

УДК 004.475: [004.652: 004.655]

Е.Е. Гринева

НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ КОРРЕКТНОГО СОЕДИНЕНИЯ ОТНОШЕНИЙ В ЗАДАЧАХ СОВМЕСТНОЙ ОБРАБОТКИ РАСПРЕДЕЛЕННЫХ ДАННЫХ

Введение. В настоящее время, когда процесс автоматизации различных видов деятельности пришел практически на каждое современное предприятие, вычислительные системы быстро накапливают большие объемы статистических данных. Такой объем информации позволяет, с одной стороны, выполнять более точные расчеты и проводить подробный анализ, с другой стороны – поиск необходимых решений представляет собой достаточно сложную аналитическую задачу.

Существующие информационные системы, использующие системы управления базами данных (СУБД), достаточно успешно решают задачи хранения и поиска информации, а также частично решают задачи анализа данных. Преодоление определенной противоречивости требований к СУБД и системам, ориентированным на анализ данных, привело к возникновению и широкому использованию концепции хранилищ данных с распределенным доступом и обработкой [1].

Совместный доступ к распределенным базам данных (БД) оказывается ключевой проблемой, решение которой актуализируется усилиями общих интеграционных тенденций развития современных методов и средств компьютеризированных информационных технологий. Как категория, интегрированная распределенная база данных определяет класс сложных и неоднозначных информационных объектов, механизмы построения и управления которыми на сегодняшний день являются наиболее актуальными. Многоаспектность интегрированных БД порождает необходимость компромиссного использования целого ряда методов, средств и компьютерных технологий для реализации корректной обработки распределенных данных.

В последние годы в теории БД и хранилищ данных сформировались новые направления научных исследований, которые получили название теории концептуального моделирования в БД и корпоративных хранилищах данных [2, 3]. Важной целью разработки таких форм моделирования является разработка методов построения интерфейса между конечным пользователем с одной стороны, и информационной системой - с другой стороны, основанных на одинаковом, непротиворечивом, содержательном представлении [4].

Современное представление технологии БД определяется тем, что в ее основу положено применение реляционной модели данных [5, 6], базирующейся на строгом аппарате реляционной алгебры и математической логики. Реляционный подход в целом рассматривается в качестве идеологии создания БД [7, 8].

Одной из основных трудностей обработки распределенных данных является поддержка корректного состояния данных при совместном доступе, то есть необходимо обеспечить корректность перехода из одного согласованного состояния БД в другое при допущении независимого ведения данных. Таким образом, целью статьи является формализация условий, выполнение которых обеспечивает адекватность данных при построении как глобального (включающего все таблицы) отношения, так и частичных представлений (храняемых запросов) при решении различных задач совместной обработки распределенных данных.

Сохранение данных для множества независимых отношений. При использовании распределенных БД как системы с совместным доступом, необходимо гарантировать согласованность и адекватность информации при выполнении динамических запросов, то есть запросов, формируемых к данным распределенной структуры и/или интегрируемой системы. Рассмотрим некоторые методы построения адекватного представления отношений посредством их интеграции в БД с единой схемой. При этом, что особенно важно, необходимо показать, что при эквивалентности схем БД эквивалентность по данным требует дополнительного подтверждения.

Введем некоторые обозначения, необходимые для решения задачи адекватного представления интегрируемых данных. Пусть $R = \{R_1(S_1), R_2(S_2), \dots, R_n(S_n)\}$ – множество всех отношений R_i со схемой S_i ($i = \overline{1, n}$), для которых необходимо построить представление и $S = \{S_1, S_2, \dots, S_n\}$ – множество схем отношений из R . При этом схемы из S принадлежат множеству $U = (S_1 \cup S_2 \cup \dots \cup S_n)$, которое будем называть универсальной схемой. Обозначим через R^U – множество возможных отношений со схемой U (множество универсальных отношений). Отметим, что в дальнейшем будем ставить в соответствие R^U множество состояний $P = \{p_1, p_2, \dots, p_m\}$, так как для каждого допустимого состояния отношений-компонент R_i ($i = \overline{1, n}$) необходимо получить допустимое состояние p_j (где $j = \overline{1, m}$) универсального отношения при условии соответствия локальным ограничениям целостности. Также обозначим через

$DB = \{dB_1, dB_2, \dots, dB_n\}$ множество всех баз данных со схемой S и возможными состояниями из P . Кроме этого, обозначим множество ограничений целостности через $I = \{F, MV, J\}$, где $F = \{f_1, f_2, \dots, f_n\}$ – множество функциональных зависимостей, $MV = \{mv_1, mv_2, \dots, mv_k\}$ – множество многозначных зависимостей, $J = \{j_1, j_2, \dots, j_m\}$ – множество зависимостей соединения.

