

Лабораторна робота №1. Застосування нормальних форм структури даних

Мета роботи.

Ознайомитись з нормальними формами та отримати практичні навички з нормалізації моделі БД.

Теоретичний матеріал.

Нормалізація являє собою процес розділення даних по окремим пов'язаним таблицям. Нормалізація усуває надлишковість даних (data redundancy) і завдяки цьому дозволяє уникнути порушення цілісності даних під час їх змінювання, тобто уникнути аномалій зміни (update anomaly).

У *ненормалізованій формі* таблиця може зберігати інформацію про дві або більше сутностей. Також вона може вміщувати стовпці що повторюються, або в стовпцях можуть зберігатись повторюванні значення.

В *нормалізованій формі* кожна таблиця зберігає інформацію тільки про одну сутність.

Ключове поняття нормалізації - функціональна залежність.

Функціональна залежність – 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 «один-до-багатьох». Функціональна залежність забезпечує основу для наукового підходу до розв'язання деяких проблем, оскільки володіє багатим набором цікавих формальних властивостей.

Нормалізація передбачає застосування нормальних форм до структури даних.

Існує 7 нормальних форм:

1. Перша нормальна форма (1НФ, 1NF) – є основою для структурованої схеми бази даних.
2. Друга нормальна форма (2НФ, 2NF) – вимагає аби дані, які зберігаються в таблицях із композитним ключем, не залежали лише від частини ключа.

3. Третя нормальна форма (3НФ, 3NF) - вимагає, аби дані в таблиці винятково від основного ключа.
4. Нормальна форма Бойса-Кодда (НФБК, BCNF) – більш суворі версія 3НФ.
5. Четверта нормальна форма (4НФ, 4NF) - потребує, аби в схемі баз даних не було нетривіальних багатозначних залежностей множин атрибутів від будь чого, окрім надмножини ключа-кандидата.
6. П'ята нормальна форма (5НФ, 5NF, PJ/NF) - вимагає, аби не було нетривіальних залежностей об'єднання, котрі б не витікали із обмежень ключів.
7. Нормальна форма домен/ключ (DKNF)– ця нормальна форма вимагає, аби в схемі не було інших обмежень окрім ключів та доменів. Зазвичай ця форма ототожнюється з шостою нормальною формою (6НФ, 6NF).

Для звичайних баз даних у більшості випадків достатньою є відповідність її схеми третій нормальній формі. Тому в подальшому більш детально будемо розглядати критерії відповідності 1НФ, 2НФ та 3НФ.

Кожна нормальна форма (за виключенням першої) передбачає, що до даних вже було застосовано попередню нормальну форму.

Критерії першої нормальної форми:

- a. Кожна таблиця повинна мати основний ключ, мінімальний набір колонок, які ідентифікують запис.
- b. Уникнення повторень груп (категорії даних, що можуть зустрічатись різну кількість разів в різних записах) правильно визначаючи неключові атрибути.
- c. Атомарність: кожен атрибут повинен мати лише одне значення, а не множину значень.

Критерії другої нормальної форми:

- a. Схема бази даних повинна відповідати всім вимогам 1НФ.
- b. Дані, що повторно з'являються в декількох рядках, мають бути винесені в окремі таблиці.

Критерії третьої нормальної форми:

- a. Схема бази даних повинна відповідати всім вимогам другої нормальної форми.
- b. Будь-яке поле, яке залежить від основного ключа та від будь-якого іншого поля, має виноситись в окрему таблицю.

Завдання.

1. Обрати тему за варіантом.
2. Створити ненормалізовану таблицю керуючись прикладом у лабораторній роботі.
3. Провести нормалізацію, яка передбачає застосування нормальних форм структури даних. Нормалізацію виконати послідовно до 3НФ.
4. Підготувати звіт з лабораторної роботи у текстовому редакторі в якому обґрунтувати свої дії.

Приклад виконання завдання.

1. Для виконання роботи обрано тему «Створення бази даних для інтернет-магазину будівельних матеріалів».
2. Створюємо новий файл MS Excel та зберігаємо його з назвою «ЛР1-<Студент>», де замість <Студент> вказується прізвище та ініціали студента.
3. У відповідності до обраної теми створюємо ненормалізовану таблицю з даними про асортимент товарів інтернет-магазину (рис.1).

	A	B	C	D	E	F	G
1	НЕНОРМАЛІЗОВАНА ТАБЛИЦЯ						
2	Найменування	Категорія	Од. виміру	Вартість, грн.	Постачальник	Термін дії договору с постачальником	
3	Цегла керамічна	Стінові	шт	15	ТОВ "Будпостач"	21.01.2023	
	Цегла силікатна			12	ПП "Альфагруп"	08.05.2025	
4	Мінеральна вата	Утеплювач	м3	4500	ПП "Промбуд"	22.05.2024	
	Пінополістирол			3200	ФЛП Ромашкін	06.07.2022	
5	Розчинник	Будхімія	л	50	ФЛП Петренко	18.05.2024	
6	Фарба інтер'єрна	Фарба	л	120	ПП "Декор+"	01.05.2025	
7	Молоток	Інструмент	шт	250	ФЛП Шевченко	15.08.2023	
8	Цвяхи	Кріплення	кг	120	ТОВ "Будкріп"	30.01.2023	
9							

Рис. 1. Приклад ненормалізованої таблиці

4. В цій таблиці наведено дані, щодо вартості певних будівельних матеріалів, а також є інформація до яких груп з каталогу вони відносяться і дані про термін дії договору с постачальником конкретного матеріалу.
5. Одразу видно, що дана таблиця не відповідає критерію атомарності, тобто для категорій «Стінові» та «Утеплювач» в

колонках «Найменування» та «Вартість» зроблено одразу два записи. Якщо робити подібні записи в реальній базі даних, то будуть виникати помилки під час виконання операцій з вартостями матеріалу (підсумки, визначення мінімуму, середнього тощо) та буде некоректно виконуватись пошук по назві матеріалу, оскільки СКБД буде «бачити» не окремо записи «Цегла керамічна» та «Цегла силікатна», а запис «Цегла керамічна Цегла силікатна».

