

1 Функціональна залежність

Функціональна залежність (далі часто ФЗ) – концепція, що лежить в основі багатьох питань, пов’язаних з реляційними базами даних, включаючи, зокрема, їхнє проектування. Математично являє собою бінарне відношення між множинами атрибутів даного відношення і є, по суті, зв’язком типу “один-до-багатьох”. ФЗ забезпечує основу для наукового підходу до розв’язання деяких проблем, оскільки володіє багатим набором цікавих формальних властивостей.

1.1 Визначення

1.1.1 Функціональна залежність

Нехай маємо відношення r зі схемою (заголовком) R , A і B – деякі підмножини множини атрибутів відношення r . Множина B функціонально залежить від A тоді і тільки тоді, коли кожне значення множини A зв’язане точно з одним значенням множини B . Іншими словами, якщо два кортежі збігаються за атрибутами A , то вони збігаються і за атрибутами B .

$$r(R), A \subseteq R, B \subseteq R$$

$$(A \rightarrow B) \Leftrightarrow ((\forall t_1, t_2 \in r : t_1(A) = t_2(A)) \Rightarrow (t_1(B) = t_2(B)))$$

У такому разі A – детермінант, B – залежна частина.

ФЗ називається тривіальною, якщо залежна частина є підмножиною детермінанта.

$$(B \subseteq A) \Rightarrow (A \rightarrow B)$$

1.1.2 Замикання множини залежностей

Одні функціональні залежності можуть припускати інші функціональні залежності. Наприклад,

$$(A \rightarrow B) \wedge (B \rightarrow C) \Rightarrow (A \rightarrow C).$$

Множина S^+ всіх ФЗ, які припускаються даною множиною ФЗ S називається замкненням множини S .

1.1.3 Замикання множини атрибутів

Нехай Z – деяка множина атрибутів відношення r , а S – множина функціональних залежностей цього відношення. Замкненням Z^+ множини атрибутів Z в межах S називається така множина атрибутів A_i відношення r , що функціональна залежність $Z \rightarrow A_i$ є членом замкнення S^+ .

$$r(R), S, Z \subseteq R, A_i \subseteq R, i = \overline{1, n}$$

1.1.4 Незвідні множини залежностей

Нехай S_1 і S_2 – деякі множини функціональних залежностей.

- Якщо будь-яка функціональна залежність з S_1 входить і в S_2 , тоді S_2 називають покриттям множини функціональних залежностей S_1 .
- Якщо S_2 – покриття для S_1 , а S_1 – для S_2 (тобто $S_1^+ = S_2^+$), тоді такі множини називаються еквівалентними.

- Множина ФЗ S називається незвідною тоді і тільки тоді, коли виконуються наступні вимоги:
 - В кожній ФЗ залежна частина містить лише один елемент;
 - Детермінант кожної ФЗ є незвідним (ні один атрибут не може бути видаленим з детермінанта без зміни замкнення S^+);
 - Жодну ФЗ з S не можна виключити без зміни замкнення S^+ .
- Для будь-якої множини ФЗ існує не менше ніж одна еквівалентна множина, яка є незвідною. Така множина називається незвідним покриттям.

1.2 Обчислення замкнень

1.2.1 Правило виводу Армстронга

В 1974 Вільям Армстронг запропонував набір правил виводу нових ФЗ на основі даних.

Нехай у нас є відношення $r(R)$ і множини атрибутів $A, B, C, D \subseteq R$. Для скорочення запису замість $X \cup Y$ будемо писати просто XY .

- Рефлексивність:

$$(B \subseteq A) \Rightarrow (A \rightarrow B)$$

- Поповнення:

$$(A \rightarrow B) \Rightarrow (AC \rightarrow BC)$$

- Транзитивність:

$$(A \rightarrow B) \wedge (B \rightarrow C) \Rightarrow (A \rightarrow C)$$

Правила виводу Армстронга **повні** (з їхньою допомогою можна вивести решту ФЗ, що мають на увазі даною множиною) і **надійні** (“зайвих” ФЗ вивести не можна; виведена ФЗ справедлива всюди, де справедлива та множина ФЗ, з якої вона була виведена).

Крім того, з даних правил досить просто виводяться, декілька додаткових правил, які спрощують задачу виведення ФЗ.

- Самовизначення:

$$A \rightarrow A$$

- Декомпозиція:

$$(A \rightarrow BC) \Rightarrow (A \rightarrow B) \wedge (A \rightarrow C)$$

- Об’єднання:

$$(A \rightarrow B) \wedge (A \rightarrow C) \Rightarrow (A \rightarrow BC)$$

- Композиція:

$$(A \rightarrow B) \wedge (C \rightarrow D) \Rightarrow (AC \rightarrow BD)$$

- Теорема загального об’єднання Дарвена:

$$(A \rightarrow B) \wedge (C \rightarrow D) \Rightarrow (A \cup (C - B) \rightarrow BD)$$

Теорема: ФЗ $A \rightarrow B$ вивідна з даної множини ФЗ S за правилами виводу Армстронга тоді і тільки тоді, коли $B \subseteq A^+$.

1.2.2 Замкнення множини атрибутів

Якщо застосовувати правила з попереднього розділу до того часу коли утворення нових ФЗ не припиниться, то ми отримаємо замкнення для даної множини ФЗ. На практиці рідко вимагається знаходити це замкнення само по собі, частіше за все нам набагато цікавіше дізнатися, чи входить та або інша ФЗ в замкнення. Для цього нам достатньо вирахувати замкнення детермінанта. Для цього існує доволі простий алгоритм.

1. Нехай X – множина атрибутів, яка врешті-решт стане замкненням.
2. Здійснюємо пошук ФЗ виду $B_1B_2 \dots B_m \rightarrow C$, де $B_1B_2 \dots B_m \subseteq X$, а $C \notin X$. Залежну частину кожної ФЗ додаємо в X .
3. Повторюємо пункт 2, доки до множини X буде неможливо додати атрибути.
4. Множина X , до якої неможливо додати атрибути і буде замкненням.

1.3 Застосування

1.3.1 Проектування БД

ФЗ є обмеженнями цілісності і визначають семантику даних, що зберігаються. При кожному оновленні СКБД повинна перевіряти їхнє дотримання. Значить, наявність великої кількості ФЗ небажане, інакше це призводить до уповільнення роботи. Для спрощення задачі необхідно скоротити набір ФЗ до мінімально необхідного.

Якщо I є незвідним покриттям початкової множини ФЗ S , то перевірка виконання ФЗ з I автоматично гарантує виконання всіх ФЗ з S . Таким чином, задача пошуку мінімально необхідного набору зводиться до пошуку незвідного покриття множини ФЗ, яке і буде використовуватись замість початкової множини.

1.3.2 Декомпозиція відношень

Теорема Хіта

Нехай дане відношення $r(A, B, C)$.

Якщо r задовольняє функціональній залежності $A \rightarrow B$, тоді воно дорівнює поєднанню його проекцій $r[A, B]$ і $r[A, C]$.

$$(A \rightarrow B) \Rightarrow (r(A, B, C) = r[A, B] \text{ JOIN } r[A, C])$$