Очевидно, что для построения представления необходимо использовать две операции реляционной алгебры: проекцию и естественное соединение (соединение), которые не обладают свойством дистрибутивности при совместном использовании. То есть, для некоторого отношения R при выполнении проекции на некоторые схемы S_1 и S_2 , а затем соединения этих проекций может нарушиться состояние отношения R . Пример такого случая приведен на рисунке 1.

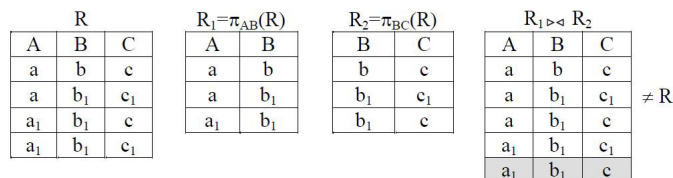


Рис. 1. Нарушение свойства дистрибутивности операции «проекция-соединение»

При проектировании схемы БД методом нормализации исходной структуры, такое свойство называется «соединение без потери информации», имеет формальный вид (1) и играет основную роль при определении ссылочной целостности [9, 10].

$$\forall R : \pi_{S_1}(R) \triangleright \triangleleft \pi_{S_2}(R) \triangleright \triangleleft \dots \triangleright \triangleleft \pi_{S_n}(R) = R, \tag{1}$$

где π – операция проекции; $\triangleright \triangleleft$ – операция соединения.

В рассматриваемой задаче свойство, определяемое выражением (1), необходимо установить и поддерживать при выполнении операций, связанных с интеграцией независимых БД.

Так как при выполнении соответствующих операций БД переходит из одного состояния в другое, определим операции проекции и соединения в терминах отображений состояний БД.

ОПРЕДЕЛЕНИЕ 1. Операция проекции π , определенная на схеме S выражается отображением, в обозначении $\xi : \pi_S(p) \rightarrow dB$, которое сопоставляет некоторое состояние базы данных p из R^U базе данных dB из DB . При этом, $\xi : \pi_S(p) \rightarrow p_i, i = \overline{1, m}$.

ОПРЕДЕЛЕНИЕ 2. Операция соединения $\triangleright \triangleleft$, определенная на схеме S выражается отображением, в обозначении $\zeta : \triangleright \triangleleft_R(dB) \rightarrow p_1 \triangleright \triangleleft p_2 \triangleright \triangleleft \dots \triangleright \triangleleft p_m$, которое сопоставляет базу данных $dB = \{p_1, p_2, \dots, p_m\}$ из DB некоторому состоянию базе данных $\triangleright \triangleleft(dB)$ из RU .

Для подтверждения введенных свойств операций проекции и соединения рассмотрим пример. Воспользуемся базой данных R , представленной на рисунке 1. Пусть $U = (A, B, C)$ и $R = \{R_1(AB), R_2(BC)\}$. Естественно, что для возможного состояния γ отношения R , база данных dB будет иметь вид, полученный в результате проекции $dB = \pi_S(p)$, и будет иметь состояние (p_1, p_2) , соответствующее отношениям R_1 и R_2 . Также видно, что при выполнении операции соединения ($R_1 \triangleright \triangleleft R_2$) получается возможное состояние p' , определяемое отношением $R' = (R_1 \triangleright \triangleleft R_2)$. Очевидно, что при выполнении операции проекции $\pi_R(p')$ результат будет соответствовать состоянию базы данных $dB = (p_1, p_2)$.

Предположим, что отношению R'_2 соответствует состояние p'_2 , представленное на рисунке 2, то есть $dB' = (p_1, p'_2)$. Тогда $R_1 \triangleright \triangleleft R'_2 = \emptyset$ над универсальным отношением RU .

