

Л6. Маніпуляційний рівень роботи з РБД

Приклад РБД



В результаті виконання етапу проектування маємо інформаційну модель БД.

ОЦІНКИ (РЕЗУЛЬТАТИ)

№зап	Студент	Дисципліна	Оцінка
1 I.C.	112	1	3
...	
8 B.B.	51	3	3

ДИСЦИПЛІНИ

Дисципліна	Назва
1	Диф.рівн.
2	Програмування
3	Мат.аналіз

СТУДЕНТИ

№студента	ПІБ	Група
112	Петренко I.C.	20
...
51	Семенов B.B.	23

ГРУПИ

№групи	Група	Староста
20	К-20	112
...
23	К-23	51

Мал. 5.1. Інформаційна модель БД

Після створення моделі РБД розглянемо маніпуляційні засоби, за допомогою яких забезпечується доступ до реляційних даних за допомогою реляційної алгебри або реляційного числення. В реалізаціях конкретних реляційних СУБД у чистому вигляді не використовується ні реляційна алгебра, ні реляційне числення. Фактично стандартом доступу до реляційних даних стала мова SQL (Structured Query Language). Мова SQL представляє собою суміш операторів реляційної алгебри та виразів реляційного числення, що застосовує синтаксис, близький до фраз англійської мови та розширений додатковими можливостями, які відсутні в реляційній алгебрі та реляційному численні. Взагалі, мова доступу до даних називається *реляційно повною*, якщо вона за впливом не поступається реляційній алгебрі

(або, що теж саме, реляційному численню), тобто будь-який оператор реляційної алгебри може бути записаний засобами цієї мови. Саме такою є мова SQL.

1. Реляційна алгебра

Замикання реляційної алгебри (замкнутість реляційної алгебри)

Реляційна алгебра представляє собою набір операторів, які використовують відношення в якості аргументів і повертають відношення в якості результату. Таким чином, реляційний оператор f виглядає як функція над відношеннями в якості аргументів:

$$R = f(R_1, R_2, \dots, R_n)$$

Реляційна алгебра є замкнутою, тобто в якості аргументів в реляційні оператори можна підставляти інші реляційні оператори, що підходять за типом:

$$R = f(f_1(R_{11}, R_{12}, \dots), f_2(R_{21}, R_{22}, \dots), \dots)$$

Таким чином, в реляційних виразах можна використовувати вкладені вирази будь-якої складності. Кожне відношення має унікальне ім'я в межах схеми бази даних. Ім'я відношення, що отримується в результаті виконання реляційної операції, визначається у лівій частині рівності. Однак можна не вимагати наявності імен від відношень, отриманих в результаті реляційних виразів, якщо ці відношення підставляються в якості аргументів в інші реляційні вирази. Такі відношення будемо називати **неіменованими відношеннями**. Неіменовані відношення реально не існують у базі даних, а тільки обчислюються в момент визначення результату реляційного оператора.

Традиційно, визначають вісім реляційних операторів, об'єднаних у дві групи:

Теоретико-множинні оператори:

1. Об'єднання
2. Перетин
3. Віднімання (різниця)
4. Декартовий добуток

Спеціальні реляційні оператори:

5. Вибірка
6. Проекція
7. З'єднання
8. Ділення

Не всі оператори є незалежними, тобто деякі з них можна виразити через інші реляційні оператори.

Відношення, сумісні за типом

Деякі реляційні оператори (наприклад, об'єднання) потребують, щоб відношення мали однакові заголовки. Дійсно, відношення складається з заголовка і тіла. Операція об'єднання двох відношень є простим об'єднанням двох множин кортежів, узятих з тіл відповідних відношень. Чи буде результат відношенням? По-перше, якщо вихідні відношення мають різну кількість атрибутів, то, зрозуміло, що множина, яка є об'єднанням таких різнотипових кортежів неможливо уявити у вигляді відношення. По-друге, нехай відношення мають однакову кількість атрибутів, але атрибути мають різні найменування. Як тоді визначити заголовок відношення, отриманого в результаті об'єднання множин кортежів? По-третє, нехай відношення мають однакову кількість атрибутів, атрибути мають однакові

найменування, але визначені на різних доменах. Тоді знов об'єднання кортежів не буде формувати відношення.

Означення 1. Будемо називати відношення **сумісними за типом**, якщо вони мають ідентичні заголовки:

- відношення мають *одну і ту ж множину імен атрибутів*, тобто для довільного атрибуту в одному відношенні знайдеться атрибут з таким же найменуванням в іншому відношенні,
- атрибути з однаковими іменами *визначені на однакових доменах*.

