АКСІОМАТИКА БАГАТОЗНАЧНИХ ЗАЛЕЖНОСТЕЙ ТАБЛИЧНИХ БАЗ ДАНИХ: ПОВНОТА ТА ЇЇ КРИТЕРІЙ

Д.Б. Буй, А.В. Пузікова
Київський національний університет імені Тараса Шевченка, Україна
buy@unicyb.kiev.ua, anna_inf@mail.ru

Аксіоматика багатозначних залежностей

Всі неозначувані тут поняття та позначення використовуються в розумінні монографії [1], зокрема, $s \mid X$ – обмеження рядка s за множиною атрибутів X.

Нехай t – таблиця, R – схема таблиці t (скінченна множина атрибутів), X, Y, W, Z – підмножини схеми R, s, s_1, s_2 – рядки таблиці t. В подальшому розгляді множина R зафіксована, крім того, фіксується універсальний домен D – множина, з якої атрибути приймають значення в інтерпретаціях.

Скажемо, що на таблиці t схеми R виконується багатозначна залежність (БЗЗ) (див., наприклад, [1]) $X \to\to Y$, якщо для двох довільних рядків s_1, s_2 таблиці t, які збігаються на множині атрибутів X, існує рядок $s_3 \in t$, який дорівнює об'єднанню обмежень рядків s_1, s_2 на множини атрибутів $X \bigcup Y$ і $R \setminus (X \bigcup Y)$ відповідно:

$$(X \to\to Y)(t) = true \overset{def}{\Leftrightarrow} \forall s_1, s_2 \in t(s_1 \mid X = s_2 \mid X \Rightarrow$$
$$\Rightarrow \exists s_3 \in t(s_3 = s_1 \mid (X \bigcup Y) \bigcup s_2 \mid R \setminus (X \bigcup Y))).$$

Отже, з семантичної точки зору БЗЗ – це предикат, заданий трьома (скінченними) множинами атрибутів X, Y і R.

Скажемо, що рядки s_1, s_2 таблиці t знаходяться у відношенні $=_X$, якщо вони збігаються на множині атрибутів X:

$$s_1 =_X s_2 \overset{def}{\Leftrightarrow} s_1 \mid X = s_2 \mid X.$$

Відношення $=_X$ розбиває множину рядків s таблиці t на класи еквівалентності, які мають наступне зображення:

$$[s]_{=_X} = \{s \mid X\} \otimes \pi_Y [s]_{=_X} \otimes \pi_{R \setminus (X \bigcup Y)} [s]_{=_X}.$$

Скажемо, що таблиця $t(R)$ є *моделлю* множини БЗЗ G, якщо кожна БЗЗ $X \rightarrow\rightarrow Y \in G$ виконується на таблиці $t(R)$:

$$t(R) \text{ модель } G \overset{def}{\Leftrightarrow}$$

$$\overset{def}{\Leftrightarrow} \forall (X \rightarrow\rightarrow Y)(X \rightarrow\rightarrow Y \in G \Rightarrow (X \rightarrow\rightarrow Y)(t) = true).$$

Вище запис $t(R)$ означає, що таблиця t має схему R.

Скажемо, що БЗЗ $X \rightarrow\rightarrow Y$ *семантично слідує* з множини БЗЗ G, якщо на кожній таблиці $t(R)$, яка є моделлю множини БЗЗ G, виконується також БЗЗ $X \rightarrow\rightarrow Y$:

$$G \models X \rightarrow\rightarrow Y \overset{def}{\Leftrightarrow} \forall t(R)(t \text{ модель } G \Rightarrow$$

$$\Rightarrow (X \rightarrow\rightarrow Y)(t) = true).$$

Лема 1 (*аксіома рефлексивності*). $\forall t(X \rightarrow\rightarrow Y)(t) = true$, де $Y \subseteq X$. □

Наслідок 1. $\varnothing \models X \rightarrow\rightarrow Y$, для $Y \subseteq X$. □

Лема 2. $\forall t(X \rightarrow\rightarrow Y)(t) = true$, де $X \bigcup Y = R$. □

Наслідок 2. $\varnothing \models X \rightarrow\rightarrow Y$, для $X \bigcup Y = R$. □

БЗЗ вигляду $X \rightarrow\rightarrow Y$, де $Y \subseteq X$ або $X \bigcup Y = R$, називається *тривіальною*.

Лема 3 (*правило повноти*).

