

УДК 681.3.06

Т.П. Колісник

Харківський національний університет внутрішніх справ, Харків

ПІДТРИМКА ЦІЛІСНОСТІ У БАЗАХ ДАНИХ З НЕОДНОРІДНОЮ СТРУКТУРОЮ

У роботі розглянуто підхід до організації підтримки цілісності при інтеграції неоднорідних розподілених баз даних. Визначено види порушень і виділені аномалії, що не можна підтримувати засобами системи управління базами даних. Якщо представити базу даних – як набір символів, обмеження – як примірник правил, а модифікацію – як додавання або видалення символу з множини, то можна забезпечити незалежність опису бази даних від фізичного представлення даних в інформаційній системі.

Ключові слова: аномалія, транзакція, база даних, порушення цілісності, підтримка даних, неоднорідність структури.

Вступ

Одним з важливих питань при організації розподілених інформаційних систем (ІС) є питання обробки транзакцій, що змінюють стан бази даних (БД), яка входить в ІС. Для коректної роботи БД необхідно забезпечити виконання операцій таким чином, щоб не порушувалася цілісність і узгодженість даних.

Постановка проблеми. Визначимо цілісність як твердження, яке може бути істинним або хибним залежно від стану бази даних, а узгодженість як стани БД, при якому виконані всі обмеження цілісності.

Якщо система управління базою даних (СУБД) не підтримує всі необхідні обмеження, то така БД може знаходитися в цілісному стані щодо інструкцій СУБД, але не буде правильною з точки зору користувача. Якщо всі обмеження виконані, то будемо говорити, що дані БД знаходяться в коректному стані.

Разом з поняттям цілісності БД виникає завдання обробки реакції на спробу порушення цілісності. Система повинна не тільки перевіряти, чи не порушуються обмеження в ході виконання різних операцій, але і якимось чином реагувати, якщо операція призводить до порушення цілісності. На практиці використовуються два типи реакції на спробу порушення цілісності:

1. Відмова виконати некоректну операцію.
2. Виконання компенсуючих дій.

Мета статті. Підтримка цілісності і узгодженості даних, з огляду на те, що розподілені системи підтримують локальну автономію і незалежність вузлів, являє собою складну задачу. Тому слід детально розглянути підхід до організації підтримки цілісності при інтеграції неоднорідних розподілених баз даних, а також визначити види порушень і виділені аномалії, що не можна підтримувати засобами системи управління базами даних. У сучасних системах синхронне змінювання даних в декількох локальних однорідних БД досягається застосуванням протоколу двофазної фіксації транзакцій. У випадку неоднорідності БД для забезпечення узгоджених змін в декількох базах даних використовують мене-

джери розподілених транзакцій [1, 2]. Якщо в БД передбачено тиражування даних, то це визначає додаткові вимоги до підтримки цілісності на вузлах, куди спрямовані потоки даних. Проблема в тому, що зміни в даних можуть ініціюватися як локально, так і ззовні, за допомогою тиражування. При цьому неминуче виникають конфлікти в процесі змін, які необхідно відстежувати і вирішувати.

У ряді робіт присвячених опису функціонування БД [1, 2, 4, 5] приділяється багато уваги ситуаціям, при яких небажані ефекти (аномалії) виникають в результаті зміни стану БД.

Виклад основного матеріалу

Види порушень цілісності. Для більшої наочності і розуміння суті проблеми приклади будемо розглядати в контексті реляційної БД.

Нехай D – множина доменів (множина допустимих значень), $D = \{D_1, D_2, \dots, D_n\}$, де D_1 – множина значень і нехай R – множина атрибутів $R = \{A_1, A_2, \dots, A_n\}$, тоді відображення виду $Dom: R \rightarrow D$ встановлює для кожного атрибуту, з якого домену вибираються значення при формуванні кортежу. Множина $T = Dom_{A_1} \times Dom_{A_2} \times \dots \times Dom_{A_n}$ визначає множини n – кортежів відносини.

Будь-яке відношення БД має, принаймні, один можливий ключ. Підмножину атрибутів $K \subseteq R$ називають ключем відношення, якщо вона володіє двома властивостями: унікальністю – ніякі допустимі значення R не містять двох різних кортежів з однаковими значеннями K і мінімальністю – жодне з підмножин R не має властивість унікальності.

