

ОБОБЩЕНИЕ ФУНКЦИОНАЛЬНЫХ ЗАВИСИМОСТЕЙ В РЕЛЯЦИОННОЙ МОДЕЛИ ДАННЫХ

Я. Деметрович, Др. Дьепеши

§ 0. Введение и основные определения

Реляционная модель данных, введенная Е.Ф. Коддом [7], является одним из самых многообразных средств обработки данных. Эта модель выдвигает на первый план не машинную эффективность, а наглядное описание данных с точки зрения пользователя. Хранящиеся данные наглядно изображаются в виде матриц. Строки матриц соответствуют записям, а столбцы – свойствам или атрибутам, характерным для данного типа записей.

Дадим теперь строгое определение. Пусть Ω – конечное, непустое множество – множество атрибутов. Для каждого $a \in \Omega$ обозначим через $D(a)$ непустое множество, являющееся множеством значений атрибута a . Подмножества декартового произведения $\prod_{a \in \Omega} D(a)$ называется реляциями над множеством Ω . Таким образом реляция R над Ω состоит из таких функций $g: \Omega \rightarrow \prod_{a \in \Omega} D(a)$, для которых $g(a) \in D(a)$, если $a \in \Omega$.

Функции, составляющие реляцию R , мы называем строками R .

Таким образом реляцию R под множеством $\Omega = \{a_1, \dots, a_n\}$ можно представить в виде двумерной таблицы:

	a_1	a_2	$\dots$	a_n
g	$g(a_1)$	$g(a_2)$	$\dots$	$g(a_n)$
$\vdots$				
h	$h(a_1)$	$h(a_2)$	$\dots$	$h(a_n)$
$\vdots$				

Понятие функциональной зависимости Е.Ф. Кодд [7] ввел следующим образом.

Пусть R -реляция над Ω и пусть $A, B \subseteq \Omega$. Будем говорить, что B функционально зависит от A в R /обозначается: $A \xrightarrow[R]{f} B$ /, если для любых $g, h \in R$

$$(1) \quad (\forall a \in A)(g(a) = h(a)) \Rightarrow (\forall b \in B)(g(b) = h(b))$$

(1) означает, что

(1') если две строки равны на всех элементах A , то они равны и на всех элементах B .

«»

На основе формального определения ((1)) и значения функциональной зависимости можно естественным образом определить еще три вида зависимостей, которые характеризуют реляции.

Заметим, что в (1) присутствуют два квантора $\forall$. Заменяя их всеми возможными способами на кванторы $\exists$ и $\forall$, мы получим формулы, имеющие смысл:

$$(2) \quad (\exists a \in A)(g(a) = h(a)) \Rightarrow (\exists b \in B)(g(b) = h(b))$$

$$(3) \quad (\exists a \in A)(g(a) = h(a)) \Rightarrow (\forall b \in B)(g(b) = h(b))$$

$$(4) \quad (\forall a \in A)(g(a) = h(a)) \Rightarrow (\exists b \in B)(g(b) = h(b))$$

(2), (3) и (4) означают соответственно следующее:

(2') если две строки равны на некотором элементе A, то они равны и на некотором элементе B.

(3') если две строки равны на некотором элементе A, то они равны на всех элементах B.

(4') если две строки равны на всех элементах A, то они равны на некотором элементе B.

По образцу определения функциональной зависимости на основе

(2), (3) и (4) дадим определение дуальной зависимости

$(A \xrightarrow[d]{R} B)$ и сильной зависимости $(A \xrightarrow[s]{R} B)$ и слабой зависимости $(A \xrightarrow[w]{R} B)$.

Если R - реляция над Ω , то пусть

$$F_R = \{(A, B) : A \xrightarrow{f}{R} B\}$$

$$D_R = \{(A, B) : A \xrightarrow{d}{R} B\}$$

$$S_R = \{(A, B) : A \xrightarrow{s}{R} B\}$$

$$W_R = \{(A, B) : A \xrightarrow{w}{R} B\}$$

Если $Y \in \{F, D, S, W\}$, $y \in \{f, d, s, w\}$ и y соответствует Y /например, если $Y = D$, то $y = d$ /, то множество вида Y_R мы называем полными y -семействами.

Естественна потребность определить для полных y -семейств наиболее общие свойства - т.е. выполняющиеся для Y_R при любых реляциях R - таким образом, чтобы эти свойства описали все полные y -семейства. Другими словами, если некоторое

$Y \subseteq P(\Omega) \times P(\Omega)$ обладает этими свойствами, то должна существовать такая реляция R над Ω , что $Y = Y_R$. Короче говоря, мы хотели аксиоматизировать y -семейства для любого $y \in \{f, d, s, w\}$.

Математическую структуру функциональных зависимостей впервые изучал В.В. Армстронг [1] . Он аксиоматизировал полные f -семейства /см. § 1, φ -аксиомы/ и доказал, что полное f -семейство определяется системой своих максимальных зависимостей /таких зависимостей, зависящая часть которых максимальна/, и даже системой своих максимальных правых частей /зависящая часть максимальной зависимости/.

Г. Цедли в [8] , аналогичным способом аксиоматизировал полные d - и s -семейства. Аксиоматизация полных w -семейств тем же путем не удалась.

В этой работе мы дадим новые аксиомы для характеристики y -семейств для любого $y \in \{f, d, s, w\}$ /в § 1 F -, D -, S - и W -аксиомы, а в § 2 F' -, D' -, S' - и W' - аксиомы/. Эти аксиомы выделяют комбинаторную структуру y -семейств, в то время как системы аксиом Армстронг и Цедли подчеркивают их логическую структуру.

Нам кажется, что комбинаторный подход более пригоден для рассмотрения возникающих здесь проблем. Кроме того, таким путем удалось аксиоматизировать полные w -семейства.

В § 1 мы приведем новые аксиомы / F -, D -, S - аксиомы/ для полных f -, d - и s -семейств, а также докажем, что они эквивалентны аксиомам, которые ввелись [1, 5, 8] .

В § 2 мы введем понятие равенственного множества реляции, охарактеризуем равенственные множества комбинаторным путем /теорема 2.1./ и, основываясь на этом понятии, введем F' -, D' -, S' - и W' - аксиомы. Будет доказано, что F' -, D' -, и S' - аксиомы эквивалентны, соответственно, φ -аксиомам ϑ и γ -аксиомам. В теореме 2.2. мы укажем на одно существенное отличие полных w -семейств от остальных полных семейств.