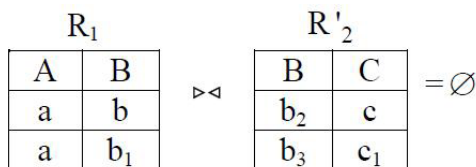


Рис. 2. Возможные состояния отношений R_1 и R'_2

Введем определение образа схемы базы данных.

ОПРЕДЕЛЕНИЕ 3. Образом множества отношений R , определенным на множестве схем S , называется проекция $\pi_S(R) = \{dB \in DB \mid dB = \pi_S(R_i), \forall R_i \in R, \forall R \subseteq DB\}$.

В связи с тем, что не каждая база данных dB является проекцией некоторого возможного состояния P , то выполняется следующая последовательность включений: $\pi_S(R) \subseteq \pi_S(R^U) \subseteq DB$. Пример включения проекций можно проследить на отношениях R_1 (рис. 1) и R'_2 (рис. 2) для базы данных $dB' = (p_1, p'_2)$.

Так как рассматриваемой задачей является построение отображения $\xi : \pi_S(p) \rightarrow dB$, то только отношения из множества $\pi_S(P)$ могут быть возможными состояниями базы данных. Очевидно, что в множество $\pi_S(P)$ входят только полностью соединимые БД со схемой S.

С другой стороны, как отмечено в [10], такое решение может быть не эффективно по трем причинам:

1. проверка полного соединения отношений БД является трудно разрешимой (NP-полной);
2. принадлежность некоторой базы данных dB к множеству $\pi_S(P)$ требует рассматривать БД как единое целое, что требует дополнительных алгоритмов проверки согласованности отношений БД;
3. неполностью соединимая БД может иметь смысл в рассматриваемой предметной области.

Пример несогласованных отношений $R \triangleright \triangleleft R_1 \triangleright \triangleleft R_2 = \emptyset$ приведен на рисунке 3.

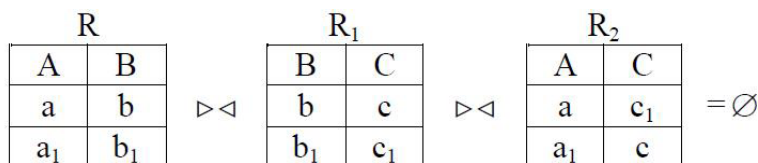


Рис. 3. Отношения несоединимые в целом

Кроме этого необходимо учитывать локальные ограничения, то есть ограничения, которые определены на БД из множества DB. Сложность этой задачи заключается в том, что их необходимо поддерживать в целом для всей DB и при этом, не проверять соответствие таким ограничениям каждое возможное состояние P.

ОПРЕДЕЛЕНИЕ 4. Частичным образом отношений R относительно схемы S будем называть множество $PF_S(R) = \{dB \in DB \mid dB = \{p_1, p_2, \dots, p_m\}, p_i \in \pi_{S_i}(R) \forall R \subseteq R^U\}$.

Другими словами PFS(R) содержит БД со схемой S, которые представляют собой проекции некоторого возможного состояния, и при этом для различных отношений БД не должны быть обязательно одними и теми же.

Покажем на примере существование частичного образа отношений некоторой БД. Очевидно, что выполняется условие $PF_S(R) \supseteq \pi_S(R)$. Пусть $U=(A, B, C)$ и $R=\{R1(S1), R2(S2)\}$, где $S1=\{AB\}$, $S2=\{BC\}$ и пусть $P=\{SAT(A \rightarrow C)\}$, где SAT – множество атрибутов всех баз данных из DB, удовлетворяющих заданным ограничениям [10, 11] (в данном случае функциональной зависимости $A \rightarrow C$), тогда база данных $dB=\{p_1, p_2\}$ принадлежит PFS(R) и не принадлежит $\pi_S(R)$ (см. табл. 1, 2).

Таблица 1

Для отношения R1 существует частичный образ

R1	
A	B
a	b
a ₁	b

Таблица 2

Для отношения R2 существует частичный образ

R2	
B	C
b	c
b	c ₁

Для представления P будем использовать БД, которые являются проекциями возможных состояний, при этом частичный образ таких БД должен быть представлен проекциями возможных состояний из P. Множество таких БД, в обозначении DBps, соответствуют выражению:

$$DB^{ps} = PF_S(R) \cap \pi_S(DB), \tag{2}$$

Из рассмотренных определений, очевидно, что $\pi_S(R) \subseteq \pi_S(DB)$. Тогда, общая схема отношения включения введенных множеств можно изобразить схемой представленной на рисунке 4.

