УДК 004.655
DOI: 10.37069/1683-4720-2021-35-12

©2021. **О.С. Сенченко**

ВЗАЄМОЗВ'ЯЗКИ РІЗНИЦІ ТА ПРОЕКЦІЇ З ІНШИМИ СИГНАТУРНИМИ ОПЕРАЦІЯМИ ТАБЛИЧНИХ АЛГЕБР

В наш час табличні (реляційні) бази даних, що засновані на реляційних принципах та моделях Е. Кодда, залишаються найбільш розповсюдженими. Табличні алгебри, які було введено В.Н. Редьком та Д.Б. Буєм, побудовані на основі реляційних алгебр Е. Кодда та суттєво їх уточнюють. Вони складають теоретичний фундамент мов запитів сучасних табличних баз даних. Елементи носія табличної алгебри уточнюють реляційні структури даних, а сигнатурні операції побудовані на базі основних табличних маніпуляцій у реляційних алгебрах та SQL-подібних мовах. Однією з актуальних задач в реляційних та табличних алгебрах є задача еквівалентного перетворення виразів з метою їх спрощення (мінімізації) або приведення до стандартного вигляду. Ця задача є одним з етапів оптимізації запитів, розв'язання цієї задачі може суттєво зменшити час обробки інформації в реляційних системах управління базами даних. Для еквівалентного перетворення виразів, що містять табличні операції, потрібно використовувати взаємозв'язки між цими операціями. Значну кількість таких взаємозв'язків вже встановлено, у більшості випадків ці взаємозв'язки для загального випадку виконуються у вигляді включень. Ця робота є етапом знаходження усіх можливих взаємозв'язків між сигнатурними операціями табличних алгебр. В роботі наведено та систематизовано взаємозв'язки різниці та проекції таблиць між собою та з іншими операціями табличних алгебр: перетином, об'єднанням, селекцією, перейменуванням, з'єднанням, насиченням та активним доповненням. Автором цієї роботи були знайдені критерії переходу деяких включень у рівності, а також критерій дистрибутивності насичення відносно різниці таблиць. Ці критерії виражені в термінах активних доменів таблиць та є природними. Важливість знайдених критеріїв рівностей для теорії табличних алгебр полягає в тому, що еквівалентні перетворення виразів, які містять табличні операції, можна здійснювати лише на основі рівностей.

MSC: 68P15.

Ключові слова: *різниця, проекція, бази даних, табличні алгебри.*

1. Вступ.

В наш час бази даних широко використовуються майже в усіх сферах людської діяльності. При цьому, найбільш розповсюдженими залишаються табличні (реляційні) бази даних, що засновані на реляційних принципах та моделях Е. Кодда [1]. З математичної точки зору реляційна база даних є скінченим набором скінчених відношень різної розмірності між заздалегідь визначеними множинами елементарних даних – доменами. Табличні алгебри, які було введено В.Н. Редьком та Д.Б. Буєм [2], побудовані на основі реляційних алгебр Е. Кодда та суттєво їх уточнюють. Вони складають теоретичний фундамент мов запитів сучасних табличних баз даних. Елементи носія табличної алгебри уточнюють реляційні структури даних, а сигнатурні операції побудовані на базі основних табличних маніпуляцій у реляційних алгебрах та SQL-подібних мовах.

Однією з актуальних задач в реляційних та табличних алгебрах є задача екві-

валентного перетворення виразів з метою їх спрощення (мінімізації) або приведення до стандартного вигляду; ця задача є одним з етапів оптимізації запитів [3], розв'язання цієї задачі може суттєво зменшити час обробки інформації в реляційних системах управління базами даних. Для еквівалентного перетворення виразів, що містять табличні операції, потрібно використовувати взаємозв'язки між цими операціями. Можна виділити п'ять типів таких взаємозв'язків:
а) вираженння однієї операції через інші – похідність операції;
б) переставність (для двох унарних операцій);
в) дистрибутивність (для двох бінарних або для бінарної та унарної операцій, причому для двох бінарних операцій існує два різних варіанта дистрибутивності);
г) аналоги відомих властивостей теоретико-множинних операцій – закони де Моргана, подвійного заперечення тощо.
д) взаємозв'язки при виконанні однієї операції декілька разів – ідемпотентність унарних операцій, асоціативність бінарних операцій (інакше кажучи, взаємозв'язки операції самої з собою).

Значну кількість таких взаємозв'язків встановлено у монографії [4], у більшості випадків ці взаємозв'язки для загального випадку виконуються у вигляді включень. Автором цієї роботи були знайдені критерії переходу деяких з цих включень у рівності, а також критерій дистрибутивності насичення відносно різниці таблиць. Ці критерії виражені в термінах активних доменів таблиць та є природними. Важливість знайдених критеріїв рівностей для теорії табличних алгебр полягає в тому, що еквівалентні перетворення виразів, які містять табличні операції, можна здійснювати лише на основі рівностей.

У цій роботі наведено всі знайдені на цей час взаємозв'язки різниці та проекції між собою та з іншими операціями табличних алгебр: перетином, об'єднанням, селекцією, перейменуванням атрибутів, з'єднанням, насиченням та активним доповненням, у певному сенсі ця робота є продовженням роботи [5], у якій наведено аналогічні взаємозв'язки для перетину та об'єднання.

2. Основні означення.

Далі розглянемо основні означення з теорії табличних алгебр за [4].

Зафіксуємо деяку непорожню множину $A = \{A'_1, \ldots, A'_n\}$, елементи якої будемо називати атрибутами. Довільну скінченну підмножину $R = \{A_1, \ldots, A_k\} \subseteq A$ назвемо схемою. Рядком s схеми R є множина пар $s = \{(A_1, d_1), \ldots, (A_k, d_k)\}$, проекція якої за першою компонентою дорівнює R. Таблицею схеми R є скінченна множина рядків схеми R. Виокремлюють дві особливі таблиці: таблицю $T_\varepsilon = \{\varepsilon\}$, де ε – порожній рядок, при цьому схема таблиці T_ε є порожнім рядком, та таблицю $T_\emptyset = \emptyset$ – порожня множина рядків довільної схеми.

На множині всіх таблиць схеми R введено такі параметричні операції:
1) перетин $\bigcap\limits_R$ таблиць схеми R – таблиця, що складається з тих і лише тих рядків, які належать одночасно всім вихідним таблицям;
2) об'єднання $\bigcup\limits_R$ таблиць схеми R – таблиця, що складається з тих і лише тих рядків, які належать хоча б одній з вихідних таблиць;

3) різниця $T_1 \underset{R}{-} T_2$ двох таблиць схеми R – таблиця, що складається з тих і лише тих рядків, які належать таблиці T_1, та не належать таблиці T_2.

