УДК 004.655

Глушко І.М., к.ф.-м.н.

Узагальнення теорем Кодда-Лакруа-Піротта

Ніжинський державний університет імені Миколи Гоголя, 16600, м. Ніжин, вул. Кропив'янського 2,
e-mail: glushkoim@gmail.com

I.M. Glushko, PhD

Generalization of Codd-Lacroix-Pirotte's theorems

Nizhyn Gogol State University, 16600, Nizhyn, Kropyiv'yanskoho str, 2,
e-mail: glushkoim@gmail.com

У статті представлено результат, який стосується еквівалентності табличної алгебри для нескінченних таблиць і відповідних реляційних числень на кортежах та доменах. При цьому поняття реляції розглядається в термінах іменних множин, а класичне реляційне числення поповнене функціональними та предикатними сигнатурами на універсальному домені.

Ключові слова: реляційні (табличні) бази даних, числення рядків, числення на домені, таблична алгебра.

In article is presented a result which concerns equivalence of table algebra for infinite tables and corresponding relational calculi. This result generalizes the classical result about the equivalence of Codd's relational algebra and tuple (domain) calculus. Concept of table (relation) is considered in terms of nominal sets. Specifying of table in terms of nominal sets is carried out by V. Redko, J. Brona, D. Buy, S. Poliakov. Traditionally the finite set of tuple is understood under the table and the authors take it into account. However, as a rule, mathematical statements about standard properties of specification of relation operations remain true for infinite relations. Further under relation we will understand any set of tuples (with common scheme), in particular infinite. Furthermore we consider only one universal domain. The classical relational calculi are filled up by functional and predicate signatures on the universal domain (while usually consider only binary predicates and functional signature is generally empty).

Key words: relation databases, tuple calculus, domain calculus, table algebra.

Статтю представив д.ф.-м.н., проф. Буй Д.Б.

Вступ

Системи управління базами даних (СУБД) в наш час використовуються практично в усіх сферах людської діяльності, які пов'язані зі збереженням та обробкою інформації. Прогрес, досягнутий в області технологій баз даних, в значній мірі базується на реляційній моделі, запропонованій Е. Коддом (E. Codd) у 70-х роках XX ст.. Реляційна алгебра розроблена Е. Коддом у вигляді сукупності операторів над таблицями [1]. Вчений також дав формальне визначення реляційного числення на кортежах та вказав алгоритм перетворення довільного виразу реляційного числення на кортежах на семантично еквівалентний йому алгебраїчний вираз [2].

Трохи пізніше М. Лакруа (M. Lacroix) з А. Піроттом (A. Pirotte) (1977 р.) запропонували альтернативну до числення на кортежах версію числення – реляційне числення зі змінними на доменах [3]. Відмінність полягає у тому, що замість змінних рядків розглядаються змінні значення із домену, які представляють компоненти рядка.

Набір операцій реляційної алгебри, запропонований Е. Коддом, з часом був розширений відповідно до потреб мов запитів. Так, Е. Клуг (A. Klug) [4] (1982 р.) розширив реляційну алгебру та реляційне числення агрегатними функціями, які дозволяють знаходити сумарні, середні, максимальні, мінімальні та інші значення елементів у стовпці таблиці, і показав еквівалентність отриманих при цьому двох формальних мов.

Три основні підходи до конструювання мов запитів, а саме реляційна алгебра, реляційне числення на кортежах та реляційне числення на доменах, розглянуто Дж. Ульманом (J.D. Ullman)

(1982) у [5]. Автор обмежує реляційні числення (на кортежах та доменах) для використання тільки скінченних відношень, тобто нескінченні відношення як такі, не розглядаються. При цьому вводяться так звані "безпечні" вирази реляційних числень, що дозволяють працювати лише з таблицями, що мають скінченну кількість рядків. Доведено еквівалентність реляційної алгебри та числень на кортежах і доменах, в яких розгляд обмежується лише безпечними виразами.

До питання еквівалентності реляційної алгебри та реляційних числень також звертається Д. Мейєр (D. Maier) (1983 р.) [6]. Автор розглядає три системи запитів: реляційне числення кортежів, реляційне числення доменів та табло. Показано, що і числення кортежів, і числення доменів еквівалентні по виразній силі реляційній алгебрі. Для кожного з цих числень представлено дві інтерпретації: необмежену та обмежену, і введено клас "безпечних" виразів, для яких обидві інтерпретації співпадають. При доведенні того, що числення кортежів при необмеженій інтерпретації еквівалентне по виразній силі реляційній алгебрі, Д. Мейєр пропонує читачу самостійно довести деякі пункти теореми, а саме для алгебраїчних виразів, які містять операції проекції, об’єднання та різниці, самостійно побудувати відповідні еквівалентні вирази реляційного числення кортежів. Також на самостійне опрацювання залишаються деталі доведення теореми про еквівалентність реляційної алгебри та числення кортежів при обмеженій інтерпретації і леми, в якій говориться про те, що для будь-якого безпечного виразу необмежена та обмежена інтерпретації співпадають. Такі пропуски частин доведень порушують повноту доведення, так як читач не має змоги перевірити правильність своїх міркувань.