$(X \rightarrow\rightarrow Y)(t) = true \Rightarrow (X \rightarrow\rightarrow R \setminus (X \bigcup Y))(t) = true$. □

Наслідок 3. $G \models X \rightarrow\rightarrow Y \Rightarrow G \models X \rightarrow\rightarrow R \setminus (X \bigcup Y)$. □

Лема 4 (правило *поповнення*). $(X \rightarrow\rightarrow Y)(t) = true \,\&\, Z \subseteq W \Rightarrow$

$\Rightarrow (X \bigcup W \rightarrow\rightarrow Y \bigcup Z)(t) = true$. □

Наслідок 4.

$G \models X \rightarrow\rightarrow Y \,\&\, Z \subseteq W \Rightarrow G \models X \bigcup W \rightarrow\rightarrow Y \bigcup Z$. □

Лема 5 (правило *транзитивності*). $(X \rightarrow\rightarrow Y)(t) = true \,\&$

$\&\, (Y \rightarrow\rightarrow Z)(t) = true \Rightarrow (X \rightarrow\rightarrow Z \setminus Y)(t) = true$. □

Наслідок 5.

$G \models X \rightarrow\rightarrow Y \,\&\, G \models Y \rightarrow\rightarrow Z \Rightarrow G \models X \rightarrow\rightarrow Z \setminus Y$. □

Наслідок 6. $G \vDash X \to\to Y \ \& \ G \vDash Y \to\to Z \ \& Z \cap Y = \varnothing \Rightarrow$

$\Rightarrow G \vDash X \to\to Z$. □

Аксіома рефлексивності і правила повноти, поповнення та транзитивності наведені у [2]. Ці правила назвемо *правилами виведення для БЗЗ*, бо вони мають синтаксичну природу.

Скажемо, що БЗЗ $X \to\to Y$ *синтаксично слідує* з множини БЗЗ G відносно схеми R ($G \vdash_R X \to\to Y$), якщо існує скінченна послідовність БЗЗ $\varphi_1, \varphi_2, ..., \varphi_{m-1}, \varphi_m$, така, що $\varphi_m = X \to\to Y$ і для $\forall i = \overline{1, m-1}$ кожна φ_i є або аксіома рефлексивності, або належить G, або отримана за яким-небудь правилом виведення для БЗЗ з попередніх у цій послідовності БЗЗ φ_j, φ_k, $j, k < i$. Послідовність $\varphi_1, \varphi_2, ..., \varphi_{m-1}, \varphi_m$ назвемо доведенням, наслідуючи традиції математичної логіки [3].

Нехай задана деяка множина БЗЗ G. *Замикання* $[G]_R$ – це множина БЗЗ, які синтаксично слідують з G відносно схеми R:

$$[G]_R \overset{def}{=} \{X \to\to Y \mid G \vdash_R X \to\to Y\}.$$

Лема 6. Виконуються властивості:

1) $G \subseteq [G]$ (*зростання*);

2) $[[G]] = [G]$ (*ідемпотентність*);

3) $G \subseteq H \Rightarrow [G] \subseteq [H]$ (*монотонність*).□

Таким чином, за термінологією [4] оператор $G \mapsto [G]$ є оператором замикання.

З описаних вище аксіоми і правил виведення можна отримати інші (похідні) правила виведення для БЗЗ.

Лема 7 (правило *псевдотранзитивності*).

$\{X \to\to Y, Y \cup W \to\to Z\} \vdash X \cup W \to\to Z \setminus (Y \cup W)$. □

Лема 8 (правила різниці).

1. $\{X \to\to Y\} \vdash X \to\to Y \setminus X$;
2. $\{X \to\to Y \setminus X\} \vdash X \to\to Y$;
3. $\{X \to\to Y\} \vdash X \to\to R \setminus Y$. □

Лема 9 (правило *адитивності*).

$\{X \to\to Y_1, X \to\to Y_2\} \vdash X \to\to Y_1 \cup Y_2$.□

Наслідок 7. Для $n = 2,3,...$ має місце твердження: $\{X \to\to Y_1,...,X \to\to Y_n\} \vdash X \to\to Y_1 \cup ... \cup Y_n$. □

Лема 10 (правила *декомпозиції*).

1. $\{X \to\to Y_1, X \to\to Y_2\} \vdash X \to\to Y_1 \cap Y_2$;
2. $\{X \to\to Y_1, X \to\to Y_2\} \vdash X \to\to Y_1 \setminus Y_2$. □

Наслідок 8. Для $n = 2,3,...$ має місце твердження: $\{X \to\to Y_1,...,X \to\to Y_n\} \vdash X \to\to Y_1 \cap ... \cap Y_n$. □

Правила псевдотранзитивності, адитивності та декомпозиції наведені у роботах [2, 5].

