и Мир 1,2 том	{1998}
Курс аналитической геометрии	{2005}
Курс аналитической геометрии	{2006}



2НФ

Переменная отношения находится в 2НФ тогда и только тогда, когда она находится в первой нормальной форме и каждый её неключевой атрибут неприводимо зависит от её первичного ключа.



Переменная отношения находится в 2НФ тогда и только тогда, когда она находится в первой нормальной форме и каждый её неключевой атрибут неприводимо зависит от её первичного ключа.

name	year_of_publishing	theme_id	publishing_house_id	language_id
1984	2 005	2	12	2
Othello	2 005	2	12	2
Анна Каренина	2 005	1	11	1
Белый Клык	2 007	2	16	1
Война и Мир 1,2 том	1 998	1	11	1
Война и Мир 3,4 том	1 998	1	11	1
Дубровский	2 005	1	15	1
Капитанская Дочка	2 005	1	15	1
Курс аналитической геометрии	2 005	5	13	1
Курс аналитической геометрии	2 006	5	13	3
Мартин Иден	2 005	2	16	1
Обломов	2 005	1	14	1
Общая физика	2 005	5	13	1
Русские сказки	2 007	4	14	1
Сердца Трёх	2 007	2	16	1
Три Сестры	2 007	1	14	1

Нормальные формы

Переменная отношения находится в 2НФ тогда и только тогда, когда она находится в первой нормальной форме и каждый её неключевой атрибут неприводимо зависит от её первичного ключа.

name	year_of_publishing	theme_id	publishing_house_id	language_id
1984	2 005	2	12	2
Othello	2 005	2	12	2
Анна Каренина	2 005	1	11	1
Белый Клык	2 007	2	16	1
Война и Мир 1,2 том	1 998	1	11	1
Война и Мир 3,4 том	1 998	1	11	1
Дубровский	2 005	1	15	1
Капитанская Дочка	2 005	1	15	1
Курс аналитической геометрии	2 005	5	13	1
Курс аналитической геометрии	2 006	5	13	3
Мартин Иден	2 005	2	16	1
Обломов	2 005	1	14	1
Общая физика	2 005	5	13	1
Русские сказки	2 007	4	14	1
Сердца Трёх	2 007	2	16	1
Три Сестры	2 007	1	14	1

Нормальные формы



Переменная отношения находится в 2НФ тогда и только тогда, когда она находится в первой нормальной форме и каждый её неключевой атрибут неприводимо зависит от её первичного ключа.

name	year_of_publishing	language_id
1984	2 005	2
Othello	2 005	2
Анна Каренина	2 005	1
Белый Клык	2 007	1
Война и Мир 1,2 том	1 998	1
Война и Мир 3,4 том	1 998	1
Дубровский	2 005	1
Капитанская Дочка	2 005	1
Курс аналитической геометрии	2 005	1
Курс аналитической геометрии	2 006	3
Мартин Иден	2 005	1
Обломов	2 005	1
Общая физика	2 005	1
Русские сказки	2 007	1
Сердца Трёх	2 007	1
Три Сестры	2 007	1

Нормальные формы



Переменная отношения находится в 2НФ тогда и только тогда, когда она находится в первой нормальной форме и каждый её неключевой атрибут неприводимо зависит от её первичного ключа.

name	theme_id	publishing_house_id
1984	2	12
Othello	2	12
Анна Каренина	1	11
Белый Клык	2	16
Война и Мир 1,2 том	1	11
Война и Мир 3,4 том	1	11
Дубровский	1	15
Капитанская Дочка	1	15
Курс аналитической геометрии	5	13
Мартин Иден	2	16
Обломов	1	14
Общая физика	5	13
Русские сказки	4	14
Сердца Трёх	2	16
Три Сестры	1	14