Суть реляційного підходу складає поняття реляції. Кодд Е. під реляцією розуміє класичне поняття скінченномісного відношення як підмножини декартового добутку множин, взятих у деякому порядку. Проте таке уточнення має ряд недоліків: інформаційний зміст реляції не залежить від порядку слідування компонент, що для відношення є суттєвим; використання стандартних імен 1, 2, … для доступу до компонент елементів реляції є обтяжливим, тому в монографії [7] були зроблені спроби розширення інтерпретацій реляцій. Редько В.Н. , Брона Ю.Й. ,Буй Д.Б., Поляков С.А. здійснили уточнення реляції в термінах іменних множин [7]. При такому уточненні розглядається один універсальний домен, що дозволяє дослідити загальні властивості табличних маніпуляцій, не вводячи до розгляду несуттєві деталі, та спростити міркування. Традиційно під реляцією розуміється скінченна множина рядків, проте, як правило, математичні твердження про стандартні властивості уточнень реляційних операцій залишаються вірними і для нескінченних реляцій. Так, детальний аналіз доведень у [7] показує, що властивість скінченності таблиць, як правило, не використовується. Тому можна піти далі та розглядати нескінчені реляції. Отже, надалі під реляцією будемо розуміти довільну множину односхемних рядків, зокрема, нескінченну. При такому розгляді постає задача встановлення еквівалентності табличної (реляційної) алгебри та реляційного числення. Вирішення цієї проблеми і пропонується у даній роботі.

Таблична алгебра нескінченних таблиць та узагальнені реляційні числення

Розглянемо дві множини: $\boldsymbol{A}$ – множину атрибутів і $\boldsymbol{D}$ – універсальний домен. Довільну (скінченну) множину атрибутів $R \subseteq \boldsymbol{A}$ назвемо *схемою*. *Рядком схеми* R називається іменна множина на парі R, $\boldsymbol{D}$, проекція якої за першою компонентою рівна R (тобто по суті розглядається функція вигляду $s : R \to \boldsymbol{D}$). Під *таблицею схеми* R розуміємо пару $\langle t, R \rangle$, де t – множина (зокрема, нескінченна) рядків вказаної схеми R. Отже, кожній таблиці приписується певна схема. Це по суті впливає тільки на випадок порожньої таблиці, оскільки за непорожньою таблицею схема відновлюється однозначно. Також внесення схеми в уточнення реляції необхідне при доведенні теорем про еквівалентність табличної алгебри та реляційного числення.

Під *табличною алгеброю нескінченних таблиць* розуміємо алгебру $\langle \boldsymbol{T}, \Omega_{P,\Xi} \rangle$, де $\boldsymbol{T}$ – множина усіх таблиць, $\Omega_{P,\Xi} = \{\bigcup_R, \bigcap_R, \setminus_R, \sigma_{p,R}, \pi_{X,R}, \underset{R_1,R_2}{\otimes}, \div_{R_2}^{R_1}, Rt_{\xi,R}, \sim_R\}$ – сигнатура, $p \in P, \xi \in \Xi$, $X, R, R_1, R_2 \subseteq \boldsymbol{A}$, P, Ξ – множини параметрів.

Означення операцій сигнатури табличної алгебри нескінченних таблиць можна знайти в [8].

Розглянемо узагальнене числення Кодда на кортежах – узагальнене числення рядків (УЧР). Так як у класичних реляційних численнях (Кодда на кортежах та Лакруа-Піротта на доменах), зазвичай, розглядають лише бінарні предикати, а функціональна сигнатура взагалі порожня, то виникає потреба у поповненні цих числень предикатними та функціональними сигнатурами на універсальному домені, що і зроблено у [8].

Вираз узагальненого числення рядків має вигляд $\{\boldsymbol{x}(R) \mid \boldsymbol{P}(\boldsymbol{x})\}$, де

1) формула $\boldsymbol{P}$ – дозволена;

2) змінна $\boldsymbol{x}$ – єдина змінна, яка входить у формулу $\boldsymbol{P}$ вільно;

3) якщо $scheme(\boldsymbol{x},\boldsymbol{P})$ визначена, то $scheme(\boldsymbol{x},\boldsymbol{P}) = R$, інакше, $attr(\boldsymbol{x},\boldsymbol{P}) \subseteq R$, де $scheme(\boldsymbol{x},\boldsymbol{P})$ – це схема рядка $\boldsymbol{x}$ у формулі $\boldsymbol{P}$, $attr(\boldsymbol{x},\boldsymbol{P})$ – це множина атрибутів, з якими рядок $\boldsymbol{x}$ зустрічається у формулі $\boldsymbol{P}$.