Аксіоматика багатозначних і функціональних залежностей

Нагадаймо, що на таблиці t виконується функціональна залежність (ФЗ) (див., наприклад, [1]) $X \to Y$, якщо для двох довільних рядків s_1, s_2 таблиці t, які збігаються на множині атрибутів X, має місце їх рівність і на множині атрибутів Y:

$$(X \to Y)(t) = true \overset{def}{\Leftrightarrow} \forall s_1, s_2 \in t\left(s_1|X = s_2|X \Rightarrow s_1|Y = s_2|Y\right)$$

.

Нехай задані множини ФЗ F і БЗЗ G. Скажемо, що таблиця $t(R)$ є *моделлю* множини $F \cup G$, якщо кожна залежність $\varphi \in F \cup G$ виконується на таблиці t:

$$t(R) \text{ модель } F \cup G \overset{def}{\Leftrightarrow} \forall \varphi(\varphi \in F \cup G \Rightarrow \varphi(t) = true).$$

ФЗ або БЗЗ φ *семантично слідує* з об'єднання множин ФЗ і БЗЗ $F \cup G$ відносно схеми R, якщо на кожній таблиці $t(R)$, яка є моделлю $F \cup G$, виконується також φ:

$$F \cup G \vDash \varphi \overset{def}{\Leftrightarrow} \forall t(R)(t(R) - \text{модель } F \cup G \Rightarrow \varphi(t) = true).$$

Лема 11. Нехай H_1 і H_2 – множини залежностей (ФЗ або БЗЗ), і нехай T_1 і T_2 множини їх відповідних моделей. Тоді виконується імплікація $H_1 \subseteq H_2 \Rightarrow T_1 \supseteq T_2$. □

Наслідок 9. Виконуються імплікації:

1. $F \vDash \varphi \Rightarrow F \cup G \vDash \varphi$;
2. $G \vDash \varphi \Rightarrow F \cup G \vDash \varphi$. □

Лема 12 (правило розширення ФЗ до БЗЗ).
$(X \to Y)(t) = true \Rightarrow (X \to\to Y)(t) = true$. □

Наслідок 10. $F \vDash X \to Y \Rightarrow F \vDash X \to\to Y$. □

Лема 13 (спільне правило для ФЗ та БЗЗ). $(X \to\to Z)(t) = true$ $\&$ $(Y \to Z')(t) = true$ $\&$ $Z' \subseteq Z$ $\&$ $Y \cap Z = \varnothing \Rightarrow$ $\Rightarrow (X \to Z')(t) = true$. □

Наслідок 11. $G \vDash X \to\to Z$ $\&$ $F \vDash Y \to Z'$ $\&$ $Z' \subseteq Z$ $\&$ $\&$ $Y \cap Z = \varnothing \Rightarrow F \cup G \vDash X \to Z'$. □

Розглянуті у лемах 12 і 13 правила наведені у [2].

Скажемо, що ФЗ або БЗЗ φ *синтаксично слідує* з об'єднання множин ФЗ і БЗЗ $F \cup G$ відносно схеми R ($F \cup G \vdash_R \varphi$), якщо існує скінченна послідовність ФЗ або БЗЗ $\varphi_1, \varphi_2, ..., \varphi_{m-1}, \varphi_m$, така, що $\varphi_m = \varphi$ і для $\forall i = \overline{1, m-1}$ кожна φ_i є або аксіома рефлексивності (для ФЗ або БЗЗ) , або належить $F \cup G$, або отримана за яким-небудь правилом виведення (повноти (для БЗЗ), поповнення (для ФЗ або БЗЗ), транзитивності (для ФЗ або БЗЗ), правила розширення ФЗ до БЗЗ або правила з леми 13) з попередніх у цій послідовності ФЗ або БЗЗ φ_j, φ_k, $j, k < i$. Як і раніше, послідовність $\varphi_1, \varphi_2, ..., \varphi_{m-1}, \varphi_m$ назвемо доведенням φ з об'єднання множин $F \cup G$.

Нехай задані деякі множини F і G.

Замикання $[F \cup G]_R$ – це множина усіх ФЗ і БЗЗ, які синтаксично слідують з $F \cup G$ відносно схеми R:

$$[F \cup G]_R \overset{def}{=} \{\varphi \mid F \cup G \vdash \varphi\}.$$

Лема 14. Виконуються властивості:

1) $F \cup G \subseteq [F \cup G]$ (*зростання*);
2) $[[F \cup G]] = [F \cup G]$ (*ідемпотентність*);
3) $F' \cup G' \subseteq F \cup G \Rightarrow [F' \cup G'] \subseteq [F \cup G]$ (*монотонність*);
4) $[F] \subseteq [F \cup G]$, $[G] \subseteq [F \cup G]$;
5) $[F] \cup [G] \subseteq [F \cup G]$.□

З властивостей 1-3 леми 14 за термінологією [4] оператор $F \bigcup G \mapsto [F \bigcup G]_R$ є оператором замикання.

