Проблемы бионики. 2001. Вып. 54. С. 3-9. ХТУРЭ

УДК 519.68

Н. Н. БУСЛИК

АКСИОМАТИЧЕСКАЯ ТЕОРИЯ ЗАВИСИМОСТЕЙ ПО СУЩЕСТВОВАНИЮ В РЕЛЯЦИОННОЙ МОДЕЛИ ДАННЫХ

Концепция универсального отношения

Одной из отличительных черт реляционной модели данных является отсутствие в ее структурном компоненте каких бы то ни было связей между отношениями (реляционными таблицами) базы данных (БД). Вместе с тем при проектировании и организации крупномасштабной БД широко используются ограничения целостности, характеризующие взаимосвязь ее отношений. Например, в реляционном языке обработки данных SQL к стандартным средствам описания схемы относятся так называемые ограничения ссылочной целостности, задаваемые с помощью чужого ключа (предложение FOREING KEY). Чужой ключ есть совокупность атрибутов отношения, сопоставляемых с атрибутами возможного ключа другого (ссылочного) отношения. Соответствующее ограничение состоит в том, что значение чужого ключа в каждом кортеже исходного отношения должно содержаться как значение возможного ключа в каком-либо кортеже ссылочного отношения.

С формальной точки зрения такое ограничение является разновидностью ограничения по существованию, накладываемого на отношение БД (зависимость по существованию) [1]. Пусть $R(A_1,..., A_k)$ и $S(B_1, ..., B_l)$ – схемы отношений r и s реляционной БД, где A_i, B_i – атрибуты отношений (точнее, их имена). Положим, набор атрибутов $X=(A_1, ..., A_m)$ отношения r объявлен как чужой ключ, ссылающийся на набор атрибутов $Y=(B_1, ..., B_m)$, который является ключом отношения s (атрибуты этих двух наборов попарно сопоставимы: A_1 – с B_1, A_2 – с B_2 и т. д.). Тогда в каждом допустимом состоянии БД отношения r и s таковы, что множество X-подкортежей отношения r (проекция r на X) является подмножеством Y-подкортежей отношения s: $r[X] \subseteq s[Y]$.

Рассмотрим более широкий класс зависимостей по существованию, не выдвигая требования к Y быть ключом отношения, (такое расширение, как увидим далее, представляет существенный интерес для теории и практики организации БД). Итак, говорим, что r зависит по существованию от s (обозначим это как $r=<s$), если $r[X] \subseteq s[Y]$ в каждом допустимом состоянии реляционной БД.

Обобщим наше определение на случай нескольких отношений. Вначале заметим, что сама возможность задания зависимости по существованию между отношениями обусловлена сопоставимостью (своего рода "равенством") атрибутов этих отношений (A_i:B_i). В целях упрощения дальнейших построений будем считать, что такие "равные" атрибуты имеют в наших отношениях одинаковые имена. Таким образом, заменив сопоставимые атрибуты некоторыми "универсальными", можем сформировать универсальное отношение (УО) как естественное соединение $r*s$ наших исходных отношений (с модифицированной схемой). Теперь для зависимости по существованию можно сформулировать эквивалентное определение: $r=<s$, если $r=(r*s)[R]$. Другими словами, зависимое отношение сохраняется в УО как множество его подкортежей.

Понятия универсального отношения и универсальных атрибутов можно распространить на любое число локальных отношений. Продолжая наши рассуждения, рассмотрим пример БД, состоящей из трех отношений со схемами: $R(A_1, A_2, A_3)$, $S(B_1, B_2, B_3)$, $T(C_1, C_2, C_3)$. Пусть заданы ограничения на таких сопоставимых наборах атрибутов: (A_1, A_2):(B_1, B_2), (B_2, B_3):(C_1, C_2) (направление зависимости не суть важно).