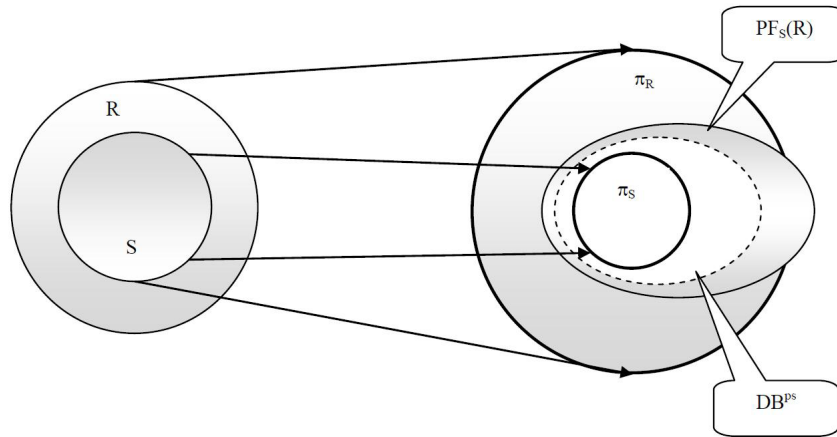


Рис. 4. Схема вложений множеств данных

Для рассмотрения всех возможных состояний, определяемых всеми базами данных из DB , введем обозначение такого множества.

ОПРЕДЕЛЕНИЕ 5. Множество возможных состояний баз данных из DB будем представлять как подмножество $d\hat{B} \subseteq DB$, в обозначении $DBprt=(\triangleright \triangleleft d\hat{B})$, и определяется как соединимое множество $\{\triangleright \triangleleft (dB_i) \mid dB_i \in DB^{prt}, i = \overline{1, n}\}$.

Аналогично теореме Риссанена о независимых проекциях отношений [12], введем определение декомпозиции множества БД на независимые компоненты.

ОПРЕДЕЛЕНИЕ 6. Декомпозиция множества отношений R относительно множества схем S на независимые компоненты корректна, если выполняются следующие условия:

C1. $\forall dB \in DB^{ps} \exists p_j \in P, \forall j = \overline{1, m} \mid \pi(p_j) = dB$.

C2. $\forall dB \in DB^{ps} \exists p_j \in P, \exists j = \overline{1, m} \mid \pi(p_j) = dB$.

Другими словами, свойство C1 определяет единственность представления, а совместно условия C1 и C2 определяют независимость компонент при декомпозиции.

Условия независимости компонент соответствует тому, что проекция состояний является отображением $\hat{\xi}: \pi_S(P) \rightarrow DB^{ps}$ с точностью до изоморфизма, то есть является взаимно-однозначным. Это подтверждается тем, что, в соответствии с условием C1, существует единственное возможное состояние $p \in P$, а также, согласно условию C2, существует однозначное представление каждого состояния из R базой данных из $DBps$.

Хотя, как было сказано, отображение $\hat{\xi}: \pi_S(P) \rightarrow DB^{ps}$ является взаимно-однозначным, обратный переход от общего представления БД к возможному состоянию является трудно вычислимым, так как проверка согласованности проекций, в общем случае, требует соединения всех отношений из рассматриваемой БД [13].

Для сохранения дистрибутивности проекции к соединению необходимо определить условия, исходя из (1), при которых операция соединения обратна к операции проекции. Соответствия между состояниями БД может быть представлена схемой на рисунке 5.

Как было отмечено, при разложении множества состояний P относительно схемы S на независимые компоненты, отображения ξ и ζ не являются обратными друг к другу.

