

ОПЕРАЦИОННАЯ СПЕЦИФИКАЦИЯ ОБЪЕКТНО-РЕЛЯЦИОННОЙ МОДЕЛИ ДАННЫХ

Статья посвящена разработке общих принципов формирования операционной спецификации объектно-реляционной модели данных. Формально определены основные операции, определен базис операционной спецификации объектно-реляционной модели данных. Предлагается использовать аппарат реляционной алгебры с рекурсивной составляющей для выполнения операций над отношениями, которые содержат вложенные неатомарные атрибуты.

Ключевые слова: объектно-реляционная модель данных, вложенные атрибуты, рекурсивная реляционная алгебра, операционная спецификация.

ВВЕДЕНИЕ

Модели данных являются ключевыми компонентами архитектуры современных систем управления базами данных (СУБД) и служат основой разработки общей методологии проектирования информационных систем.

Объективные обстоятельства, такие как: различные способы формального описания объектов в математике; разнообразие структур данных и средств манипулирования данными, развитых в языках программирования; разнообразие предметных областей, отображаемых в базах данных, способствовали неограниченному росту числа моделей данных и поддерживающих их СУБД. В 1977 году А. Макионучи предложил упростить требования, которые выдвигались к первой нормальной форме реляционной модели данных. В дальнейшем такие ученые как Л. Колби, М. Левен, Х. Ли, С. Абитебол, Н. Бидоит, и др. развивали данное направление и предложили ряд моделей данных, которые основывались на этом упрощении.

Подход на основе слабо нормализованных отношений использовался для решения таких задач как обработка текста, проектирование автоматизированных систем документооборота, обработку составных пользовательских типов данных и многих других.

Преимущества такой нормализации отношения позволяет преодолеть ряд ограничений, которые присутствуют в первой нормальной форме. Использование ННФ позволяет представить данные об объекте в одном отношении, а не распределять по нескольким. Одним из главных преимуществ данного подхода является минимальное использование операции соединения отношений, на которую приходится большие временные затраты при выполнении запросов.

Модель базы данных, основанная на принципах ННФ, или объектно-реляционная модель данных [1, 5] позволяет отношению хранить атрибуты, которые содержат неатомарные значения, таким образом, такие атрибуты сами являются отношениями (подсхемами).

ЦЕЛЬ ИССЛЕДОВАНИЙ

Статья посвящена разработке общих принципов формирования операционной спецификации объектно-реляционной модели данных. Предлагается использовать аппарат реляционной алгебры с рекурсивной составляющей для выполнения операций над отношениями, которые содержат вложенные неатомарные атрибуты.

ИССЛЕДОВАНИЕ ОСОБЕННОСТЕЙ ОБЪЕКТНО-РЕЛЯЦИОННОЙ МОДЕЛИ

Модели ННФ были разработаны и разделены на две категории. Модели первой категории назывались нерекурсивными моделями [1, 3, 4], а модели второй категории – рекурсивными моделями [1, 2]. Основным отличием является рекурсивный и нерекурсивный характер операторов модели. Рекурсивные операторы могут неоднократно применяться к вложенным отношениям (подсхемам) на его различных уровнях. Нерекурсивные операторы не обладают такими свойствами. Модели, которые поддерживают вложенность или рекурсию имеют ряд преимуществ по сравнению с нерекурсивными, так как являются универсальными для реляционной модели и объектно-реляционной модели. Этот вывод основан на следующих утверждениях:

– для модели, которая содержит вложенные отношения, основным принципом является сокращение числа обрабатываемых кортежей при выполнении операций;

– традиционные методы оптимизации запросов могут использоваться в операциях рекурсивной алгебры. Кроме того, такие методы могут применяться к запросам, которые выражены с помощью рекурсивной алгебры с рекурсивными операциями на любом уровне вложенности;

– в операциях нерекурсивной алгебры действия производятся только на всем кортеже. Рекурсивная алгебра решает прямые манипуляции с кортежем, как на верхнем уровне, так и на нижнем уровне вложенных отношений [2].

Для определения операций объектно-реляционной модели данных используем аппарат рекурсивной реляционной алгебры.