В наступній лемі вкажемо ще одне спільне правило для БЗЗ і ФЗ.

Лема 15. $\{X \to\to Y, X \bigcup Y \to Z\} \vdash X \to Z \setminus Y$. □

Замиканням $[X]_{F \bigcup G, R}$ множини X (відносно множини $F \bigcup G$ і схеми R) називається сім'я усіх множин правих частин БЗЗ, які синтаксично слідують з множини $F \bigcup G$:

$$[X]_{F \bigcup G, R} \overset{def}{=} \{Y \mid X \to\to Y \in [F \bigcup G]_R\}.$$

Лема 16. Виконуються властивості:

1. $[X]_F \in [X]_{F \bigcup G, R}$;
2. $[X]_{F \bigcup G, R} = [[X]_F]_{F \bigcup G, R}$. □

Зауважимо, що оператор $X \mapsto [X]_{F \bigcup G, R}$ не є оператором замикання; для обґрунтування достатньо вказати, що даний оператор не володіє властивістю ідемпотентності (поняття множини атрибутів і замикання множини атрибутів відносно об'єднання множин ФЗ і БЗЗ $F \bigcup G$ і схеми R мають різну природу, тому вираз $[[X]_{F \bigcup G}]_{F \bigcup G, R}$ не має сенсу в означених термінах).

Базисом $[X]^{bas}_{F \bigcup G, R}$ множини X відносно множини $F \bigcup G$ і схеми R називається підсім'я замикання $[X]_{F \bigcup G, R}$, яка володіє властивостями:

1. $\forall W (W \in [X]^{bas}_{F \bigcup G, R} \Rightarrow W \neq \varnothing)$;
2. $\forall W_i, W_j (W_i, W_j \in [X]^{bas}_{F \bigcup G, R} \Rightarrow W_i \bigcap W_j = \varnothing)$;
3. $\forall Y (Y \in [X]_{F \bigcup G, R} \Rightarrow \exists i, 1 \leq i \leq n (Y = \bigcup_{W_i \in [X]^{bas}_{F \bigcup G, R}} W_i))$

Лема 17. Виконуються властивості.

1. $\bigcup_{W \in [X]^{bas}_{F \bigcup G, R}} W = R$, для $X \subseteq R$;
2. $A \in [X]_F \Rightarrow A \in [X]^{bas}_{F \bigcup G, R}$. □

Нехай φ – ФЗ або БЗЗ.

Твердження 1 (коректність аксіоматики БЗЗ і ФЗ). Якщо φ синтаксично виводиться з множини $F \bigcup G$, то φ виводиться з $F \bigcup G$ семантично:

$$F \bigcup G \vdash \varphi \Rightarrow F \bigcup G \vDash \varphi . \square$$

Твердження 2 (повнота аксіоматики БЗЗ і ФЗ). Якщо φ семантично виводиться з множини $F \bigcup G$, то φ виводиться з $F \bigcup G$ синтаксично:

$$F \bigcup G \vDash \varphi \Rightarrow F \bigcup G \vdash \varphi . \square$$

Твердження про повноту аксіоматики наводиться в [2].

Теорема 1. Відношення семантичного та синтаксичного слідування для аксіоматики БЗЗ та ФЗ збігаються:

$$F \bigcup G \vDash \varphi \Leftrightarrow F \bigcup G \vdash \varphi . \square$$

Критерій повноти аксіоматики багатозначних і функціональних залежностей

Аналіз доведення основного результату про збіжність відношень синтаксичного та семантичного слідування для аксіоматики БЗЗ та ФЗ показує, що воно проведено в припущенні: $|D| \geq 2$ та $|R| \geq 2$. Для повноти викладення залишається розглянути збіжність відношень $\vDash$ та $\vdash$ у випадках $|D| < 2$ або $|R| < 2$.

Лема 18. Усі БЗЗ виконуються на порожній таблиці. БЗЗ вигляду $X \rightarrow\rightarrow \varnothing$, зокрема, БЗЗ вигляду $\varnothing \rightarrow\rightarrow \varnothing$ виконується на довільній таблиці. □

Лема 19. Довільна таблиця є моделлю порожньої множини БЗЗ. □

Лема 20. Множина тривіальних БЗЗ замкнена відносно правил повноти, поповнення та транзитивності. □

Лема 21. Всі БЗЗ виконуються на довільній однорядковій таблиці. □

Нехай G – множина БЗЗ, через $(G)_{ntr}$ (None TRivial) позначимо підмножину множини G, що отримується вилученням тривіальних БЗЗ.