Окремі відношення не є сумісними за типом, але стають такими після деякого перейменування атрибутів. Для того, щоб такі відношення можна було використовувати в реляційних операторах, вводиться допоміжний **оператор перейменування атрибутів**.

Оператор перейменування атрибутів.

Оператор перейменування атрибутів має наступний синтаксис:

R RENAME Atr₁, Atr₂,... AS NewAtr₁, NewAtr₂,...

де

R - відношення,

Atr₁, Atr₂,... - вихідні імена атрибутів,

NewAtr₁, NewAtr₂,... - нові імена атрибутів.

В результаті застосування оператора перейменування атрибутів отримуємо нове відношення, з зміненими іменами атрибутів.

Приклад 1. Наступний оператор повертає неіменоване відношення, в якому атрибут City_Num перейменований у CityId:

City RENAME City_Num AS CityId

Теоретико-множинні оператори

- Об'єднання

Означення 2. **Об'єднанням** двох сумісних за типом відношень *A* и *B* називається відношення з тим же заголовком, що і у відношень *A* та *B*, і тілом, що складається з кортежів, які належать або *A*, або *B*, або обом відношенням.

Синтаксис операції об'єднання: *A UNION B*

Зауваження. Результат об'єднання, як і будь-яке інше відношення, не може містити однакових кортежів. Тому, якщо деякий кортеж входить і у відношення *A*, і у відношення *B*, то у об'єднання він входить лише один раз.

Приклад 2. Нехай дано два відношення *A* та *B* з інформацією про співробітників:

Таблиця 1 (відношення *A*)

<i>Табельний номер</i>	ПІБ	Зарплатня
<i>1</i>	Івашко	1000
<i>2</i>	Шевчук	2000
<i>3</i>	Сидорчук	3000

Таблиця 2 (відношення *B*)

<i>Табельний номер</i>	ПІБ	Зарплатня
<i>1</i>	Івашко	1000

2	Петренко	2500
4	Сидорчук	3000

Об'єднання відношень A та B буде мати вигляд:

Таблиця 3 (відношення $A \cup B$)

Табельний номер	ПІБ	Зарплатня
1	Івашко	1000
2	Шевчук	2000
3	Сидорчук	3000
2	Петренко	2500
4	Сидорчук	3000

Зауваження. Як видно з наведеного прикладу, потенційні ключі, які були в відношеннях A та B не наслідуються об'єднанням цих відношень. Тому, в об'єднанні відношень A та B атрибут "Табельний номер" може містити дублікати значень. Якщо б це було не так і ключі наслідувалися, то це б суперечило поняттю об'єднання як "об'єднання множин". Зрозуміло, об'єднання відношень A та B має, як і будь-яке інше відношення, потенційний ключ, наприклад, складений з усіх атрибутів.

- Перетин

Означення 3. Перетином двох сумісних за типом відношень A та B називається відношення з тим же заголовком, що й у відношень A та B , і тілом, що складається з кортежів, які належать одночасно обом відношенням A та B .

Синтаксис операції перетину: $A \cap B$

Пример 3. Для відношень A та B , що є в попередньому прикладі, перетин має вигляд:

Таблиця 4 (відношення $A \cap B$)

Табельний номер	ПІБ	Зарплатня
1	Івашко	1000

Зауваження. На перший погляд, здається, що на відміну від операції об'єднання, потенційні ключі могли б наслідуватися перетином відношень. Однак це не так. Взагалі, жодні реляційні оператори не передають результуючому відношенню ніяких даних про потенційні ключі. В якості причини цього можна було б навести тривіальну думку, що так відбувається більш просто та симетрично - усі оператори працюють однаково. Насправді, причина більш глибока, і вона полягає в тому, що потенційний ключ - семантичне поняття, що відбиває особливість об'єктів предметної області. Наявність потенційних ключів не впливає з структури відношення, а явно задається для кожного відношення, виходячи з його змісту. Реляційні оператори, натомість, є формальними операціями над відношеннями і виконуються однаково, незалежно від змісту даних, що містяться у відношеннях. Тому, реляційні оператори нічого не можуть "знати" про зміст даних. Трамбування результату реляційних операцій - справа користувача.

- Віднімання (різниця) відношень

Означення 4. Різницею двох сумісних за типом відношень A та B називається відношення з тим самим заголовком, що й у відношень A та B , і тілом, що складається з кортежів, які належать відношенню A і не належать відношенню B .