Применение операций рекурсивной реляционной алгебры не ограничено количеством уровней вложенности отношений. Такие операции как объединение, разность, пересечение, проекция, селекция, декартово произведение, θ -соединение формально определены и используют рекурсивное определение вложенных отношений. Для каждой рекурсивной операции базовым является определение аналогичной нерекурсивной. Рекурсивное определение может быть переведено в одно нерекурсивное, где отношения не содержат других вложенных атрибутов. Рекурсивные операции для вложенных отношений так же формально определены.

ФОРМАЛИЗАЦИЯ ОПЕРАЦИЙ РЕКУРСИВНОЙ РЕЛЯЦИОННОЙ АЛГЕБРЫ

Для рекурсивных вложенных операций объединения, разности и пересечения необходимо наличие двух вложенных (в общем случае) отношений r и q со схемами отношений R и Q соответственно. Необходимым условием для выполнения операций является так же идентичность схем отношений:

$$R = Q; R = \{S(R), R_1, R_2, \dots, R_n\}, \quad (1)$$

где $S(R)$ – множество ключей вложенных атрибутов и все атомарные атрибуты R и Q (одинаковые для двух отношений); $\{R_1, R_2, \dots, R_n\}$ – неключевые вложенные атрибуты R и Q .

$Attr(R)$ – множество всех атрибутов (атомарных и вложенных) двух схем отношений R и Q , t_r и t_q – кортежи отношений r и q соответственно в итоговом отношении.

Каждое определение для рекурсивных операций реляционной алгебры состоит из двух частей: для плоских отношений и вложенных. При этом, нерекурсивная опе-

рация является частным случаем рекурсивной. Выполнения операций рекурсивной реляционной алгебры рассматриваются на примерах схем отношений с вложенными атрибутами (рис. 1).

Рекурсивная операция объединения (\cup)

Определение операции объединения двух отношений r и q , $r \cup q$ является комплексным – состоит из двух определений:

- 1) нерекурсивная операция объединения для плоских отношений $r \cup q$;
- 2) рекурсивная операция объединения для вложенных отношений $r \cup q$:

$$r \cup q = \{t \mid ((\exists t \in r) \wedge (\forall t_q \in q)(t[S(R)] \neq t_q[S(R)])) \wedge ((\exists t \in r) \wedge (\forall t_r \in q)(t[S(R)] = t_r[S(R)])) \wedge ((\exists t_r \in r)(\exists t_q \in q)((t[S(R)] = t_r[S(R)] = t_q[S(R)])) \wedge (t[R_1] = t_r[R_1] \cup t_q[R_1]) \wedge \dots \wedge (t[R_n] = t_r[R_n] \cup t_q[R_n]))\}. \quad (2)$$

Пример 1. Объединение отношений R и Q .

$T_1 = R \cup Q$. Результирующее отношение T_1 имеет вид (рис. 2).

A	B		
	C	D	
		M	N
01	a1	1	F
		2	G
02	a2	4	F
		2	G
		1	H
03	a3	1	H
	a4	2	H
04	a5	3	K
		4	H
	a6	1	T
		2	P

Рис. 2. Результат операции объединения вложенных отношений

A	B		
	C	D	
		M	N
01	a1	1	F
		2	G
02	a2	4	F
		1	H
03	a3	1	H
	a4	2	H

а) Отношение R

Z	L	K	
		K	
		X	Y
01	a1	1	F
		2	G
02	a2	4	F
		2	G
04	a5	3	K
		4	H
	a6	1	T
		2	P

б) Отношение Q

F	K	
	P	O
	abc	a
2		
def	n	5
	m	
ert	y	1
	x	5

в) Отношение D

Рис. 1. Схемы отношений R, Q, D

Разность двух отношений r и q , $r \dot{=} q$ определяется как совокупность определения нерекурсивной разности для плоских отношений ($r - q$) и определение рекурсивной разности для вложенных отношений ($r \dot{=} q$):

$$r \dot{=} q = \{t \mid ((\exists t \in r)(\forall t_q \in q)(t[S(R)] = t_r[S(R)] - t_q[S(R)]) \wedge t[R_1] = t_r[R_1] \wedge \dots \wedge (t[R_n] = t_r[R_n])) \vee ((\exists t_r \in r), (\exists t_q \in q)(t[S(R)] \neq t_r[S(R)])) \wedge ((\exists t_r \in r)(\exists t_q \in q)(t[S(R)] = t_r[S(R)] = t_q[S(R)]) \wedge (t[R_1] = t_r[R_1] \dot{=} t_q[R_1]) \wedge \dots \wedge (t[R_n] = t_r[R_n] \dot{=} t_q[R_n])))\} \quad (3)$$

Пример 2. Операция разности для вложенных отношений имеет вид $T_2 = R \dot{=} Q$ (рис. 3).

