



ПРАВИЛА КОДДА НОРМАЛІЗАЦІЯ ДАНИХ

ЛЕКЦІЯ 3 КУРСУ «ОРГАНІЗАЦІЯ БАЗ ДАНИХ» («БАЗИ ДАНИХ»)

АВТОР ГОЛУБ Б.Л. 2016 Р.

ПРАВИЛА 0 ТА 1

0. **ФУНДАМЕНТАЛЬНЕ ПРАВИЛО** (*FOUNDATION RULE*)

РЕЛЯЦІЙНА СУБД ПОВИННА БУТИ ЗДАТНА ПОВНІСТЮ КЕРУВАТИ БАЗОЮ ДАНИХ, ВИКОРИСТОВУЮЧИ ЗВ'ЯЗКИ МІЖ ДАНИМИ.

1. **ІНФОРМАЦІЙНЕ ПРАВИЛО** (*INFORMATION RULE*)

ІНФОРМАЦІЯ ПОВИННА БУТИ ПРЕДСТАВЛЕНА У ВИГЛЯДІ ДАНИХ, ЩО ЗБЕРІГАЮТЬСЯ У КОМІРКАХ. ДАНІ, ЩО ЗБЕРІГАЮТЬСЯ У КОМІРКАХ, ПОВИННІ БУТИ АТОМАРНИМИ. ПОРЯДОК РЯДКІВ В РЕЛЯЦІЙНІЙ ТАБЛИЦІ НЕ ПОВИНЕН ВПЛИВАТИ НА ЗМІСТ ДАНИХ.

ПРАВИЛО ІНФОРМАЦІЇ. ВСЯ ІНФОРМАЦІЯ В БАЗІ ДАНИХ ПОВИННА БУТИ ПРЕДСТАВЛЕНА ВИКЛЮЧНО НА ЛОГІЧНОМУ РІВНІ І ЛИШЕ ОДНИМ СПОСОБОМ – У ВИГЛЯДІ ЗНАЧЕНЬ, ЩО МІСТЯТЬСЯ В ТАБЛИЦЯХ. ФАКТИЧНО ЦЕ **НЕФОРМАЛЬНЕ ВИЗНАЧЕННЯ РЕЛЯЦІЙНОЇ БАЗИ ДАНИХ.**

ПРАВИЛО 2

ПРАВИЛО ГАРАНТОВАНОГО ДОСТУПУ (GUARANTEED ACCESS RULE)

ДОСТУП ДО ДАНИХ ПОВИНЕН БУТИ ВІЛЬНИМ ВІД ДВОЗНАЧНОСТІ. ДО КОЖНОГО ЕЛЕМЕНТУ ДАНИХ ПОВИНЕН БУТИ ГАРАНТОВАНИЙ ДОСТУП ЗА ДОПОМОГОЮ КОМБІНАЦІЇ ІМЕНІ ТАБЛИЦІ, ПЕРВИННОГО КЛЮЧА РЯДКУ Й ІМЕНІ СТОВПЦЯ.

ПРАВИЛО 2 ВКАЗУЄ НА РОЛЬ ПЕРВИННОГО КЛЮЧА ПРИ ПОШУКУ ІНФОРМАЦІЇ В БАЗІ ДАНИХ. ІМ'Я ТАБЛИЦІ ДОЗВОЛЯЄ ЗНАЙТИ НЕОБХІДНУ ТАБЛИЦЮ, ІМ'Я СТОВПЦЯ ДОЗВОЛЯЄ ЗНАЙТИ ПОТРІБНИЙ СТОВПЕЦЬ, А ПЕРВИННИЙ КЛЮЧ ДОЗВОЛЯЄ ЗНАЙТИ РЯДОК, ЩО МІСТИТЬ ШУКАНИЙ ЕЛЕМЕНТ ДАНИХ.

ПРАВИЛО 3

СИСТЕМАТИЧНА ОБРОБКА NULL-ЗНАЧЕНЬ (SYSTEMATIC TREATMENT OF NULL VALUES)

НЕВІДОМІ ЗНАЧЕННЯ NULL, ВІДМІННІ ВІД БУДЬ-ЯКОГО ВІДОМОГО ЗНАЧЕННЯ, ПОВИННІ ПІДТРИМУВАТИСЯ ДЛЯ ВСІХ ТИПІВ ДАНИХ ПРИ ВИКОНАННІ БУДЬ-ЯКИХ ОПЕРАЦІЙ. НАПРИКЛАД, ДЛЯ ЧИСЛОВИХ ДАНИХ НЕВІДОМІ ЗНАЧЕННЯ НЕ ПОВИННІ РОЗГЛЯДАТИСЯ ЯК НУЛІ, А ДЛЯ СИМВОЛЬНИХ ДАНИХ – ЯК ПОРОЖНІ РЯДКИ.