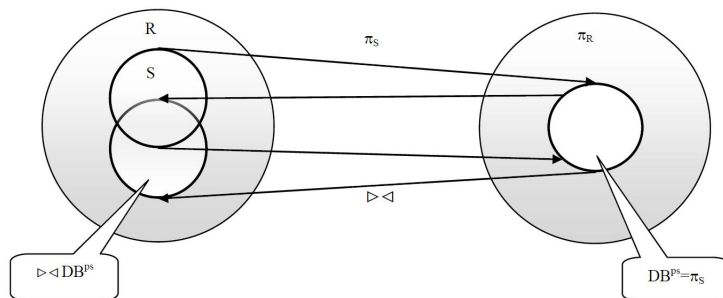


Рис. 5. Схема перехода состояний при выполнении операций проекции и соединения

Пример, приведенный на рис. 6, определяет разложение отношения R в базу данных $dB = \{p_1, p_2\}$ с отношениями R1 и R2 соответственно, но при этом $\triangleright \triangleleft (dB) \neq R$.

A	B	C
a	b	c
a ₁	b	c ₁

B	C
b	c
b	c ₁

A	B
a	b
a ₁	b

A	B	C
a	b	c
a	b	c ₁
a ₁	b	c
a ₁	b	c ₁

$\neq R$

Рис. 6. Разложение, не соответствующее взаимно-однозначному отображению

Вспользуемся определением адекватности декомпозиции отношения, приведенным в работе [14] и определим условия разложения отношения, при котором сохраняется адекватная информация.

ОПРЕДЕЛЕНИЕ 7. Разложение отношения R будет сохранять информацию, если множество возможных состояний обладают следующими свойствами:

CA1. $\forall dB \in DB^{ps} \exists p_j \in P, \forall j=1 \mid \pi_S(p_j) = dB$

CA2. Если $\forall p \in DB$ выполняется $\pi_S(p) \in DB^{ps}$, то $p \in P$.

Как и в определении 6, свойство CA1 определяет единственность представления, а свойство CA2 – сохранение ограничений. Соответственно, совместное выполнение свойств CA1 и CA2 определяют условия сохранения адекватной информации при разложении отношения.

Для обобщения свойств, которыми должно обладать множество P и обеспечения соединения без потери информации, расширим определение проекции состояний.

ОПРЕДЕЛЕНИЕ 8. Прообразом множества баз данных \overline{dB} как подмножества $\pi_S(DB)$, в обозначении $\overline{DB^{prt}}$, будем называть множество возможных состояний из DB^{prt} , определяемых как $\overline{DB^{prt}} = \{p \in DB \mid \pi_S(p) \in \overline{DB^{prt}}\}$.

Для множества частичных образов баз данных DB^{ps} также справедливо определение 8, таким образом, рассмотрим свойства подмножества $\overline{DB^{ps}}$. Структурно вложенность множеств с учетом введенных определений можно представить схемой на рисунке 7.

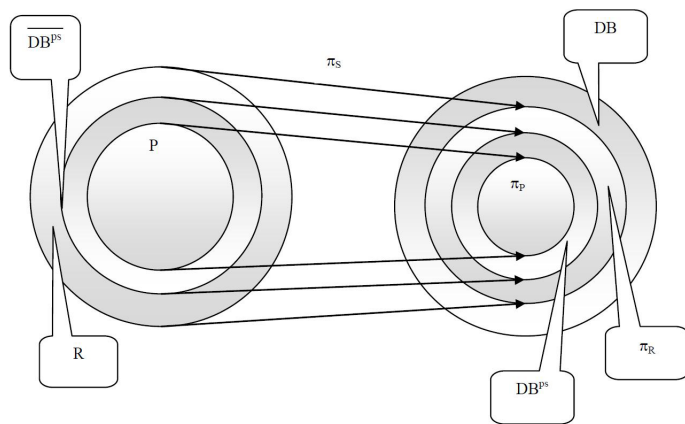


Рис. 7. Схема вложенности множеств состояний

Покажем некоторые соответствия между состояниями БД и определим условия их эквивалентности.

УТВЕРЖДЕНИЕ 1. Равенство вида $\overline{DB^{ps}} = DB$ эквивалентно условию CA2 из определения 7.

ДОКАЗАТЕЛЬСТВО. Проведем доказательство утверждения в двух направлениях. Предположим, что условие CA2 выполняется. Тогда, так как из введенных выше определений вытекает, что $\forall p \in \overline{DB^{ps}} \Rightarrow \pi_S(p) \in DB^{ps}$, то $p \in DB$. Таким образом, $\overline{DB^{ps}} \subset DB$, откуда следует, что $\overline{DB^{ps}} = DB$.