Якщо на R задано множини обмежень цілісності I , то $Sch(R, I)$ будемо називати схемою відносини (або, в загальному, випадки схемою БД, якщо R представляє сукупність атрибутів всіх відносин БД).

Серед множини обмежень цілісності для початку виділімо внутрішні обмеження, тобто обмеження, які визначає модель даних. У нашому випадку це обмеження ключа, що не допускає однакових зна-

чень в заданих атрибутах і обмеження на значення, що визначаються доменами кожного атрибута. У роботі [5] такі обмеження називаються K – залежність і D – залежність відповідно.

Нехай $Sch(R, I)$ – схема відносини і $r \subseteq T$ – допустимий стан Sch – тобто така структура БД, і набір даних при якому виконані всі обмеження цілісності, і нехай t – довільний кортеж, заданий на множині R . Кортеж t називається сумісним зі станом r зі схемою $Sch(R, I)$, якщо:

- 1) $t[K] \neq r[K]$ для будь якого K ;
- 2) $t[A_i] \subset \text{Dom}A_i$ для кожного $A_i \in R, i = \overline{1, n}$.

Іншими словами, кортеж t називається сумісним з r , якщо він може бути доданий до відношення без порушення D - і K - залежностей.

База даних буде мати аномалію, якщо r – дозволений стан, t – кортеж, сумісний з r , але $r[t]$ – новий стан (з включеним кортежем), не є допустимим.

У загальному випадку аномалія – це протиріччя між деякою концептуальною схемою предметної галузі (ПрГ) і схемою БД. Концептуальну схему (КС) визначимо як спосіб представлення знань про ПрГ, тобто такий узгоджений набір пропозицій, що виражають твердження, істинні для будь-якого стану ПрГ. При цьому, схему БД можна визначити як формальний спосіб представлення знань про ПрГ засобами конкретного додатка. Всі знання про ПрГ фіксуються у вигляді набору пропозицій, можливо на природній мові, які виражають ствердження про об'єкти, їх взаємозв'язки, класифікації об'єктів, правила і обмеження, тобто описуються КС.

Аномалії виникають в тому випадку, коли знання про ПрГ виявляються невимовними в схемі БД або, принаймні, входять з нею в суперечність. Згідно з [6] схема відносин відповідає концептуальній схемі ПрГ, а обмеження, що підтримуються СУБД – схемі БД. Таким чином, аномалії виникають через те, що обмеження, які входять у схему відносин, неможливо виразити через обмеження, які підтримуються СУБД.

Визначимо інформаційну базу (ІБ) як набір пропозицій, узгоджених один з одним і з КС, що виражають твердження, істинні для даного стану ПрГ.

Нехай L – множина всіх можливих пропозицій ІБ. Кожній пропозиції ставиться у відповідність оцінка істинності. Правила, що входять до КС, виділяють підмножину допустимих пропозицій $I \subseteq L$. Пропозиції з $L - I$ є відповідно хибними. Безліч пропозицій, яким у поточному стані ПрГ відповідає оцінка "Істина" називають актуальними пропозиціями ІБ в позначенні L_A пропозиції, що належать множині $L - L_A$, називають неактуальними пропозиціями ІБ.

Очевидно, що множина L_A буде збільшуватися в міру функціонування інформаційної системи. Схематично ця ситуація представлена на рис. 1.

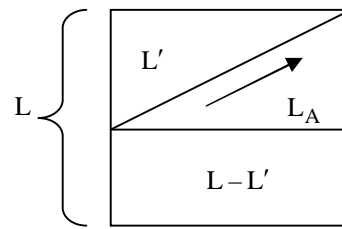


Рис. 1. Структура інформаційної бази

Розглянемо аналогічну ситуацію на рівні БД. Нехай T – множина всіх можливих пропозицій БД, що описують ПрГ. Згідно з обмеженнями у схемі БД, обумовленими КС, виділяється підмножина T' допустимих кортежів БД. Очевидно, що кортежі з підмножини $T - T'$ є недопустимими.