Іншими словами, операції перетину, об'єднання і різниці таблиць є обмеженням відповідно теоретико-множинних перетину, об'єднання і різниці на множині таблиць однакової схеми.

Проекцією за множиною атрибутів $X \subseteq R$ називається унарна параметрична операція π_X, значенням якої є таблиця, що складається з обмежень за X усіх рядків вихідної таблиці: $\pi_X(T) = \{s \| X | s \in T\}$ (тут обмеження розуміємо стандартно: $s \| x = s \bigcap (X \times pr_2 s)$, де $pr_2 s$ – проекція рядка s за другою компонентою). Селекцією за предикатом $P : S \to \{true, false\}$, де S – множина всіх рядків, є унарна параметрична операція σ_P, яка таблиці співставляє її підтаблицю, що містить рядки, на яких предикат P принимає значення ІСТИНА.

Введемо визначення операції перейменування атрибутів. Нехай ξ – ін'єктивне та функціональне бінарне відношення на множині R. Покладемо η – таке поповнення відношення ξ парами вигляду (x, x), що η є всюдивизначеним відношенням на R: $\eta = \{\xi \bigcup (x, x) | x \in (R - Dom\xi)\}$. Перейменування рядка $s = \{(A_1, d_1), \dots, (A_k, d_k)\}$ схеми R відношенням ξ є рядок $\xi(s) = \{(\eta(A_1), d_1), \dots, (\eta(A_k), d_k)\}$, а перейменування таблиці T схеми R відношенням ξ є таблиця $Rn(T)_\xi$, яка є об'єднанням перейменувань рядків таблиці T відношенням ξ; перейменування атрибутів схеми R позначатимемо через $\xi(R)$. Змістовно кажучи, операція перейменування зводиться до перейменування атрибутів таблиць у відповідності до відображення-параметру ξ, тобто, перейменування таблиці є перейменуванням перших компонент пар – елементів рядків.

Для визначення операції з'єднання введемо одне допоміжне поняття. Бінарні відношення ρ і τ називаються сумісними (позначається $\rho \approx \tau$), якщо $\rho \| X = \tau \| X$, де $X = pr_1(\rho) \bigcap pr_1(\tau)$. З'єднанням називається бінарна операція $\otimes$, значенням якої є таблиця, що складається з усіх можливих об'єднань сумісних рядків вихідних таблиць, тобто $T_1 \otimes T_2 = \{s_1 \bigcup s_2 \mid s_1 \in T_1 \wedge s_2 \in T_2 \wedge s_1 \approx s_2\}$. При цьому, якщо R_1 – схема таблиці T_1, а R_2 – схема таблиці T_2, то схемою таблиці $T_1 \otimes T_2$ є множина $R_1 \bigcup R_2$.

Для визначення операції насичення введемо одне важливе поняття. Активним доменом атрибуту A відносно таблиці T називається множина $D_{A,T} = \{d | \exists s \in T \wedge (A, d) \in s\}$. Змістовно кажучи, $D_{A,T}$ містить усі значення атрибуту A в таблиці T. Насиченням $C(T)$ називається таблиця $\prod\limits_{A \in R} D_{A,T}$, де R – схема таблиці T, а $\prod$ – оператор прямого (декартового) добутку, що відповідає індексуванню $A \mapsto D_{A,T}$, $A \in R$. Активним доповненням таблиці T схеми R називається таблиця $\tilde{T} = C(T) \underset{R}{-} T$.

Табличною алгеброю називають часткову алгебру з носієм – множиною всіх таблиць довільної схеми, наведеними вище операціями (причому у [4] насичення вважається допоміжною операцією), а також операцією ділення таблиць, яка, як показано у [4, с. 83], може бути вираженою через проекцію, різницю та з'єднання;

в цій роботі операція ділення не використовується.

3. Взаємозв'язки між різницею та іншими табличними операціями.

Далі, для компактності викладання результатів, покладемо, що всі зазначені таблиці мають непорожню схему R з кількістю атрибутів $k > 0$ (окрім випадків, коли про схему таблиці та потужність схеми сказано явно).

Операція перетину таблиць є дистрибутивною відносно різниці: $T_1 \bigcap_R (T_2 \underset{R}{-} T_3) = (T_1 \bigcap_R T_2) \underset{R}{-} (T_1 \bigcap_R T_3)$. Доведення наведено у [5].

Виконується включення $T_1 \underset{R}{-} (T_2 \bigcap_R T_3) \supseteq (T_1 \underset{R}{-} T_2) \bigcap_R (T_1 \underset{R}{-} T_3)$, яке перетворюється у рівність тоді і лише тоді, коли $T_1 \bigcap_R T_2 = T_1 \bigcap_R T_3$. Доведення наведено у [5].

Оскільки різниця таблиць (на відміну від їх перетину) не є комутативною операцією, розглянемо ще один варіант дистрибутивності між різницею та перетином.

Лема 1. *Виконується рівність* $(T_1 \bigcap_R T_2) \underset{R}{-} T_3 = (T_1 \underset{R}{-} T_3) \bigcap_R (T_2 \underset{R}{-} T_3)$.

Доведення. Нехай $s \in (T_1 \bigcap_R T_2) \underset{R}{-} T_3$. Тоді, за означенням перетину та різниці, виконуються приналежності $s \in T_1$, $s \in T_2$, $s \notin T_3$. Тому $s \in (T_1 \underset{R}{-} T_3)$, $s \in (T_2 \underset{R}{-} T_3)$, звідки випливає $s \in (T_1 \underset{R}{-} T_3) \bigcap_R (T_2 \underset{R}{-} T_3)$, тобто $(T_1 \bigcap_R T_2) \underset{R}{-} T_3 \subseteq (T_1 \underset{R}{-} T_3) \bigcap_R (T_2 \underset{R}{-} T_3)$.

Нехай тепер $s \in (T_1 \underset{R}{-} T_3) \bigcap_R (T_2 \underset{R}{-} T_3)$. Тоді, за означенням перетину та різниці, виконуються приналежності $s \in T_1$, $s \in T_2$, $s \notin T_3$, тому $s \in (T_1 \bigcap_R T_2) \underset{R}{-} T_3$. $\Box$

Виконується включення $T_1 \bigcup_R (T_2 \underset{R}{-} T_3) \supseteq (T_1 \bigcup_R T_2) \underset{R}{-} (T_1 \bigcup_R T_3)$, яке перетворюється у рівність тоді і лише тоді, коли $T_1 = T_\emptyset$. Доведення наведено у [5].

