

# Модели данных

## А3. Ограничения целостности



Московский государственный технический университет  
имени Н.Э. Баумана

**Факультет ИБМ**

Июль 2024 года

Москва

Артемьев Валерий Иванович © 2024

# 5. Ограничения целостности транзакционных данных

- Что такое ограничения целостности базы данных?
- Ограничения целостности на уровне СУБД
- Целостность данных (Data Integrity)
- Домен данных
- Регулярные выражения
- Ссылочная целостность (Reference Integrity)
- Описание ограничений целостности в ЛМД
- Пример описания ограничений целостности в ЛДМ «Домашняя библиотека»
- Описание ограничений целостности в ФМД
- Пример описания ограничений целостности в ЛДМ «Домашняя библиотека»
- Целостность транзакций (Transaction Integrity)
- В качестве резюме

# Целостность базы данных

*Совокупность характеристик качества элементов данных в БД:*

- *полнота,*
- *допустимость,*
- *целостность ссылок,*
- *согласованность,*
- *уникальность,*
- *непрерывность.*

Существуют разные возможности обеспечения качества БД.

# Характеристики качества базы данных



# Возможности обеспечения качества базы данных

1. *Упрощение ввода данных* с помощью выбора из списка в любом приложении
2. *Контроль качества данных при вводе* значений атрибутов и добавлении и редактировании записей
3. *Использование возможностей систем управления базами данных (СУБД)* в учётно-операционных приложениях
4. Контроль и очистка данных производится *при ведении справочников и реестров*
5. Контроль и очистка *при интеграции данных из разных источников в аналитические БД.*

# Средства СУБД для контроля качества данных

1. *Декларативные ограничения целостности* атрибутов, ключей и записей, а также утверждения.
2. *Транзакции* обеспечивают непрерывность операций.
3. *Триггеры* – процедуры событийной проверки данных.
4. *Хранимые процедуры* обеспечивают сложные алгоритмы контроля данных в СУБД.

# Ограничения целостности баз данных

*Накладываемые на элементы данных условия, ложные значения которых говорят о нарушении целостности состояния базы данных и обрабатываются определённым образом.*

Ограничения целостности определяются

- *содержанием данных,*
- *структурой данных,*
- *смыслом данных и бизнес-правилами.*

# Какие бывают ограничения целостности?

Ограничения целостности обеспечиваются на разных уровнях:

- *Ограничения целостности атрибутов*
- *Ограничения целостности ключей*
- *Ограничения записи*
- *Ограничения таблицы*
- *Ограничения схемы*

Проверки бывают *немедленными до сохранения данных* или *отложенные после завершения ряда операций*, составляющих транзакцию.



# Ограничения целостности атрибутов

Для определения области допустимых значений задают:

- Тип данных атрибута.
- Разрядность/ длина атрибута.
- Обязательность значения (NOT NULL).
- Задание диапазона значений.
- Уникальность значения атрибута.
- Формат (шаблон) атрибута.

Все эти ограничения могут быть поименованы и представлять собой *домен данных (domain)* или *пользовательские типы данных (User Data Type, UDT)*

Примеры простых доменов:

Тип\_ИНН\_юрлиц или Цифры(10)

Деньги или Десятичное(19,4)

Охват: “малый”; “средний”; “большой”

# Примеры ограничений атрибутов

Значение положительное

**CHECK (value > 0) SET DEFAULT 1**

Диапазон от 1 до 1000 включительно

**CHECK (value BETWEEN 1 AND 1000)**

Список городов: “Москва”, “Воронеж”, “Санкт-Петербург”

**CHECK (city IN ('Москва', 'Воронеж', 'Санкт-Петербург'))**

Имя не начинается на “Г-н”, “Г-жа”

**CHECK (name NOT LIKE 'Г-н %' AND NOT LIKE 'Г-жа %')**

В шаблоне LIKE имеются 2 простых метасимвола:

*% любая последовательность символов*

*\_ любой одиночный символ*

# Регулярные выражения (regex)

*Механизм расширенного поиска и замены текста на основе специального шаблона.*

Появился в системе Unix в виде команды `grep/egrep` для поиска, фильтрации и замены в текстовых файлах.

