УДК 004.655

Д.Б. БУЙ, Ю.О. ГРИШКО

Київський національний університет імені Тараса Шевченка, Україна

ЗАСТОСУВАННЯ ТЕОРІЇ МУЛЬТИМНОЖИН

У даній статті розглянуті питання застосування теорії мультимножин в табличних базах даних та ДНК-обчисленнях. Наведені основні означення теорії мультимножин. Введені операції над мультимножинами та досліджено основні властивості цих операцій. Побудовано решітку мультимножин, яку вкладено у дві повні решітки. Уточнено обчислюваність на мультимножинах. Також у статті продемонстровано застосування отриманих результатів у сучасних СУБД, при побудові денотаційної семантики рекурсивної форми CTE-виразів у сучасних SQL-подібних мовах, обчисленнях на ДНК у біоінформатиці.

Ключові слова: *мультимножина, решітка, обчислюваність, CTE, SQL, ДНК-обчислення.*

Вступ

Існує ряд практичних задач інформатики та її застосувань, особливістю яких є множинність і повторюваність даних [1 – 6].

Репрезентативними прикладами таких задач є: уточнення таблиць із дублікатами рядків та маніпуляцій над таблицями у сучасних табличних (реляційних) базах даних; уточнення обчислень на ДНК; конкурсний відбір пропозицій на виконання робіт, заснований на багатокритеріальних оцінках заявок, зроблених декількома експертами; діагностика захворювань із подібними симптомами, при наявності висновків різних лікарів; соціологічні опитування різних груп населення стосовно певних проблем. Зручною моделлю для подання та дослідження багатокритеріальних об'єктів є мультимножини – змістовно кажучи, сукупності з повтореннями.

Можливість присутності у мультимножині довільного, у тому числі й нескінченного, числа елементів і довільної скінченної кількості екземплярів кожного з елементів принципово відрізняє мультимножину від множини. Ця особливість породжує цілий ряд властивостей мультимножини як якісно нового математичного об'єкту [7, 8].

Мета статті – розглянути питання застосування теорії мультимножин в табличних базах даних та ДНК-обчисленнях.

1. Основні означення

Наведемо формальне означення мультимножини. Для цього позначимо: $\mathrm{N} \overset{\mathrm{def}}{=} \{0,1,2,...\}$ – множина натуральних чисел з нулем і $\mathrm{N}^{+} \overset{\mathrm{def}}{=} \{1,2,...\}$ – множина натуральних чисел без нуля.

Нехай U – деяка множина (в класичному канторівському розумінні). Тоді *мультимножина* α *з основою* U – це функція вигляду: $\alpha : \mathrm{U} \to \mathrm{N}^{+}$. Основу мультимножини α будемо позначати як U_{α}.

Мультимножина називається *порожньою* і позначається як $\varnothing_{\mathrm{m}}$, якщо її основа – порожня множина. Мультимножини, областю значень яких є одноелементна множина виду $\{1\}$, назвемо *1-мультимножинами*.

Зафіксуємо D – універсум елементів основ мультимножин; тоді булеан $\mathrm{P(D)}$ є універсумом основ мультимножин.

Нехай задана мультимножина α з основою $\mathrm{U} = \mathrm{dom}\,\alpha$. Зазначимо, що $\mathrm{dom}\,\alpha$ – область означеності мультимножини як функції.

Характеристичною функцією мультимножини α називається функція вигляду $\chi_{\alpha} : \mathrm{D} \to \mathrm{N}$, значення якої задається наступною кусковою схемою:

$$\chi_{\alpha}(\mathrm{d}) = \begin{cases} \alpha(\mathrm{d}), \text{ якщо } \mathrm{d} \in \mathrm{dom}\,\alpha, \\ 0, \text{ інакше}; \end{cases}$$

для усіх $\mathrm{d} \in \mathrm{D}$ [12].

Мультимножини α і β називаються *рівними*, якщо вони рівні як функції.

Мультимножина β *включається у мультимножину* α ($\beta \preceq \alpha$), якщо:

$$\beta \preceq \alpha \overset{\mathrm{def}}{\Leftrightarrow} \mathrm{U}_{\beta} \subseteq \mathrm{U}_{\alpha} \;\&\; \forall \mathrm{d}\left(\mathrm{d} \in \mathrm{U}_{\beta} \Rightarrow \beta(\mathrm{d}) \le \alpha(\mathrm{d})\right).$$

Безпосередньо з означення випливає, що бінарне відношення включення мультимножин є відношенням часткового порядку.