Дозволеність формули – це вимога погодження схем рядка у підформулах та вимога того, щоб змінна, зв'язана квантором, входила вільно у формулу, яка слідує за квантором.

Реляційне числення зі змінними на доменах відрізняється від числення рядків тим, що замість змінних рядків розглядаються змінні значення із доменів, які представляють компоненти рядка. Оскільки в даній роботі розглядається лише один універсальний домен, то будемо говорити про узагальнене числення на домені (УЧД).

Вираз числення на домені має вигляд $\{\boldsymbol{x}_1^{A_1},...,\boldsymbol{x}_n^{A_n} \mid \boldsymbol{P}(\boldsymbol{x}_1^{A_1},...,\boldsymbol{x}_n^{A_n})\}$, де

1) формула $\boldsymbol{P}$ – дозволена, а $\boldsymbol{x}_1^{A_1},...,\boldsymbol{x}_n^{A_n}$ – всі вільні змінні, які входять у формулу $\boldsymbol{P}$, причому кожна змінна $\boldsymbol{x}_i^{A_i}$ асоційована з атрибутом A_i, де всі A_i попарно різні, $i = 1,\dots,n$;

2) $R = \{A_1,...,A_n\}$ – схема виразу.

Поняття дозволеної формули УЧД визначається аналогічно як і в УЧР.

Еквівалентність табличної алгебри нескінченних таблиць та відповідних числень

Доведення еквівалентності табличної алгебри нескінченних таблиць та відповідних числень проведено наступним чином: спочатку доведемо, що для довільного виразу табличної алгебри нескінченних таблиць можна побудувати еквівалентний вираз УЧР, потім покажемо, що для виразу УЧР можна побудувати еквівалентний вираз УЧД і, нарешті, доведемо, що для будь-якого виразу УЧД можна побудувати еквівалентний вираз табличної алгебри нескінченних таблиць. Вираз УЧР E еквівалентний виразу табличної алгебри нескінченних таблиць F, якщо значення виразу E співпадає із значенням виразу F за умови, що табличні константи і табличні змінні інтерпретуються однаково в цих двох виразах.

Теорема 1. Якщо F – вираз табличної алгебри нескінченних таблиць, то можна ефективно побудувати еквівалентний йому вираз E узагальненого числення рядків.

Доведення. При доведенні теореми розглядаємо вирази табличної алгебри нескінчених таблиць, які містять тільки операції об'єднання, різниці, селекції, проекції, з'єднання та перейменування, оскільки операції перетину, ділення та активного доповнення можна виразити через вказані операції [8]. Доведення проведемо індукцією за числом операцій в F.

База індукції (немає операцій). Можливі два випадки. По-перше, $F = \langle t, R\rangle$ – стала таблиця, де t – нескінченна множина рядків схеми R. Покладемо $E = \{\boldsymbol{x}(R) \mid t_R(\boldsymbol{x})\}$. По-друге, $F = \langle X, R\rangle$ – змінна таблиця. У даному випадку покладемо $E = \{\boldsymbol{x}(R) \mid X_R(\boldsymbol{x})\}$.

Крок індукції. Припустимо, що твердження теореми виконується для будь-якого виразу табличної алгебри нескінченних таблиць, який містить менше $k = 1,2,3...$ операцій. Нехай вираз F містить k операцій.

1. $F = F_1 \cup_R F_2$. Тоді вирази F_1 і F_2 мають менше k операцій, тому існують, згідно індуктивного припущення, вирази УЧР $\{\boldsymbol{x}(R) \mid \boldsymbol{P}(\boldsymbol{x})\}$ і $\{\boldsymbol{x}(R) \mid \boldsymbol{Q}(\boldsymbol{x})\}$, еквівалентні виразам F_1 і F_2 відповідно. Покладемо $E = \{\boldsymbol{x}(R) \mid \boldsymbol{P}(\boldsymbol{x}) \vee \boldsymbol{Q}(\boldsymbol{x})\}$.

2. $F = F_1 \setminus_R F_2$. Тоді, як і у випадку 1, існують вирази УЧР $\{\boldsymbol{x}(R) \mid \boldsymbol{P}(\boldsymbol{x})\}$ і $\{\boldsymbol{x}(R) \mid \boldsymbol{Q}(\boldsymbol{x})\}$, еквівалентні виразам F_1 і F_2 відповідно. Покладемо $E = \{\boldsymbol{x}(R) \mid \boldsymbol{P}(\boldsymbol{x}) \wedge \neg\boldsymbol{Q}(\boldsymbol{x})\}$.