Как видим, здесь атрибут B_2 сопоставим с одной стороны с A_2, с другой – с C_1. Пока ничто не препятствует возможности присвоить одно универсальное имя всем трем атрибутам (что, кстати, даст возможность сопоставить между собой отношения R и T). Выполним возможные переименования, заменив A_1 и B_1 на A; A_2, B_2 и C_1 на B; B_3 и C_2 на C. Получим такую схему универсального отношения: $U(A, B, A3, C, C3)$.

Очевидно, в качестве левой и правой частей зависимостей по существованию теперь могут выступать множества отношений, образующие на уровне экстенсионала (экземпляров отношений) естественные соединения. При этом "естественность" соединения определяется в контексте нашего УО ($u=r*s*t$). Рассмотрим, например, зависимость вида $r=<st$. Такое ограничение означает выполнение требования $r=u[R]$.

Попутно заметим, что при отсутствии общих атрибутов, отношений их естественное соединение вырождается в декартово произведение; каждое из локальных отношений сохраняется в таком соединении.

Вернемся к проблеме формирования УО. Предположение об УО (универсальное допущение) широко используется в теории БД как постулат, служащий для обоснования эквивалентных преобразований БД (ее схемы) [1, 2]. Известны различные трактовки этого допущения, что обусловлено не столько различием в подходах к проблеме эквивалентных преобразований, сколько различием учитываемых видов ограничений целостности БД (наличие которых только и дает возможность выполнять эквивалентные преобразования). Исходим из того, что схема УО (и, соответственно, способ его формирования) целиком и полностью определяется заданными зависимостями по существованию между исходными локальными отношениями. Трудность вычисления УО здесь связана с тем обстоятельством (хорошо известным из практики), что одни и те же отношения могут находиться в нескольких связях и характеризоваться различными зависимостями по существованию.

Пусть, например, между отношениями $R(A_1, A_2, A_3)$ и $S(B_1, B_2, B_3)$ заданы две зависимости: $r(A_2){=}{<}s(B_1)$ и $r(A_3){=}{<}s(B_1)$ (здесь в скобках указаны сопоставимые атрибуты). Атрибут B_1 сопоставляется как с A_2 из отношения R, так и с A_3 из того же отношения. При этом, разумеется, о "равенстве" атрибутов A_2 и A_3 говорить не приходится. Какое "естественное" соединение следует тогда выполнить для получения УО? Ответ на поставленный вопрос хорошо известен (см., например, [3]): поскольку одно отношение S выступает одновременно в двух ролях, оно используется для формирования УО дважды – всякий раз со своим отличным набором атрибутов. Таким образом, вместо одного "физического" отношения получим два "виртуальных", имеющих, например, схемы $S1(B1_1, B1_2, B1_3)$ и $S2(B2_1, B2_2, B2_3)$. Теперь можем выполнить "универсализацию" атрибутов по изложенной выше методике.

Как видим, в УО все атрибуты S вынужденно дублируются. Такое же дублирование наблюдается во всех связях S с другими отношениями. Отсюда число "виртуальных" отношений может многократно превосходить число "реальных". В то же время наша абстракция в виде УО не ведет к усложнению спецификации ограничений целостности в схеме БД: пользователь (проектировщик БД) характеризует связи между "реальными", а не "виртуальными" отношениями; все остальные преобразования осуществляются "автоматически".

Более детальное изложение концепций УО выходит за рамки настоящей статьи (для ознакомления здесь можно порекомендовать упомянутые выше источники, а также [4]). В дальнейшем не будем интересоваться природой заданных ограничений, характером отношений "реальных" или "виртуальных", на которых они определены. Основной целью нашего исследования является выяснение фундаментальных (аксиоматических) свойств зависимостей по существованию. О побудительных мотивах и значимости такого исследования будет сказано несколько ниже. Пока же отметим, что в теории БД известны и другие подходы к определению зависимостей по существованию. Так, во [2] рассматриваются зависимости по существованию между атрибутами отношения (Е-зависимости): атрибут A зависит от B ("требует" B), если в каждом кортеже отношения из того, что значение A определено (не "пусто"), следует, что значение B также определено. Аналогично вводится понятие зависимости между множествами атрибутов, которое, как видим, в чем-то сходно с нашими определениями. Действительно, в нашем естественном соединении каждый кортеж зависимого отношения является подкортежем соединения и имеет в качестве "партнера" подкортеж ссылочного отношения в том же кортеже соединения.