Над мультимножинами визначено аналоги стандартних теоретико-множинних операцій: об'єднання, перетину, різниці, симетричної різниці, доповнення, прямого з'єднання. Крім цього, над мультимножинами визначено операції, що використовують специфіку мультимножин, і тому незастосовні до абстрактних множин: додавання, добуток, множення числа на мультимножину (табл. 1).

Таблиця 1

Основні операції над мультимножинами

Назва операції	Позначення	Характеристична функція мультимножини результату
Об'єднання	$\cup_{\mathrm{All}}$	$\max(\chi_\alpha(d), \chi_\beta(d))$
	$\cup_1$	$sg(\max(\chi_\alpha(d), \chi_\beta(d)))$
Перетин	$\cap_{\mathrm{All}}$	$\min(\chi_\alpha(d), \chi_\beta(d))$
	$\cap_1$	$sg(\min(\chi_\alpha(d), \chi_\beta(d)))$
Додавання	$+_{\mathrm{All}}$	$\chi_\alpha(d) + \chi_\beta(d)$
	$+_1$	$sg(\chi_\alpha(d) + \chi_\beta(d))$
Різниця	$\backslash_{\mathrm{All}}$	$\chi_\alpha(d) \dot{-} \chi_\beta(d)$
	$\backslash_1$	$sg(\chi_\alpha(d)) \dot{-} sg(\chi_\beta(d))$
Симетрична різниця	Δ_{All}	$\lvert \chi_\alpha(d) - \chi_\beta(d) \rvert$
	Δ_1	$\lvert sg(\chi_\alpha(d)) - sg(\chi_\beta(d)) \rvert$
Добуток	$\bullet_{\mathrm{All}}$	$\chi_\alpha(d) \cdot \chi_\beta(d)$
	$\bullet_1$	$sg(\chi_\alpha(d) \cdot \chi_\beta(d))$
Множення числа на мультимножину	$k \cdot \alpha$	$k \cdot \chi_\alpha(d)$
Пряме з'єднання	$\otimes$	$\chi(\langle d_1, d_2 \rangle) = \chi_\alpha(d_1) \cdot \chi_\beta(d_2), \quad \forall d_1, d_2 \in D$; основа вихідної мультимножини – декартовий добуток основ мультимножин аргументів $U_\alpha \times U_\beta$
Доповнення	$\overline{\alpha}$	$\chi_{\overline{\alpha}}(d) = \chi_Z(d) \dot{-} \chi_\alpha(d), \quad \forall d \in D$, $\overline{\alpha} = Z \backslash_{\mathrm{All}} \alpha$; мультимножина Z – параметр операції; вона відіграє роль універсальної множини

Розглянуто властивості введених операцій: характеристика відношення включення мультимножин у термінах операцій перетину та об'єднання мультимножин; ідемпотентність, комутативність, асоціативність, дистрибутивність, закони поглинання, аналоги законів подвійного заперечення та де Моргана.

Можливі наступні застосування отриманих результатів: мультимножини дозволяють уточнювати таблиці з дублікатами рядків, а операції над таблицями в SQL-подібних мовах уточнюються операціями над мультимножинами [7, 12]. Таким чином, отримана низка співвідношень між операціями над мультимножинами може використовуватися в оптимізаційних блоках процесорів SQL-подібних мов.

Оптимізатори запитів — найбільш складні та найбільш цікаві компоненти СУБД.

Сформулюємо проблеми оптимізації SQL-запитів. Мова SQL декларативна. У формулюваннях SQL-запитів вказується, якими властивостями мають володіти дані, що хоче отримати користувач. Але нічого не сказано про те, як система повинна інтерпретувати запит. Проблема полягає у тому, щоб за декларативною специфікацією запиту побудувати програму (так званий план виконання запиту), яка б виконувалася максимально ефективно та видавала результати, що відповідають вказаним у запиті властивостям. Тобто, основне утруднення полягає в тому, що потрібно вміти будувати усі можливі програми, результати яких відповідають вказаним властивостям, та обирати із множини цих програм саме ту програму, виконання якої було б найбільш ефективним.