Теперь предположим, что выполняется равенство $\overline{DB^{ps}} = DB$. Возьмем проекцию $\pi_S(p)$ некоторого состояния $p \in DB$, лежащую в области множества DB^{ps} . Для такого состояния выполняется условие $p \in \overline{DB^{ps}}$, из чего вытекает, что $p \in DB$. Таким образом, условие CA2 выполняется.

Утверждение доказано.

Утверждение 1 определяет тот факт, что из $\overline{DB^{ps}} = DB$ следует равенство $DB^{ps} = \pi_S(DB)$. Данный факт определяет следствие, связывающее условия C1, C2 (см. определение 6) и условия CA1 и CA2 (см.

определение 7). Так как условие CA2 влечет условие C2, то условия CA1 и CA2 влекут условия C1 и C2. При этом обратное утверждение неверно.

Кроме того, из свойств CA1 и CA2 вытекает не только то, что возможное состояние p_i , соответствующее некоторому отношению R, можно представить базой данных $dB_i \in DB^{ps}$, но и то, что никакое возможное состояние $p_j \in (DB - R)$ (разность множеств атрибутов DB и R) не может быть представлено той же базой данных $dB_i \in DB^{ps}$.

Критерием декомпозиции отношения R без потерь информации на схему базы данных S является выражение (1). Воспользуемся результатами, полученными в работе [10], и введем выражению (1) альтернативное определение свойства соединения без потерь информации через отображение «проекция-соединение» или PJ-отображение.

ОПРЕДЕЛЕНИЕ 9. PJ-отображение, определяемое на множестве схем S, в обозначении m_S , представляет собой функцию на отношениях из R со схемой S, имеющую вид:

$$m_S(R) = \pi_{S_1}(R) \bowtie \pi_{S_2}(R) \bowtie \dots \bowtie \pi_{S_n}(R). \tag{3}$$

Согласно теории реляционной модели данных, любое отношение БД является конечным и, соответственно, операционная спецификация реляционной модели строит конечные отношения в результате ее использования. Таким образом, можно говорить о неподвижной точке при выполнении последовательности операций проекции и соединения.

ОПРЕДЕЛЕНИЕ 10. Отношение R является неподвижной точкой функции $m_S(R)$, если выполняется равенство $m_S(R) = R$.

Таблица 3

Состояние БД, принадлежащее неподвижной точке

R				
A	B	C	D	E
a	b	c	d	e
a	b_1	c	d_1	e
a	b_1	c	d_1	e_1
a	b	c	d_1	e_1
a	b	c	d_1	e

Таблица 4

Состояние БД, не принадлежащее неподвижной точке

R				
A	B	C	D	E
a	b	c	d	e
a	b_1	c	d_1	e
a	b_1	c	d_1	e_1
a	b	c	d_1	e_1

Например, если состояние p базы данных со схемой $U(A, B, C, D, E)$ представлено отношением $R = \{ABD, BC, ADE\}$ (табл. 3), то R не является неподвижной точкой.

Если p соответствует состоянию, представленному в таблице 4, то база данных R входит в $FIX(R)$, где $FIX(R)$ - множество всех неподвижных точек согласно [10].

Очевидно, что неподвижных точек может быть конечное множество, в зависимости от множества состояний БД (состояний БД также конечное множество, см. выше).

Из выше сказанного можно определить условие сохранения информации с обратным отображением к проекции.

ОПРЕДЕЛЕНИЕ 11. Для множества состояний P отношений R со схемой S соединение (без потерь данных) будет выполняться корректно, если выполняется условие:

$$F1. P \subseteq FIX(R).$$

Другими словами, для множества состояний P базы данных R выполняется условие (1), при этом справедливо выражение (4).

Связь между свойством соединения без потери данных (F1) и единственностью представления (C1)

определяется утверждением о том, что из F1 вытекает C1.