В теореме 2.3. мы покажем, что F' -, D' -, S' - и W' - аксиомы действительно описывают полные f -, d -, s - и w -семейства.

В дальнейшем часто употребляем следующие обозначения:

Если в область определения строки g ($g \in R$) $A \subseteq \Omega$, то обозначим через $g \upharpoonright_A$ те строки, область определения которой совпадает с A и для которой $(\forall a \in A) g \upharpoonright_A(a) = g(a)$.

В заключении сделаем несколько практических замечаний. Функциональные, дуальные, сильные и слабые зависимости являются такими ограничениями, знание которых дает возможность охарактеризовать реляции с помощью ограничений, налагаемых на некоторое подмножество полного множества столбцов.

Проектировщики реляционных баз данных знакомы с некоторыми зависимостями, имеющими место в базах данных. Обычно эти зависимости не образуют полного семейства, то есть в базе данных наверняка имеют место и другие зависимости. Их определение и экономичное хранение является одной из основных проблем.

Отметим также, что теория зависимостей тесно связана с одной исключительно трудной комбинаторной проблемой: с определением мощности так называемых Δ -систем.

§ 1. Старые и новые аксиомных разных зависимостей.

В этом параграфе, мы опишем известные аксиомы функциональных [1], двойственных и сильных систем зависимостей [8], а затем приведем новые аксиомы /F-, D-, S - аксиома/ эквивалентные им. Дальнейший анализ новых аксиом приведет к абстрактному описанию слабых систем зависимостей.

На основе функциональных зависимостей / φ - аксиом/ казалось, что проблемы теории разных - функциональных, двойственных, сильных и слабых - зависимостей имеют логический характер [2,3,11] и по нашему мнению их легче рассматривать на основе F-аксиомы.

Сходство F-, D-, S- аксиом легко описывается, поэтому нетрудно перенести на случай функциональных, двойственных зависимостей понятия и результаты, известные для функциональных зависимостей. Согласно F'- D'- S'- W' аксиомам /§ 2/ то же самое можно сделать и для слабых зависимостей.

Начнем с определения функциональных, двойственных, сильных и слабых зависимостей.

Определение 1.1. Пусть даны подмножества A, B ($A, B \subseteq \Omega$) и ре-
ляция R над множеством Ω . Мы говорим, что множество B

- (i) функционально /обозначается: $A \xrightarrow{f}_R B$ /
- (ii) двойственно /обозначается: $A \xrightarrow{d}_R B$ /
- (iii) сильно /обозначается: $A \xrightarrow{s}_R B$ /
- (iv) слабо /обозначается: $A \xrightarrow{w}_R B$ /

зависит от множества A в ре-ляции R , если соответственно выполняется следующее условие:

- (i) $(\forall g, h \in R) (g \upharpoonright_A = h \upharpoonright_A \Rightarrow g \upharpoonright_B = h \upharpoonright_B)$
- (ii) $(\forall g, h \in R) ((\exists a \in A) (g(a) = h(a)) \Rightarrow (\exists b \in B) (g(b) = h(b)))$
- (iii) $(\forall g, h \in R) ((\exists a \in A) (g(a) = h(a)) \Rightarrow g \upharpoonright_B = h \upharpoonright_B)$
- (iv) $(\forall g, h \in R) (g \upharpoonright_A = h \upharpoonright_A \Rightarrow (\exists b \in B) (g(b) = h(b)))$.

Обозначим зависимости типа $A \xrightarrow[R]{f} B$, $A \xrightarrow[R]{d} B$, $A \xrightarrow[R]{s} B$, $A \xrightarrow[R]{w} B$ соответственно через

$$F_R = \{(A, B) : A \xrightarrow[R]{f} B\}, \quad D_R = \{(A, B) : A \xrightarrow[R]{d} B\},$$

$$S_R = \{(A, B) : A \xrightarrow[R]{s} B\}, \quad W_R = \{(A, B) : A \xrightarrow[R]{w} B\}$$

и назовем каждое из этих множеств полным семейством данного типа относительно реляции R .

Для простоты введем следующие обозначения:

$$Y_R \in \{F_R, D_R, S_R, W_R\}.$$

Тогда

$$Y_R = \{(A, B) : A \xrightarrow[R]{y} B\}, \text{ где } y \in \{f, d, s, w\}.$$

Семейство Y_R назовем полным семейством типа y .

Разные зависимости до сих пор изучались во многих работах [4,10,12,13]. Основной целью этих работ являлась аксиоматизация полных семейств этих зависимостей [1,6] и изучение их свойств [3,4,5,9]. Аксиоматизация функциональных зависимостей велась в двух основных направлениях: отыскание протых аксиом, с которыми легко работать [1], а также разработка минимальных систем аксиом [6].

Армстронг в работе [1] аксиоматизировал полные семейства функционального типа и в работе [8] были даны некоторые системы аксиом для полных семейств двойственного и сильного типа. В дальнейшем мы приведем системы аксиом для полных семейств функционального, двойственного, слабого и сильного типа, а

также покажем, что эти новые аксиомы эквивалентны старым. Для полноты опишем старые аксиомы.

Пусть дано семейство подмножества Y , где $Y \subseteq P(\Omega) \times (\Omega)$ и $P(\Omega)$ семейство всех подмножеств над множеством Ω .