ПРАВИЛО 3 ВИМАГАЄ, ЩОБ ВІДСУТНІ ДАНІ МОЖНА БУЛО УЯВИТИ ЗА ДОПОМОГОЮ НЕДІЙСНИХ ЗНАЧЕНЬ (NULL).

ПРАВИЛО 4

ПРАВИЛО ДОСТУПУ ДО СИСТЕМНОГО КАТАЛОГУ НА ОСНОВІ РЕЛЯЦІЙНОЇ МОДЕЛІ (DYNAMIC ON-LINE CATALOG BASED ON THE RELATIONAL MODEL)

СЛОВНИК ДАНИХ ПОВИНЕН ЗБЕРІГАТИСЯ У ФОРМІ РЕЛЯЦІЙНИХ ТАБЛИЦЬ, І СУБД ПОВИННА ПІДТРИМУВАТИ ДОСТУП ДО НЬОГО ЗА ДОПОМОГОЮ СТАНДАРТНИХ МОВНИХ ЗАСОБІВ, ТИХ САМИХ, ЯКІ ВИКОРИСТОВУЮТЬСЯ ДЛЯ РОБОТИ З РЕЛЯЦІЙНИМИ ТАБЛИЦЯМИ, ЩО МІСТЯТЬ ДАНІ КОРИСТУВАЧА.

ПРАВИЛО 4 СВДЧИТЬ, ЩО РЕЛЯЦІЙНА БАЗА ДАНИХ ПОВИННА САМА СЕБЕ ОПИСУВАТИ. ІНШИМИ СЛОВАМИ, БАЗА ДАНИХ ПОВИННА МІСТИТИ НАБІР СИСТЕМНИХ ТАБЛИЦЬ, ЩО ОПИСУЮТЬ СТРУКТУРУ САМОЇ БАЗИ ДАНИХ.

ПРАВИЛО 5

ПРАВИЛО ПОВНОТИ ПІДМОВИ МАНІПУЛЮВАННЯ ДАНИМИ (COMPREHENSIVE DATA SUBLANGUAGE RULE)

СИСТЕМА УПРАВЛІННЯ РЕЛЯЦІЙНИМИ БАЗАМИ ДАНИХ ПОВИННА ПІДТРИМУВАТИ ХОЧА Б ОДНУ РЕЛЯЦІЙНУ МОВУ, ЯКА

А) МАЄ ЛІНІЙНИЙ СИНТАКСИС,

Б) МОЖЕ ВИКОРИСТОВУВАТИСЯ ЯК ІНТЕРАКТИВНОЮ, ТАК І В ПРИКЛАДНИХ ПРОГРАМАХ,

В) ПІДТРИМУЄ ОПЕРАЦІЇ ВИЗНАЧЕННЯ ДАНИХ, ВИЗНАЧЕННЯ УЯВЛЕНЬ, МАНІПУЛЮВАННЯ ДАНИМИ (ІНТЕРАКТИВНІ ТА ПРОГРАМНІ), ОБМЕЖУВАЧІ ЦІЛІСНОСТІ, УПРАВЛІННЯ ДОСТУПОМ ТА ОПЕРАЦІЇ УПРАВЛІННЯ ТРАНЗАКЦІЯМИ (BEGIN, COMMIT І ROLLBACK).

ПРАВИЛО 5 ВИМАГАЄ, ЩОБ СУБД ВИКОРИСТОВУВАЛА МОВУ РЕЛЯЦІЙНОЇ БАЗИ ДАНИХ. ТАКА МОВА ПОВИННА ПІДТРИМУВАТИ ВСІ ОСНОВНІ ФУНКЦІЇ СУБД – СТВОРЕННЯ БАЗИ ДАНИХ, ЧИТАННЯ І ВВЕДЕННЯ ДАНИХ, РЕАЛІЗАЦІЮ ЗАХИСТУ БАЗИ ДАНИХ І Т.Д.