Обидві частини проблеми є нетривіальними. Передусім, необхідно виявити усі коректні плани виконання запиту, або, принаймні, не пропустити який-небудь план, що є найбільш ефективним. Далі, для полегшення вирішення другої частини проблеми необхідно максимально скоротити простір коректних планів, залишивши лише ті плани, які претендують на максимальну ефективність. Обидві ці задачі не можна повністю формалізувати, адже відсутні точні математичні критерії вибору. Зазвичай, вирішення таких задач опирається на евристичні алгоритми [9].

Підсумовуючи вищесказане, відмітимо, що проблема оптимізації є досить нетривіальною. Для більш ефективного її вирішення виникає потреба у формалізації цього процесу.

Одним із можливих способів формалізації є визначення існуючих операцій над таблицями в термінах операцій над мультимножинами.

Відмітимо, що операціям $+_{All}$, $\cap_{All}$ та $\setminus_{All}$, які були визначені вище, відповідають наступні ключові слова в SQL-подібних мовах: UNION ALL, INTERSECT ALL, MINUS ALL. А операціям $+_1$, $\cap_1$, $\setminus_1$, призначеним для побудови 1-мультимножин, відповідають ключові слова DISTINCT UNION, DISTINCT INTERSECT, DISTINCT MINUS відповідно [2].

Таким чином, використовуючи формальне визначення відповідних операцій та їхніх властивостей (ідемпотентність, комутативність, асоціативність, дистрибутивність), можна реалізовувати першу частину процесу оптимізації, а саме побудову планів виконання запиту на основі математичного підходу. Це збільшить достовірність та ефективність знаходження найбільш оптимальних планів виконання запиту.

2. Побудова решітки мультимножин

Розглянемо наступні властивості операцій об'єднання $\cup_{All}$ та перетину $\cap_{All}$.

Лема 1. Операції $\cup_{All}$ та $\cap_{All}$ ідемпотентні, комутативні, асоціативні.

Лема 2. Для довільних мультимножин α і β виконуються наступні закони поглинання $\alpha \cap_{All} (\alpha \cup_{All} \beta) = \alpha$, $\alpha \cup_{All} (\alpha \cap_{All} \beta) = \alpha$.

Виходячи з леми 1, можна розглядати дві комутативні ідемпотентні напівгрупи $\langle M, \cup_{All} \rangle$ и $\langle M, \cap_{All} \rangle$, де M – сім'я мультимножин (відповідного універсума D).

Використовуючи добре відомий результат теорії решіток (див., наприклад, [13, § 8, с. 151, теорема 1]), можна напівгрупу за об'єднанням перетворити у верхню піврешітку, а напівгрупу за перетином – у нижню. Часткові порядки верхньої та нижньої піврешіток задаються відповідно наступними визначеннями: $\alpha \overline{\leq} \beta \overset{def}{\Leftrightarrow} \alpha \cup_{All} \beta = \beta$, $\alpha \overline{\trianglelefteq} \beta \overset{def}{\Leftrightarrow} \alpha \cap_{All} \beta = \alpha$, причому:

$$\sup_{\overline{\leq}} \{\alpha, \beta\} = \alpha \cup_{All} \beta, \ \inf_{\overline{\trianglelefteq}} \{\alpha, \beta\} = \alpha \cap_{All} \beta.$$

Можна безпосередньо перевірит, що ці порядки співпадають із порядком включення мультимножин $\preceq$. Отже, встановлено теорему 1.

Теорема 1. Частково впорядкована множина $\langle M, \preceq \rangle$ є решіткою, причому $\sup_{\preceq} \{\alpha, \beta\} = \alpha \cup_{All} \beta$, $\inf_{\preceq} \{\alpha, \beta\} = \alpha \cap_{All} \beta$.