Недавно ещё популярный язык PERL наиболее полно поддерживал регулярные выражения.

Многие языки программирования и СУБД ограниченно поддерживают регулярные выражения.

Даже в Word (но почему-то не в Excel) есть поиск и замена с минимальными возможностями regex.

# Синтаксис регулярных выражений

## Метасимволы, квантификаторы и операторы привязки

Символ	Описание
^	Привязка выражения к началу строки
\$	Привязка выражения к концу строки
*	Символ встречается 0 или более раз
?	Символ встречается 0 или 1 раз
+	Символ встречается 1 и более раз
{m}	Встречается ровно m раз
{m,}	Встречается минимум m раз
{m, n}	Встречается минимум m раз, но не более n раз
	Разделяет альтернативные варианты
( )	Группирует подвыражения для альтернативы или квантификатора
[char]	Обозначает список символов
[^char]	Все символы, кроме указанных
[0-9]	Диапазон символов

# Примеры регулярных выражений (1)

RGB-код цвета **#XXXXXX** (6 шестнадцатерич. цифр)

**# [0-9A-Fa-f] {6}**

# просто символ решётки

[0-9A-Fa-f] шестнадцатеричные цифры

{6} повторить 6 раз предшествующую группу

**Нумерация разделов, пунктов и подпунктов**

1 или 1.2 или 1.2.3 (одна цифра уровня)

**[1-9] (\.[1-9]){0,2}** или **[1-9]|([1-9](\.[1-9])){1,2}**

[1-9] цифры от 1 до 9

( ) группа символов

\. просто точка, \ экранирует метасимвол точку

{0,2} повторить от 0 до 2 раз

# Примеры регулярных выражений (2)

Телефонные номера в Москве 8(999)999-99-99

**`(\+7|8)\([0-9]{3}\)[0-9]{3}(-[0-9][0-9]){2}`**

`(\+7|8)` скобки группируют коды страны

`\` символ экранирования метасимволов

`\+` просто плюс

`|` метасимвол альтернативы +7 или 8

`\( \)` открывающаяся и закрывающаяся скобки

`[0-9]` диапазон символов, здесь цифры

`{3}` квантификатор: 3 раза повторить цифру

`(-[0-9][0-9])` группа из минуса и 2-х цифр

`{2}` повторить 2 раза предшествующую группу

Онлайн  
сервис  
**Regex101**

Ограничение формата номера телефона

**CHECK (tel REGEXP '8\([0-9]{3}\)[0-9]{3}(-[0-9][0-9]){2}')**

# Ограничения целостности ключей

## Ограничения целостности *первичных* ключей:

- явно указывается, что простой атрибут является *первичным ключом, т.е. непустым и уникальным*
- указывается, что список атрибутов является *первичным составным ключом*.

## Ограничения целостности *альтернативных* ключей:

- указывается, что значения атрибута *уникально*
- или совокупность значений атрибутов *уникальна*

# Примеры ограничений целостности ключей

Простой атрибут id является *первичным ключом*:

**id INTEGER PRIMARY KEY**

Список атрибутов name, DoB является *первичным составным ключом*:

**PRIMARY KEY (name, DoB)**

Значения атрибута employ\_NR *уникально*:

**employ\_NR CHAR(5) NOT NULL UNIQUE**

*Совокупность значений атрибутов уникальна*:

**UNIQUE (dept\_NR, unit\_NR)**



# Ограничения ссылочной целостности

- Явно указывается, что простой атрибут является *внешним ключом для таблицы и первичного ключа*
- Задаётся действие при нарушении ограничения:
  - *сделать ссылку пустой*
  - *ничего не делать*
  - *при удалении «родителя» каскадом удалить «детей»*
  - *при изменении ключа «родителя» каскадом обновить ссылки у «детей»*

# Примеры ограничений ссылочной целостности

*Простой атрибут является внешним ключом для таблицы и первичного ключа:*

**FOREIGN KEY (city\_ID) REFERENCES Cities(city\_ID)**

*Задаётся действие при нарушении ограничения:*

- *сделать ссылку пустой* **SET NULL**
- *ничего не делать* **NO ACTION**
- *при удалении «родителя» каскадом удалить «детей»* **ON DELETE CASCADE**
- *при изменении ключа «родителя» каскадом обновить ссылки у «детей»* **ON UPDATE CASCADE**