Для доказательства этого утверждения рассмотрим два возможных состояния $p_1, p_2 \in P$. Равенство $\pi_S(p_1) = \pi_S(p_2)$ влечет равенство $\triangleright \triangleleft (\pi_S(p_1)) = \triangleright \triangleleft (\pi_S(p_2))$, тогда F1 влечет равенство $p_1 = \triangleright \triangleleft (\pi_S(p_1))$, откуда вытекает, что $p_1 = p_2$, следовательно, справедливо условие C1. Как следствие, можно утверждать, что условие F1 совместно с условием CA2 влечет свойство CA1.

Рассматривая множество введенных свойств как систему, можно доказать, что условие CA1 совместно с условием CA2 влечет за собой свойство F1.

Предположим, что все условия сохранения информации (т. е. C2 и CA2) выполняются. Для возможного состояния $p \in P$ построим базу данных $dB = \pi_S(p)$, тогда $\triangleright \triangleleft (dB)$ такое состояние, что $\triangleright \triangleleft (dB) = dB$. Так как $dB \in DB^{ps}$ и выполняется $\triangleright \triangleleft (dB)$, то базы данных, для которых выполняется CA2, входят во множество R. Так как $\pi_S(p) = \pi_S(\triangleright \triangleleft (dB))$, то, согласно свойству CA1, имеем $\triangleright \triangleleft (dB) = p$. Из этого следует равенство $p = \triangleright \triangleleft (dB) = \triangleright \triangleleft (\pi_S(p))$, что соответствует условию F1.

Промежуточным итогом проведенных исследований, определяющих систему условий для обеспечения корректности данных при построении универсального отношения, является следующий факт: если для некоторой базы данных DB справедливо свойство CA2, то свойства CA1 и F1 эквивалентны относительно DB.

Другими словами, для условия сохранения данных на некотором состоянии P обратным отображением к ограничению проекции ξ будет отображение соединения ζ (рис. 8).

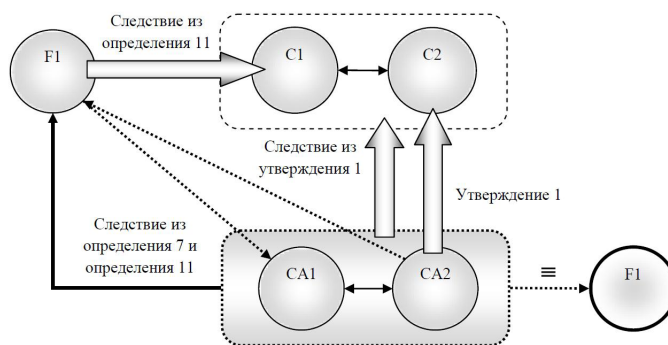


Рис. 8. Схема доказательства сохранения данных при построении универсального отношения

Для демонстрации данного утверждения рассмотрим пример. Пусть для некоторой базы данных $R = \{R_1(S_1), R_2(S_2)\}$, где $S_1 = \{AB\}$ и $S_2 = \{BC\}$ имеют универсальную схему $U = (ABC)$, и пусть множество состояний удовлетворяет $P = \{SAT(B \rightarrow C)\}$. Здесь функциональная зависимость $B \rightarrow C$ гарантирует выполнение свойства F1. При условии, что некоторое состояние p удовлетворяет функциональной зависимости $B \rightarrow C$, проекция $\pi_{BC}(p_i)$ также удовлетворяет зависимости $B \rightarrow C$. Также справедливо обратное, если $\pi_{BC}(p_i)$ удовлетворяет $B \rightarrow C$, то и p_i будет удовлетворять $B \rightarrow C$. Из этого вытекает, что, если $\pi_{BC}(p_i) \in DB^{ps}$ для некоторого p , то $p \in P$, и при этом для P выполняется условие CA2, то такое состояние обладает свойством сохранения данных.

Выводы. Анализ отечественных и зарубежных источников показал, что исследования не рассматривают формальную сторону проблемы, позволяющую выстроить единую концепцию совместного использования неоднородных БД при локальной независимости вычислительных систем. В связи с этим, становится актуальной проблема создания формализованного системного подхода к решению задач разработки эффективных методов управления информационными распределенными ресурсами вычислительных систем, развитие методов и средств решения практических задач обработки данных. Эти вопросы и составили основу исследований в данной работе.