Нехай T_A – множина кортежів БД, яким у поточному стані БД відповідає оцінка "Істина". Такі кортежі будемо називати актуальними кортежами БД. Відповідно кортежі, що належать множині $T - T_A$, будемо називати неактуальними кортежами БД.

Схематично розглянуті визначення представлені на рис. 2.

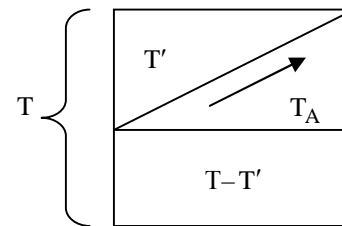


Рис. 2. Інформаційна структура бази даних

Завдання підтримки даних у цілісному стані.

Розглядаючи аномалії в термінах введених визначень можна відзначити, що вони нерозривно пов'язані з поняттям КС та її ІБ.

Суперечності можуть виникати при відображенні допустимого стану ІБ в неприпустимий стан БД або навпаки.

Важливо відзначити, що аномалії виникають в результаті виконання різних операцій маніпулювання даними в БД, а також при зміні стану ПрГ, що призводить до необхідності внести зміну в БД.

У роботі [6] визначено два види аномалій у відповідності з відображеннями ІБ в БД.

Аномалії першого роду – це відображення ІБ в БД, при якому допустима пропозиція ІБ відображається в неприйнятні кортежі БД.

Аномалії другого роду – це відображення, при якому допустима актуальна пропозиція ІБ є допустимим актуальним кортежем БД, а зворотне відображення переводить його в неактуальну пропозицію ІБ.

Аномалії першого роду виникають в тому випадку, коли операції над даними неможливо виконати через обмеження у схемі БД. Аномалії другого роду виникають в тому випадку, коли дані введені

правильно, але операції над БД можуть призвести до неправильних результатів.

Промислові СУБД, що використовуються на практиці, дозволяють підтримувати аномалії першого роду, але не враховують аномалії другого роду.

Враховуючи сучасні вимоги інтеграції неоднорідних БД, що використовують не тільки різні структури, але і різні типи документів, необхідно мати можливість уніфікованого представлення даних та їх інтерпретацію незалежно від їх фізичного представлення.

Модель представлення неоднорідних даних.

Будемо розглядати ІБ деякої ПрГ в термінах формальної логіки предикатів першого порядку. Використання такого підходу дозволить описати функціонування БД, не прив'язуючись до структури зберігання і обробки даних.

Скориставшись матеріалом статті [8] будемо розглядати БД як множину літералів разом з множиною обмежень цілісності у вигляді набору правил, що визначають допустимі стани БД.

Нехай фіксований деякий кінцевий алфавіт A , елементами якого є символи (літери, цифри, розділові знаки, знаки операцій, зв'язки тощо). Елементи множини A^n утворюють слова довжини n в алфавіті A , тобто, слово – це кінцева послідовність символів у A .

Базу даних будемо визначати як множину R -символів з A разом з множиною обмежень L . Обмеження цілісності L будемо висловлювати правилом виду $l \leftarrow l_1, l_2, \dots, l_k$, де l_1, l_2, \dots, l_k – символи з A та $k \geq 1$.

На змістовному рівні R представляє факти реального світу, а L представляє властивості, яким ці факти задовольняють. Допускається, що R змінюється при змінах реального світу, в той час як L може і не мінятися. Обмеження цілісності полягають в тому, що якщо всі символи l_1, l_2, \dots, l_k входять в R , то й символ l також повинен входити в R .

Іншими словами підтримка цілісності виражається в тому, що якщо змінилися деякі правила, що описують вимоги до даних, то необхідно змінити існуючі правила, що визначають цілісність цих даних.

Множину R будемо називати сумісною, якщо вона не містить символу одночасно з його запереченням. Сумісність обмежень визначається такою властивістю, що почавши з будь-якого символу не можна породити який-небудь символ і його заперечення. Під модифікацією R розуміється операція додавання або видалення символу, при виконанні якої R залишається сумісним.