6. Приведемо цю таблицю до 1НФ розділивши проблемні записи (рис. 2). Також, для зручності, було додано унікальний ідентифікатор кожного запису (строки),

10	ТАБЛИЦЯ У 1НФ						
11	МатеріалID	Матеріал	Категорія	Од. виміру	Вартість, грн.	Постачальник	Термін дії договору с постачальником
12	1	Цегла керамічна	Стінові	шт	15	ТОВ "Будпостач"	21.01.2023
13	2	Цегла силікатна	Стінові	шт	12	ПП "Альфагруп"	08.05.2025
14	3	Мінеральна вата	Утеплювач	м3	4500	ПП "Промбуд"	22.05.2024
15	4	Пінополістирол	Утеплювач	м3	3200	ФЛП Ромашкін	06.07.2022
16	5	Розчинник	Будхімія	л	50	ФЛП Петренко	18.05.2024
17	6	Фарба інтер'єрна	Фарба	л	120	ПП "Декор+"	01.05.2025
18	7	Молоток	Інструмент	шт	250	ФЛП Шевченко	15.08.2023
19	8	Цвяхи	Кріплення	кг	120	ТОВ "Будкріп"	30.01.2023
20							

Рис. 2. Таблиця у 1НФ

7. Отримана нами таблиця не відповідає 2НФ, оскільки атрибут «Категорія» повторюється декілька разів. Розіб'ємо дану таблицю на дві (рис. 3).
8. Для того, щоб наші дані відповідали вимогам другої нормальної форми ми розбили початкову таблицю на дві таблиці «Матеріали» та «Категорії», і додали третю таблицю в якій відобразили до яких категорій відносяться певні матеріали.

ТАБЛИЦЯ У 2НФ					
МатеріалID	Матеріал	Од. виміру	Вартість, грн.	Постачальник	Термін дії договору с постачальником
1	Цегла керамічна	шт	15	ТОВ "Будпостач"	21.01.2023
2	Цегла силікатна	шт	12	ПП "Альфагруп"	08.05.2025
3	Мінеральна вата	м3	4500	ПП "Промбуд"	22.05.2024
4	Пінополістирол	м3	3200	ФЛП Ромашкін	06.07.2022
5	Розчинник	л	50	ФЛП Петренко	18.05.2024
6	Фарба інтер'єрна	л	120	ПП "Декор+"	01.05.2025
7	Молоток	шт	250	ФЛП Шевченко	15.08.2023
8	Цвяхи	кг	120	ТОВ "Будкріп"	30.01.2023

КатегоріяID	Категорія	МатеріалID	КатегоріяID
1	Стінові	1	1
2	Утеплювач	2	1
3	Будхімія	3	2
4	Фарба	4	2
5	Інструмент	5	3
6	Кріплення	6	4
		7	5
		8	6

Рис. 3. Приведення до 2НФ

9. Таблиця «Матеріали», яку ми отримали після перетворення не відповідає вимогам 3НФ, оскільки вона містить в собі атрибут «Термін дії договору с постачальником», який жодним чином не залежить від основного ключа таблиці, а є властивістю атрибуту «Постачальник». Це може привести до аномалій оновлення, коли під час додавання нових матеріалів буде записано термін дії нового договору, а в старих записах термін залишиться без змін.
10. Для уникнення помилок в базі розіб'ємо таблицю «Матеріали» наступним чином (рис. 4).

ТАБЛИЦІ У 3НФ							
МатеріалID	Матеріал	Од. виміру	Вартість, грн.	ПостачальникID	ПостачальникID	Постачальник	Термін дії договору с постачальником
1	Цегла керамічна	шт	15	1	1	ТОВ "Будпостач"	21.01.2023
2	Цегла силікатна	шт	12	2	2	ПП "Альфагруп"	08.05.2025
3	Мінеральна вата	м3	4500	3	3	ПП "Промбуд"	22.05.2024
4	Пінополістирол	м3	3200	4	4	ФЛП Ромашкін	06.07.2022
5	Розчинник	л	50	5	5	ФЛП Петренко	18.05.2024
6	Фарба інтер'єрна	л	120	6	6	ПП "Декор+"	01.05.2025
7	Молоток	шт	250	7	7	ФЛП Шевченко	15.08.2023
8	Цвяхи	кг	120	8	8	ТОВ "Будкріп"	30.01.2023

Рис. 4. Приведення таблиці «Матеріали» до 3НФ

11. Таким чином в результаті перетворень початкової таблиці ми отримали набір пов'язаних між собою таблиць. В результаті чого, наші данні мають чітку структуру і захищені від помилок під час оновлення.