

Лекция 7

Элементы теории реляционных
баз данных:
функциональные зависимости
и декомпозиция без потерь

Функциональные зависимости

Пусть задана переменная отношения R , и X и Y являются произвольными подмножествами заголовка R («составными» атрибутами).

В значении переменной отношения R *атрибут Y функционально зависит от атрибута X* в том и только в том случае, если каждому значению X соответствует в точности одно значение Y . В этом случае говорят также, что атрибут X *функционально определяет* атрибут Y (X является детерминантом (*определителем*) для Y , а Y является зависимым от X).

Будем обозначать это как $R.X \rightarrow R.Y$.

Функциональные зависимости

Для примера будем использовать отношение СЛУЖАЩИЕ_ПРОЕКТЫ {СЛУ_НОМ, СЛУ_ИМЯ, СЛУ_ЗАРП, ПРО_НОМ, ПРОЕКТ_РУК}. Очевидно, что если СЛУ_НОМ является первичным ключом отношения СЛУЖАЩИЕ, то для этого отношения справедлива функциональная зависимость (Functional Dependency – FD) СЛУ_НОМ→СЛУ_ИМЯ.

СЛУ_НОМ	СЛУ_ИМЯ	СЛУ_ЗАРП	ПРО_НОМ	ПРОЕКТ_РУК
2934	Иванов	22400.00	1	Иванов
2935	Петров	29600.00	1	Иванов
2936	Сидоров	18000.00	1	Иванов
2937	Федоров	20000.00	1	Иванов
2938	Иванова	22000.00	1	Иванов
2939	Сидоренко	18400.00	2	Иваненко
2940	Федоренко	20400.00	2	Иваненко
2941	Иваненко	22600.00	2	Иваненко

Функциональные зависимости

Для тела отношения СЛУЖАЩИЕ_ПРОЕКТЫ выполняются следующие FD:

$СЛУ_НОМ \rightarrow СЛУ_ИМЯ$

$СЛУ_НОМ \rightarrow СЛУ_ЗАРП$

$СЛУ_НОМ \rightarrow ПРО_НОМ$

$СЛУ_НОМ \rightarrow ПРОЕКТ_РУК$

$\{СЛУ_НОМ, СЛУ_ИМЯ\} \rightarrow СЛУ_ЗАРП$

$\{СЛУ_НОМ, СЛУ_ИМЯ\} \rightarrow ПРО_НОМ$

$\{СЛУ_НОМ, СЛУ_ИМЯ\} \rightarrow \{СЛУ_ЗАРП, ПРО_НОМ\}$

$ПРО_НОМ \rightarrow ПРОЕКТ_РУК$

и т. д.

Эти FD должны быть верны для любого допустимого значения переменной отношения СЛУЖАЩИЕ_ПРОЕКТЫ и могут рассматриваться как *инварианты*, или *ограничения целостности* этой переменной отношения.

Функциональные зависимости

Поскольку имена всех служащих различны, то выполняются и такие FD :

$СЛУ_ИМЯ \rightarrow СЛУ_НОМ$

$СЛУ_ИМЯ \rightarrow СЛУ_ЗАРП$

$СЛУ_ИМЯ \rightarrow ПРО_НОМ$

и т. д.

Т. к. никакие двое служащих, участвующие в разных проектах, не получают одинаковую зарплату, выполняется FD:

$СЛУ_ЗАРП \rightarrow ПРО_НОМ$

Если атрибут A отношения R является возможным ключом, то для любого атрибута B этого отношения всегда выполняется FD $A \rightarrow B$.

Функциональные зависимости

FD $A \rightarrow B$ называется *тривиальной*, если $A \supseteq B$ (т. е. множество атрибутов A включает множество B или совпадает с множеством B).

Любая тривиальная FD всегда выполняется. Например, в отношении

СЛУЖАЩИЕ_ПРОЕКТЫ всегда выполняется FD

$\{\text{СЛУ_ЗАРП}, \text{ПРО_НОМ}\} \rightarrow \text{СЛУ_ЗАРП}$

Тривиальные FD нельзя трактовать как ограничения целостности.

Замыкание множества функциональных зависимостей

Замыканием множества FD S является множество FD S^+ , включающее все FD , логически выводимые из FD множества S .

Два примера FD , из которых следуют (или *выводятся*) другие FD :

Для отношения СЛУЖАЩИЕ_ПРОЕКТЫ выполняется, например, FD $СЛУ_НОМ \rightarrow \{СЛУ_ЗАРП, ОТД_НОМ\}$. Из этой FD выводятся FD $СЛУ_НОМ \rightarrow СЛУ_ЗАРП$ и $СЛУ_НОМ \rightarrow ОТД_НОМ$.

Имеется также пара FD $СЛУ_НОМ \rightarrow ОТД_НОМ$ и $ОТД_НОМ \rightarrow ПРОЕКТ_РУК$. Из них выводится FD $СЛУ_НОМ \rightarrow ПРОЕКТ_РУК$.

FD вида $СЛУ_НОМ \rightarrow ПРОЕКТ_РУК$ называются транзитивными, поскольку $ПРОЕКТ_РУК$ зависит от $СЛУ_НОМ$ «транзитивно», через $ПРО_НОМ$.

FD $A \rightarrow C$ называется *транзитивной*, если существует такой атрибут B , что имеются функциональные зависимости $A \rightarrow B$ и $B \rightarrow C$ и отсутствует функциональная зависимость $C \rightarrow A$.

Аксиомы Армстронга

Пусть A , B и C являются (в общем случае, составными) атрибутами отношения R . Множества A , B и C могут иметь непустое пересечение. Для краткости будем обозначать через AB $A \cup B$. Тогда:

- если $B \subseteq A$, 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).