Визначимо семантику R з обмеженнями L як множину S всіх символів породжених множиною R за допомогою правил з L . Очевидно, що потужність множини S буде не менше ніж потужність R ,

тобто $|S| \geq |R|$. Цей факт впливає з того, що правила, що обмежують допустимий стан БД можуть бути ширше, ніж стан БД в конкретний момент часу. Залежність стану БД від обмежень можна виразити схемою представленою на рис. 3.

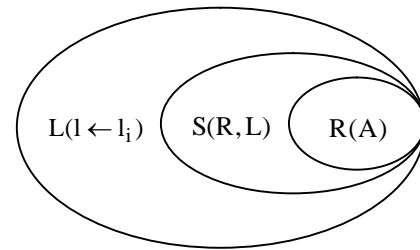


Рис. 3. Структура покриття обмежень, що впливають на стан БД

Нехай R – сумісна БД, L – обмеження цілісності цієї БД і S – її семантика. Модифікацію R розуміємо як виконання операцій: додавання або видалення l . При додаванні, l повинен входити в модифіковану семантику S' , а при видаленні, l повинен, бути відсутній в S' .

Виходячи з того, що L є більш широким поняттям, ніж можливий стан R , допускаються різні уявлення семантики для однієї предметної області.

Наприклад. Нехай $R = \{a, b\}$ – БД з обмеженнями $L = \{c \leftarrow a, b; d \leftarrow c, b\}$. У даному випадку можливі дві семантики $S_1 = \{a, b\}$ і $S_2 = \{a, b, c\}$. Як видно обидві семантики є сумісними. Нехай необхідно додати $-c$. Тоді модифікованої БД буде відповідати множина $R' = \{a, b, -c\}$, при цьому модифікована семантика може бути представлена двома множинами $S'_1 = \{a, b, -c, d\}$ і $S'_2 = \{a, b, c, -c, d\}$. Очевидно, що семантика S'_2 є несумісною і виключається з розгляду.

У загальному випадку завдання можна сформулювати наступним чином. Для деякої БД і запиту, що задає зміни потрібно таким чином змінити стан БД, щоб виконати запит і зберегти семантику модифікованої БД сумісною.

Для того щоб модифікована БД була допустима, її семантика S' повинна задовольняти наступним умовам [7]:

1. S' повинна містити l або не містити у разі додавання або видалення відповідно.
2. S' повинна бути сумісна.
3. S' повинна бути замкнута щодо L .
4. Літерали в S' повинні бути виведені (з використанням правил L) з літералів R які залишилися після модифікації БД.
5. S' повинна відповідати мінімальній зміні семантики вихідної множини R .

Будь-яка множина літералів S' , що задовольняє ці умови, є припустимою семантикою модифікованої БД.

Мінімальність зміни семантики можна визначити, використовуючи симетричну різницю вихідної і модифікованої семантик, що виражається формулою $S \div S' = (S - S') \cup (S' - S)$, де різниця $S - S'$ представляє множину літералів видалених з S , а $S' - S$ – множина доданих літералів. Таким чином, $S \div S'$ представляє загальне змінення при модифікації БД з максимальним набором символів, що виводяться з R' .

Наприклад. Нехай $R = \{a, b\}$ – БД з обмеженнями $L = \{d \leftarrow a, c\}$ і семантикою $S = \{a, b\}$, і нехай в R додано символ c . Тоді $R = \{a, b, c\}$ і припустимі дві семантики $S_1 = \{a, b, c, d\}$ і $S_2 = \{a, c, d\}$, так як за правилом L можлива ситуація коли символ b надлишковий для функціонування БД. Але з точки зору всіх можливих вимог до стану БД необхідно вибрати задовільну семантику. Знайдемо

$$\begin{aligned} S \div S'_1 &= (\{a, b\} - \{a, b, c, d\}) \cup \\ &\cup (\{a, b, c, d\} - \{a, b\}) = \{c, d\} \\ \text{і } S \div S'_2 &= (\{a, b\} - \{a, c, d\}) \cup \\ &\cup (\{a, c, d\} - \{a, b\}) = \{d\}. \end{aligned}$$

Таким чином, необхідно вибрати семантику S_1 , так як вона зберігає більше символів, ніж S_2 .