Виконується включення $T_1 \underset{R}{-} (T_2 \bigcup_R T_3) \subseteq (T_1 \underset{R}{-} T_2) \bigcup_R (T_1 \underset{R}{-} T_3)$, яке перетворюється у рівність тоді і лише тоді, коли $T_1 \bigcap_R T_2 = T_1 \bigcap_R T_3$. Доведення наведено у [5].

Оскільки, на відміну від об'єднання, різниця таблиць не є комутативною операцією, розглянемо ще один варіант дистрибутивності між різницею та об'єднанням.

Лема 2. *Виконується рівність* $(T_1 \bigcup_R T_2) \underset{R}{-} T_3 = (T_1 \underset{R}{-} T_3) \bigcup_R (T_2 \underset{R}{-} T_3)$.

Доведення. Нехай $s \in (T_1 \bigcup_R T_2) \underset{R}{-} T_3$. Тоді, за означенням об'єднання та різниці, виконується $s \notin T_3$, а також хоча б одна з приналежностей: $s \in T_1$ або $s \in T_2$. Тоді, з $s \in T_1$ випливає $s \in (T_1 \underset{R}{-} T_3)$, а з $s \in T_2$ випливає $s \in (T_2 \underset{R}{-} T_3)$, у будь-якому випадку $s \in (T_1 \underset{R}{-} T_3) \bigcup_R (T_2 \underset{R}{-} T_3)$, тому $(T_1 \bigcup_R T_2) \underset{R}{-} T_3 \subseteq (T_1 \underset{R}{-} T_3) \bigcup_R (T_2 \underset{R}{-} T_3)$.

Нехай тепер $s \in (T_1 \underset{R}{-} T_3) \bigcup_R (T_2 \underset{R}{-} T_3)$. Тоді, за означенням об'єднання та різниці, виконуються $s \notin T_3$, а також хоча б одна з приналежностей $s \in T_1$ та $s \in T_2$, звідки випливає $s \in (T_1 \bigcup_R T_2) \underset{R}{-} T_3$. $\Box$

Неважко бачити, що різниця є не комутативною та не асоціативною операцією, проте справедлива

Лема 3. *Виконується включення* $(T_1 \underset{R}{-} T_2) \underset{R}{-} T_3 \subseteq T_1 \underset{R}{-} (T_2 \underset{R}{-} T_3)$. *Це включення перетворюється у рівність тоді і лише тоді, коли* $T_1 \underset{R}{\bigcap} T_3 = T_\emptyset$.

Доведення. Нехай $s \in (T_1 \underset{R}{-} T_2) \underset{R}{-} T_3$. Тоді, за означенням різниці, виконується $s \in T_1$, $s \notin T_2$, $s \notin T_3$. Тоді, з $s \notin T_2$ випливає $s \notin (T_2 \underset{R}{-} T_3)$, та, з урахуванням $s \in T_1$, за означенням різниці, $s \in T_1 \underset{R}{-} (T_2 \underset{R}{-} T_3)$, що доводить включення.

Нехай тепер виконується $T_1 \underset{R}{\bigcap} T_3 = T_\emptyset$ та $s \in T_1 \underset{R}{-} (T_2 \underset{R}{-} T_3)$. За означенням різниці $s \in T_1$ та $s \notin (T_2 \underset{R}{-} T_3)$. З $s \in T_1$ та $T_1 \underset{R}{\bigcap} T_3 = T_\emptyset$ випливає $s \notin T_3$, з чого, з урахуванням $s \notin (T_2 \underset{R}{-} T_3)$, випливає $s \notin T_2$, тому $s \in (T_1 \underset{R}{-} T_2)$ та $s \in (T_1 \underset{R}{-} T_2) \underset{R}{-} T_3$, що доводить достатність критерію.

Доведемо необхідність критерію. Нехай виконується рівність $(T_1 \underset{R}{-} T_2) \underset{R}{-} T_3 = T_1 \underset{R}{-} (T_2 \underset{R}{-} T_3)$. Від супротивного, припустимо, що $T_1 \underset{R}{\bigcap} T_3 \neq T_\emptyset$ та нехай $s \in T_1 \underset{R}{\bigcap} T_3$. Тоді, з $s \in T_3$ випливає $s \notin (T_2 \underset{R}{-} T_3)$, з чого, з урахуванням $s \in T_1$, випливає $s \in T_1 \underset{R}{-} (T_2 \underset{R}{-} T_3)$. З припущення про виконання рівності $(T_1 \underset{R}{-} T_2) \underset{R}{-} T_3 = T_1 \underset{R}{-} (T_2 \underset{R}{-} T_3)$ випливає $s \in (T_1 \underset{R}{-} T_2) \underset{R}{-} T_3$. Проте, за означенням різниці, ця приналежність не може виконуватися через $s \in T_3$, що доводить необхідність критерію. □

Виконується включення $\pi_X(T_1) \underset{X}{-} \pi_X(T_2) \subseteq \pi_X(T_1 \underset{R}{-} T_2)$, що доведено у [4, с. 42]. Знайдемо необхідні та достатні умови, за яких це включення перетворюється у рівність. Далі, у теоремі 1 покладемо T_1, T_2 – таблиці схеми R, $X = \{X_1, \ldots, X_m\}$; а елементи множини $R - X$ позначимо $O_1, \ldots, O_p$.

Теорема 1. *Рівність* $\pi_X(T_1) \underset{X}{-} \pi_X(T_2) = \pi_X(T_1 \underset{R}{-} T_2)$ *виконується тоді та лише тоді, коли для будь-якого рядка* $s = \{(X_1, x_1), \ldots, (X_m, x_m)\} \in \pi_X(T_1) \underset{X}{\bigcap} \pi_X(T_2)$ *та всіх значень* $o_1 \in D_{O_1, T_1}, \ldots, o_p \in D_{O_p, T_1}$, *з* $s' = \{(X_1, x_1), \ldots, (X_m, x_m), (O_1, o_1), \ldots, (O_p, o_p)\} \in T_1$ *випливає* $s' \in T_2$.