3. $F = \sigma_{\tilde{p},R}(F_1)$. Нехай $\{\boldsymbol{x}(R) \mid \boldsymbol{P}(\boldsymbol{x})\}$ – вираз УЧР, еквівалентний виразу F_1. Покладемо $E = \{\boldsymbol{x}(R) \mid \boldsymbol{P}(\boldsymbol{x}) \wedge \boldsymbol{p}(\boldsymbol{x}(A_1),...,\boldsymbol{x}(A_m))\}$, де $R = \{A_1,...,A_m\}$ схема таблиці, яка є значенням виразу F_1. Тут припускається, що предикат-параметр селекції заданий так: $\tilde{p}(s) = true \Leftrightarrow \boldsymbol{p}(s(A_1),...,s(A_m)) = true$, $s \in S(R)$, де $\boldsymbol{p}$ – m-арний сигнатурний предикатний символ.

4. $F = \pi_{X,R}(F_1)$. Нехай $\{\boldsymbol{x}(R) \mid \boldsymbol{P}(\boldsymbol{x})\}$ – вираз УЧР, еквівалентний виразу F_1. Покладемо $E = \{\boldsymbol{y}(X \cap R) \mid \exists \boldsymbol{x}(R)(\boldsymbol{P}(\boldsymbol{x}) \wedge \bigwedge_{A \in X \cap R} \boldsymbol{y}(A) = \boldsymbol{x}(A))\}$.

Особливий випадок: якщо $X \bigcap R = \varnothing$, то $E = \{\boldsymbol{y}(\varnothing) \mid \exists \boldsymbol{x}(R)\boldsymbol{P}(\boldsymbol{x})\}$.

5. $F = F_1 \underset{R_1,R_2}{\otimes} F_2$. Нехай $\{\boldsymbol{x}(R_1) \mid \boldsymbol{P}(\boldsymbol{x})\}$ і $\{\boldsymbol{y}(R_2) \mid \boldsymbol{Q}(\boldsymbol{y})\}$ – вирази УЧР, еквівалентні виразам F_1 та F_2 відповідно. Покладемо $E = \{\boldsymbol{z}(R_1 \bigcup R_2) \mid \exists \boldsymbol{x}(R_1)\exists \boldsymbol{y}(R_2)(\boldsymbol{P}(\boldsymbol{x}) \wedge$ $\wedge \boldsymbol{Q}(\boldsymbol{y}) \wedge \underset{A\in R_1}{\wedge} \boldsymbol{z}(A) = \boldsymbol{x}(A) \wedge \underset{A\in R_2}{\wedge} \boldsymbol{z}(A) = \boldsymbol{y}(A))\}$.

6. $F = Rt_{\xi,R}(F_1)$, де $\xi : \boldsymbol{A} \tilde{\to} \boldsymbol{A}$ – ін'єктивна (взагалі кажучи, часткова) функція, що здійснює перейменування атрибутів. Тоді, згідно індуктивного припущення, існує вираз УЧР $\{\boldsymbol{x}(R) \mid \boldsymbol{P}(\boldsymbol{x})\}$, еквівалентний виразу F_1. Покладемо $E = \{\boldsymbol{y}(R_2) \mid \exists \boldsymbol{x}(R)(\boldsymbol{P}(\boldsymbol{x}) \wedge$ $\underset{C\in R\setminus dom\xi}{\wedge} \boldsymbol{y}(C) = \boldsymbol{x}(C) \wedge \underset{A\in R\bigcap dom\xi}{\wedge} \boldsymbol{x}(A) = \boldsymbol{y}(\xi(A)))\}$, де $R_2 = (R \setminus dom\xi) \bigcup \xi[R]$ (зауважимо, що це схема перейменованої таблиці).

Теорема 1 встановлює, що УЧР не менш виразне, ніж таблична алгебра нескінченних таблиць (використовуючи термінологію [6]).

Покажемо, що для виразу УЧР можна побудувати еквівалентний вираз УЧД. Задамо відображення $\varphi : E \to H$, яке кожному виразу УЧР E ставить у відповідність еквівалентний вираз УЧД H.