Оказывается, что Е-зависимости обладают теми же аксиоматическими свойствами, что и функциональные зависимости атрибутов: рефлексивность, транзитивность и др. (см. ниже). Это дает возможность "логически" выводить одни зависимости из других, анализировать избыточность множества заданных зависимостей и т. д. К сожалению, в нашем случае непосредственно применить имеющуюся аксиоматику функциональных зависимостей невозможно, поскольку ограничения (и их аксиоматические свойства) характеризуют связи между отношениями, а не атрибутами. Следующие два раздела статьи посвящены анализу аксиоматических свойств зависимостей по существованию между отношениями реляционной БД. В последнем разделе обсуждаются некоторые прикладные аспекты использования полученных результатов.

Исследование свойств зависимостей по существованию

Как уже отмечалось, при изучении свойств зависимостей по существованию можно опереться на известные аксиоматические свойства функциональных зависимостей атрибутов. Напомним, это ограничение состоит в том, что всякому значению атрибута A_1 в отношении соответствует не более одного значения атрибута A_2. Декларативная спецификация функциональной зависимости атрибута B от атрибута A имеет вид $A \rightarrow B$. Аналогично, если X и Y – подмножества атрибутов отношения, то $X \rightarrow Y$ означает функциональную зависимость Y от X (при этом не обязательно, чтобы $X \cap Y = \varnothing$).

Существуют и широко используются несколько полных и надежных аксиоматических систем функциональных зависимостей. В частности, одной из популярных является приведенная ниже система аксиом Делобеля [5].

FD1. Рефлексивность: $X \rightarrow X$.

FD2. Проективность: $X \rightarrow YZ \Rightarrow X \rightarrow Y$.

FD3. Пополнение: $X \rightarrow Y \Rightarrow XZ \rightarrow Y$.

FD4. Аддитивность: $X \rightarrow Y,\ X \rightarrow Z \Rightarrow X \rightarrow YZ$.

FD5. Транзитивность: $X \rightarrow Y,\ Y \rightarrow Z \Rightarrow X \rightarrow Z$.

(Здесь запись вида XY обозначает объединение множеств атрибутов X и Y).

Для нашего дальнейшего исследования важным является то обстоятельство, что аксиоматика F-зависимостей может быть использована не только в приложении собственно к зависимостям, но и к пересечениям произвольных множеств; известно, что отношение включения пересечений множеств обладает теми же аксиоматическими свойствами [6].

Продолжим анализ зависимостей по существованию. Формально, рассмотрим множество отношений, заданных структурной схемой БД $U = \{r_1 (R_1),\ ...,\ r_q\ (R_q)\}$, где r_i – имя отношения; R_i – носитель отношения; $\bigcup_{i=1}^{q} R_i$ – носитель УО. На множестве L^* подмножеств U определим бинарное отношение, такое, при котором пара $(L_i\ , L_j)$ принадлежит этому отношению, если задана зависимость $L_i =< L_j$, т. е. зависимость по существованию вида $r_{i1}...r_{im} =< r_{j1}...r_{jn}$, где r_{ik} – отношение, входящее в L_i; r_{jl} – отношение входящее в L_j.

Покажем, что зависимости по существованию наподобие рассмотренных выше F-зависимостей, обладают рядом аксиоматических свойств, позволяющих дедуктивно выводить одни зависимости из других. Перечислим эти свойства, одновременно доказывая их значимость (корректность) для модели реляционной БД. Предварительно введем следующие обозначения. Объединение множеств L и M обозначим LM; при этом подразумеваем, что если на уровне экстенсионала $L = r_1 * ... * r_m$, $M = s_1 * ... * s_n$, то $LM = r_1 * ... * r_m * s_1 * ... s_n$. Носитель L обозначим как $[L]$, подразумевая, что если $R_1,\ ...,\ R_m$ – носители отношений $r_i\ (r_i \in L)$, то $[L] = \bigcup_{i=1}^{m} R_i$. Пересечение носителей $[L]$ и $[M]$ обозначим $[L, M]$.