Доведення. Необхідність. Нехай виконується рівність $\pi_X(T_1) \underset{X}{-} \pi_X(T_2) = \pi_X(T_1 \underset{R}{-} T_2)$. Від супротивного, припустимо, що критерій не виконується, тобто існують рядок $s = \{(X_1, x_1), \ldots, (X_m, x_m)\} \in \pi_X(T_1) \underset{R}{\bigcap} \pi_X(T_2)$ та значення $o_1 \in D_{O_1, T_1 \underset{R}{\bigcup} T_2}, \ldots, o_p \in D_{O_p, T_1 \underset{R}{\bigcup} T_2}$, що $s' = \{(X_1, x_1), \ldots, (X_m, x_m), (O_1, o_1), \ldots, (O_p, o_p)\} \in T_1$ та $s' \notin T_2$. Тоді $s' \in T_1 \underset{R}{-} T_2$, та, за визначенням проекції, $\{(X_1, x_1), \ldots, (X_m, x_m)\} \in \pi_X(T_1 \underset{R}{-} T_2)$, тобто $s \in \pi_X(T_1 \underset{R}{-} T_2)$. З іншого боку, з $s \in \pi_X(T_1) \underset{R}{\bigcap} \pi_X(T_2)$ випливає $s \in \pi_X(T_1)$ та $s \in \pi_X(T_2)$, тому, за визначенням різниці таблиць, $s \notin \pi_X(T_1) \underset{X}{-} \pi_X(T_2)$, з чого випливає не виконання рівності $\pi_X(T_1) \underset{X}{-} \pi_X(T_2) = \pi_X(T_1 \underset{R}{-} T_2)$. Це суперечить припущенню та доводить необхідність.

Доведемо достатність. Нехай виконується критерій. Від супротивного, припустимо, що рівність $\pi_X(T_1) \underset{X}{-} \pi_X(T_2) = \pi_X(T_1 \underset{R}{-} T_2)$ не виконується, це є можливим лише при існуванні такого рядка $s = \{(X_1, x_1), \dots, (X_m, x_m)\}$, що $s \in \pi_X(T_1 \underset{R}{-} T_2)$ та $s \notin \pi_X(T_1) \underset{X}{-} \pi_X(T_2)$. Оскільки $s \in \pi_X(T_1 \underset{R}{-} T_2)$, то в таблиці $T_1 \underset{R}{-} T_2$ існує рядок $s' = \{(X_1, x_1), \dots, (X_m, x_m), (O_1, o_1), \dots, (O_p, o_p)\}$, для деяких значень $o_1 \in D_{O_1, T_1 \underset{R}{-} T_2}, \dots, o_p \in D_{O_p, T_1 \underset{R}{-} T_2}$. З $T_1 \underset{R}{-} T_2 \subseteq T_1$ випливає $\forall A \in R(D_{A, T_1 \underset{R}{-} T_2} \subseteq D_{A, T_1})$, тому $o_1 \in D_{O_1, T_1}, \dots, o_p \in D_{O_p, T_1}$. За визначенням різниці таблиць, $s' \in T_1$ та $s' \notin T_2$. За визначенням проекції, з $s' \in T_1$ випливає $\{(X_1, x_1), \dots, (X_m, x_m)\} \in \pi_X(T_1)$, тобто $s \in \pi_X(T_1)$. За припущенням $s \notin \pi_X(T_1) \underset{X}{-} \pi_X(T_2)$, що, зважаючи на $s \in \pi_X(T_1)$, може бути лише в тому випадку, коли $s \in \pi_X(T_2)$, тому $s \in \pi_X(T_1) \underset{X}{\bigcap} \pi_X(T_2)$. Тоді, з припущення виконання критерію випливає $s' \in T_2$. Одержана суперечність доводить хибність припущення про невиконаність рівності $\pi_X(T_1) \underset{X}{-} \pi_X(T_2) = \pi_X(T_1 \underset{R}{-} T_2)$ та доводить достатність. $\square$

У [4, с. 39] доведено дистрибутивність селекції відносно різниці:
$\sigma_P(T_1 \underset{R}{-} T_2) = \sigma_P(T_1) \underset{R}{-} \sigma_P(T_2) = \sigma_P(T_1) \underset{R}{-} T_2$.

У [4, с. 90] доведено дистрибутивність перейменування відносно різниці:
$Rn(T_1 \underset{R}{-} T_2)_\xi = Rn(T_1)_\xi \underset{\xi(R)}{-} Rn(T_2)_\xi$.

Розглянемо взаємозв'язки між різницею та з'єднанням. Оскільки різниця є частковою операцією, яка визначена лише у випадку, коли схеми обох таблиць-аргументів є однаковими, а схема таблиці-з'єднання є об'єднанням схем вихідних таблиць, при дослідженні можливої дистрибутивності існують певні обмеження на схеми вихідних таблиць. У [4, с. 52] доведено дистрибутивність з'єднання відносно різниці: нехай R_1 – схема таблиці T_1, R – схема таблиці T_2. Тоді, якщо R – схема таблиці T_3, виконується рівність: $T_1 \otimes (T_2 \underset{R}{-} T_3) = (T_1 \otimes T_2) \underset{R \bigcup R_1}{-} (T_1 \otimes T_3)$. Зауважимо, що у випадку, за яким схеми таблиць T_2 та T_3 не збігаються, ліва та права частини наведеної вище рівності не можуть одночасно бути визначеними, оскільки ліва частина обов'язково буде невизначеною.

Для дослідження можливої дистрибутивності різниці відносно з'єднання (зважаючи на некомутативність різниці будемо розглядати два варіанти такої дистрибутивності) звернемо увагу, що ліва та права частина наведених нижче виразів одночасно будуть визначеними лише за випадку рівності схем вихідних таблиць. Справедливе твердження.

Лема 4. *Нехай R – схема таблиць T_1, T_2, T_3. Тоді:*

а) виконується рівність $(T_1 \otimes T_2) \underset{R}{-} T_3 = (T_1 \underset{R}{-} T_3) \otimes (T_2 \underset{R}{-} T_3)$;

б) виконується включення $T_1 \underset{R}{-} (T_2 \otimes T_3) \supseteq (T_1 \underset{R}{-} T_2) \otimes (T_1 \underset{R}{-} T_3)$;

в) рівність $T_1 \underset{R}{-} (T_2 \otimes T_3) = (T_1 \underset{R}{-} T_2) \otimes (T_1 \underset{R}{-} T_3)$ виконується тоді та лише тоді, коли $T_1 \underset{R}{\bigcap} T_2 = T_1 \underset{R}{\bigcap} T_3$.