Синтаксис операції віднімання (різниці) відношень: $A \text{ MINUS } B$

Приклад 4. Для відношень A та B , що є у попередньому прикладі, різниця (віднімання) $A - B$ має вигляд:

Таблиця 5 (відношення A MINUS B)

<i>Табельний номер</i>	ПІБ	Зарплатня
2	Шевчук	2000
3	Сидорчук	3000

- Декартовий добуток відношень

Означення 5. Декартовим добутком двох відношень A та B називається відношення, заголовок якого є **зчепленням заголовків** відношень A та B : $(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$, а тіло складається з кортежів, які є **зчепленням кортежів** відношень A та B : $(a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_m)$, таких, що $(a_1, a_2, \dots, a_n) \in A, (b_1, b_2, \dots, b_m) \in B$.

Синтаксис операції декартового добутку: $A \text{ TIMES } B$

Зауваження 1. Потужність добутку $A \text{ TIMES } B$ дорівнює добутку потужностей A та B , тобто кожен кортеж відношення A з'єднується з кожним кортежем відношення B

Зауваження 2. Якщо у відношеннях A та B є атрибути з однаковими найменуваннями, то перед виконанням операції декартового добутку такі атрибути необхідно перейменувати.

Зауваження 3. Множити можна довільні два відношення, сумісність за типом при цьому не потрібна.

Таблиця 6 (відношення A TIMES B)

Приклад 5. Нехай задано два відношення A та B з даними про працівників і деталі:

<i>Табельний номер</i>	ПІБ	Зарплатня
1	Івашко	1000
2	Шевчук	2000
3	Сидорчук	3000

<i>Номер деталі</i>	Найменування
1	Болт
2	Гайка
3	Шайба

Декартовий добуток відношень A та B буде мати вигляд

<i>Табельний номер</i>	ПІБ	<i>Номер деталі</i>	Найменування
1	Івашко	1	Болт
1	Івашко	2	Гайка
1	Івашко	3	Шайба
2	Шевчук	1	Болт
2	Шевчук	2	Гайка
2	Шевчук	3	Шайба
3	Сидорчук	1	Болт

3	Сидорчук	2	Гайка
3	Сидорчук	3	Шайба

Зауваження. Сама по собі операція декартового добутку не дуже важлива, тому що вона не дає ніякої нової інформації порівняно з вихідними відношеннями. Для реальних запитів ця операція практично ніколи не використовується. Однак операція декартового добутку важлива для виконання спеціальних реляційних операцій, про які будемо говорити далі.

Спеціальні реляційні оператори

Вибірка (обмеження, селекція)

Означення 6. Вибіркою (обмеженням, селекцією) на відношенні A з умовою s будемо називати відношення з тим самим заголовком, що й у відношенні A , і тілом, що складається з кортежів, значення атрибутів яких при підстановці в умову s дають значення ІСТИНА. Умова s представляє собою логічний вираз, в який можуть входити атрибути відношення A і (або) скалярні вирази.

У найпростішому випадку умова s має вигляд $X\Theta Y$, где Θ - один з операторів порівняння ($=, !=, >, <, >=, <=$ і т.д.), а X і Y - атрибути відношення A або скалярні значення. Такі вибірки називаються **Θ -вибірками (тета-вибірками)** або **Θ -обмеженнями, Θ -селекціями.**

Синтаксис операції вибірки: A WHERE s або A WHERE $X\Theta Y$

Приклад 6. Нехай задано два відношення A з даними про співробітників

Табельний номер	ПІБ	Зарплатня
1	Івашко	1000
2	Шевчук	2000
3	Сидорчук	3000

Результат вибірки A WHERE $Зарплатня < 3000$ буде мати вигляд:

Табельний номер	ПІБ	Зарплатня
1	Івашко	1000
2	Шевчук	2000

Зміст операції вибірки зрозумілий - вибирати кортежі відношення, які задовольняють деякій умові. Таким чином, операція вибірки дає "горизонтальний зріз" відношення з певною умовою.

Проекція

Означення 7. Проекцією відношення A за атрибутами X, Y, \dots, Z , де кожен з атрибутів належить відношенню A , називається відношення з заголовком X, Y, \dots, Z і тілом, що містить множину кортежів виду (x, y, \dots, z) , таких, для яких у відношенні A знайдуться кортежі зі значенням атрибуту X рівним x , значенням атрибуту Y рівним y , ..., значенням атрибуту Z рівним z .

Синтаксис операції проекції: $A[X, Y, \dots, Z]$.

Зауваження. Операція проекції дає "вертикальний зріз" відношення, в якому видалені усі наявні при такому зрізі дублікати кортежів.