ED0. Вырожденная выполнимость: $[L, M] = \varnothing \Rightarrow L =< M$.

Доказательство. Ввиду отсутствия общих атрибутов у любых двух отношений r и s их естественное соединение вырождается в декартово произведение. По определению зависимости получаем $r =< (r \times s)$, $s =< (r \times s)$.

Обоюдная зависимость означает согласованность отношений. В данном случае имеем вырожденную согласованность, не представляющую интереса для анализа модели БД.

ED1. Псевдорефлексивность: $LM =< L$.

Доказательство. Как видим, аксиома не сопровождается никакими дополнительными синтаксическими условиями, т. е. она, как будет показано, справедлива для любых L, M независимо от состояния носителей. Действительно, для любых отношений r и s справедливо $r*s*r = r*s$, откуда по определению $rs =< r$.

Частный случай аксиомы – рефлексивность: $L =< L$.

ED2. Пополнение: $L =< M \Rightarrow L =< LM$.

Доказательство. Рассмотрим отношения r, s, такие, при которых $r =< s$. В силу того что $r*s = r*r*s$, имеем $r =< rs$.

ED3. Проективность: $L =< MN \Rightarrow L =< M$

Доказательство. Как видим, аксиома должна быть справедливой безусловно, т. е. при любых носителях рассматриваемых отношений. Пусть r, s, t – отношения, при которых $r =< st$. Тогда соединение $r*s*t$ содержит все R-кортежи, а значит, и $r*s$ содержит все R-кортежи.

ED4. Транзитивность: $L=<M,\ LM=<N \Rightarrow L=<N$.

Доказательство. Пусть отношения r, s, t таковы, что $r=<s$ и $rs=<t$. По свойству ED2 отсюда вытекают зависимости $r=<rs$ и $rs=<rst$. Следовательно, $r{*}s{*}t$ сохраняет все RS-кортежи, а rs все R-кортежи, откуда $r=<rst$ и по свойству ED3 $r=<t$.

ED5. Условная транзитивность: $L=<M,\ M=<N \Rightarrow L=<N$, если $[L, N] \subseteq [M]$.

Доказательство. Пусть r, s, t – такие отношения, при которых $r=<s$ и $s=<t$. Проведем полную индукцию возможных "синтаксических ситуаций".

Случай 1. Пусть не выполняется ни одно из включений: $[R, S] \subseteq [T]$, $[R, T] \subseteq [S]$, $[S, T] \subseteq [R]$. Тогда существует БД, в которой выполняются $r=<s$, $s=<t$, но не выполняется $r=<t$. Пример такой БД показан на рис. 1.

```
r(A B C)    s(B C D)    t(A D)    [R,S]=BC
a b c       b c d       a d       [R,T]=A
x b c       y c d       o h       [S,T]=D
```

Рис. 1

Случай 2. Пусть выполняется включение $[R, S] \subseteq [T]$ и не выполняются включения $[R, T] \subseteq [S]$, $[S, T] \subseteq [R]$. Тогда существует БД, в которой не выполняется $r=<t$ (рис. 2).

```
r(A B E)    s(A C D)    t(A C B)    [R,S]=A
a b e       a c d       a c b       [R,T]=AB
x y e       x c d       x c b       [S,T]=AC
                        o c b
```

Рис. 2

Случай 3. Пусть выполняется включение $[S, T] \subseteq [R]$, и не выполняются $[R, T] \subseteq [S]$, $[R, S] \subseteq [T]$. Пример БД, в которой не выполняется $r=<t$ (рис. 3).

```
r(A B C)   s(A B D)   t(B C)    [R,S]=AB
  a b c      a b d      b c     [R,T]=BC
  a b z      x b d      y c     [S,T]= B
```