Доведення. У [4, с. 52] доведено, що у випадку, за яких схеми таблиць T та T'

збігаються та дорівнюють R, $T \otimes T' = T \bigcap_R T'$. Тоді, з рівності $(T_1 \bigcap_R T_2) \underset{R}{-} T_3 = (T_1 \underset{R}{-} T_3) \bigcap_R (T_2 \underset{R}{-} T_3)$, яку доведено у лемі 1 цієї роботи, безпосередньо випливає рівність (а), з включення $T_1 \underset{R}{-} (T_2 \bigcap_R T_3) \supseteq (T_1 \underset{R}{-} T_2) \bigcap_R (T_1 \underset{R}{-} T_3)$, доведенного у [5], безпосередньо випливає включення (б), а з того, що критерієм перетворення включення $T_1 \underset{R}{-} (T_2 \bigcap_R T_3) \supseteq (T_1 \underset{R}{-} T_2) \bigcap_R (T_1 \underset{R}{-} T_3)$ у рівність є $T_1 \bigcap_R T_2 = T_1 \bigcap_R T_3$, що доведено в [5], безпосередньо випливає пункт (в) цієї леми. □

Знайдемо взаємозв'язки між різницею та насиченням. Розглянемо такий приклад.

Приклад 1. Нехай $R = \{A, B\}$, $T_1 = \begin{array}{cc} \mathbf{A} & \mathbf{B} \\ 0 & 0 \\ 1 & 1 \\ 2 & 1 \end{array}$ та $T_2 = \begin{array}{cc} \mathbf{A} & \mathbf{B} \\ 0 & 1 \\ 2 & 1 \end{array}$.

Тоді $T_1 \underset{R}{-} T_2 = \begin{array}{cc} \mathbf{A} & \mathbf{B} \\ 0 & 0 \\ 1 & 1 \end{array}$, $C(T_1) = \begin{array}{cc} \mathbf{A} & \mathbf{B} \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 2 & 0 \\ 2 & 1 \end{array}$ та $C(T_2) = \begin{array}{cc} \mathbf{A} & \mathbf{B} \\ 0 & 1 \\ 2 & 1 \end{array}$.

Отже, $C(T_1 \underset{R}{-} T_2) = \begin{array}{cc} \mathbf{A} & \mathbf{B} \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{array}$, $C(T_1) - C(T_2) = \begin{array}{cc} \mathbf{A} & \mathbf{B} \\ 0 & 0 \\ 1 & 0 \\ 1 & 1 \\ 2 & 0 \end{array}$.

З цього прикладу можна побачити, що насичення не є дистрибутивним відносно різниці, більш того, для загального випадку $C(T_1 \underset{R}{-} T_2) \not\subseteq C(T_1) \underset{R}{-} C(T_2)$ та $C(T_1 \underset{R}{-} T_2) \not\supseteq C(T_1) \underset{R}{-} C(T_2)$. Знайдемо критерії виконання включення $C(T_1 \underset{R}{-} T_2) \subseteq C(T_1) \underset{R}{-} C(T_2)$.

Лема 5. *Включення $C(T_1 \underset{R}{-} T_2) \subseteq C(T_1) \underset{R}{-} C(T_2)$ виконується тоді та лише тоді, коли існує такий індекс i, що $D_{A_i, T_1 \underset{R}{-} T_2} \bigcap D_{A_i, T_2} = \emptyset$.*

Доведення. Необхідність. Нехай виконується включення $C(T_1 \underset{R}{-} T_2) \subseteq C(T_1) \underset{R}{-} C(T_2)$. Від супротивного, припустимо, що критерій не виконується, тобто $\forall i$ $D_{A_i, T_1 \underset{R}{-} T_2} \bigcap D_{A_i, T_2} \neq \emptyset$. Нехай $x_1 \in D_{A_1, T_1 \underset{R}{-} T_2} \bigcap D_{A_1, T_2}, \ldots, x_k \in D_{A_k, T_1 \underset{R}{-} T_2} \bigcap D_{A_k, T_2}$. Позначимо $s = \{(A_1, x_1), \ldots, (A_k, x_k)\}$. З $x_1 \in D_{A_1, T_1 \underset{R}{-} T_2}, \ldots, x_k \in D_{A_k, T_1 \underset{R}{-} T_2}$ за означенням насичення випливає $s \in C(T_1 \underset{R}{-} T_2)$. З іншого боку, з $x_1 \in D_{A_1, T_2}, \ldots, x_k \in D_{A_k, T_2}$ за означенням насичення випливає $s \in C(T_2)$, тому $s \notin C(T_1) \underset{R}{-} C(T_2)$, тобто

$C(T_1 \underset{R}{-} T_2) \nsubseteq C(T_1) \underset{R}{-} C(T_2)$, ця суперечність доводить необхідність критерію.

Доведемо достатність. Нехай виконується критерій. Від супротивного, припустимо, що $C(T_1 \underset{R}{-} T_2) \nsubseteq C(T_1) \underset{R}{-} C(T_2)$, тобто існує такий рядок $s = \{(A_1, x_1), \ldots, (A_k, x_k)\}$, що $s \in C(T_1 \underset{R}{-} T_2)$ та $s \notin C(T_1) \underset{R}{-} C(T_2)$. Тоді $x_1 \in D_{A_1, T_1 \underset{R}{-} T_2}, \ldots, x_k \in D_{A_k, T_1 \underset{R}{-} T_2}$. За означенням насичення, з $s \in C(T_1 \underset{R}{-} T_2)$ випливає існування в таблиці $T_1 \underset{R}{-} T_2$ таких рядків $s_1, \ldots, s_k$, що $(A_1, x_1) \in s_1, \ldots, (A_k, x_k) \in s_k$, тому, $s_1, \ldots, s_k \in T_1$ та $s \in C(T_1)$. У цьому випадку $s \notin C(T_1) \underset{R}{-} C(T_2)$ може виконуватися лише за умови $s \in C(T_2)$, з чого, за означенням насичення, випливає $x_1 \in D_{A_1, T_2}, \ldots, x_k \in D_{A_k, T_2}$. Одержали $x_1 \in D_{A_1, T_1 \underset{R}{-} T_2} \bigcap D_{A_1, T_2}, \ldots, x_k \in D_{A_k, T_1 \underset{R}{-} T_2} \bigcap D_{A_k, T_2}$. Це суперечить припущенню про виконуваність критерію та доводить його достатність. □

Знайдемо критерії виконання включення $C(T_1) \underset{R}{-} C(T_2) \subseteq C(T_1 \underset{R}{-} T_2)$. Справедливе твердження.