Пример7. Нехай задано відношення A з інформацією про постачальників, що містить їх найменування та місто розташування:

<i>№ постачальника</i>	Найменування	Місто
1	TDK	Tokio
2	Iptv	Київ
3	Eram	Київ
4	Eram	Мінськ

Проекція $A[\text{Місто}]$ буде мати вигляд:

Місто
Tokio
Київ
Мінськ

З'єднання

Операція з'єднання відношень разом з операціями вибірки та проекції є однією з найбільш важливих реляційних операцій.

Традиційно розглядається декілька видів операції з'єднання:

1. Загальна операція з'єднання
2. Θ - з'єднання (тета - з'єднання)
3. Екві-з'єднання
4. Природне з'єднання

Найбільш важливим з цих видів з'єднання є операція природного з'єднання. Усі види з'єднання є частковими випадками загальної операції з'єднання.

- Загальна операція з'єднання

Означення 8. **З'єднанням** відношень A і B за умовою s називається відношення $(A \text{ TIMES } B) \text{ WHERE } s$, де s представляє собою логічний вираз, в який можуть входити атрибути відношень A та B і (або) скалярні вирази.

Таким чином, операція з'єднання є результатом послідовного застосування операцій декартового добутку та вибірки. Якщо у відношеннях A і B є атрибути з однаковими найменуваннями, то перед виконанням з'єднання такі атрибути необхідно перейменувати.

- Θ -з'єднання

Означення 9. Нехай відношення A містить атрибут X , відношення B містить атрибут Y , а Θ - один з операторів порівняння ($>$, $>=$, $<$, $<=$, $=$, \neq і т.д.). Тоді Θ - **з'єднанням** відношення за атрибутом X з відношенням за атрибутом Y називають відношення $(A \text{ TIMES } B) \text{ WHERE } X\Theta Y$.

Це частковий випадок операції загального з'єднання.

Іноді для операції Θ - з'єднання застосовують такий, більш короткий синтаксис:

$$A [X \Theta Y] B$$

- Екві-з'єднання

Найбільш важливим частковим випадком Θ -з'єднання є випадок, коли операцією Θ є рівність.

Синтаксис *екви-соединения*:

$$A [X=Y] B$$

Недоліком екві-з'єднання є те, що якщо з'єднання відбувається за атрибутами з однаковими найменуваннями (таке буває досить часто), то у результуючому відношенні з'являються два атрибути з однаковими значеннями. Позбавитися цього недоліку можна, взявши проєкцію по всіх атрибутах, крім одного, що дублюються.

- Природне з'єднання

Означення 10. Нехай дано відношення $A(A_1, A_2, \dots, A_m, X_1, X_2, \dots, X_p)$ і $B(X_1, X_2, \dots, X_p, B_1, B_2, \dots, B_m)$, які мають однакові атрибути X_1, X_2, \dots, X_p (тобто атрибути з однаковими іменами та визначені на однакових доменах).

Тоді **природнім з'єднанням** відношень A і B називається відношення з заголовком $(A_1, A_2, \dots, A_m, X_1, X_2, \dots, X_p, B_1, B_2, \dots, B_m)$ і тілом, що містить множину кортежів $(a_1, a_2, \dots, a_m, x_1, x_2, \dots, x_p, b_1, b_2, \dots, b_m)$, таких, що $(a_1, a_2, \dots, a_m, x_1, x_2, \dots, x_p) \in A$ і $(x_1, x_2, \dots, x_p, b_1, b_2, \dots, b_m) \in B$.

Природне з'єднання настільки важливе, що для нього використовують спеціальний синтаксис:

$$A \text{ JOIN } B$$

Зауваження 1. У синтаксисі природнього з'єднання не вказується, за якими атрибутами відбувається з'єднання. Природне з'єднання відбувається *за всіма* однаковими атрибутами.

Зауваження 2. Природне з'єднання еквівалентне такій послідовності реляційних операцій:

- Перейменувати однакові атрибути у відношеннях
- Виконати декартовий добуток відношень
- Виконати вибірку за співпадаючими значеннями атрибутів, які мають однакові імена
- Виконати проєкцію, видаляючи атрибути, що повторюються
- Перейменувати атрибути, повернувши їм початкові імена

Зауваження 3. Можна виконувати послідовне природне з'єднання декількох відношень. Нескладно перевірити, що природне з'єднання (як і з'єднання загального вигляду) має властивість **асоціативності** (такі з'єднання можна записувати, опускаючи дужки).