Можна виділити два основних етапи приведення БД до нового стану при її модифікації: ревізія та оновлення [7]. Ревізія визначається як застосування деякої формули до R і S , і опис результуючого стану у вигляді іншої модифікованої структури R' і S'_i ($i = \overline{1, n}$, де n – кількість можливих семантик). На етапі оновлення визначається структура R' і, враховуючи умови яким повинна задовольняти модифікована семантика (наприклад, умові мінімальної зміни) вибирається S' .

Модель модифікації БД можна представити у вигляді правила “подія, умова \rightarrow дії”:
 $p_i(u_1, u_2, \dots, u_k, l_1, l_2, \dots, l_m) \rightarrow d, d_2, \dots, d_m$, де u_i – подія (запит на оновлення), l_i – оновлені символи, d_i – дії (вставка або вилучення). Нехай деяка

подія u_i впливає на стани БД r_0 . Ця подія активізує умову l_i яка відповідає правилу p_i і виконує дії d_1, d_2, \dots, d_n , які генерують нову подію u_j і змінюють стан БД на r_1 тощо до тих пір, поки не буде одержано стан r_n , що задовольняє цілісності.

Висновки

Як було зазначено вище, питання вибору стратегії модифікації БД є важким завданням, якщо вирішувати його прямим перебором усіх правил при яких стан БД або обмеження будуть сумісними. Очевидним продовженням розглянутого підходу є пошук “найкращого” способу виводити допустимі семантики та ефективного методу модифікації БД. При цьому необхідно враховувати можливі циклічні і тупикові ситуації.

Список літератури

1. Руденко Д.А. Модель и средства поддержки данных в задачах интеграции неоднородных информационных систем / Д.А. Руденко, В.В. Тулунов // Системы обработки информации. – Х.: ХУПС, 2007. – Вып. 4 (62). – С. 110-114.
2. Дейт К. Введение в системы баз данных: пер. с англ. / К. Дейт. – 7-е изд. – М.: Издательский дом “Вильямс”, 2001. – 1072 с.
3. Гарсиа-Молина Г. Системы баз данных. Полный курс: пер. с англ. / Г. Гарсиа-Молина, Дж. Ульман, Дж. Уидом. – М.: Изд. дом “Вильямс”, 2003. – 1088 с.
4. Codd E. A relational model of data for large shared data banks / E. Codd // SACM 13. – 1970. – No. 6. – P. 1958-1982.
5. Ульман Дж. Основы систем баз данных / Дж. Ульман; пер. с англ. М.Р. Козаловского и В.В. Козутовского; под ред. М.Р. Козаловского. – М.: Финансы и статистика, 1983. – 334 с.
6. Fagin R.A. Normal form for relation databases that is based on domains and keys / R.A. Fagin // ACM Transaction on database systems. – 1981. – V.6, #3. – P. 387-415.
7. Дехтярь М.И. Восстановление ограниченной целостности за счет наименьших достаточных изменений / М.И. Дехтярь, А.Я. Диковский, Н. Спиратос // Программирование. – 1998. – № 2. – С. 27-37.
8. Лоран Д. Детерминированное поддержание ограниченной целостности / Д. Лоран, Н. Спиратос, Д. Статт // Программирование. – 1998. – № 2 – С. 38-57.

Надійшла до редколегії 21.02.2011

Рецензент: д-р техн. наук, проф. С.Г. Удовенко, Харківський національний університет радіоелектроніки, Харків.

ПОДДЕРЖКА ЦЕЛОСТНОСТИ В БАЗАХ ДАННЫХ С НЕОДНОРОДНОЙ СТРУКТУРОЙ

Т.П. Колесник

В статье рассмотрен подход к организации поддержки целостности при интеграции неоднородных распределенных баз данных. Определены виды нарушений и выделены аномалии, которые нельзя поддержать средствами управления базами данных.

Ключевые слова: аномалия, транзакция, база данных, нарушение целостности, поддержка данных, неоднородность структуры.

SUPPORT OF INTEGRITY IN DATABASES WITH A HETEROGENEOUS STRUCTURE

T.P. Kolesnik

The article describes the approach to integrity support organization when integrating heterogeneous distributed databases. The types of faults were determined as well as the abnormalities were defined which can not be supported by the database control tools.

Keywords: abnormality, transaction, database, integrity fault, data support, structural imperfection.