Аналогично представлено определение рекурсивной операции пересечения:

$$r \dot{\cap} q = \{t \mid (\exists t_r \in r)(\exists t_q \in q)((t[S(R)] = t_r[S(R)] \cap t_q[S(R)]) \wedge t[R_1] = t_r[R_1] \dot{\cap} t_q[R_1] \wedge \dots \wedge (t[R_n] = t_r[R_n] \dot{\cap} t_q[R_n]))\} \quad (4)$$

Пример 3. Производная операция пересечения $T_3 = R \dot{\cap} Q$ для отношений с комплексной схемой представлена на рис. 4.

Операция вложенной проекции может быть представлена с помощью нерекурсивно и рекурсивной составляющей.

Пусть r вложенное отношение со схемой R .

Пусть так же, $\{R_{a1}, \dots, R_{ak}\}$ – подмножество атомарных атрибутов верхнего уровня R , над которыми должна выполняться операция проекции и $\{R_{n1}, \dots, R_{nm}\}$ – подмножество вложенных атрибутов R , которые могут участвовать в операции проекции полностью или частично, при этом ($k, m \geq 0$).

Прежде чем определить операцию проекции, необходимо дать определение списку проекции.

A	B		
	C	D	
		M	N
02	a2	1	H
03	a3	1	H
	a4	2	H

Рис. 3. Результат выполнения операции разности для вложенных отношений

A	B		
	C	D	
		M	N
01	a1	1	F
		2	G
02	a2	4	F
03	a3	1	H
	a4	2	H

Рис. 4. Операция пересечения двух отношений с вложенными атрибутами

В общем случае, список проекций – это список путей атрибутов, которые могут входить в проекцию.

Определение операции проекции $\tilde{\pi}(rL_n)$ во вложенном отношении r :

$$\tilde{\pi}_{(R_{a1}, \dots, R_{ak}, R_{n1}, \dots, R_{nm}, L_{nm})}(r) = \{t \mid (\exists t_r \in r)((t[R_{a1}] = t_r[R_{a1}]) \wedge \dots \wedge (t[R_{ak}] = t_r[R_{ak}]) \wedge (t[R_{n1}] = \tilde{\pi}(t_r[R_{n1}L_{n1}]) \wedge \dots \wedge (t[R_{nm}] = \tilde{\pi}(t_r[R_{nm}L_{nm}])))\}, \quad (5)$$

где t_r – кортеж отношения r , t – кортеж результирующего отношения, полученного в результате операции проекции.

Пример 4. Операция проекции для отношения R имеет вид $T_4 = \tilde{\pi}_{(A, D, (N))}(R)$ (рис. 5).

Операция селекции для вложенных отношений имеет ряд особенностей.

Пусть r вложенное отношение со схемой R . Пусть так же,

$\{R_{a1}, \dots, R_{ak}\}$ – подмножество атомарных атрибутов верхнего уровня R , над которыми должна выполняться операция селекции и $\{R_{n1}, \dots, R_{nm}\}$ – подмножество вложенных атрибутов R , где k и m меньше или равно числу атомарных или вложенных атрибутов верхнего уровня отношения R соответственно. Пусть c – множество условий R . При этом c имеет вид $\{c_a, c_n\}$, где $c_a = \{c_{a1}, \dots, c_{ak}\}$ – это множество условий, которые должны быть истинными для каждого атрибута R_{a1}, \dots, R_{ak} отношения R , соответственно. И $c_n = \{c_{n1}, \dots, c_{nm}\}$ – множество условий, которые должны выполняться для вложенных атрибутов R_{n1}, \dots, R_{nm} отношения R . Когда применяются два множества условий, то результатом является множество, полученное в результате выполнения операции пересечения результатов двух множеств условий.

Вложенная операция селекции может быть определена как:

Вложенная операция селекции может быть определена как:

$$\tilde{\delta}_{c_{a1}, \dots, c_{nm}}(rL_\delta) = \{t \mid (\exists t_r \in r)((t[Atr(R) - \{R_{n1}, \dots, R_{nm}\}] = t_r[Atr(R) - \{R_{n1}, \dots, R_{nm}\}]) \wedge ((t[R_{n1}] = \tilde{\delta}(t_r[R_{n1}L_{n1}] \neq \emptyset) \wedge \dots \wedge ((t[R_{nm}] = \tilde{\delta}(t_r[R_{nm}L_{nm}] \neq \emptyset)))\}, \quad (6)$$

где L_δ – список селекции (аналогично определению списка проекции).