Будем говорить, что семейство Y удовлетворяет φ -аксиомам, тогда и только тогда, когда для каждого из множеств A, B, C, D ($A, B, C, D \subseteq \Omega$) выполняются следующие условия:

- (F1) $(A, A) \in Y$
- (F2) $(A, B) \in Y, (B, C) \in Y \Rightarrow (A, C) \in Y$
- (F3) $(A, B) \in Y, C \supseteq A, D \subseteq B \Rightarrow (C, D) \in Y$
- (F4) $(A, B) \in Y, (C, D) \in Y \Rightarrow (A \cup C, B \cup D) \in Y$

Будем говорить, что семейство Y удовлетворяет ϑ -аксиомам тогда и только тогда, когда для каждого из множеств A, B, C, D выполняются следующие условия:

- (D1) $(A, A) \in Y$.
- (D2) $(A, B) \in Y, (B, C) \in Y \Rightarrow (A, C) \in Y$
- (D3) $(A, B) \in Y, C \subseteq A, B \subseteq D \Rightarrow (C, D) \in Y$
- (D4) $(A, B) \in Y, (C, D) \in Y \Rightarrow (A \cup C, B \cup D) \in Y$
- (D5) $(A, \emptyset) \in Y \Rightarrow A = \emptyset$

Будем говорить, что семейство Y удовлетворяет γ -аксиомам, тогда и только тогда, когда для каждого из множеств A, B, C, D и для каждого $a \in \Omega$ выполняются следующие условия:

- (S1) $(\{a\}, \{a\}) \in Y$
- (S2) $(A, B) \in Y, (B, C) \in Y, B \neq \emptyset \Rightarrow (A, C) \in Y$
- (S3) $(A, B) \in Y, C \subseteq A, D \subseteq B \Rightarrow (C, D) \in Y$
- (S4) $(A, B) \in Y, (C, D) \in Y \Rightarrow (A \cap C, B \cup D) \in Y$
- (S5) $(A, B) \in Y, (C, D) \in Y \Rightarrow (A \cup C, B \cap D) \in Y$

φ - ϑ - и γ -аксиомы описывают свойства системы функциональных, двойственных и сильных зависимостей, независимые от специальной структуры реляций. Это означает, что во-первых для всякой реляции R над множеством Ω полное семейство функциональных зависимостей F_R удовлетворяет φ -аксиомам [1,6], полное семейство двойственных зависимостей D_R удовлетворяет ϑ -аксиомам [6] и полное семейство сильных зависимостей S_R удовлетворяет γ -аксиомам [8]; а во-вторых, если семейство F удовлетворяет φ -аксиомам, то существует такая реляция R , что $F = F_R$ [1,10]; если семейство D удовлетворяет ϑ -аксиомам, то существует такая реляция R , что $D = D_R$ [8] и если S удовлетворяет γ -аксиомам, то существует такая реляция R , что $S = S_R$ [8].

В дальнейшем нам понадобятся несколько технических лемм.

Лемма 1.1. Пусть семейство $F \subseteq P(\Omega) \times P(\Omega)$ такое, что если $(A, B) \in F$ и $B \neq \emptyset$, то $A \neq \emptyset$. Тогда семейство F удовлетворяет φ -аксиомам тогда и только тогда, когда семейство $D = \{(B, A) : (A, B) \in F\}$ удовлетворяет ϑ -аксиомам.

Доказательство. Утверждение леммы легко вытекает из φ и ϑ аксиом. Из-за аксиомы (D5) надо предполагать, что $((A, B) \in F, B \neq \emptyset \rightarrow A \neq \emptyset)$ условие выполняется. $\square$

Замечание. Пусть семейство F удовлетворяет φ -аксиомам. Обозначим через $F' = F \setminus \{(\emptyset, B) : B \neq \emptyset\}$. Тогда семейство F' удовлетворяет условиям леммы 1.1. и если $A \subseteq \Omega, A \neq \emptyset$, тогда $(\forall B \subseteq \Omega)((A, B) \in F \leftrightarrow (A, B) \in F')$. Заметим еще, что если семейство F удовлетворяет φ -аксиомам, то к F естественным образом можно поставить в соответствия некоторое семейство $D(F)$, для которого выполняются ϑ -аксиомы, таким образом, что $D(F) = \{(B, A) : (A, B) \in F \text{ \& } A \neq \emptyset \rightarrow B = \emptyset\}$. Естественно, что при этом "одно и то же D может соответствовать более чем одному семейству F ".

В дальнейшем вместо системы аксиом φ , ϑ и γ мы определим новые системы аксиом и дадим систему аксиом для полных семейств слабого типа, которые не содержат зависимости, где первое подмножество B пустое т.е. $(\emptyset \xrightarrow[R]{W} B)$.

F - аксиома:

Семейство $F \subseteq P(\Omega) \times P(\Omega)$ удовлетворяет F-аксиомам тогда и только тогда, когда для каждого $(X, Y) \in P(\Omega) \times P(\Omega) \setminus F$ существует такое множество $E (E \subseteq \Omega)$, что выполняются следующие условия:

- (i) $X \subseteq E$ и $Y \not\subseteq E$
- (ii) если $(A, B) \in F$ и $A \subseteq E$ то $B \subseteq E$.

D - аксиома.

Семейство $D \subseteq P(\Omega) \times P(\Omega)$ удовлетворяет D-аксиомам тогда и только тогда, когда если для каждого $(X, Y) \in P(\Omega) \times P(\Omega) \setminus D$ существует такое множество $E (E \subseteq \Omega)$, что выполняются следующие условия:

- (i) $X \cap E \neq \emptyset$ и $Y \cap E = \emptyset$
- (ii) если $(A, B) \in D$ и $A \cap E \neq \emptyset$, то $B \cap E \neq \emptyset$

S - аксиома

Семейство $S \subseteq P(\Omega) \times P(\Omega)$ удовлетворяет S-аксиомам тогда и только тогда, когда если для каждого $(X, Y) \in P(\Omega) \times P(\Omega) \setminus S$ существует такое множество $E (E \subseteq \Omega)$, что выполняются следующие условия:

- (i) $X \cap E \neq \emptyset$ и $Y \not\subseteq E$
- (ii) если $(A, B) \in S$ и $A \cap E \neq \emptyset$, то $B \subseteq E$.

W - аксиома.

Семейство $W \subseteq P(\Omega) \times P(\Omega)$ удовлетворяет W-аксиомам тогда и только тогда, когда если для каждого $(X, Y) \in P(\Omega) \times P(\Omega) \setminus W$ существует такое множество $E (E \subseteq \Omega)$, что выполняются следующие условия:

$$(i) \quad X \subseteq E \text{ и } Y \cap E = \emptyset$$

$$(ii) \text{ если } (A, B) \in W \text{ и } A \subseteq E, \text{ то } B \cap E \neq \emptyset.$$

Покажем, что новые аксиомы эквивалентны старым.

Теорема 1.1.