В статье предложен набор необходимых и достаточных условий, позволяющих решать определенный класс задач, связанных с построением адекватных представлений для интегрированных, многомерных и распределенных данных. Выполнение этих условий позволяет обеспечить корректное выполнение запросов в среде гетерогенных баз данных. Кроме этого, отношения, для которых рассмотренные условия выполняются, образуют множество эквивалентных БД относительно данных, что дает возможность интегрировать локально независимые информационные системы с гетерогенными структурами данных.

Практическая значимость полученных результатов заключается в том, что поддержка корпоративных систем, в большинстве случаев, осуществляется с помощью программных средств, реализующих бизнес-логику предприятия, что значительно сокращает жизненный цикл информационных систем. Предлагаемый метод обработки данных, основанный на выполнении условий, приведенных в работе,

позволяет поддерживать совместный доступ к распределенным данным при минимальном использовании дополнительного программного кода.

Эквивалентность баз данных не ограничивается вопросами адекватных отображений структур данных. Одним из наиболее важных вопросов является построение эффективного метода проверки целостности данных, которые не всегда могут быть согласованы. Таким образом, в классе рассматриваемых задач, дальнейшие исследования необходимо направить на разработку методов поддержки целостности локально независимых БД при построении единого информационного пространства в интегрированных и многомерных вычислительных системах.

ЛИТЕРАТУРА:

1. Яковлев С.А., Арсеньев Б.П., Ильин В.П. Интеграция распределенных баз данных на основе сетевых технологий: Учебн. Пособие: СПб.: Изд.-полигр.центр СПбГЭТУ (ЛЭТИ). - 1998. – 68 с.
2. Тальхайм Б. О концептуальном моделировании / Б. Тальхайм // Интеллектуальные системы. – 2006. – Т. 10. – Вып. 14. – С. 303 – 342.
3. Thalheim B. Conceptual Application Domain Modelling / B. Thalheim, K.-D. Schewe, Hui Ma // Conceptual Modelling 2009, Sixth Asia-Pacific Conference on Conceptual Modelling (APCCM 2009). – Wellington, New Zealand, January 20-23 2009. – P. 49-57.
4. Yao H. Cohesion Metrics for Ontology Design and Application / Haining Yao, Anthony Mark Orme, Letha Etzkorn // Journal of Computer Science. – 2005. – №1(1). – P. 107 – 113.
5. Дейт К. Дж. Введение в системы баз данных / К. Дж. Дейт. – М.: Диалектика, 1998. – 784 с.
6. Codd E. F. A Database Sublanguage Founded on the Relational Calculus / E. F. Codd // ACM SIGFIDET Workshop on Data Description, Access and Control. – 1971. – P. 35 – 61.
7. Гарсия-Молина Г. Системы баз данных. Полный курс / Г. Гарсия-Молина, Дж. Ульман, Дж. Уидом.— М.: Вильямс, 2003. — 1088 с.
8. Востров Г.Н. Современные проблемы развития баз данных / Г.Н. Востров, Е.В. Малахов // Тр. Одес. политехн. ун-та. – 1999. – Вып. 2(8). – С. 162 – 164.
9. Ульман Дж. Основы систем баз данных / Дж. Ульман. – М.: Финансы и статистика, 1983. – 334 с.
10. Мейер Д. Математическая теория реляционных баз данных / Д. Мейер. – М.: Мир, 1981. – 476 с.
11. Буслик Н.Н. Глобальные схемы реляционных баз данных: концепция и методы построения / Н.Н. Буслик. – Х.: ХНУРЭ, 2002. – 68 с.
12. Rissanen J. Independent components of relations / J. Rissanen // ACM Transactions on Database Systems – Vol. 2, № 4. – 1977. – P. 317 – 325
13. Таянский С.С. Модель согласования данных при интеграции информационных систем / С.С. Таянский. // Радиоэлектроника и информатика – 2006. – № 2. – С. 84 – 91.
14. Arora A.K. The information preserving properties of relation databases transformation / A.K. Arora, C.R. Carlson // VLDB IV, West Berlin, Germany. – ACM IEEE, 1978. – P. 352-359.
15. УДК 004.475: [004.652: 004.655]

ГРИНЕВА Елена Евгеньевна, методист 1 кат., Харьковский национальный университет радиоэлектроники.

Научные интересы: поддержка целостности распределенных данных при динамических изменениях структурных компонент схемы базы данных.