A	D
	N
01	F
	G
02	F
	H
03	H
	H

Рис. 5. Отношение T_4 – результат выполнения операции проекции

В общем случае, результатом двух операций селекции является операция пересечение двух селекций:

$$\bar{\delta}_c(r L_\delta) = \bar{\delta}_{c_{a1}, \dots, c_{ak}, c_{n1}, \dots, c_{nm}}(r L_\delta) = \bar{\delta}_{c_{a1}, \dots, c_{ak}}(r) \bar{\delta}_{c_{n1}, \dots, c_{nm}}(r L_\delta). \quad (7)$$

Пример 5. На рис. 6 представлен результат выполнения операции селекции $T_5 = \bar{\delta}_{(A=02 \text{ AND } B(D(M)) < 3)}(R)$.

Для отношений с вложенными атрибутами определена операция декартового произведения.

Пусть r и q два вложенных (в общем случае) отношения со схемами R и Q соответственно. Пусть $\text{Attr}(R)$ – все атрибуты (вложенные и атомарные) R , а $\text{Attr}(Q)$ – все атрибуты Q , а L – путь соединения в R . Определим, также, что R_i – вложенные атрибуты R , L_i – путь R_i , t_r – кортеж в отношении r , t_q – кортеж отношения q , t – кортеж результирующего отношения, тогда:

$$\begin{aligned} (r L \bar{\times} q) &= (r(R_i L_i) \bar{\times} q) \equiv (q \bar{\times} r(R_i L_i)) = \\ &= \{t \mid (\exists t_r \in r)((t[\text{Attr}(R)] - \{R_i\}) \wedge t[R_i]_i = \\ &= (t_r[R_i] L_i \bar{\times} q))\}. \end{aligned} \quad (8)$$

Для данной операции выполняется свойство коммутативности. Следовательно, верно равенство:

$$(r(R_i L_i) \bar{\times} q) \equiv (q \bar{\times} r(R_i L_i)). \quad (9)$$

Особенностью данной операции декартового произведения является возможность ее выполнения на разных уровнях вложенности, в зависимости от пути атрибутов L . Для двух вложенных отношений операция $\bar{\theta}$ -соединения базируется на операторах сравнения, где $\theta = \{>, <, =, \leq, \geq, \neq\}$.

Частным случаем является операция естественного соединения, где в качестве оператора сравнения используется равенство (=).

Операция $\bar{\theta}$ -соединения выполняется путем применения операции селекции к результатам операции декартового произведения двух отношений. Операция сводится к сравнению двух атрибутов в результирующем отношении. Два атрибута должны иметь одинаковый уровень вложенности в итоговом отношении.

Пусть r и q два вложенных (в общем случае) отношения со схемами R и Q соответственно. Пусть X и Y – два атомарных атрибута, которые принадлежат отношениям R и Q соответственно, $\bar{\theta}$ – условие, которое должно выполняться. Без потери общности, $L_{\delta Y' \rightarrow Y}$ – путь селекции Y , который начинается с Y' , который, в свою очередь, принадлежит такому же уровню вложенности как и X .

A	B		
	C	D	
		M	N
02	a2	1	H

Рис. 6. Операция селекции для вложенных отношений

Тогда, рекурсивное вложенное соединение двух отношений r и q можно определить как:

$$r \bar{\theta} q = \bar{\delta}((r \times q)_{X \theta Y} L_{\delta Y' \rightarrow Y}). \quad (10)$$

Пример 6. Операция естественного соединения для отношений R и D представлена в виде $T_8 = R(B, (C, (D))) [M \equiv O] D(F, (K))$, при этом, на первом этапе, выполняется операция декартового произведения $R(B, (C, D)) \bar{\times} D(F, (K))$, затем – операция селекции, согласно условию. В результате отношение T_8 имеет вид (рис. 7).