(i) Пусть дано семейство F . F удовлетворяет φ -аксиомам тогда и только тогда, когда F удовлетворяет F-аксиомам.

(ii) Пусть дано семейство D . D удовлетворяет ϑ аксиомам тогда и только тогда, когда D удовлетворяет D-аксиомам.

(iii) Пусть дано семейство S . S удовлетворяет γ -аксиомам тогда и только тогда, когда S удовлетворяет S-аксиомам.

Доказательство

(i) Легко проверить что, если семейство F удовлетворяет F-аксиомам, тогда F удовлетворяет φ -аксиомам тоже. Например, покажем, что если семейство F удовлетворяет F-аксиомам, то аксиома (F2) выполняется. Доказательство проведем от противного. Пусть $(A, B) \in F$, $(B, C) \in F$ и $(A, C) \notin F$. Тогда исходя из того, что семейство F удовлетворяет F-аксиоме, существует такое множество $E (E \subseteq \Omega)$, что $A \subseteq E$, $C \not\subseteq E$. А из того, что $(A, B) \in F$ и $(B, C) \in F$ следует, что

$$((A, B) \in F, A \subseteq E \rightarrow B \subseteq E), (B, C) \in F, B \subseteq E \rightarrow C \subseteq E).$$

А это противоречие.

Предположим, что семейство F удовлетворяет φ -аксиомам и пусть $(x, y) \in P(\Omega) \times P(\Omega) \setminus F$. Покажем, что тогда

- (1) существует такое множество $E \supseteq X$, что
 $(E, Y) \in P(\Omega) \times P(\Omega) \setminus F$ и если $E' \supset E$, то $(E', Y) \in F$.

Действительно, согласно аксиомы (F1) $(\Omega, \Omega) \in F$, поэтому из аксиомы (F3) следует, что $(\Omega, Y) \in F$. Таким образом $(x, y) \in P(\Omega) \times P(\Omega) \setminus F$ и $(\Omega, Y) \in F$ т.е. существует такое максимальное $E \supseteq X$, для которого $(E, Y) \in F$. Тем самым утверждение (1) доказано.

Пусть F удовлетворяет φ -аксиомам и пусть $(x, y) \in P(\Omega) \times P(\Omega) \setminus F$. Мы утверждаем, что множество E ($E \supseteq X$) удовлетворяет F -аксиоме из-за утверждения (1). Если $Y \subseteq E$, то из аксиом (F1) и (F3) следует, что $(E, Y) \in F$, в то время, как согласно утверждению (1) $(E, Y) \notin F$; таким образом $Y \not\subseteq E$.

Пусть $(A, B) \in F$. Предположим, что $A \subseteq E$ и $B \not\subseteq E$. Тогда $E' = E \cup B \supset E$ и значит $(E', Y) \in F$. Согласно аксиоме (F1) $(E, E) \in F$, поэтому из аксиомы (F4) следует, что $(E, E') = (E \cup A, E \cup B) \in F$. Но тогда, согласно аксиоме (F2) $((E, E') \in F, (E', Y) \in F \rightarrow (E, Y) \in F)$, что является противоречием. Таким образом $B \subseteq E$ и тем самым доказательство (i) закончено.

(ii) Пусть $F = \{(A, B) : (B, A) \in D\}$. Согласно лемме 1.1. семейство F удовлетворяет φ -аксиомам тогда и только тогда, когда семейство D удовлетворяет ϑ -аксиомам. Вследствие пункта (i) достаточно доказать, что семейство F удовлетворяет F -аксиоме тогда и только тогда, когда семейство D удовлетворяет D -аксиоме. Предположим, что для семейства F справедлива F -аксиома. Согласно F -аксиоме, для $(x, y) \in P(\Omega) \times P(\Omega) \setminus F$

существует такое множество E ($E \subseteq \Omega$), что $X \subseteq E$, $Y \subseteq E$ и если $(A, B) \in F$ и $A \subseteq E$, то $B \subseteq E$. В дальнейшем обозначим через $E(X, Y)$, выше определенное множество E , т.е. $E(X, Y) = E$. Тогда для $(Y, X) \in P(\Omega) \times P(\Omega) \setminus D$ в множестве $(\Omega \setminus E(X, Y))$ такое будет для которого выполняется D -аксиома. Поскольку $(X, Y) \in P(\Omega) \times P(\Omega) \setminus F \Leftrightarrow (Y, X) \in P(\Omega) \times P(\Omega) \setminus D$, то из этого следует, что D удовлетворяет D -аксиоме.

Аналогично доказывается, что если семейство D удовлетворяет D -аксиоме, то семейство F удовлетворяет F -аксиоме. Таким образом, пункт (ii) полностью доказан.

(iii) Предположим, что семейство S удовлетворяет S -аксиоме.

Нетрудно проследить, что тогда S удовлетворяет γ -аксиоме, аналогично тому, как это было сделано для пункта (i). Остановимся подробнее на доказательстве аксиомы (S_1). Предположим, что существует такое $a(a \in \Omega)$, для которого $(\{a\}, \{a\}) \notin S$. Тогда согласно S -аксиоме существует такое $E(E \subseteq \Omega)$, что $\{a\} \cap E \neq \emptyset$ и $(\Omega \setminus E) \cap \{a\} \neq \emptyset$. А это противоречит тому, что $|\{a\}| = 1$. Следовательно, $(\{a\}, \{a\}) \in S$.

Предположим теперь, что семейство S удовлетворяет γ -аксиомам и пусть $(X, Y) \in P(\Omega) \times P(\Omega) \setminus S$. Покажем, что тогда (2) существуют $a(a \in X)$ и множество $E(E \subseteq \Omega)$ такие, что $a \in E$, $(\{a\}, E) \in S$ и если $E' \supset E$, то $(\{a\}, E') \in P(\Omega) \times P(\Omega) \setminus S$.