Рис. 3

Случай 4. Пусть выполняется включение $[R, S] \subseteq [T]$, $[S, T] \subseteq [R]$, но не выполняется $[R, T] \subseteq [S]$. Тогда существует БД, в которой не выполняется $r=<t$ (рис. 4).

```
r(A B C)    s(B D)    t(B C E)    [R,S]=B
  a b c       b d       b c e     [R,T]=BC
  a y c       y d       y z e     [S,T]=B
```

Рис. 4

Случай 5. Пусть выполняется $[R, T] \subseteq S$. Ввиду того что $r=<s$, все $[R, T]$-подкортежи отношения r сохраняются в rs, а из-за того что $[R, T] \subseteq S$, эти подкортежи сохраняются и в s. Так как $s=<t$, то все $[R, T]$-подкортежи s сохраняются в st, а следовательно, и все $[R, T]$-подкортежи r сохраняются в st. Таким образом, $r=<st$ и по свойству ED3 $r=<t$ выполняется.

Рассмотрев все возможные "синтаксические ситуации" приходим к выводу, что $[R, T] \subseteq S$ является как достаточным, так и необходимым синтаксическим условием.

ED6. Условная аддитивность: $L=<M,\ L=<N \Rightarrow L=<MN$, если $[M, N] \subseteq [L]$.

Доказательство. Схема доказательства аналогична схеме доказательства ED5. Поэтому не будем рассматривать каждый возможный случай включения пересечений носителей. На рис.5. приведен показательный контрпример. Остановимся на ситуации, когда отношения r, s, t таковы, что $r=<s$ и $r=<t$

выполняются, и выполняется *[S, T]* $\subseteq R$. Пусть s' – подмножество кортежей *s*, которые сохраняют *r* (т.е. участвуют в соединении со всеми кортежами *r*), t' – подмножество *t*, сохраняющее *r*. Поскольку *[R, S]*–подкортежи содержат в себе *[S, T]*–подкортежи (*[S, T]* $\subseteq$ *[R, S]* следует из *[S, T]* $\subseteq R$ и *[S, T]* $\subseteq S$ по аксиоме FD4) и *[R, T]*-подкортежи содержат *[S, T]*-подкортежи, то s' и t' содержат все *[S, T]*-подкортежи отношения *r*. Рассмотрим теперь соединение *s*t*. Очевидно, для носителей *r* и *s*t* справедливо *[R, ST]=[R, S]* $\cup$ *[R, T]*. Поскольку в s' и t' содержатся все необходимые *[S, T]*-подкортежи, являющиеся подкортежами всех необходимых *[R, S]*- и *[R, T]*-подкортежей, постольку $s'*t'$ будет содержать все необходимые *[R, ST]*-подкортежи, т. е. все подкортежи отношения *r*. Отсюда следует, что *r* сохраняется в *r*s*t*.

```
r(A B C)      s(A B D)      t(B C D)      s*t(A B C D)      r=<s,r=<t
  a b c         a b h         b c d           a b z h       [S,T]⊄[R]
                x b d         b z h           x b c d       ¬(r=<st)
```

Рис. 5

Итак, мы показали, что синтаксическое условие *[S, T]* $\subseteq$ *[T]* для ED6 является не только достаточным, но и необходимым.

Полнота и непротиворечивость теории

Свойства ED0–ED6 будем рассматривать как единую систему аксиом (ED0, ED1) и правил вывода (ED2–ED6) формальной теории зависимостей по существованию. Множество E^+ зависимостей, выводимых в системе ED0–ED6 из заданного множества *E*, назовем замыканием *E*.

Покажем теперь, что система ED0-ED6 надежна, непротиворечива и полна. Свойство надежности (значимости) подразумевает, что все зависимости из E^+ выполняются в любой БД *b*, удовлетворяющей множеству *E* ($b \in SAT(E)$).