Пример иллюстрирует, что если список селекции пуст, то итоговое отношение так же не содержит значений атрибутов на всех уровнях вложенности. Таким образом, в отношении T_8 исключен атрибут $A = \{02\}$, т. к. в данном кортеже не найдено подстроку, удовлетворяющим условию $R(\dots) [M \equiv O] D(\dots)$. Таким образом, очевидно, что результат операции $\bar{\theta}$ -соединения зависит от списка селекции, и соответственно, от результата выполнения промежуточной операции декартового произведения.

ВЫВОДЫ

В качестве математического аппарата для описания операционной спецификации объектно-реляционной модели данных предлагается использовать рекурсивную реляционную алгебру. В статье исследованы и усовершенствованы основные операции, с помощью которых может обеспечиваться выполнение действий над вложен-

A	C	B(C, D, K, F)				F
		D		K		
		M	N	P	O	
01	a1	1	F	a	1	abc
		2	G	a	2	
		1	F	a	2	
		2	G	a	1	
02	a2	4	F	n	5	def
		1	H	m	5	
		4	F	n	5	
		1	H	m	5	
03	a3	1	H	y	1	ert
		1	H	x	5	
	a4	2	H	y	1	ert
		2	H	x	5	

a) $R(B, (C, D)) \bar{\times} D(F, (K))$

A	C	B(C, D, K, F)				F
		D		K		
		M	N	P	O	
01	a1	1	F	a	1	abc
03	a3	1	H	y	1	ert

б) $T_8 = R(B, (C, (D))) [M \equiv O] D(F, (K))$

Рис. 7. Отношение T_8 – результат операции $\bar{\theta}$ -соединения

ними отношениями. Предложена общая интерпретация операции декартового произведения, и соответственно, производных операций. Существует ряд СУБД, в которых реализован объектно-реляционный подход, но такая реализация не содержит комплексной поддержки всех компонент объектно-реляционной модели, в том числе и операционной спецификации. В данной работе операционная спецификация объектно-реляционной модели представлена в общем виде, сохраняя общность с реляционной моделью данных, и не ассоциирована с конкретной реализацией в СУБД.

СПИСОК ЛИТЕРАТУРЫ

1. *Makinouchi, A.* A consideration on Normal Form of Not-Necessarily-Normalized Relations in the Relational Data Model [Text] // in Proc. of the 3rd International Conference on Very Large Data Bases. – Tokyo, 1977. – P. 447–453.
2. *Garani, G.* Generalized Relation Data Model [Text]: // International Journal of Computer Systems Science and Engineering (IJCSSES). – 2007. – vol. 4, N 1. –P. 43–59.
3. *Мейер, Д.* Теория реляционных баз данных [Текст] : пер. с англ. – М. : Мир, 1987. – С. 45–48.
4. *Тиори, Т.* Проектирование структур баз данных: в 2-х кн. [Текст] / Т. Тиори, Дж. Фрай. – М. : Мир, 1985. – Кн. 1. – С. 287 : Кн. 2 – С. 320.
5. *Чапланова, Е. Б.* Об одном подходе к построению объектно-реляционной модели данных [Текст] / Н. В. Касаткина, С. С. Таянский, Е. Б. Чапланова // Збірник наукових праць

Військового інституту Київського національного університету імені Тараса Шевченка. – К. : ВІКНУ, 2009. – Вип. 20. – С. 141–146.

Стаття надійшла до редакції 22.02.2012.

Чапланова О. Б.

ОПЕРАЦІЙНА СПЕЦИФІКАЦІЯ ОБ'ЄКТНО-РЕЛЯЦІЙНОЇ МОДЕЛІ ДАНИХ

Стаття присвячена розробці загальних принципів формування операційної специфікації об'єктно-реляційної моделі даних. Формально визначено основні операції, визначено базис операційної специфікації об'єктно-реляційної моделі даних. Пропонується використовувати апарат реляційної алгебри з рекурсивної складової для виконання операцій над відносинами, які містять вкладені неатомарні атрибути.

Ключові слова: об'єктно-реляційна модель даних, вкладені атрибути, рекурсивна реляційна алгебра, операційна специфікація.

Chaplanova E.

OPERATING SPECIFICATIONS OF THE OBJECT-RELATIONAL DATA MODEL

The article proposed the development of general principles of formation of the operating specifications of object-relational data model. Formally defined the basic operations, and defined basis of the operating specifications of object-relational data model. It is proposed to use the apparatus of the relational algebra with a recursive component to perform operations on relations, which contain nested non-atomic attributes.

Key words: object-relational data model, nested attributes recursive relational algebra, the operating specification.