Действительно, если для любого $a \in X$ справедливо, что $(\{a\}, Y) \in S$, то последовательное применение аксиомы (S_5) дает нам, что $(X, Y) \in S$. Таким образом существует такое $a \in X$, что $(\{a\}, Y) \in P(\Omega) \times P(\Omega) \setminus S$. Если для всякого $b(b \in Y)$ выполняется, что $(\{a\}, \{b\}) \in S$, то последовательно применяя аксиомы (S_4) получаем, что $(\{a\}, Y) \in S$; то есть существует такое $b \in Y$, что $(\{a\}, \{b\}) \in P(\Omega) \times P(\Omega) \setminus S$. Из ак-

сиом (S1) и (S2) следует, что существует такое множество E , что $a \in E$ и $(\{a\}, E) \in S$, причем E является максимальным множеством обладающим этим свойством. Тем самым утверждение (2) доказано.

Пусть семейство S удовлетворяет γ -аксиомам и пусть $(X, Y) \in P(\Omega) \times P(\Omega) \setminus S$. Мы утверждаем, что тогда семейство S удовлетворяет S -аксиоме и для множества E , определенного в утверждении (2), выполняются условия (i) и (ii) S -аксиомы.

Из $a \in X \cap E$ следует, что $A \cap E \neq \emptyset$. Кроме того из $(\{a\}, E) \in S$, $(\{a\}, \{b\}) \notin S$ и аксиомы (S3) вытекает, что $Y \not\subseteq E$, так как $b \in Y \setminus E$.

Пусть $(A, B) \in S$ и предположим, что $A \cap E \neq \emptyset$ и $B \not\subseteq E$. Пусть $c \in A \cap E$ и $d \in B \cap \Omega \setminus E$. Из аксиомы (S3) следует, что $(\{c\}, \{d\}) \in S$, а из аксиомы (S1) вытекает, что $(\{c\}, \{c\}) \in S$. Согласно аксиоме (S5) и $(\{a\}, E) \in S$ имеем, что $(\{a, c\}, \{c\}) \in S$, а используя аксиомы (S3) получаем, что $(\{a\}, \{c\}) \in S$. Из $(\{a\}, \{c\}) \in S$, $(\{c\}, \{d\}) \in S$ и из аксиомы (S2) следует, что $(\{a\}, \{d\}) \in S$. И наконец на основе (S4) имеем: $(\{a\} \cup \{a\}, E \cup \{d\}) = (\{a\}, E') \in S$, что является противоречием, так как $E' \supset E$. Таким образом $(A, B) \in S$, $A \cap E \neq \emptyset$ следует, что $B \subseteq E$. Тем самым утверждение (iii) и теорема 1.1. доказаны.

§ 2. Равенственное множество

Пусть дана реляция R над множеством Ω .

Определение 2.1. Определим равенственное множество ϵ_R реляции R следующим образом:

Пусть $h, g \in R$, введем обозначение

$$E(h, g) = \{a \in \Omega : h(a) = g(a)\} ;$$

Тогда

$$\epsilon_R = \{E(h, g) : h, g \in R, h \neq g\} .$$

Определение 2.2. Пусть дано семейство подмножеств A над множеством Ω . Назовем A Δ -системой, если для любых множеств A, B, C, D ($A, B, C, D \subseteq \Omega$) таких что $A \neq B$ и $C \neq D$, выполняется соотношение $A \cap B = C \cap D$.

Замечание: легко заметить, что семейство множеств A является Δ -системой тогда и только тогда, когда для любых A, B ($A \neq B, A, B \in A$) $A \cap B = \cap A$.

Теорема 2.1. Пусть дана реляция R над множеством Ω .

- (i) если $h, f, g \in R$, то $\{E(h, g), E(h, f), E(g, f)\}$ является Δ -системой
- (ii) если $\epsilon = \{E_{ij} : 1 \leq i < j \leq k\}$ такое, что для каждого i, j, ℓ ($1 \leq i < j < \ell \leq k$) семейство $\{E_{ij}, E_{i\ell}, E_{j\ell}\}$ является Δ -системой, то существует такая реляция R для которой $\epsilon = \epsilon_R$.

Доказательство.

- (i) Из-за симметрии достаточно показать, что $a \in E(h, g) \cap E(g, f) \rightarrow a \in E(h, f)$. Действительно,

$$a \in E(h, g) \cap E(g, f) \longleftrightarrow h(a) = g(a) = f(a) \rightarrow \\ \rightarrow h(a) = f(a) \longleftrightarrow a \in E(h, f).$$

(ii) Для доказательства достаточно сконструировать по данному семейству ϵ такую реляцию R , что $\epsilon = \epsilon_R$, причем реляция R должна содержать ровно k элементов, т.е.

$$R = \{h_1, h_2, \dots, h_k\}.$$

Конструкцию проведем индуктивно. Пусть дан произвольный элемент $h_1 \in R$, и предположим, что мы уже определили элементы $h_1, h_2, \dots, h_n$ так, что если $1 \leq i < j \leq n$, то $E(h_i, h_j) = E_{ij}$, где $n < k$.

$$\text{Пусть } h_{n+1}(a) = \begin{cases} h_i(a), & \text{если существует такое } i (1 \leq i \leq n), \text{ что } a \in E_{i, n+1} \\ \max(h_i(b) : b \in \Omega \quad i \leq 1 \leq n) + 1 & \text{иначе.} \end{cases}$$

Тогда

(a) h_{n+1} является хорошо определенным элементом. Мы покажем, что если $a \in E_{i, n+1} \cap E_{j, n+1}$ ($i, j \leq n$), то $h_i(a) = h_j(a)$. Это действительно так, поскольку семейство множеств $\{E_{i, j}, E_{i, n+1}, E_{j, n+1}\}$ по предположению является Δ -системой. Таким образом

$$a \in E_{i, n+1} \cap E_{j, n+1} \rightarrow a \in E_{i, j} = E(h_j, h_j)$$

по индуктивному предположению.

(b) если $1 \leq i \leq n$ и $a \notin E_{i, n+1}$, то $h_i(a) \neq h_{n+1}(a)$.

Если

$$a \notin \bigcup_{j=1}^n E_{j, n+1}, \text{ то } h_i(a) \neq h_{n+1}(a)$$

по определению h_{n+1} . А если существует такое j ($1 \leq j \leq n$), что $a \in E_{j, n+1}$, тогда по пункту (a) $h_j(a) = h_{n+1}(a)$.