ПРАВИЛО 6

ПРАВИЛО МОДИФІКАЦІЇ ПРЕДСТАВЛЕНЬ (VIEW UPDATING RULE)

КОЖНЕ ПОДАННЯ МАЄ ПІДТРИМУВАТИ ВСІ ОПЕРАЦІЇ МАНІПУЛЮВАННЯ ДАНИМИ, ЯКІ ПІДТРИМУЮТЬ РЕЛЯЦІЙНІ ТАБЛИЦІ: ОПЕРАЦІЇ ВИБІРКИ, ВСТАВКИ, МОДИФІКАЦІЇ І ВИДАЛЕННЯ ДАНИХ.

ПРАВИЛО 6 СТОСУЄТЬСЯ УЯВЛЕНЬ, ЯКІ Є ВІРТУАЛЬНИМИ ТАБЛИЦЯМИ, ЯКІ ДОЗВОЛЯЮТЬ ПОКАЗУВАТИ РІЗНИМ КОРИСТУВАЧАМ РІЗНІ ФРАГМЕНТИ СТРУКТУРИ БАЗИ ДАНИХ.

ПРАВИЛО 7

ПРАВИЛО ВИСОКОРІВНЕВИХ ОПЕРАЦІЙ МОДИФІКАЦІЇ ДАНИХ (HIGH-LEVEL INSERT, UPDATE, AND DELETE)

ОПЕРАЦІЇ ВСТАВКИ, МОДИФІКАЦІЇ І ВИДАЛЕННЯ ДАНИХ ПОВИННІ ПІДТРИМУВАТИСЯ НЕ ТІЛЬКИ ПО ВІДНОШЕННЮ ДО ОДНОГО РЯДКУ РЕЛЯЦІЙНОЇ ТАБЛИЦІ, АЛЕ ПО ВІДНОШЕННЮ ДО БУДЬ-ЯКОЇ БЕЗЛІЧІ РЯДКІВ.

ПРАВИЛО 7 АКЦЕНТУЄ УВАГУ НА ТОМУ, ЩО БАЗИ ДАНИХ ЗА СВОЄЮ ПРИРОДОЮ ОРІЄНТОВАНІ НА МНОЖИНИ. ВОНО ВИМАГАЄ, ЩОБ ОПЕРАЦІЇ ДОДАВАННЯ, ВИДАЛЕННЯ І ОНОВЛЕННЯ МОЖНА БУЛО ВИКОНУВАТИ НАД МНОЖИНАМИ РЯДКІВ.

ПРАВИЛО 8

ПРАВИЛО ФІЗИЧНОЇ НЕЗАЛЕЖНОСТІ ДАНИХ (PHYSICAL DATA INDEPENDENCE)

ДОДАТКИ НЕ ПОВИННІ ЗАЛЕЖАТИ ВІД ВИКОРИСТОВУВАНИХ СПОСОБІВ ЗБЕРІГАННЯ ДАНИХ НА НОСІЯХ, ВІД АПАРАТНОГО ЗАБЕЗПЕЧЕННЯ КОМП'ЮТЕРІВ, НА ЯКИХ ЗНАХОДИТЬСЯ РЕЛЯЦІЙНА БАЗА ДАНИХ.

ПРИКЛАДНІ ПРОГРАМИ І УТИЛІТИ ДЛЯ РОБОТИ З ДАНИМИ ПОВИННІ НА ЛОГІЧНОМУ РІВНІ ЗАЛИШАТИСЯ НЕДОТОРКАНИМИ ЗА БУДЬ-ЯКИХ ЗМІНАХ СПОСОБІВ ЗБЕРІГАННЯ ДАНИХ АБО МЕТОДІВ ДОСТУПУ ДО НИХ.

ПРАВИЛО 9

ПРАВИЛО ЛОГІЧНОЇ НЕЗАЛЕЖНОСТІ ДАНИХ (LOGICAL DATA INDEPENDENCE)

ПРЕДСТАВЛЕННЯ ДАНИХ В ДОДАТКУ НЕ ПОВИННО ЗАЛЕЖАТИ ВІД СТРУКТУРИ РЕЛЯЦІЙНИХ ТАБЛИЦЬ. ЯКЩО В ПРОЦЕСІ НОРМАЛІЗАЦІЇ ОДНА РЕЛЯЦІЙНА ТАБЛИЦЯ РОЗДІЛЯЄТЬСЯ НА ДВІ, ПОДАННЯ ПОВИННЕ ЗАБЕЗПЕЧИТИ ОБ'ЄДНАННЯ ЦИХ ДАНИХ, ЩОБ ЗМІНА СТРУКТУРИ РЕЛЯЦІЙНИХ ТАБЛИЦЬ НЕ ПОЗНАЧАЛОСЯ НА РОБОТІ ДОДАТКІВ.