Нехай $E = \{\boldsymbol{y}(R) \mid P(\boldsymbol{y})\}$ – вираз УЧР, змінна $\boldsymbol{y}$ – єдина змінна, яка входить у формулу $\boldsymbol{P}$ вільно, а $R = \{A_1,...,A_n\}$. Використовуємо $\boldsymbol{y}$ як синтаксичну змінну, областю зміни якої є змінні УЧР. Застосуємо спочатку відображення φ до термів УЧР:

1) кожну предметну константу $\boldsymbol{d}$ УЧР замінимо на константу $\boldsymbol{d}^A$ УЧД, де $A \in R$ – атрибут асоційований з даною константою;

2) кожен терм $\boldsymbol{x}(A_i)$ УЧР замінимо змінною $\boldsymbol{x}_i^{A_i}$ УЧД;

3) кожен терм УЧР $\boldsymbol{f}(\boldsymbol{v}_1,...,\boldsymbol{v}_n)$, де $\boldsymbol{v}_i$ – терми УЧР, $i = 1,...,n$, замінимо на терм УЧД $\boldsymbol{f}(\boldsymbol{u}_1,...,\boldsymbol{u}_n)$, де $\boldsymbol{u}_j$ – терми УЧД, $j = 1,...,n$, отримані в результаті попередніх замін.

Застосуємо відображення φ до атомів УЧР:

а1) кожен атом $t_{R'}(\boldsymbol{z})$ УЧР замінимо на атом УЧД $t_{R'}(\boldsymbol{z}_1^{A_1},\ldots,\boldsymbol{z}_m^{A_m})$, де $\boldsymbol{z}_1^{A_1},\ldots,\boldsymbol{z}_m^{A_m}$ – змінні УЧД, $R' = \{A_1,...,A_m\}$ – схема сталої таблиці $\langle t, R' \rangle$;

а2) кожен атом $X_{R'}(\boldsymbol{z})$ УЧР замінимо на атом УЧД $X_{R'}(\boldsymbol{z}_1^{A_1},\ldots,\boldsymbol{z}_m^{A_m})$, де $\boldsymbol{z}_1^{A_1},\ldots,\boldsymbol{z}_m^{A_m}$ – змінні УЧД, $R' = \{A_1,...,A_m\}$ – схема змінної таблиці $\langle X, R' \rangle$;

а3) кожен атом УЧР $\boldsymbol{p}(\boldsymbol{v}_1,...,\boldsymbol{v}_m)$, де $\boldsymbol{v}_i$ – терми УЧР, $i = 1,...,m$, замінимо на атом числення на домені $\boldsymbol{p}(\boldsymbol{u}_1,...,\boldsymbol{u}_m)$, де $\boldsymbol{u}_j$ – терми УЧД, $j = 1,...,m$, отримані в результаті попередніх замін.

Кожну формулу $\boldsymbol{P}$ УЧР перетворимо на формулу $\boldsymbol{P}'$ УЧД, де формула $\boldsymbol{P}'$ – це формула $\boldsymbol{P}$, у якій для кожного атома виконані заміни а1-а3, а кожна вільна змінна $\boldsymbol{z}$ схеми $R = \{A_1,...,A_n\}$ УЧР замінена n змінними $\boldsymbol{z}_1^{A_1},...,\boldsymbol{z}_n^{A_n}$ УЧД.

- Формулу з квантором УЧР $\exists \boldsymbol{z}(R_2)\boldsymbol{G}$, $R_2 = \{A_1,...,A_m\}$ перетворимо на формулу УЧД вигляду $\exists \boldsymbol{z}_1^{A_1}(A_1)...\exists \boldsymbol{z}_m^{A_m}(A_m)\boldsymbol{G}'$.
- Формулу з квантором УЧР $\forall \boldsymbol{z}(R_2)\boldsymbol{G}$, $R_2 = \{A_1,...,A_m\}$ перетворимо на формулу УЧД вигляду $\forall \boldsymbol{z}_1^{A_1}(A_1)...\forall \boldsymbol{z}_m^{A_m}(A_m)\boldsymbol{G}'$.

Виконавши вказані заміни, отримаємо вираз УЧД $H = \{\boldsymbol{y}_1^{A_1},...,\boldsymbol{y}_n^{A_n} \mid \boldsymbol{P}(\boldsymbol{y}_1^{A_1},...,\boldsymbol{y}_n^{A_n})\}$. В результаті проведених замін очевидно, що кожна змінна $\boldsymbol{z}_i^{A_i}$ УЧД може приймати точно ті самі значення, що і $\boldsymbol{z}(A_i)$. Отже, значення виразу H співпадає із значенням виразу E. Тому має місце наступна теорема.

Теорема 2. Якщо E – вираз узагальненого числення рядків, то можна ефективно побудувати еквівалентний йому вираз H узагальненого числення на домені.

Таким чином, УЧД є не менш виразним, ніж УЧР.

Теорема 3. Для кожного виразу узагальненого числення на домені H можна ефективно побудувати еквівалентний йому вираз F табличної алгебри нескінченних таблиць.