# Пример ограничений целостности в ЛМД «Домашняя библиотека»

ТЕМА: тема\_ID **целое(2), ключ**,  
тема\_название **строка(40), непустая, уникальная**

КНИГА: книга\_ID **целое (3), ключ < 10000**,  
книга\_название **строка (100), непустая**,  
книга\_аннотация **строка (1000), непустая**,  
издатель\_ID **целое(3), непустое, ссылка на (ИЗДАТЕЛЬ.издатель\_ID)**,

...

СОЗДАТЕЛЬ: создатель\_ID **целое (3), ключ**,  
создатель\_ФИО **строка (50), непустая, уникальная**

СОЗДАТЕЛЬ\_РОЛЬ: роль\_ID **целое(3), ключ**,  
роль\_название **строка(20), непустая, уникальная**

УЧАСТИЕ: книга\_ID **целое(3), непустое, ссылка на (КНИГА.книга\_ID)**,  
создатель\_ID **целое(3), непустое, ссылка на (СОЗДАТЕЛЬ.создатель\_ID)**,  
роль\_ID **целое(3), непустое, ссылка на (СОЗДАТЕЛЬ\_РОЛЬ.роль\_ID)**,  
**ключ (книга\_ID, создатель\_ID, роль\_ID)**

# Пример ограничений целостности в ФМД «Домашняя библиотека»

Topics (topic\_ID SMALLINT PRIMARY KEY,  
topic\_title VARCHAR(40) NOT NULL UNIQUE),

Books (book\_ID SMALLINT PRIMARY KEY CHECK (VALUE < 10000),  
book\_title VARCHAR(100) NOT NULL,  
book\_annotation VARCHAR(1000) NOT NULL,  
publisher\_ID SMALLINT,  
FOREIGN KEY (publisher\_ID) REFERENCES Publishers(publisher\_ID)

...

Creators (creator\_ID SMALLINT PRIMARY KEY,  
creator\_name VARCHAR(50) NOT NULL UNIQUE)

Creator\_roles (role\_ID SMALLINT PRIMARY KEY,  
role\_title VARCHAR(20) NOT NULL UNIQUE)

Parties (book\_ID SMALLINT, creator\_ID SMALLINT, role\_ID SMALLINT,  
PRIMARY KEY (book\_ID, creator\_ID, role\_ID),  
FOREIGN KEY book\_ID REFERENCES Books(book\_ID)  
ON DELETE CASCADE ON UPDATE CASCADE,  
FOREIGN KEY creator\_ID REFERENCES Creators(creator\_ID) ...  
FOREIGN KEY role\_ID REFERENCES Creator\_roles(role\_ID) ...

# Целостность транзакций

Транзакция – неделимая последовательность операторов манипулирования данными (чтения, удаления, вставки, модификации), приводящая к одному из двух возможных результатов:

- либо последовательность выполняется, если все операторы правильные,
- либо вся транзакция откатывается, если хотя бы один оператор не может быть успешно выполнен.

Механизм транзакций обеспечивает *непрерывность БД*, а изолированность данных поддерживает *доступность данных при параллельной работе* пользователей.

# В качестве резюме

- Ограничения целостности БД – важный механизм обеспечения качества данных в транзакционных системах и при ведении справочников и реестров.
- Нужно учитывать наличие контроля данных при вводе, не дублировать его в СУБД.
- Ограничения целостности создают накладные расходы и снижают производительность.
- В аналитических системах и хранилищах данных используют внешние механизмы интеграции и контроля данных, не используют ограничения целостности БД.
- Обратите внимание на регулярные выражения.

**Терпения и удачи всем, кто связан  
с моделированием данных**

**Спасибо за внимание!**

**Валерий Иванович Артемьев**

**Департамент статистики и управления данными  
Банк России**

**Тел.: +7(495) 753-96-25**

**e-mail: [avi@cbr.ru](mailto:avi@cbr.ru)**