Лема 6. *Включення $C(T_1) \underset{R}{-} C(T_2) \subseteq C(T_1 \underset{R}{-} T_2)$ виконується тоді та лише тоді, коли виконується одна з двох взаємовиключних умов:*

1) $\forall i (D_{A_i, T_1 \underset{R}{-} T_2} = D_{A_i, T_1})$;

2)$\forall i (D_{A_i, T_1 \underset{R}{-} T_2} \neq D_{A_i, T_1} \rightarrow (\forall j \neq i \quad D_{A_j, T_1} \subseteq D_{A_j, T_2}))$.

Доведення. Необхідність. Нехай виконується включення $C(T_1) \underset{R}{-} C(T_2) \subseteq C(T_1 \underset{R}{-} T_2)$ та не виконується умова 1. Покажемо, що у такому випадку повинна виконуватися умова 2. Від супротивного, припустимо, що це не так, тобто існує такий індекс i, що $D_{A_i, T_1 \underset{R}{-} T_2} \neq D_{A_i, T_1}$, та для деякого $j \neq i$ виконується $D_{A_j, T_1} \nsubseteq D_{A_j, T_2}$. Нехай $x \in D_{A_j, T_1}$ та $x \notin D_{A_j, T_2}$. З $D_{A_i, T_1 \underset{R}{-} T_2} \neq D_{A_i, T_1}$, зважаючи на $T_1 \underset{R}{-} T_2 \subseteq T_1$, та, відповідно, $D_{A_i, T_1 \underset{R}{-} T_2} \subseteq D_{A_i, T_1}$, випливає існування в D_{A_i, T_1} такого y, що $y \in D_{A_i, T_1}$ та $y \notin D_{A_i, T_1 \underset{R}{-} T_2}$. Тоді, за означенням активного домену, $\exists s \in T_1 | (A_i, y) \in s$. Нехай $s = \{(A_1, d_1), \ldots, (A_{j-1}, d_{j-1}), (A_j, d_j), (A_{j+1}, d_{j+1}), \ldots, (A_k, d_k), (A_i, y)\}$. Позначимо $s' = \{(A_1, d_1), \ldots, (A_{j-1}, d_{j-1}), (A_j, x), (A_{j+1}, d_{j+1}), \ldots, (A_k, d_k), (A_i, y)\}$. З $s \in T_1$ та $x \in D_{A_j, T_1}$ випливає $s' \in C(T_1)$, з $x \notin D_{A_j, T_2}$ випливає $s' \notin C(T_2)$, з $y \notin D_{A_i, T_1 \underset{R}{-} T_2}$ випливає $s' \notin C(T_1 \underset{R}{-} T_2)$, що неможливо, зважаючи на припущення про виконуваність включення $C(T_1) \underset{R}{-} C(T_2) \subseteq C(T_1 \underset{R}{-} T_2)$. Одержана суперечність доводить необхідність.

Доведемо достатність. Нехай виконується умова 1. Тоді, за визначенням насичення, $C(T_1 \underset{R}{-} T_2) = C(T_1)$, тому, за визначенням різниці, має місце включення $C(T_1) \underset{R}{-} C(T_2) \subseteq C(T_1 \underset{R}{-} T_2) = C(T_1)$.

Нехай виконується умова 2. Від супротивного, припустимо, що $C(T_1) \underset{R}{-} C(T_2) \nsubseteq$

$C(T_1 \underset{R}{-} T_2)$, тобто існує такий рядок $s = \{(A_1, d_1), \ldots, (A_k, d_k)\}$, що $s \in C(T_1) \underset{R}{-} C(T_2)$ та $s \notin C(T_1 \underset{R}{-} T_2)$, з чого випливає $\exists i | (A_i, d_i) \in s \wedge d_i \notin D_{A_i,(T_1 \underset{R}{-} T_2)}$. При цьому, з $s \in C(T_1) \underset{R}{-} C(T_2)$ випливає $d_i \in D_{A_i,T_1}$, отже, $D_{A_i,T_1 \underset{R}{-} T_2} \neq D_{A_i,T_1}$. Крім того, з $s \in C(T_1) \underset{R}{-} C(T_2)$ випливає також $s \in C(T_1)$ та $s \notin C(T_2)$, тобто $\forall z \quad d_z \in D_{A_z,T_1}$. За означенням насичення, з $d_i \in D_{A_i,T_1}$ випливає існування у таблиці T_1 такого рядка s', що $(A_i, d_i) \in s'$, а з $d_i \notin D_{A_i,(T_1 \underset{R}{-} T_2)}$, з урахуванням того, що таблиця $T_1 \underset{R}{-} T_2$ є результатом вилучення з таблиці T_1 тих рядків, що належать таблиці T_2, випливає $s' \in T_2$, тому $d_i \in D_{A_i,T_2}$. Оскільки за припущенням умова 2 виконується, тобто $\forall j \neq i \quad D_{A_j,T_2} \subseteq D_{A_j,T_2}$, то з $d_j \in D_{A_j,T_1}$ випливає $d_j \in D_{A_j,T_2}$. Таким чином, за означенням насичення, маємо $s \in C(T_2)$. Одержана суперечність доводить хибність припущення про невиконуваність включення $C(T_1) \underset{R}{-} C(T_2) \subseteq C(T_1 \underset{R}{-} T_2)$ за виконання умови 2. $\square$

Об'єднаємо результати, що одержано у лемах 5 та 6, у наступному твердженні.

Теорема 2. *Рівність $C(T_1) \underset{R}{-} C(T_2) = C(T_1 \underset{R}{-} T_2)$ виконується тоді та лише тоді, коли існує такий індекс i, що $D_{A_i,T_1 \underset{R}{-} T_2} \bigcap D_{A_i,T_2} = \emptyset$, та, крім того, виконується одна з двох взаємовиключних умов:*

1) $\forall i D_{A_i,T_1 \underset{R}{-} T_2} = D_{A_i,T_1}$;

2)$\forall i (D_{A_i,T_1 \underset{R}{-} T_2} \neq D_{A_i,T_1} \to (\forall j \neq i \quad D_{A_j,T_1} \subseteq D_{A_j,T_2}))$.

У якості взаємозв'язків між різницею та активним доповненням можна розглянути розповсюдження відомої з теорії множин рівності $A - B = A \wedge \bar{B}$ на табличні операції.

У [4, с. 37] доведено включення $T_1 \underset{R}{-} T_2 \supseteq T_1 \underset{R}{\bigcap} \tilde{T}_2$. Знайдемо критерії перетворення цього включення у рівність.