Доведення. Нехай $H = \{\boldsymbol{x}_1^{A_1},...,\boldsymbol{x}_n^{A_n} \mid \boldsymbol{P}(\boldsymbol{x}_1^{A_1},...,\boldsymbol{x}_n^{A_n})\}$ – вираз УЧД, значенням якого є таблиця схеми $R = \{A_1,...,A_n\}$. Побудуємо для кожної підформули $\boldsymbol{G}$ формули $\boldsymbol{P}$ еквівалентний табличний вираз F_G.

Доведемо індукцією по числу операторів у підформулі $\boldsymbol{G}$ з $\boldsymbol{P}$, що якщо $\boldsymbol{G}$ містить вільні змінні на універсальному домені $\boldsymbol{y}_1^{A_1},...,\boldsymbol{y}_m^{A_m}$, то для виразу $\{\boldsymbol{y}_1^{A_1},...,\boldsymbol{y}_m^{A_m} \mid \boldsymbol{G}(\boldsymbol{y}_1^{A_1},...,\boldsymbol{y}_m^{A_m})\}$ існує еквівалентний вираз табличної алгебри F_G. Тоді, якщо $\boldsymbol{G}$ є сама $\boldsymbol{P}$, отримаємо табличний вираз для $\{\boldsymbol{x}_1^{A_1},...,\boldsymbol{x}_n^{A_n} \mid \boldsymbol{P}(\boldsymbol{x}_1^{A_1},...,\boldsymbol{x}_n^{A_n})\}$. Припускається, що $\boldsymbol{y}_1^{A_1},...,\boldsymbol{y}_m^{A_m}$ – всі вільні змінні в $\boldsymbol{G}$, і значенням виразу $\{\boldsymbol{y}_1^{A_1},...,\boldsymbol{y}_m^{A_m} \mid \boldsymbol{G}(\boldsymbol{y}_1^{A_1},...,\boldsymbol{y}_m^{A_m})\}$ є таблиця схеми $R_G = \{A_1,...,A_m\}$.

Замінимо змінні у формулі $\boldsymbol{P}$ таким чином, щоб жодна змінна не входила одночасно і вільно, і зв'язано у $\boldsymbol{P}$ та не була зв'язана квантором у двох різних місцях. Тепер кожна змінна асоційована з атрибутом, який або входить у квантор $\forall \boldsymbol{x}^A(A)$ чи $\exists \boldsymbol{x}^A(A)$, якщо змінна зв'язана у формулі $\boldsymbol{P}$, або, якщо змінна вільна в $\boldsymbol{P}$, співставленим змінній $\boldsymbol{x}_i^{A_i}$, що знаходиться у виразі H зліва від вертикальної лінії (тобто в останньому випадку асоційований атрибут – це атрибут, співставлений змінній $\boldsymbol{x}_i^{A_i}$). Будь-якому атрибуту A співставимо константну таблицю $[\boldsymbol{D}]_A$ схеми $\{A\}$. Значенням цієї таблиці є одноатрибутна таблиця $\langle t, \{A\}\rangle$, яка містить всі рядки $s = \{\langle A, d\rangle\}$, де $d \in \boldsymbol{D}$ – довільний елемент множини $\boldsymbol{D}$.

База індукції (немає операторів, тобто логічних зв'язок і кванторів). Підформула $\boldsymbol{G}$ – атом вигляду $\boldsymbol{p}(\boldsymbol{u}_1,...,\boldsymbol{u}_m)$, або $t_R(\boldsymbol{a}_1^{A_1},...,\boldsymbol{a}_m^{A_m})$, або $X_R(\boldsymbol{z}_1^{A_1},\ldots,\boldsymbol{z}_m^{A_m})$.

1. Нехай $\boldsymbol{G}$ має вигляд $\boldsymbol{p}(\boldsymbol{u}_1,...,\boldsymbol{u}_m)$, де $\boldsymbol{u}_i$ – терми УЧД, причому $\boldsymbol{y}_1^{A_1},...,\boldsymbol{y}_k^{A_k}$ – всі змінні цих термів, а $A_1,\ldots,A_k$ – атрибути співставлені цим змінним. Покладемо $F_G = \sigma_{\widetilde{p},\{A_1,\ldots,A_k\}}\left([\boldsymbol{D}]_{A_1} \underset{R_1,R_2}{\otimes} \cdots \underset{R_1\cup\ldots\cup R_{k-1},R_k}{\otimes} [\boldsymbol{D}]_{A_k}\right)$, де $R_i = \{A_i\}$, $i = 1,\ldots,k$. Це робиться в припущенні, що предикат-параметр селекції заданий наступним чином $\widetilde{p}(s) = true \Leftrightarrow \boldsymbol{p}(s(A_{k_1}),...,s(A_{k_m})) = true$, $s \in S(R)$, де $\boldsymbol{p}$ – m-арний сигнатурний предикатний символ.