Твердження 1. Виконуються наступні твердження:

1) порожня мультимножина $\varnothing_m$ (її характеристична функція є константною функцією, яка всюди дорівнює нулю) − найменший елемент в $\langle M, \preceq \rangle$;

2) $\inf \mu = \alpha$ для довільної непорожньої множини мультимножин $\mu \subseteq M$; тут характеристична функція точної нижньої грані (інфімуму) α задається виразом $\chi_\alpha(d) = \min\limits_{\beta \in \mu} \chi_\beta(d)$;

3) для довільної, зокрема, порожньої сім'ї мультимножин μ:

точна верхня грань (супремум) μ існує $\Leftrightarrow$ μ обмежено зверху;

4) $\sup \mu = \alpha$, де μ − довільна сім'я мультимножин, яка має точну верхню грань, а характеристична функція мультимножини α задається виразом $\chi_\alpha(d) = \max\limits_{\beta \in \mu} \chi_\beta(d)$.

Теорема 2. Ч.в.м. $\langle M, \preceq \rangle$ є умовно повною множиною та повною піврешіткою, при цьому точні грані знаходять відповідно до формул твердження 1.

Поповнимо частково впорядковану множину (ч. в. м.) $\langle M, \preceq \rangle$ найбільшим елементом T. Отриману ч. в. м. позначимо через $\langle M \cup \{T\}, \preceq \rangle$.

Наслідок 1. Ч. в. м. $\langle M \cup \{T\}, \preceq \rangle$ є повною решіткою з найменшим елементом $\varnothing_m$ та найбільшим елементом T.

Ч. в. м. $\langle M, \preceq \rangle$ можна також вкласти в іншу повну решітку. Для цього розширимо поняття мультимножини.

Із цією метою поповнимо множину натуральних чисел без нуля N^+ зі стандартним порядком найбільшим елементом ∞ та покладемо $N^+_\infty = N^+ \cup \{\infty\}$. Під мультимножиною будемо розуміти функцію вигляду $\alpha : U_\alpha \to N^+_\infty$. Сім'ю усіх таких мультимножин позначимо через M_∞. Порядок на множині N^+_∞ позначимо через $\leq_\infty$, тоді порядок на мультимножинах $\preceq$ розширимо так: $\alpha \preceq_\infty \beta \overset{def}{\Leftrightarrow} U_\alpha \subseteq U_\beta \;\&\; \forall d \left(d \in U_\alpha \Rightarrow \alpha(d) \leq_\infty \beta(d)\right)$, $\alpha, \beta \in M_\infty$. Зауважимо, що по суті порядки $\preceq$, $\preceq_\infty$ будуються прямим добутком ч. в. м. $\langle N, \leq \rangle$ та $\langle N_\infty, \preceq_\infty \rangle$ відповідно, де $N_\infty = N \cup \{\infty\}$, ∞ – найбільший елемент.

Теорема 3. Ч. в. м. $\langle M_\infty, \preceq_\infty \rangle$ є повною решіткою з найменшим елементом $\varnothing_m$ та найбільшим елементом T_∞, де $T_\infty : D \to \{\infty\}$, $T_\infty(d) = \infty$ для всіх $d \in D$. Точні нижні грані знаходять за формулами твердження 1. Для точних верхніх граней виконується формула $\sup \mu = \alpha$, де характеристична функція мультимножини α така – $\chi_\alpha(d) = \sup_{\leq_\infty, \beta \in \mu} \chi_\beta(d)$.

Повну решітку мультимножин, побудовану узагальненням поняття мультимножини, можна використовувати для задання денотаційної семантики рекурсивних CTE-запитів (Common Table Expression) у сучасних SQL-подібних мовах.

Розглянемо опис семантики оператора CTE. Спочатку зробимо декілька зауважень загального порядку.

Засоби задання семантики об’єктів різноманітних предметних областей можна розділити на явні (конотативні, операційні) та неявні (денотативні). Явні засоби задання характеризуються наявністю ефективних процедур задання, тоді як у неявних засобах задання об’єкти визначаються за допомогою специфікації їхніх властивостей.

Денотативні засоби задання є досить потужними та зручними у застосуванні. В математиці та інформатиці ці засоби подаються у вигляді методу нерухомої точки.

У цьому випадку об’єкти задаються як найменші, відносно певного порядку, рішення рівнянь виду $x = F(x)$, $x \in D$, де D – частково впорядкована множина певного класу.