Лема 22. Виконуються наступні твердження:

1) $G \vdash \varphi$, якщо $\varphi \in G$;
2) $G \vDash \varphi$, якщо $\varphi \in G$;

3) $G \vdash \varphi \Leftrightarrow (G)_{ntr} \vdash \varphi$;

4) $G \vDash \varphi \Leftrightarrow (G)_{ntr} \vDash \varphi$;

5) $G \vdash \varphi$, якщо φ – тривіальна БЗЗ;

6) $G \vDash \varphi$, якщо φ – тривіальна БЗЗ. □

Залежність збіжності відношень синтаксичного та семантичного слідування для *аксіоматики БЗЗ при різних* значеннях потужностей множин R та D вказана у табл. 1. Символ "+" (відповідно "–") в комірці означає, що при вказаних припущеннях відношення $\vDash$ та $\vdash$ збігаються (не збігаються).

Таблиця 1. Всі варіанти потужностей множин R та D для аксіоматики багатозначних залежностей

D \ R	$\|R\|=0$	$\|R\|=1$	$\|R\|\geq 2$
$\|D\|=0$	+	+	–
$\|D\|=1$	+	+	–
$\|D\|\geq 2$	+	+	+

Твердження 3. Заповнення табл. 1 для аксіоматики БЗЗ коректне. □

Лема 23. Виконується імплікація: $X \to Y$ – тривіальна $\Rightarrow$ $X \dashrightarrow\dashrightarrow Y$ – тривіальна. □

Наслідок 12. Множина тривіальних ФЗ і БЗЗ замкнена відносно правил виведення аксіоматики ФЗ і БЗЗ. □

Зауважимо, що спільне для ФЗ і БЗЗ правило $\dfrac{X \to\to Z,\ Y \to Z',\ Z' \subseteq Z,\ Y \cap Z = \varnothing}{X \to Z'}$ не може бути застосовано до двох тривіальних залежностей; це забезпечується умовами $Y \cap Z = \varnothing$ і $Z' \subseteq Z$, з яких випливає, що $Y \cap Z' = \varnothing$, отже, ФЗ $Y \to Z'$ не є тривіальною.

Залежність збіжності відношень $\vDash$ та $\vdash$ для аксіоматики БЗЗ та ФЗ при різних значеннях потужностей множин R та D вказана у табл. 2 (усі позначення ті ж самі, що і для табл. 1).

Таблиця 2. Всі варіанти потужностей множин R та D для аксіоматики БЗЗ та ФЗ

D \ R	$\|R\|=0$	$\|R\|=1$	$\|R\|\geq 2$
$\|D\|=0$	+	–	–
$\|D\|=1$	+	–	–
$\|D\|\geq 2$	+	+	+

Твердження 4. Заповнення табл. 2 для аксіоматики БЗЗ та ФЗ коректне. □

Теорема 2. Відношення $\models$ та $\vdash$ для аксіоматики БЗЗ і ФЗ співпадають тоді і тільки тоді, коли $|D| \geq 2$ або $|R| = 0$. □

Висновки

В роботі розглянута аксіоматика багатозначних залежностей в табличних базах даних і аксіоматика функціональних та багатозначних залежностей; встановлені критерії повноти цих аксіоматик в термінах потужностей універсального домену та множини атрибутів. Наступний крок полягає в дослідженні незалежності аксіом та правил виведення для розглянутих аксіоматизацій.

Список використаних джерел

1. Реляційні бази даних: табличні алгебри та SQL-подібні мови / В.Н. Редько, Ю.Й. Брона, Д.Б. Буй, С.А. Поляков. – Київ: Видавничий дім «Академперіодика», 2001. – 198 с.
2. Beeri C. A complete axiomatization for functional and multivalued dependencies / C. Beeri, R. Fagin, J. Howard // Proceedings of the ACM-SIGMOD Conference, August 3-5, 1977, Toronto, Canada. – P. 47-61.
3. Линдон Р. Заметки по логике / Р. Линдон. – Москва: Мир, 1968. – 128 с.
4. Скорняков Л.А. Элементы теории структур / Л.А. Скорняков. – Москва: Наука, 1982. – 160 с.
5. Мейер Д. Теория реляционных баз данных: [пер. с англ.] / Д. Мейер. – Москва: Мир, 1987. – 608 с.