

ПОСТРОЕНИЕ БЛОЧНО-АЦИКЛИЧЕСКОЙ СХЕМЫ НЕСОГЛАСОВАННОЙ БАЗЫ ДАННЫХ

Шахова О. И.

Научный руководитель – к.т.н., доц. Руденко Д.А.

Харьковский национальный университет радиоэлектроники
(61166, Харьков, пр. Ленина, 14, каф. информатики, тел. (057) 702-14-19
Wwwradistka.ru@mail.ru

The article deals with one class of the circuits of databases, namely - class of the not cyclic circuits having several desirable properties will be considered. These properties are formulated as conditions, which should be carried out for all databases with such circuit.

Важным свойством при проверке целостности данных для некоторых схем баз данных (БД) является возможность обойтись без универсального отношения (УО). Определим эти свойства. Пусть $R = \{R_1, R_2, \dots, R_p\}$ – некоторая схема БД и $d = \{r_1, r_2, \dots, r_p\}$ – БД над R . Известно, что проверка свойства полноты соединения для отношений r_1, r_2, \dots, r_p относится к вычислительно трудным задачам [1]. В некоторых случаях эту проблему можно обойти. Говорят, что d согласована в целом (СЦ), если r_1, r_2, \dots, r_p обладают свойством полноты соединения. Базу данных d называют попарно - согласованной (ПС), если свойством полноты соединения обладает произвольная пара отношений r_i и r_j ($i \neq j$). Сложность проверки свойства ПС полиномиально зависит от количества отношений в БД. При этом СЦ с необходимостью влечет за собой ПС (СЦ \rightarrow ПС), а ПС не всегда достаточно для СЦ (ПС \rightarrow СЦ).

В качестве примера рассмотрим две схемы БД $R^A = \{ABC, BCD, CDF\}$ и $R^B = \{ABC, BCD, CE, DE\}$. В данном случае свойство ПС \rightarrow СЦ выполняется для БД со схемой R^A . Видно, что это свойство зависит от того, каким образом атрибуты отношений пересекаются между собой. Основываясь на этом факте, рассмотрим алгоритм, проверяющий выполнение свойства ПС \rightarrow СЦ.

Алгоритм редукции Грэхема [2] состоит в последовательном выполнении шагов, на каждом из которых к схеме БД применяется одно из двух правил редукции - до тех пор, пока не окажется, что ни одно из них применить больше нельзя. Правила редукции имеют следующую последовательность шагов:

- удаление схемы отношения (УС) – если схема отношения собственным образом содержится в некоторой другой схеме отношения;
- удаление атрибута (УА) – если атрибут принадлежит не более чем одной схеме отношения.

Будем говорить, что редуктивный алгоритм Грэхема завершился успешно, если в результате его применения к схеме получается пустое множество. В противном случае имеем набор атрибутов, образующих связный

блок, то есть каждый атрибут принадлежит более чем одной схеме, и ни одна схема не входит собственным образом в другую.

Схемы, для которых алгоритм Грэхема не выполняется, будем называть циклическими, а схемы, для которых алгоритм завершается успешно - ациклическими.

Для успешного завершения алгоритма Грэхема необходимо, чтобы схема была ациклическая, в противном случае ее необходимо привести к ациклическому виду.

Пусть $d = \{r_1, \dots, r_p\}$ - БД со схемой $R = \{R_1, \dots, R_p\}$, и пусть $d' = \{r_i, \dots, r_j\}$ - связной блок со схемой $R' = \{R_i, \dots, R_j\}$, тогда отношение $b = r_i > \dots > r_j$ со схемой $B = R_i \cup \dots \cup R_j$, в обозначении $b(B)$, будем называть сильно связным блоком для $d(R)$.

Замена в циклической схеме связных блоков на сильно связные приводит ее к ациклическому виду. Ациклические схемы с сильно связным блоком будем называть блочно - ациклическими.

Пусть база d имеет схему $R = \{AB, BCD, CDE, DEK, EKM, KMN\}$. В результате применения алгоритма Грэхема остается связный блок со схемой $R' = \{BCD, CDE, DEK\}$. Естественное соединение отношений со схемой R' порождает сильно связный блок $\mathcal{R} = \{BCDEK\}$. Таким образом, схема $R^{block} = \{AB, \mathcal{R}, EKM, KMN\} = \{AB, BCDEK, EKM, KMN\}$ является ациклической.

Итогом полученных результатов является следующее утверждение.

Утверждение: Для того чтобы циклическая схема удовлетворяла свойству ПС \rightarrow СЦ, необходимо, чтобы она содержала хотя бы один сильно связный блок.

Доказательство этого утверждения следует из редуктивного алгоритма Грэхема и определения сильно связного блока.

Таким образом, сведение циклической схемы к блочно - ациклическому виду дает значительное снижение временных затрат на поддержку целостности при ведении БД за счет проверки попарной согласованности.

Пусть в БД входит k отношений, каждое из которых содержит n кортежей. Тогда количество соединений E' при полном соединении определяется формулой: $E' = n^k$; при попарном соединении количество соединений

$$E^p \text{ вычисляется формулой: } E^p = \sum_1^{\frac{k(k-1)}{2}} n^2 = \frac{k(k-1)}{2} n^2 = pn^2, \text{ где } p = \frac{k(k-1)}{2}.$$

Список источников:

1. Буслик Н.Н. Глобальные схемы реляционных баз данных: концепция и методы построения.: Монография. - Харьков: ХНУРЭ, 2002. - 68 с.
2. Мейер Д. Теория реляционных баз данных. - М.: Мир, 1987. - 608 с.