Лема 7. *Рівність $T_1 \underset{R}{-} T_2 = T_1 \underset{R}{\bigcap} \tilde{T}_2$ виконується тоді та лише тоді, коли для будь-якого $A \in R$ виконується включення $D_{A,T_1} \subseteq D_{A,T_2}$.*

Доведення. Необхідність. Нехай виконується рівність $T_1 \underset{R}{-} T_2 = T_1 \underset{R}{\bigcap} \tilde{T}_2$. Від супротивного, припустимо, що критерій не виконується, тобто існують такі $A \in R$, $x \in D_{A,T_1}$, що $x \notin D_{A,T_2}$. За визначенням активного домену, $\exists s \in T_1 | (A, x) \in s$. Оскільки $x \notin D_{A,T_2}$, то $s \notin T_2$, тому $s \in T_1 \underset{R}{-} T_2$. З іншого боку, за визначенням насичення, з $x \notin D_{A,T_2}$ випливає $s \notin C(T_2)$, тому $s \notin \tilde{T}_2 = C(T_2) - T_2$, з чого випливає $s \notin T_1 \underset{R}{\bigcap} \tilde{T}_2$, що суперечить припущенню про виконання рівності та доводить необхідність критерію.

Доведемо достатність. Нехай виконується критерій. Від супротивного, припустимо, що рівність $T_1 \underset{R}{-} T_2 = T_1 \underset{R}{\bigcap} \tilde{T}_2$ не виконується, що, з урахуванням включення $T_1 \underset{R}{-} T_2 \supseteq T_1 \underset{R}{\bigcap} \tilde{T}_2$ еквівалентно існуванню такого рядка $s = \{(A_1, d_1), \ldots, (A_k, d_k)\} \in$

$T_1 \underset{R}{-} T_2$, що $s \notin T_1 \underset{R}{\bigcap} \tilde{T_2}$. З $s \in T_1 \underset{R}{-} T_2$ випливає $s \in T_1$ та $s \notin T_2$, з умов критерію випливають приналежності $d_1 \in D_{A_1,T_2}, \ldots, d_k \in D_{A_k,T_2}$, тому $s \in C(T_2)$, з чого, з урахуванням $s \notin T_2$, отримаємо $s \in \tilde{T_2}$, тому $s \in T_1 \underset{R}{\bigcap} \tilde{T_2}$. Одержана суперечність доводить достатність. □

4. Взаємозв'язки між проекцією та іншими табличними операціями.

У [4, с. 42] доведено включення $\pi_X(\bigcap_i T_i) \subseteq \bigcap_i \pi_X(T_i)$. Це включення перетворюється у рівність тоді та лише тоді, коли для кожного рядка $s = \{(X_1, x_1), \ldots, (X_m, x_m)\} \in \bigcap_i \pi_X(T_i)$ існують такі значення $o_1 \in D_{O_1, \bigcap_i T_i}, \ldots, o_p \in D_{O_p, \bigcap_i T_i}$, що $s' = \{(X_1, x_1), \ldots, (X_m, x_m), (O_1, o_1), \ldots, (O_p, o_p)\} \in \bigcap_i T_i$, де $\{O_1, \ldots, O_p\} = R - X$. Доведення наведене в [5].

У [4, с. 42] доведено дистрибутивність проекції відносно об'єднання: $\pi_X(\bigcup_i T_i) = \bigcup_i \pi_X(T_i)$.

У [4, с. 42] доведено рівність $\pi_X(\pi_Y(T)) = \pi_{X \bigcap Y}(T)$, з якої безпосередньо випливає ідемпотентність проекції: $\pi_X(\pi_X(T)) = \pi_X(T)$.

У [4, с. 46] доведено критерій переставності селекції та проекції: для будь-якої таблиці T схеми R рівність $\pi_X(\sigma_P(T)) = \sigma_P(\pi_X(T))$ виконується тоді та лише тоді, коли істиність предиката P на множині R співпадає з істинністю P на X.

У [4, с. 90] доведено переставність перейменування та проекції:
$Rn(\pi_X(T))_\xi = \pi_{\xi(X)}(Rn(T)_\xi)$.

У [4, с. 59] доведено включення $\pi_X(T_1 \otimes T_2) \subseteq \pi_X(T_1) \otimes \pi_X(T_2)$. У цій же монографії на ст. 60 наведено критерій, за яким це включення перетворюється у рівність: нехай R_1 та R_2 – схеми таблиць T_1 та T_2 відповідно. Рівність $\pi_X(T_1 \otimes T_2) = \pi_X(T_1) \otimes \pi_X(T_2)$ виконується тоді та лише тоді, коли $R_1 \bigcap R_2 \subseteq X$.

У [4, с. 45] доведено переставність насичення та проекції: $C(\pi_X(T)) = \pi_X(C(T))$.

У [4, с. 45] доведено включення $(\widetilde{\pi_X(T)}) \subseteq \pi_X(\tilde{T})$. Неважко бачити, що при $T = T_\emptyset$ та $T = T_\varepsilon$ це включення перетворюється у рівність. Знайдемо необхідні та достатні умови, за яких це включення перетворюється у рівність у випадку $T \neq T_\emptyset$ та $T \neq T_\varepsilon$. Далі, у теоремі 3 покладемо T – таблиця схеми R, $X = \{X_1, \ldots, X_m\}$; а елементи множини $R - X$ позначимо $O_1, \ldots, O_p$.

Теорема 3. *Рівність* $(\widetilde{\pi_X(T)}) = \pi_X(\tilde{T})$ *виконується тоді та лише тоді, коли для будь-якого рядка* $s = \{(X_1, x_1), \ldots, (X_m, x_m)\} \in \pi_X(T)$ *та будь-яких значень* $o_1 \in D_{O_1,T}, \ldots, o_p \in D_{O_p,T}$ *виконується:* $s' = \{(X_1, x_1), \ldots, (X_m, x_m), (O_1, o_1), \ldots, (O_p, o_p)\} \in T$.