Элемент $a \notin E_{ij}$, потому что $\{E_{i, j}, E_{i, n+1}, E_{j, n+1}\}$ является Δ -системой, по этому по индуктивному предположению

$h_i(a) \neq h_j(a)$, т.е. $h_i(a) \neq h_{n+1}(a)$. Из-за пунктов (а) и (b) $E(h_i, h_{n+1}) = E_{i, n+1}$ ($1 \leq i \leq n$), то есть индукционный шаг работает.

Пусть $R = \{h_1, h_2, \dots, h_k\}$. Тогда легко заметить, что $\epsilon_R = \epsilon$. тем самым теорема 2.1. доказана.

По теореме 2.1. можно естественным образом аксиоматизировать полные семейства типа f-d-s-w.

Пусть даны семейства F, D, S и W над множеством Ω .

F' - аксиома

Семейство F удовлетворяет F'-аксиоме тогда и только тогда, когда существует такое k и множество $\{E_{i,j} : 1 \leq i < j \leq k, E_{i,j} \subseteq \Omega\}$, что выполняются следующие условия:

- (i) если $(X, Y) \in P(\Omega) \times P(\Omega) \setminus F$, то существуют такие i, j ($1 \leq i < j \leq k$), что $X \subseteq E_{i,j}$ и $Y \not\subseteq E_{i,j}$;
- (ii) если $(A, B) \in F$ и $A \subseteq E_{i,j}$, то $B \subseteq E_{i,j}$;
- (iii) для любых i, j, l ($1 \leq i < l \leq k$) семейство $\{E_{i,j}, E_{i,l}, E_{j,l}\}$ является Δ -системой.

D' - аксиома

Семейство D удовлетворяет D'-аксиоме тогда и только тогда, когда существует такое k и множество $\{E_{i,j} : 1 \leq i < j \leq k, E_{i,j} \subseteq \Omega\}$, что выполняются следующие условия:

- (i) если $(X, Y) \in P(\Omega) \times P(\Omega) \setminus D$, то существуют такие i, j ($1 \leq i < j \leq k$), что $X \cap E_{i,j} \neq \emptyset$ и $Y \cap E_{i,j} = \emptyset$;

(ii) Если $(A, B) \in D$ и $A \cap E_{i,j} \neq \emptyset$, то $B \cap E_{i,j} \neq \emptyset$;

(iii) То же условие, что и в пункте (iii) F' -аксиомы.

S' - аксиома.

Семейство S удовлетворяет S' -аксиоме тогда и только тогда, когда существует такое k и множество

$\{E_{i,j} : 1 \leq i < j \leq k, E_{i,j} \subseteq \Omega\}$, что выполняются следующие условия:

(i) Если $(x, y) \in P(\Omega) \times P(\Omega) \setminus S$, то существуют такие i, j ($1 \leq i < j \leq k$), что $x \cap E_{i,j} \neq \emptyset$ и $y \cap E_{i,j} = \emptyset$;

(ii) Если $(A, B) \in S$ и $A \cap E_{i,j} \neq \emptyset$, то $B \subseteq E_{i,j}$;

(iii) То же условие, что и в пункте (iii) F' -аксиомы.

W' - аксиома

Семейство W удовлетворяет W' -аксиоме тогда и только тогда, когда существует такое k и $\{E_{i,j} : 1 \leq i < j \leq k, E_{i,j} \subseteq \Omega\}$, что выполняются следующие условия:

(i) Если $(x, y) \in P(\Omega) \times P(\Omega) \setminus W$, то существуют такие i, j ($1 \leq i < j \leq k$), что $x \subseteq E_{i,j}$ и $y \cap E_{i,j} = \emptyset$;

(ii) Если $(A, B) \in W$ и $A \subseteq E_{i,j}$, то $B \cap E_{i,j} \neq \emptyset$;

(iii) То же условие, что и в пункте (iii) F' -аксиомы .

Замечание. Множество $E_{i,j}$ определенное в F' -аксиоме является максимальным элементом [1], т.е. если $(A, B) \in F$ и $A \subseteq E_{i,j}$, то $B \subseteq E_{i,j}$. Аналогичным образом можно определить максимальные элементы для полных семейств типа $d-s-w$, и легко получаются результаты, аналогичные результатам Армстронга [1].

Теорема 2.2.

(i) Пусть $Y \subseteq P(\Omega) \times P(\Omega)$ и $Y \in \{F, D, S\}$. Семейство Y удовлетворяет Y -аксиоме тогда и только тогда, когда удовлетворяет и Y' -аксиоме.

(ii) Для любого множества Ω ($|\Omega| \geq 3$) существует семейство W такое, что W удовлетворяет W -аксиоме и не удовлетворяет W' -аксиоме.

Доказательство.

(i) Пусть $Y = F, Y' = F$ и предположим, что F удовлетворяет F -аксиоме. Покажем, что F удовлетворяет и F' -аксиоме. Действительно, пусть $(X, Y) \in P(\Omega) \times P(\Omega) \setminus F$ тогда $E(X, Y) \subseteq \Omega$ то множество, которое гарантирует F -аксиома. Далее, пусть $\{E_2, E_3, \dots, E_k\} = \{E(X, Y) : (X, Y) \in P(\Omega) \times P(\Omega) \setminus F\}$.

Пусть $E_{1,j} = E_j$, если $1 < j \leq k$, и $E_{i,j} = E_i \cap E_j$, если $1 < i < j \leq k$. Мы утверждаем, что тогда семейство множества $\{E_{i,j} : 1 \leq i < j \leq k\}$ показывает, что семейство F удовлетворяет F' -аксиоме. Очевидно, что пункт (i) F' -аксиомы выполняется.

Пусть $(A, B) \in F$ и предположим, что $A \subseteq E_{i,j}$ ($1 \leq i < j \leq k$), т.е. $A \subseteq E_j$, если $i = 1$ и $A \subseteq E_i \cap E_j$, если $i > 1$. По пункту (ii) F -аксиомы и определению множества $E_2, E_3, \dots, E_k$ $B \subseteq E_{i,j}$. Это означает, что пункт (ii) F' -аксиомы выполняется.