ПРАВИЛА 8 І 9 ОЗНАЧАЮТЬ ВІДДІЛЕННЯ КОРИСТУВАЧА ТА ПРИКЛАДНОЇ ПРОГРАМИ ВІД НИЗЬКОРІВНЕВОЇ РЕАЛІЗАЦІЇ БАЗИ ДАНИХ.

ПРАВИЛО 10

ПРАВИЛО НЕЗАЛЕЖНОСТІ КОНТРОЛЮ ЦІЛІСНОСТІ (INTEGRITY INDEPENDENCE)

ВСЯ ІНФОРМАЦІЯ, НЕОБХІДНА ДЛЯ ПІДТРИМКИ ЦІЛІСНОСТІ, ПОВИННА БУТИ В СЛОВНИКУ ДАНИХ. МОВА ДЛЯ РОБОТИ З ДАНИМИ ПОВИННА ВИКОНУВАТИ ПЕРЕВІРКУ ВХІДНИХ ДАНИХ І АВТОМАТИЧНО ПІДТРИМУВАТИ ЦІЛІСНІСТЬ ДАНИХ.

ПРАВИЛО 10 ГОВОРИТЬ, ЩО МОВА БАЗИ ДАНИХ ПОВИННА ПІДТРИМУВАТИ ОБМЕЖУВАЛЬНІ УМОВИ, ЩО НАКЛАДАЮТЬСЯ НА ДАНІ, ЯКІ ВВОДЯТЬСЯ, І ДІЇ, ЯКІ МОЖУТЬ БУТИ ВИКОНАНІ НАД ДАНИМИ.

ПРАВИЛО 11

ПРАВИЛО НЕЗАЛЕЖНОСТІ ВІД РОЗМІЩЕННЯ (DISTRIBUTION INDEPENDENCE)

БАЗА ДАНИХ МОЖЕ БУТИ РОЗПОДІЛЕНОЮ, МОЖЕ ПЕРЕБУВАТИ НА ДЕКІЛЬКОХ КОМП'ЮТЕРАХ, І ЦЕ НЕ ПОВИННО ВПЛИВАТИ НА ДОДАТКИ. ПЕРЕНЕСЕННЯ БАЗИ ДАНИХ НА ІНШИЙ КОМП'ЮТЕР НЕ ПОВИННО ВПЛИВАТИ НА ДОДАТКИ.

ПРАВИЛО 11 ГОВОРИТЬ, ЩО МОВА БАЗИ ДАНИХ ПОВИННА ЗАБЕЗПЕЧУВАТИ МОЖЛИВІСТЬ РОБОТИ З РОЗПОДІЛЕНИМИ ДАНИМИ, РОЗТАШОВАНИМИ НА ІНШИХ КОМП'ЮТЕРНИХ СИСТЕМАХ.

ПРАВИЛО 12

ПРАВИЛО УЗГОДЖЕНОСТІ МОВНИХ РІВНІВ (THE NONSUBVERSION RULE)

ЯКЩО ВИКОРИСТОВУЄТЬСЯ НИЗЬКОРІВНЕВА МОВА ДОСТУПУ ДО ДАНИХ, ВОНА НЕ ПОВИННА ІГНОРУВАТИ ПРАВИЛА БЕЗПЕКИ І ПРАВИЛА ЦІЛІСНОСТІ, ЯКІ ПІДТРИМУЮТЬСЯ МОВОЮ БІЛЬШ ВИСОКОГО РІВНЯ.

ПРАВИЛО 12 ЗАПОБІГАЄ ВИКОРИСТАННЮ ІНШИХ МОЖЛИВОСТЕЙ ДЛЯ РОБОТИ З БАЗОЮ ДАНИХ КРІМ МОВИ БАЗИ ДАНИХ, ОСКІЛЬКИ ЦЕ МОЖЕ ПОРУШИТИ ЇЇ ЦІЛІСНІСТЬ.