Теорема 1. Система ED0–ED6 надежна.

Доказательство. Для каждой аксиомы ED0–ED6 её значимость уже доказана. Разумеется, их совместное применение также не может порождать незначимую зависимость.

Теорема 2. Система ED0–ED6 непротиворечива.

Доказательство. Другими словами, для любых множества *E* и зависимости *h* всегда существует БД *b* из *SAT(E)*, такая, что $b \in SAT(h)$. По-другому, $SAT(E^{\cup}\{h\})$ – непустое множество. В предельном случае БД может удовлетворять всем возможным зависимостям. В качестве примера такой БД можно рассмотреть множество локальных отношений, носители которых не имеют общих атрибутов. Нетрудно убедиться в том, что в силу ED0 БД удовлетворяет всем возможным зависимостям.

Говорим, что теория полна, если для зависимости, не выводимой из *E* ($h \notin E^-$), найдется БД *b*, в которой все зависимости из *E* выполняются, а зависимость *h* не выполняется ($b \notin SAT(h)$).

Теорема 3. Система ED0–ED6 полна.

Доказательство. Пусть $h= r=<s \notin E^+$. Построим БД *b*, отвечающую этому условию (рис. 6). Здесь, очевидно, выполняются все зависимости между отношениями *r*, t_i и их соединениями, а также между отношениями *s*, t_i и их соединениями. В то же время наше условие предопределяет отсутствие ограничений вида $rt_i=<s$, так как вне зависимости от вида отношений БД по правилу ED4 из $r=<t_i$ и $rt_i=<s$ следовало бы $r=<s$. Аналогично, невозможно выполнение зависимостей вида $r=<st_i$, поскольку в силу ED2 из них следует $r=<s$. Таким образом, в БД *b* выполняются все возможные зависимости по существованию на множестве ее отношений. Пусть *E* – произвольное множество зависимостей, такое, что $r=<s \notin E^+$. Тогда с учетом непротиворечивости теории $b \in SAT(E)$, что и требовалось доказать.

```
r(A B)    t1(B B1)  . . .  tn(B Bn)    s(A C)
  a b       b 1              b n         a c
  x b
```

Рис. 6

Независимой (избыточной) назовем систему аксиом, в которой ни одна аксиома не может быть выведена из множества других. В нашем случае можем говорить об избыточности аксиом ED2 и ED4.

Вначале покажем избыточность безусловной транзитивности (ED4). Носители отношений, упоминающихся в аксиоме, удовлетворяют синтаксическим условиям ED5 и, таким образом, ED4 – частный случай ED5.

Можно также "упростить" ED1, заменив её аксиомой рефлексивности: $L{=<}L$. Действительно, из $LM{=<}LM$ по аксиоме ED3 следует $LM{=<}L$.

Исключим теперь ED2. Исходной в ней является зависимость $L{=<}M$. С учетом $L{=<}L$ по аксиоме ED6 имеем $L{=<}LM$ (применение ED6 законно, поскольку LM удовлетворяет синтаксическим условиям аксиомы).

Таким образом, избыточность в ED-системе введена исключительно из соображений удобства её использования в процессе вывода тех или иных зависимостей.

Прикладные аспекты теории

Кратко остановимся на вопросах применения теории зависимостей по существованию к задачам проектирования БД. Следует прежде всего заметить, что поддержка зависимостей по существованию как ограничений целостности БД является вычислительно емкой задачей (строго говоря, – NP-полной). Отсюда становится очевидной необходимость сокращения числа зависимостей, подлежащих контролю в процессе ведения БД. Наличие надежной и полной системы аксиом дает возможность решить задачу вычисления базиса произвольного множества зависимостей E, т. е. такого неизбыточного множества H, при котором $H^{+}=E^{+}$. Общая схема вычислений сводится к следующему: для каждой зависимости множества E последовательно проверяем, выводится ли она из остальных; если да – исключаем эту зависимость из нашего множества. По окончании просмотра всех элементов E получим базис для E.