Пусть $1 \leq i < j < \ell \leq k$, тогда рассмотрим два случая:

(a) Пусть $i = 1$. Тогда $E_{i,j} = E_j$, $E_{i,\ell} = E_\ell$ и $E_{j,\ell} = E_j \cap E_\ell$ поэтому пересечение любых двух множеств семейства $\{E_{i,j}, E_{i,\ell}, E_{j,\ell}\}$ равно $E_j \cap E_\ell$. Это означает, что $\{E_{i,j}, E_{i,\ell}, E_{j,\ell}\}$ является Δ -системой.

(в) Пусть $i > 1$. Тогда $E_{i,j} = E_i \cap E_j$, $E_{i,\ell} = E_i \cap E_\ell$, $E_{j,\ell} = E_j \cap E_\ell$ и поэтому пересечение любых двух элементов множеств $\{E_{i,j}, E_{i,\ell}, E_{j,\ell}\}$ равно $E_i \cap E_j \cap E_\ell$. Отсюда вытекает, что семейство $\{E_{i,j}, E_{i,\ell}, E_{j,\ell}\}$ является Δ -системой.

Если семейство F удовлетворяет F' -аксиоме, то очевидно, что оно удовлетворяет и F -аксиоме. Тем самым случай $Y = F$ полностью доказаны.

Пусть $Y = D$ и $Y = D'$, и предположим, что семейство D удовлетворяет D -аксиоме. Покажем, что тогда D удовлетворяет и D' -аксиоме.

Действительно, пусть $(X, Y) \in P(\Omega) \times P(\Omega) \setminus D$ тогда $E(X, Y) \subseteq \Omega$, то множество, которое гарантирует D -аксиома. Далее, пусть $\{E_1, E_2, \dots, E_k\} = \{E(X, Y) : (X, Y) \in P(\Omega) \times P(\Omega) \setminus D\}$. Если $1 \leq i \leq k$, то $E_{2i-1, 2i} = E_i$, а если $1 \leq i < j \leq 2k$ и множество $E_{i,j}$ не определено, то $E_{i,j} = \emptyset$.

Мы утверждаем, что семейство множеств $\{E_{i,j} : 1 \leq i < j \leq 2k\}$ показывает, что семейство D удовлетворяет D' -аксиоме. Действительно, легко заметить, что пункт (i) D' -аксиомы выполняется. Заметим, что если множество $E_{i,j}$ не совпадает ни с одним из множеств $E(X, Y)$, то $E_{i,j} = \emptyset$ и поэтому $((A, B) \in D \& A \cap E_{i,j} \neq \emptyset \rightarrow B \cap E_{i,j} \neq \emptyset)$. Отсюда вытекает, что пункт (ii) D' -аксиомы также выполняется. А множество семейств $\{E_{i,j}, E_{i,\ell}, E_{j,\ell}\}$ ($1 \leq i < j < \ell \leq k$) является Δ -системой, потому что по крайней мере два из этих элементов пустые.

Легко заметить, что если семейство D удовлетворяет D' -аксиоме, то оно удовлетворяет и D -аксиоме.

В случае $Y = S$ доказательство проводится аналогично случаю $Y = F$.

(ii) Для простоты будем считать, что $\Omega = \{a, b, c\}$. (В общем случае считается, что $a, b \in \Omega$ и элемент $\{c\}$ надо заменить всеми элементами $\Omega \setminus \{a, b\}$.)

Пусть $W = \{(A, B) : A \subseteq \{a\} \rightarrow a \in B \text{ и } A \subseteq \{b\} \rightarrow b \in B\}$. Заметим, что тогда семейство W удовлетворяет D-аксиоме. Действительно, если $(X, Y) \in P(\Omega) \times P(\Omega) \setminus W$, то либо $X \subseteq \{a\}$ и $a \notin Y$, либо $X \subseteq \{b\}$ и $b \notin Y$. В первом случае множество $E = \{a\}$, а во втором случае множество $E = \{b\}$ показывает, что W-аксиома выполняется.

Мы утверждаем, что для только что определенного семейства W W' -аксиома не выполняется. Доказательство проведем от противного. Предположим, что семейство множеств $\epsilon = \{E_{i,j} : 1 \leq i < j \leq k\}$ удовлетворяет W' -аксиоме для семейства W . Тогда

(1) $\{a\} \in \epsilon$ и $\{b\} \in \epsilon$,

так как $(\{a\}, \Omega \setminus \{a\}) \notin W$ и $(\{b\}, \Omega \setminus \{b\}) \notin W$;

(2) $\emptyset \notin \epsilon$ и $\{c\} \notin \epsilon$,

так как $(\emptyset, \Omega) \in W$ и $(\{c\}, \Omega \setminus \{c\}) \in W$.

Рассмотрим два случая:

(а) Пусть $\{a\} = E_{i,j}$ и $\{b\} = E_{i,j}$. Тогда, поскольку $\{E_{i,j}, E_{i,\ell}, E_{j,\ell}\}$ является Δ -системой, то либо $E_{j,\ell} = \emptyset$, либо $E_{j,\ell} = \{c\}$, а это противоречит тому, что $\{c\}, \emptyset \notin \epsilon$ по (2).

(в) Пусть $\{a\} = E_{i,j}$, $\{b\} = E_{\ell,m}$ и $| \{i, j, \ell, m\} | = 4$.

В дальнейшем можно предположить, что $(i, j) = (1, 2)$ и $(\ell, m) = (3, 4)$, потому что мы изучаем случай, когда

$$\epsilon = \{E_{i,j}, E_{i,\ell}, E_{i,m}, E_{j,\ell}, E_{j,m}, E_{\ell,m}\}.$$

/Действительно, семейство ϵ может содержать и другие элементы тоже./

Исследуем множество $E_{1,3}$.

Если $E_{1,3} = \{a\}$ или $E_{1,3} = \{b\}$, то мы приходим к случаю (а). Разумеется, $E_{1,3} \neq \{c\}$ и $E_{1,3} \neq \emptyset$ из-за (2). Если $E_{1,3} = \{b, c\}$, то $E_{1,2} \cap E_{1,3} = \emptyset$ и, поскольку $\{E_{1,2}, E_{1,3}, E_{2,3}\}$ является Δ -системой, то $E_{2,3} = \emptyset$, что противоречит (2). Следовательно, $a \in E_{1,3}$. Значит $a \in E_{1,3} \cap E_{1,2}$, а поскольку $\{E_{1,2}, E_{1,3}, E_{2,3}\}$ — Δ -система, то $a \in E_{2,3}$.