2. Нехай $\boldsymbol{G}$ має вигляд $t(\boldsymbol{a}_1^{A_1},...,\boldsymbol{a}_m^{A_m})$, де $\boldsymbol{a}_i^{A_i}$ – константа або змінна на універсальному домені $\boldsymbol{D}$. Нехай $R = \{A_1,...,A_m\}$ – схема сталої таблиці $\langle t, R\rangle$. Покладемо $F_G = \pi_{Y,R}\left(\sigma_{\widetilde{p},R}\left(\langle t, R\rangle\right)\right)$, де $\widetilde{p}$ – предикат-параметр селекції, який є кон'юнкцією рівностей вигляду $A_i = \boldsymbol{a}_i^{A_i}$ по всіх i, для яких $\boldsymbol{a}_i^{A_i}$ – константа; крім того $Y = \{A_j \mid \boldsymbol{a}_j^{A_j}$ – змінна$\}$.

3. Нехай $\boldsymbol{G}$ має вигляд $X(\boldsymbol{a}_1^{A_1},...,\boldsymbol{a}_m^{A_m})$, де $\boldsymbol{a}_i^{A_i}$ – константа або змінна на універсальному домені $\boldsymbol{D}$. Нехай $R = \{A_1,...,A_m\}$ – схема змінної таблиці $\langle X, R\rangle$. Покладемо $F_G = \pi_{Y,R}\left(\sigma_{\widetilde{p},R}\left(\langle X, R\rangle\right)\right)$, де $\widetilde{p}$ – предикат-параметр селекції, який є кон'юнкцією рівностей вигляду $A_i = \boldsymbol{a}_i^{A_i}$ по всіх i, для яких $\boldsymbol{a}_i^{A_i}$ – константа; крім того $Y = \{A_j \mid \boldsymbol{a}_j^{A_j}$ – змінна$\}$.

Крок індукції. Припустимо, що підформула $\boldsymbol{G}$ містить принаймні один оператор і що гіпотеза індукції справедлива для всіх підформул формули $\boldsymbol{P}$, що мають менше операторів, ніж $\boldsymbol{G}$.

1. $\boldsymbol{G} = \neg\boldsymbol{Q}$. Нехай F_Q – табличний вираз для $\boldsymbol{Q}$, а значенням виразу F_Q є таблиця схеми R_Q. Покладемо $F_G = \sim_{R_Q} F_Q$.

2. $\boldsymbol{G} = \boldsymbol{Q} \vee \boldsymbol{Q}'$. Нехай вільні змінні з $\boldsymbol{Q}$ – це $z_1^{B_1},...,z_k^{B_k}, v_1^{C_1},...,v_p^{C_p}$, а з $\boldsymbol{Q}'$ – це $z_1^{B_1},...,z_k^{B_k}, w_1^{K_1},...,w_q^{K_q}$, де змінні $v_1^{C_1},...,v_p^{C_p}$ та $w_1^{K_1},...,w_q^{K_q}$ попарно не співпадають. Змістовно кажучи, $z_1^{B_1},...,z_k^{B_k}$ – всі спільні змінні виразів $\boldsymbol{Q}$ та $\boldsymbol{Q}'$. Нехай F_Q та $F_{Q'}$ – табличні вирази для $\boldsymbol{Q}$ і $\boldsymbol{Q}'$ відповідно, причому C_i, $i = 1,..., p$ – атрибути, асоційовані зі змінними $v_1^{C_1},...,v_p^{C_p}$, а K_j, $j = 1,...,q$ – атрибути асоційовані із змінними $w_1^{K_1},...,w_q^{K_q}$. Покладемо $F_1 = F_Q \underset{R_Q,\{K_1\}}{\otimes} [\boldsymbol{D}]_{K_1} \underset{R_Q\cup\{K_1\},\{K_2\}}{\otimes} \cdots \underset{R_Q\cup\{K_1,...,K_{q-1}\},\{K_q\}}{\otimes} [\boldsymbol{D}]_{K_q}$, а $F_2 = F_{Q'} \underset{R_{Q'},\{C_1\}}{\otimes} [\boldsymbol{D}]'_{C_1} \underset{R_{Q'}\cup\{C_1\},\{C_2\}}{\otimes} \cdots \underset{R_{Q'}\cup\{C_1,...,C_{p-1}\},\{C_p\}}{\otimes} [\boldsymbol{D}]'_{C_p}$, де R_Q та $R_{Q'}$ – схеми таблиць, які є значеннями табличних виразів F_Q та $F_{Q'}$ відповідно. Позначимо через R_{F_1} та R_{F_2} – схеми таблиць, які є значеннями табличних виразів F_1 та F_2 відповідно. За побудовою виразів F_1 та F_2 їхні схеми R_{F_1} та R_{F_2} рівні. Тому покладемо $F_G = F_1 \cup_{R_{F_1}} F_2$.