Конкретизуючи питання щодо денотаційної семантики CTE-виразів, приходимо до розгляду рівняння виду: $x = t +_{All} F(x)$, де t – таблиця-параметр (початкове наближення, початковий стан); F – функція над таблицями, що є семантикою відповідної частини CTE-запиту; область значень невідомого x – множина таблиць, що уточнюються як скінченні мультимножини, основами яких виступають множини рядків однієї схеми.

У загальному випадку це рівняння може не мати розв’язку в класі скінченних таблиць. Зауважимо, що функція F повинна зберігати порожню таблицю та бути дистрибутивною відносно операції $+_{All}$ [4].

Якщо зняти вимогу скінченності для основ та кількості дублікатів (розмірності та потужності мультимножини) у таблицях, то при вказаних достатніх умовах на праву частину даного рівняння, це рівняння завжди буде мати найменший розв’язок виду: $\underset{i=0,1,\ldots}{+_{All}} t_i$, де $t_0 \overset{def}{=} t$, а $t_{i+1} \overset{def}{=} t +_{All} F(t_i)$ [4, 8].

Повна решітка мультимножин, яка отримана зняттям обмежень скінченності, саме і призначена для обґрунтування цих конструкцій.

3. Обчислюваність на множинах та мультимножинах

Обчислюваність вводиться як нумераційна обчислюваність за А.І. Мальцевим. Апаратом для задання класу обчислюваних функцій виступають примітивні програмні алгебри (ППА), введені В.Н. Редьком. Головна особливість сигнатурних операцій ППА полягає в їх загальнозначності та абстрактності: операції використовують тільки властивості бути функцією чи предикатом; вони інваріантні відносно природи множини, якій належать аргументи та значення функцій (на відміну від операцій класичної теорії рекурсії [11]).

Наведемо основні означення теорії ППА. Носієм такої алгебри виступає множина багатомісних часткових функцій та предикатів, а сигнатура складається із операцій параметричної суперпозиції функції у функцію (предикат) $S^{(m+1)}$, $m = 1, 2, \ldots$, тернарного розгалуження функцій (предикатів) за предикатом $\Diamond^{(3)}$ та параметричного циклування функцій за предикатом $*^{(n+1)}$, $n = 1, 2, \ldots$

У роботі [6] побудовано повну систему (систему породжуючих) арифметичної ППА σ_N, яка складається із предикату $x < y$, функції слідування $s(x)$, функції додавання $x + y$, функції множення $x \cdot y$, селекторів I_m^n, $n = 1, 2, \ldots$, $m = 1, 2, \ldots, n$ та константної функції $0(x)$, фіксуючої 0: $\sigma_N = \left\{ x < y, s(x), x + y, x \cdot y, I_m^n, 0(x) \right\}_{\substack{n=1,2,\ldots \\ m=1,2,\ldots,n}}$.

Далі будується нова, більш зручна в технічному плані, система породжуючих арифметичної ППА.

На основі побудованої системи породжуючих арифметичної ППА будуються системи породжуючих множинної та мультимножинної (арифметичної) ППА. Системи породжуючих множинної та мультимножинної ППА містять предикат рівності та селектори. Специфічність множинної (мультимножинної) ППА проявляється в тому, що її система породжуючих, крім вказаних елементів, містить функції об'єднання множин, додавання множин, різниці множин та константі функції, що фіксують порожню множину та сінглтон $\{1\}$ (відповідно функції об'єднання мультимножин, додавання мультимножин, різниці мультимножин, константні функції, що фіксують порожню мультимножину та мультимножинний сінглтон $\{1^1\}$, а також специфічну функцію, таку, що $\{n_1^1\},\{k_1^1\} \overset{\phi}{\mapsto} \{n_1^{k_1}\}$) [3].

Отриманий алгебраїчний опис класу обчислюваних функцій над мультимножинами застосовано для моделювання ДНК-обчислень у біоінформатиці та алгоритмічних маніпуляцій над таблицями у сучасних СУБД.

Частково рекурсивні мультимножинні функції є моделями ДНК-обчислень у біоінформатиці.

ДНК-обчислення – це форма обчислень, яка використовує ДНК, біохімію та молекулярну біологію натомість традиційним комп'ютерним технологіям, які побудовано на кремнію [1]. Відповідно, ДНК-комп'ютер – обчислювальна система, що використовує обчислювальні можливості молекул ДНК.