Разумеется, для вычисления базиса необходимо уметь эффективно решать задачу выводимости произвольной зависимости из множества других. Как известно, аксиоматические теории прямого решения подобной проблемы (так называемой проблемы членства) не дают. Алгоритм решения проблемы членства для зависимостей по существованию в рамках настоящей статьи не рассматривается. В то же время относительно нашей теории справедливо утверждение о ее разрешимости. Действительно, замыкание всякого множества E конечно (как конечно и само множество E). Следовательно, вычислив замыкание для E, мы можем для произвольной зависимости h указать, принадлежит ли она E^{+}.

Другой важной проблемой проектирования, связанной с поддержкой зависимостей по существованию, является отыскание согласованных в целом максимальных подмножеств отношений БД. Обоюдная зависимость отношений говорит об их согласованности. Поддержка согласованности как ограничения целостности БД вызывает серьезные затруднения. Действительно, если схемой БД определено, что одно отношение зависит по существованию от другого и в то же время второе зависит от первого, то не представляется возможным добавить новый кортеж в одно из отношений без нарушения указанного ограничения (подразумевается, что пересечение носителей отношений не пусто и новый кортеж содержит новые значения атрибутов пересечения). В таких случаях целесообразно объединять схемы обоих локальных отношений в одну схему и хранить в БД одно отношение вместо двух исходных. Аналогичное решение может быть принято и в случае согласованности нескольких отношений.

Говорим, что множество отношений согласованно в целом, если каждое из них сохраняется в естественном соединении всех отношений этого множества. Другими словами, каждое отношение зависит по существованию от множества остальных. Кроме того, множество отношений должно быть связно. Связность понимается таким образом, что носитель каждого отношения имеет непустое пересечение с объединением носителей других отношений. Такое требование обусловлено отсутствием необходимости поддерживать зависимость по существованию между несвязными отношениями (см. аксиому ED0). Согласованное в целом множество по изложенным выше соображениям целесообразно объединить в одно отношение (СЦ-блок). В общем случае все множество исходных локальных отношений БД разбивается на СЦ-блоки, количество которых должно быть, очевидно, минимальным. Таким образом, при наличии в схеме БД ограничений по существованию между отношениями желательно отыскивать максимальные (не расширяемые) согласованные в целом множества локальных отношений.

Поставленная задача решается в несколько этапов [4]. На первом из них исходя из заданного множества зависимостей находят пары согласованных отношений, для чего относительно каждой пары решается описанная выше проблема членства. Очевидно, попарная согласованность некоторого множества отношений является необходимым условием согласованности в целом для этого множества. Дальнейшие построения состоят в формировании множеств из отношений, являющихся попарно согласованными и удовлетворяющих синтаксическому условию так называемой квазистягиваемости соединения (детальное описание процедур этого этапа, известных как алгоритм Грэхема, можно найти в [2]).

Таким образом, полученная система аксиом и правил вывода для зависимостей по существованию служит надежной теоретической основой для разработки алгоритмов анализа множеств ограничений целостности в виде таких зависимостей и выбора наиболее эффективных вариантов схемы реляционной БД.

Список литературы: 1. *Цикритзис Д., Лоховски Ф.* Модели данных. М.: Мир, 1985. 2. *Мейер Д.* Теория реляционных баз данных. М.: Мир, 1987. 3. *Дейт К.* Введение в системы баз данных. К.: Диалектика, 1998. 4. *Буслик Н.Н.* Глобальные схемы реляционных баз данных: концепция и методы построения. Харьков, 2000. Деп. в ГНТБ Украины, № 25-Ук2000. 5. *Armstrong W., Delobel C.* Decomposition and functional dependencies in relations // ACM Transactions on Database Systems. 1980. Vol. 5, № 4. 6. *Sagiv Y.* et al. An equivalence between relational database dependencies and a fragment of propositional logic // Journal of the ACM. 1982. Vol. 29, № 5.

Поступила в редколлегию 14.03.2001