3. $\boldsymbol{G} = \boldsymbol{Q} \wedge \boldsymbol{Q}'$. Даний вираз УЧД можна виразити через операції заперечення та

диз'юнкції. За законом де Моргана $\boldsymbol{G} = \neg(\neg\boldsymbol{Q} \vee \neg\boldsymbol{Q}')$.

4. $\boldsymbol{G} = \exists \boldsymbol{x}^{A}(A)\boldsymbol{Q}$. Нехай F_Q – табличний вираз для $\boldsymbol{Q}$. Покладемо F_G рівним $\pi_{X \setminus \{A\}, X}(F_Q)$, де X – схема таблиці, яка є значенням табличного виразу F_Q.

5. $\boldsymbol{G} = \forall \boldsymbol{x}^{A}(A)\boldsymbol{Q}$. Даний вираз УЧД можна виразити через операцію заперечення та квантор існування $\forall \boldsymbol{x}^{A}(A)\boldsymbol{Q} = \neg(\exists \boldsymbol{x}^{A}(A)(\neg\boldsymbol{Q}))$.

Отже, доведено, що таблична алгебра нескінченних таблиць є не менш виразною, ніж УЧД.

Враховуючи теореми 1, 2 та 3, встановлено основний результат.

Теорема 4. Таблична алгебра нескінченних таблиць, узагальнене числення рядків і узагальнене числення на домені еквівалентні.

Висновки

У статті подано огляд літератури з питання еквівалентності реляційної алгебри, числення на кортежах та числення на доменах. Запропоновано узагальнення теорем Кодда-Лакруа-Піротта для випадку нескінченних таблиць. При цьому можна піти далі, обмежуючи розгляд лише скінченними таблицями, та показати еквівалентність отриманих при цьому формалізмів.

.

Список використаних джерел

1. *Codd E.F.* A Relational Model of Data for Large Shared Data Banks / E.F. Codd // Comm. of ACM. – 1970.– 13, № 6. – P. 377-387.
2. *Codd E.F.* Relational Completeness of Data Base Sublanguages / E.F. Codd // Data Base Systems. – New York: Prentice-Hall. – 1972. – P. 65-93.
3. *Lacroix M.* Domain-oriented Relational Languages / M. Lacroix, A. Pirotte // Proc. 3rd Int. Conf. on Very Large Data Bases. – Tokyo, October, 1977. – P. 370-378.
4. *Klug A.* Equivalence of Relational Algebra and Relational Calculus Query Languages Having Aggregate Functions / A. Klug // J. ACM 29. – July 1982. – № 3. – P. 699-717.
5. *Ullman J.D.* Principles of database systems / J.D. Ullman. – Rockville, Maryland: Computer Science Press, 1982. – 484 p.
6. *Maier D.* The theory of relational databases / D. Maier. – Rockville, Maryland: Computer Science Press, 1983. – 637 p.
7. 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
8. *Глушко І.М.* Числення та розширення сигнатур табличних алгебр: дис. … кандидата фіз.-мат. наук: 01.05.01 / І.М. Глушко. – К., 2012. – 142с.

References

1. CODD, E.F. (1970) A Relational Model of Data for Large Shared Data Banks. *Comm. of ACM*. 13(6). pp. 377-387.
2. CODD, E.F. (1972) Relational Completeness of Data Base Sublanguages. *Data Base Systems*. New York: Prentice-Hall. pp. 65-93.
3. LACROIX, M. and PIROTTE, A. (1977) Domain-oriented Relational Languages. In *3rd Int. Conf. on Very Large Data Bases*. Tokyo, 6-8 October. Tokyo: IEEE Computer Society. pp. 370-378.
4. KLUG, A. (1982) Equivalence of Relational Algebra and Relational Calculus Query Languages Having Aggregate Functions *J. ACM*. 29(3). pp. 699-717.
5. ULLMAN, J.D. (1982) Principles of database systems. Rockville, Maryland: Computer Science Press.
6. MAIER, D. (1983) The theory of relational databases. Rockville, Maryland: Computer Science Press.
7. REDKO, V.N., BRONA, YU.J., BUY, D.B. AND POLIAKOV, S.A. (2001) Reliatsiini bazy danykh: tablychni alhebry ta SQL-podibni movy. Kyiv: Vydavnychyi dim «Akademperiodyka».
8. GLUSHKO, I.M. (2012) Chyslennia ta rozshyrennia sygnatur tablychnykh algebr. A Thesis for the Degree of kandydat fizyko-matematychnykh nauk. Kyiv: Kyivskyi natsionalnyi universytet imeni Tarasa Shevchenka.

Надійшла до редколегії 11.08.2014