Аксиомы Армстронга

Дейт по практическим соображениям предложил расширить базовый набор правил вывода еще пятью правилами:

- $A \rightarrow A$ (самодетерминированность);
- если $A \rightarrow BC$, то $A \rightarrow B$ и $A \rightarrow C$ (декомпозиция);
- если $A \rightarrow B$ и $A \rightarrow C$, то $A \rightarrow BC$ (объединение);
- если $A \rightarrow B$ и $C \rightarrow D$, то $AC \rightarrow BD$ (композиция);
- если $A \rightarrow BC$ и $B \rightarrow D$, то $A \rightarrow BCD$ (накопление) .

Замыкание множества атрибутов

Пусть заданы отношение R , множество Z атрибутов этого отношения (подмножество заголовка R , или составной атрибут R) и некоторое множество FD S , выполняемых для R . Тогда **замыканием Z над S** называется наибольшее множество Z^+ таких атрибутов Y отношения R , что FD $Z \rightarrow Y$ входит в S^+ .

Суперключом отношения R называется любое подмножество K заголовка R , включающее, по меньшей мере, хотя бы один возможный ключ R .

Одно из следствий этого определения состоит в том, что подмножество K заголовка отношения R является суперключом тогда и только тогда, когда для любого атрибута A (возможно, составного) заголовка отношения R выполняется FD $K \rightarrow A$. В терминах замыкания множества атрибутов K является суперключом тогда и только тогда, когда K^+ совпадает с заголовком R .

Минимальное покрытие множества FD

Множество FD S_2 называется **покрытием множества FD** S_1 , если любая FD, выводимая из S_1 , выводится также из S_2 .

S_2 является покрытием S_1 тогда и только тогда, когда $S_1^+ \subseteq S_2^+$. Два множества FD S_1 и S_2 называются **эквивалентными**, если каждое из них является покрытием другого, т. е. $S_1^+ = S_2^+$.

Множество FD S называется минимальным в том и только в том случае, когда удовлетворяет следующим свойствам:

- правая часть любой FD из S является множеством из одного атрибута (простым атрибутом);
- детерминант каждой FD из S обладает свойством **минимальности**; это означает, что удаление любого атрибута из детерминанта приводит к изменению замыкания S^+ , т. е. порождению множества FD, не эквивалентного S ;
- удаление любой FD из S приводит к изменению S^+ , т. е. порождению множества FD, не эквивалентного S .

Декомпозиция без потерь и FD

Считаются правильными такие декомпозиции отношения, которые обратимы, т. е. имеется возможность собрать исходное отношение из декомпозированных отношений без потери информации.

Такие декомпозиции называются ***декомпозициями без потерь***.

Корректные и некорректные декомпозиции отношений

СЛУЖАЩИЕ_ПРОЕКТЫ				
СЛУ_НОМЕР	СЛУ_ИМЯ	СЛУ_ЗАРП	ПРО_НОМ	ПРОЕКТ_РУК
2934	Иванов	22000.00	1	Иванов
2941	Иваненко	22000.00	2	Иваненко

Декомпозиция (1). Отношения СЛУЖ и СЛУ_ПРО

СЛУ_НОМ	СЛУ_ИМЯ	СЛУ_ЗАРП
2934	Иванов	22000.00
2941	Иваненко	22000.00

СЛУ_НОМ	ПРО_НОМ	ПРОЕКТ_РУК
2934	1	Иванов
2941	2	Иваненко

Декомпозиция (2). Отношения СЛУЖ и ЗАРП_ПРО

СЛУ_НОМ	СЛУ_ИМЯ	СЛУ_ЗАРП
2934	Иванов	22000.00
2941	Иваненко	22000.00

СЛУ_ЗАРП	ПРО_НОМ	ПРОЕКТ_РУК
22000.00	1	Иванов
22000.00	2	Иваненко

Теорема Хита

Пусть задано отношение r $\{A, B, C\}$ (A, B и C , в общем случае, являются составными атрибутами) и выполняется $FD A \rightarrow B$.

Тогда $r = (r \text{ ПРОЕКТ}\{A, B\}) \text{ NATURAL JOIN } (r \text{ ПРОЕКТ}\{A, C\})$.

СЛУЖАЩИЕ_ОТДЕЛЫ_ПРОЕКТЫ		
СЛУ_НОМ	СЛУ_ОТД	ПРО_НОМ
2934	630	1
2941	631	1
2934	630	2
2941	631	2

Декомпозиция.
Отношения СЛУЖ_ОТДЕЛЫ и СЛУЖ_ПРОЕКТЫ

СЛУ_НОМ	СЛУ_ОТД
2934	630
2941	631

СЛУ_НОМ	ПРО_НОМ
2934	1
2941	1
2934	2
2941	2

Минимальные FD

Атрибут B минимально зависит от атрибута A , если выполняется минимальная слева FD $A \rightarrow B$.

Минимальной слева называется FD с минимальным детерминантом.

Например, в отношении СЛУЖАЩИЕ_ПРОЕКТЫ выполняются FD $СЛУ_НОМ \rightarrow СЛУ_ЗАРП$ и $\{СЛУ_НОМ, СЛУ_ИМЯ\} \rightarrow СЛУ_ЗАРП$. Первая FD является минимальной слева, а вторая — нет.

Поэтому $СЛУ_ЗАРП$ минимально зависит от $СЛУ_НОМ$, а для $\{СЛУ_НОМ, СЛУ_ИМЯ\}$ свойство минимальной зависимости не выполняется.

Диаграммы функциональных зависимостей