Доведення. Необхідність. Нехай виконується рівність $(\widetilde{\pi_X(T)}) = \pi_X(\tilde{T})$. Від супротивного, припустимо, що критерій не виконується, тобто існують рядок $s = \{(X_1, x_1), \ldots, (X_p, x_p)\} \in \pi_X(T)$ та значення $o_1 \in D_{O_1,T}, \ldots, o_p \in D_{O_p,T}$, що $s' = \{(X_1, x_1), \ldots, (X_p, x_p), (O_1, o_1), \ldots, (O_p, o_p)\} \notin T$. Тоді з $s \in \pi_X(T)$ випливає $s \notin (\widetilde{\pi_X(T)})$. З іншого боку, з $s \in \pi_X(T)$ та $o_1 \in D_{O_1,T}, \ldots, o_p \in D_{O_p,T}$ випливає

$s' \in C(T)$, що, з урахуванням $s' \notin T$, дає $s' \in \tilde{T}$, тому, $s \in \pi_X(\tilde{T})$, тобто рівність $(\widetilde{\pi_X(T)}) = \pi_X(\tilde{T})$ не виконується. Це суперечить припущенню та доводить необхідність.

Доведемо достатність. Нехай виконується критерій. Від супротивного, припустимо, що рівність $(\widetilde{\pi_X(T)}) = \pi_X(\tilde{T})$ не виконується, це є можливим лише при існуванні такого рядка $w = \{(X_1, w_1), \ldots, (X_p, w_p)\}$, що $w \in \pi_X(\tilde{T})$ та $w \notin (\widetilde{\pi_X(T)})$. Тоді, з $w \in \pi_X(\tilde{T})$, за визначенням проекції випливає існування в таблиці $\tilde{T}$ рядка $w' = \{(X_1, w_1), \ldots, (X_p, w_p), (O_1, o_1), \ldots, (O_p, o_p)\}$. За визначенням активного доповнення з $w' \in \tilde{T}$ випливає $w' \in C(T)$, $w' \notin T$ та $o_1 \in D_{O_1,T}, \ldots, o_p \in D_{O_p,T}$. За припущенням критерій виконується, що можливо лише за випадку $w \notin \pi_X(T)$. З $w' \in C(T)$ випливає $w \in \pi_X(C(T))$, а з переставності проекції на насичення випливає $w \in C(\pi_X(T))$, тому, за означенням активного доповнення одержимо $w \in (\widetilde{\pi_X(T)})$. Одержана суперечність доводить хибність припущення про невиконаність рівності $(\widetilde{\pi_X(T)}) = \pi_X(\tilde{T})$ та доводить достатність. □

5. Висновки.

У роботі розглянуто та систематизовано взаємозв'язки різниці та проекції таблиць між собою та з іншими сигнатурними операціями табличних алгебр: перетином, об'єднанням, селекцією, перейменуванням, з'єднанням, насиченням та активним доповненням. Знайдено критерії, за яких деякі взаємозв'язки, що у загальному випадку є включеннями, стають рівностями, а також критерій дистрибутивності насичення відносно різниці таблиць. Ці результати можуть бути використаними для перетворення виразів, що містять табличні операції, задля їх мінімізації або переходу до стандартного вигляду. У подальшому планується розглянути взаємозв'язки між іншими табличними операціями, в першу чергу операції ділення з рештою сигнатурних табличних операцій, а також знаходження ймовірностей виконання знайдених у цій роботі критеріїв для довільних таблиць та таблиць із певними фіксованими параметрами.

Цитована література

1. *Codd E.F.* A Relational Model of Data for Large Shared Data Banks // Communications of the ACM. – 1970. – Vol. 13, Iss. 6. – P. 377–387.
2. *Red'ko V.N., Bui D.B.* Foundations of the theory of relational database models // Cybernetics and System Analysis. – 1996. – Vol. 32, Iss. 4. – P. 471–478.
3. *Garcia-Molina H., Ullman J.D., Widom J.* Database Systems. The Complete Book, 2th ed. Pearson Prentice Hall, 2009. 1203 p.
4. *Редько В.Н., Брона Ю.Й., Буй Д.Б., Поляков С.А.* Реляційні бази даних: табличні алгебри та SQL-подібні мови. Київ: Академперіодика, 2001. 198 с.
5. *Сенченко А.С.* Взаимосвязи между пересечением, объединением и другими сигнатурными операциями в табличных алгебрах // Праці Інституту прикладної математики і механіки НАН України. – 2018. – том 32. – С. 121–132.

References

1. Codd E.F. (1970). A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, *13*(6), 377-387. https://doi.org/10.1145/362384.362685

2. Red'ko V.N. & Bui D.B. (1996). Foundations of the theory of relational database models. *Cybernetics and System Analysis*, *32*(4), 471-478. https://doi.org/10.1007/BF02366768
3. Garcia-Molina H., Ullman J.D. & Widom J. (2009). Database Systems. The Complete Book, 2th ed. Pearson Prentice Hall.
4. Red'ko V.N., Brona Yu. Y., Bui D.B. & Polyakov S.A. (2001). Relational databases: table algebras and SQL-like languages. Kyiv: Academperiodica (in Ukrainian).
5. Senchenko A.S. (2018). Vzaimosvyazi mezhdu peresecheniem, obyedineniem i drugimi signaturnymi operatsiyami v tablichnykh algebrakh. *Proceedings of the Institute of Applied Mathematics and Mechanics of NAS of Ukraine*, *32*, 121-132 (in Russian). https://doi.org/10.37069/1683-4720-2018-32-13

A.S. Senchenko

The interrelations between difference and projection and other signature operations in table algebra.

Currently, databases are widely used in almost all areas of human activity. For all variety of different types of databases the most common are relational (table) databases, mathematical model of which was proposed by E. Codd. From mathematical point of view, a relational database is a finite set of finite relations between different predefined sets of basic data. Table algebra introduced by V.N. Red'ko and D.B. Buy is based on Codd's relational algebra and significantly improves it. It formed the theoretical foundation of modern database query language. Elements of the carrier of table algebra specify relational data structures, and signature operations are based on the basic table manipulations in relational algebra and SQL-like languages. One of the most actual tasks in relational and table algebras is the problem of equivalent transformation of expressions in order to minimize or reduce them to a standard form; it is one of the stages of query optimization, and can also significantly reduce the processing time of information in relational database management systems. For the decision of this problem the interrelations between the basic table operations are used. In the present, a significant number of such interrelations have been established, most of which for the general case are performed as inclusions. The author has found criteria for the transition of some such inclusions into equalities. These criteria are expressed in terms of the active domains of the tables and are natural. In this paper, the interrelations of the difference and the projection of tables with other signature operations of table algebras: intersection, union, selection, renaming of attributes, join, saturation, active complement are considered.

Keywords: *difference, projection, databases, table algebra.*

Інститут прикладної математики і механіки НАН України, Слов'янськ
`Senchenko.a76@gmail.com`

Отримано 01.11.21