С другой стороны $a \notin E_{3,4}$ и значит $a \notin E_{3,4} \cap E_{2,3}$. Поскольку $\{E_{2,3}, E_{3,4}, E_{2,4}\}$ является Δ -системой и $a \in E_{2,3}$, то $a \notin E_{2,4}$. Отсюда следует, что $E_{2,4} \subseteq \{b, c\}$. В то же время $E_{2,4} \neq \{b\}$, так как это приводит к случаю (а). Кроме того, $E_{2,4} \neq \emptyset$ и $E_{2,4} \neq \{c\}$ из-за (2). Следовательно, $E_{2,4} = \{b, c\}$. Отсюда получаем, что $b \in E_{2,4} \cap E_{3,4}$ и значит $b \in E_{2,3}$. Таким образом $\{a, b\} \subseteq E_{2,3}$. Поскольку $b \notin E_{1,2}$, то отсюда следует, что $b \notin E_{1,3}$. Сравнивая этот результат с полученными выше ограничениями, получаем, что $E_{1,3} = \{a, c\}$. Но тогда $E_{1,3} \cap E_{3,4} = \emptyset$ и $E_{1,3} \cup E_{3,4} = \Omega$, то есть $E_{1,4} = \emptyset$, что удовлетворяет утверждению (2).

Замечание: Теорема 2.2. указывает на существующую разницу между слабыми и всеми другими зависимостями. Дело в том, что если первое множество некоторой зависимости пустое, то это представляет проблему только для слабых зависимостей.

Теорема 2.3. Пусть дано семейство $Y \subseteq P(\Omega) \times P(\Omega)$ и $y \in \{F, D, S, W\}$. Если семейство Y удовлетворяет Y' -аксиоме, то существует реляция R над множеством Ω такая, что $Y = Y_R$. Наоборот: если дана реляция R над множеством Ω , то семейство Y_R удовлетворяет Y' -аксиоме; где $y' \in \{F', D', S', W'\}$.

Доказательство. Покажем, что семейство Y удовлетворяет F' -аксиоме, а это проверяется на множестве $\epsilon = \{E_{i,j} : 1 \leq i < j \leq k\}$. Тогда из-за пункта (iii) Y' -аксиомы и из-за пункта (ii) теоремы 2.1. существует такая реляция R над Ω , что $Y = Y_R$.

Естественно, что из-за пунктов (i) и (ii) Y' -аксиомы $Y = Y_R$.

Наоборот, если дана реляция R над множеством Ω , то семейство множеств $\epsilon_R = \{E_{i,j} : 1 \leq i < j \leq k\}$ указывает на то, что семейство Y_R действительно удовлетворяет Y' -аксиоме, где $R = \{h_1, h_2, \dots, h_k\}$ и $E_{i,j} = E(h_i, h_j)$. Тем самым теорема доказана.

Л И Т Е Р А Т У Р А

- [1] Armstrong, W.W., Dependency structures of data base relationships, Information Processing 74, North-Holland Publ.Co., (1974) 580-583.
- [2] Armstrong, W.W., On the generation of dependency structures of relational data bases, Publication #272, Université de Montréal. (1977).
- [3] Beeri, C. - Fagin, R. - Howard, J.H., A complete axiomatization for functional and multivalued dependencies in database relations, Proc. ACM SIGMOD Int. Conf. on Management of Data, Toronto (1977) 47-61.
- [4] Békéssy, A. - Demetrovics, J., Contribution to the theory of data base relations, Discrete Math. 27(1979) 1-10.
- [5] Békéssy, A. - Demetrovics, J. - Hannák, L. - Frankl, P. - Katona, G., On the number of maximal dependencies in a data base relation of fixed order, Discrete Math. 29 (1980).
- [6] Codd, E.F., A relational model of data for large shared data banks, Comm. ACM. 13(1970) 377-387.
- [7] Codd, E.F., Further normalization of the data base relational model, Courant Computer Science Symposia 6 Data Base System, Prentice Hall, Englewood Cliffs, N.J. (1971) 33-64.
- [8] Czédli, G., Függőségek relációs adatbázis modellben, Alk. Mat. Lapok (1980)
- [9] Demetrovics, J., Candidate keys and antichains, SIAM J. on Algebraic and Discrete Methods 1 (1980) 1, 92.
- [10] Demetrovics, J., On the equivalence of candidate keys with Sperner systems, Acta Cybernetica 4 (1979) 247-252
- [11] Fagin, R., Multivalued dependencies and a new normal form for relational databases, ACM Trans. on Database System 2:3 (Sept. 1977), 262-278.
- [12] Rissanen, J., Theory of relations for databases - a tutorial survey, Proc. 7th Symp. on Foundations of Computer Science, Poland 1978, Lecture Notes in Computer Science 64, Springer Verlag, pp. 537-551.
- [13] Sagiv, Y., An algorithm for inferring multivalued dependencies that works also for a subset of propositional logic, Department of Computer Science, University of Illinois at Urban-Champaign, 1979.

ÖSSZEFOGLALÁS

A relációs adatmodell funkcionális függőségeinek általánosítása

Demetrovics J. - Gyepesi Gy.

Ebben a cikkben relációs adatmodellek funkcionális függőségeivel és ennek lehetséges általánosításával; a duális, erős és gyenge függőségekkel foglalkozunk. Axiomatizáljuk ezen függésfajták teljes családjait /a funkcionális függőségek teljes családjaira [1]-ben adtak először axiómarendszert/.

Axiómáink mátrixok egyenlőségalmazának egy kombinatorikus jellemzésén alapulnak. Rámutatunk egy a gyenge függés és a többi között fennálló, fontos különbségre.

A generalisation of the functional dependencies in the relational data base

Demetrovics, J. - Gy. Gyepesi,

In this paper we deal with the functional dependencies of relational data models and their three generalizations: the dual, strong and weak dependencies. We axiomatize the full families of these dependencies of any kind /full families of functional dependencies were firstly axiomatized in [1] and that of dual and strong dependencies in [8] as well; the axiomatization of full families of weak dependencies is a new result/.

Our axioms are based on a combinatorial characterization of equality-sets of matrices. We prove an important difference between the weak dependencies and the rest.