ПОСЛІДОВНІСТЬ НОРМАЛЬНИХ ФОРМ

- ПЕРША НОРМАЛЬНА ФОРМА (1NF);
- ДРУГА НОРМАЛЬНА ФОРМА (2NF);
- ТРЕТЯ НОРМАЛЬНА ФОРМА (3NF);
- НОРМАЛЬНА ФОРМА БОЙСА–КОДДА (BCNF);
- ЧЕТВЕРТА НОРМАЛЬНА ФОРМА (4NF);
- П'ЯТА НОРМАЛЬНА ФОРМА, АБО НОРМАЛЬНА ФОРМА ПРОЕКЦІЇ–З'ЄДНАННЯ (5NF АБО PJ/NF).

ОСНОВНІ ВЛАСТИВОСТІ НОРМАЛЬНИХ ФОРМ



КОЖНА НАСТУПНА НОРМАЛЬНА ФОРМА В ДЕЯКОМУ СЕНСІ КРАЩА ЗА ПОПЕРЕДНЮ



ПРИ ПЕРЕХОДІ ДО НАСТУПНОЇ НОРМАЛЬНОЇ ФОРМИ ВЛАСТИВОСТІ ПОПЕРЕДНІХ НОРМАЛЬНИХ ВЛАСТИВОСТЕЙ ЗБЕРІГАЮТЬСЯ

ПЕРША НОРМАЛЬНА ФОРМА

Перша нормальна форма (1NF) – значення всіх атрибутів відносин атомарні

ФУНКЦІОНАЛЬНІ ЗАЛЕЖНОСТІ

Функціональна залежність

• У відношенні R атрибут Y функціонально залежить від атрибута X (X і Y можуть бути складеними) в тому і тільки в тому випадку, якщо кожному значенню X відповідає в точності одне значення Y : $R.X \rightarrow R.Y$

Повна функціональна залежність

• Функціональна залежність $R.X \rightarrow R.Y$ називається повною, якщо атрибут Y не залежить функціонально від будь-якої точної підмножини X

Транзитивна функціональна залежність

• Функціональна залежність $R.X \rightarrow R.Y$ називається транзитивною, якщо існує такий атрибут Z , що є функціональні залежності $R.X \rightarrow R.Z$ і $R.Z \rightarrow R.Y$ і відсутня функціональна залежність $R.Z \rightarrow R.X$

АТРИБУТИ

Неключовий атрибут

- Неключовим атрибутом називається будь-який атрибут відносини, що не входить до складу первинного ключа

Взаємно незалежні атрибути

- Два або більше атрибута є взаємно незалежними, якщо жоден з цих атрибутів не є функціонально залежним від іншого

ДРУГА НОРМАЛЬНА ФОРМА

Відношення **R** знаходиться в другій нормальній формі (**2NF**) в тому і тільки в тому випадку, коли знаходиться в **1NF**, і кожен неключовий атрибут повністю залежить від первинного ключа.

ТРЕТЯ НОРМАЛЬНА ФОРМА

Відношення **R** знаходиться в третій нормальній формі (**3NF**) в тому і тільки в тому випадку, якщо знаходиться в **2NF** і кожен неключовий атрибут нетранзитивно залежить від первинного ключа

НОРМАЛЬНА ФОРМА БОЙСА–КОДДА

Детермінант – будь-який атрибут, від якого повністю функціонально залежить деякий інший атрибут

Відношення **R** знаходиться в нормальній формі Бойса–Кодда (**BCNF**) в тому і тільки в тому випадку, якщо кожен детермінант є можливим ключем

ЧЕТВЕРТА НОРМАЛЬНА ФОРМА

У відношенні $R(A, B, C)$ існує багатозначна залежність $R.A \twoheadrightarrow R.B$ в тому і тільки в тому випадку, якщо безліч значень B , яке відповідає парі значень A і C , залежить тільки від A і не залежить від C

Відношення R знаходиться в четвертій нормальній формі (4NF) в тому і тільки в тому випадку, якщо в разі існування багатозначної залежності $A \twoheadrightarrow B$ всі інші атрибути R функціонально залежать від A

П'ЯТА НОРМАЛЬНА ФОРМА

Залежність з'єднання

Відношення $R(X, Y, \dots, Z)$ задовольняє залежності з'єднання $*$ (X, Y, \dots, Z) в тому і тільки в тому випадку, коли R відновлюється без втрат шляхом з'єднання своїх проєкцій на X, Y, \dots, Z

Відношення R знаходиться в п'ятій нормальній формі (нормальній формі проєкції-з'єднання – PJ / NF) в тому і тільки в тому випадку, коли будь-яка залежність з'єднання в R впливає з існування деякого можливого ключа в R