Розвиток цієї предметної області було розпочато у 1994 р. роботою Леонарда Адлемана (Leonard Adleman), в якій було показано як за допомогою пробірки з ДНК можна досить ефективно вирішити класичну задачу комівояжера. Також відомі роботи Еуди Шапіро (Ehud Shapiro) із реалізації скінченних автоматів, Ерика Вінфрі (Erik Winfree) із синтезу різноманітних поверхонь (зокрема, такої відомої фрактальної структури як килима Серпинського) [10].

Молекулу ДНК формально можна подати у вигляді пари слів у чотирьохлітерному алфавіті $\aleph \overset{\text{def}}{=} \{А, Т, Ц, Г\}$, кожний символ якого позначає відповідну основу нуклеотидів: $(\xi_1,...,\xi_n,\eta_1,...,\eta_n)$, де $\xi_1,...,\xi_n \in \aleph$, $\eta_1,...,\eta_n \in \aleph$. Якщо компліментарність уточнити у вигляді бієкції $\psi:\aleph \to \aleph$, $\psi(А)=Т$, $\psi(Т)=А$, $\psi(Ц)=Г$, $\psi(Г)=Ц$, тоді мають виконуватися рівності $\psi(\xi_i)=\eta_i$ для усіх $i=1,..,n$. Таким чином, друге слово пари однозначно відновлюється за першим словом.

Тому, обчислення на ДНК уточнюються як обчислення над словами у вказаному алфавіті. Формальна модель такої словарної обчислюваності була побудована в роботі [5] у термінах ППА.

Більш складний випадок, коли, по-перше, ланцюжки ДНК мають, взагалі кажучи, різну довжину, та, по-друге, враховується "зсув" одного ланцюжка відносно іншого (зсуви виникають із-за подовження, доповнення, вкорочення, розрізу, модифікації або зшивки ланцюжків).

Очевидно, що виникає необхідність у загальних моделях молекулярних обчислень, які б дозволяли планувати нові експерименти та узагальнювати існуючі. Однією з таких моделей є модель паралельної фільтрації (Parallel Filtering Model). Основою цієї моделі виступає "пробірка", формальною моделлю якої, у свою чергу, виступає мультимножина рядків над алфавітом $\aleph = \{А, Т, Ц, Г\}$.

Відповідно до цієї моделі над "пробіркою" визначаються операції злиття, розмноження, виявлення, добування, розділення за довжиною та розділення за префіксом.

Перелічені операції можна отримати із системи породжуючих мультимножинної ППА. Цей факт дає підставу стверджувати, що частково рекурсивні мультимножинні функції виступають формальними моделями ДНК-обчислень.

Крім цього, як визначалося вище, мультимножини рядків виступають уточненнями таблиць з дублікатами рядків в табличних базах даних. Отже, частково рекурсивні мультимножинні функції є також моделями алгоритмічних маніпуляцій над таблицями в сучасних СУБД.

Висновки

В роботі наведені елементи теорії мультимножин. Продемонстровано застування отриманих теоретичних результатів для: (1) уточнення таблиць з дублікатами рядків в сучасних СУБД та маніпуляцій над такими таблицями; (2) побудови денотаційної семантики рекурсивної форми CTE-виразів (Common Table Expressions) сучасних SQL-подібних мов; (3) уточнень обчислень на ДНК у біоінформатиці.

Література

1. *DNA computing. [Електронний ресурс]: Wikipedia, the free encyclopedia. – Режим доступу: http://en.wikipedia.org/wiki/DNA_computing.*

2. *SQL: Операции [Електронний ресурс]. – Режим доступу: http://articles. org.ru/docum/sql_ oper.php.*

3. *Богатирьова, Ю.О. Примітивні програмні алгебри функцій множинних (мультимножинних) аргументів та значень [Текст] / Ю.О. Богатирьова, Д.Б. Буй, В.Н. Редько // Доповіді НАН України. №9. – 2011. – С. 32-35.*

4. *Буй, Д.Б. Композиційна семантика рекурсивних запитів в SQL-подібних мовах [Текст] /*

Д.Б. Буй, С.А. Поляков // Вісник Київського національного університету імені Тараса Шевченка. Сер.: фіз.-мат. науки. – 2010. – № 1. – С. 45−50.

5. *Буй, Д.Б. Примитивные программные алгебры целочисленных и словарных функций [Текст] / Д.Б. Буй, В.Н. Редько // Докл. АН УССР. Сер.: А. физ.-мат. и тех. науки. – 1984. – № 10. – С. 69–71.*

6. *Буй, Д.Б. Примитивные программные алгебры. [Текст] / Д.Б. Буй, В.Н. Редько // Кибернетика. – 1984. – №4. – С. 1–7.*

7. *Буй, Д.Б. Теория мультимножеств: библиография, применение в табличных базах данных [Текст] / Д.Б. Буй, Ю.А. Богатырёва // Радіоелектронні і комп'ютерні системи. – № 7(48). – 2010. – С. 56–62.*

8. *Буй, Д.Б. Теорія програмних алгебр композиційного типу та її застосування: дис. ... доктора фіз.-мат. наук: 01.05.03 / Буй Дмитро Борисович. – К., 2002. – 365 с.*

9. *Кузнецов, С.Д. Оптимизация запросов: вечнозеленая область [Електронний ресурс] / С.Д. Кузнецов. – Режим доступу: http://citforum.ru/database/articles/sql_optimization.shtml.*

10. *Малинецкий, Г.Г. Вычисления на ДНК. Эксперименты. Модели. Алгоритмы. Инструментальные средства [Електронний ресурс] / Г.Г. Малинецкий, С.А. Науменко. – Режим доступу: http://www.keldysh.ru/papers/2005/prep57/prep2005_57.html.*

11. *Мальцев, А.И. Алгоритмы и рекурсивные функции [Текст] / А.И. Мальцев. – Москва: Наука, 1986. – 367 с.*

12. *Редько, В.Н. Реляційні бази даних: табличні алгебри та SQL-подібні мови [Текст] / В.Н. Редько, Ю.Й. Брона, Д.Б. Буй, С.А. Поляков. – Київ: Видавничий дім "Академперіодика", 2001. – 198 с.*

13. *Скорняков, Л.А. Элементы алгебры [2-е изд.] [Текст] / Л.А. Скорняков. – Москва: Наука, 1986. – 240 с.*

Надійшла до редакції 21.02.2013, розглянута на редколегії 13.03.2013

Рецензент: д-р техн. наук, проф., проф. каф. інженерії програмного забезпечення Б.М. Конорев, Національний аерокосмічний університет ім. М.Є. Жуковського «ХАІ», Харків, Україна.

ПРИМЕНЕНИЕ ТЕОРИИ МУЛЬТИМНОЖЕСТВ

Д.Б. Буй, Ю.А. Гришко

В статье рассмотрены вопросы применения теории мультимножеств в табличных базах данных и ДНК-вычислениях. Приведены основные определения теории мультимножеств. Введены операции над мультимножествами и исследованы основные свойства этих операций. Построено решетку мультимножеств, которую вложено в две полные решетки. Уточнена вычислимость на мультимножествах. Продемонстрировано применения полученных результатов в современных СУБД, при построение денотационной семантики рекурсивной формы CTE-выражений в современных SQL-подобных языках, вычислениях на ДНК в биоинформатике.

Ключевые слова: мультимножество, решетка, вычислимость, CTE, SQL, ДНК-вычисления.

APPLICATIONS OF THE MULTISETS THEORY

D.B. Buy, J.A. Grishko

In the article the questions of application multise theory are considered in tabular databases and DNA-calculations. The main definitions of the multisets theory are given. Operations over a multisets are entered and the main properties of these operations are investigated. It is constructed a lattice of multisets which it is enclosed in two complete lattices. The computability over multisets is specified. Applications of the received results are shown: in modern DBMS, for construction of the denotational semantics of a recursive form of CTE-expressions in modern SQL-like languages, DNA-calculations in bioinformatics.

Key words: multiset, lattice, computability, CTE, SQL, DNA calculations.

Буй Дмитро Борисович – доктор фіз.-мат. наук, професор, професор кафедри теорії та технології програмування Київського національного університету імені Тараса Шевченка, e-mail: buy@unicyb.kiev.ua.

Гришко Юлія Олександрівна – кандидат фіз.-мат. наук, асистент кафедри інформаційних систем Київського національного університету імені Тараса Шевченка, e-mail: jul.grishko@gmail.com.