УДК 004.652.43+004.652.44

А. Григорович[1], В. Григорович [2]
[1]Дрогобицький державний педагогічний університет імені Івана Франка,
кафедра інформаційних систем і технологій;
[2]Національний університет “Львівська політехніка”,
кафедра інформаційних систем та мереж

ЕКВІВАЛЕНТНІСТЬ ВИРАЗІВ РЕЛЯЦІЙНОГО ЧИСЛЕННЯ ДОМЕНІВ ДЛЯ НЕНОРМАЛІЗОВАНИХ ВІДНОШЕНЬ ТА РОЗШИРЕНОЇ РЕЛЯЦІЙНОЇ АЛГЕБРИ

Введено означення безпечного виразу реляційного числення доменів для ненормалізованих відношень, доведено теореми про еквівалентність виразів числення доменів для ненормалізованих відношень та виразів розширеної реляційної алгебри. Для доведення теорем використано метод математичної індукції.

Ключові слова: ненормалізовані відношення, розширена реляційна алгебра, числення доменів для ненормалізованих відношень, безпечний вираз.

The article introduces the definition of safe expression of the domain relational calculus for nested relations, proves theorems on equivalence of the expressions of the domain relational calculus for the nested relations and of the extension of relational algebra. There is used the method of mathematical induction for proving theorems.

Key words: nested relations, extended relational algebra, domain relational calculus for nested relations, safe expression.

Постановка проблеми у загальному вигляді

Розроблена Е. Коддом реляційна модель даних – це логічна модель даних, що описує структуру даних у вигляді відношень, операції над даними та спеціальні правила, що забезпечують цілісність даних [1]. До складу реляційної моделі разом із описом структури даних входить визначення множини операторів над даними, що передбачає визначення методів перетворення відношень та отримання на їх основі нових відношень. Ці методи можна задавати реляційною алгеброю або реляційним численням. Еквівалентність реляційної алгебри Кодда та реляційного числення доведена в [2], доведення стосувалося відношень з атомарними атрибутами (надалі – *плоских відношень*). Проте не доведена еквівалентність розширеної на ненормалізовані відношення реляційної алгебри та числення доменів для ненормалізованих відношень.

Аналіз останніх досліджень

Ненормалізовані відношення вперше введено у [3]. У [4] наведено інтуїтивне означення ненормалізованих відношень: “скрізь, де дозволяються атомарні (скалярні) значення, дозволяються також відношення, отже, допускаються відношення в складі відношень. При цьому зберігається основа реляційної моделі, хоч і порушується вимога першої нормальної форми”. У [5] досліджено ненормалізовану реляційну модель баз даних; залежності даних у ненормалізованій реляційній моделі; алгебраїчні властивості відношень із гніздуванням; властивості залежностей даних на реляційних структурах з гніздуванням та питання нормалізації відношень із гніздуванням; розширену на ненормалізовані відношення реляційну алгебру, в якій визначено оператори упакування NEST та розпакування UNNEST. Оператор NEST, аргументом якого є плоске відношення, утворює вкладені відношення над однаковими значеннями даних у деякій підмножині атрибутів плоского відношення. Оператор UNNEST є зворотним до оператора NEST. Він перетворює ненормалізоване відношення на плоске.

Використовуючи методологію [6] та числення предикатів другого степеня [7], у [8] за допомогою алфавіту, атомарних виразів та формул означено реляційне числення зі змінними на доменах для ненормалізованих відношень. Проте залишилася не доведеною еквівалентність побудованого реляційного числення доменів для ненормалізованих відношень та розширеної реляційної алгебри, що і зумовлює актуальність вибраної теми дослідження.

Цілі статті

Мета статті – довести еквівалентність виразів реляційного числення доменів для ненормалізованих відношень та розширеної реляційної алгебри.

Еквівалентними називатимемо вирази реляційного числення доменів для ненормалізованих відношень та розширеної реляційної алгебри, які дають однаковий результат, якщо в них аргументами є ті самі відношення. Еквівалентність позначатимемо символом ≡.

Основний матеріал

З метою виключення результатів, що не мають змісту чи не можуть бути реалізовані, для реляційного числення існує обмеження скінченності відношень, оскільки в базах даних можна реалізувати лише відношення зі скінченною кількістю кортежів.

З метою унеможливлення представлення за допомогою реляційних числень нескінченних відношень (наприклад, $\{ x_1 \dots x_n \mid \neg R(x_1, \dots, x_n) \}$) у [2, с.105] введено поняття "безпечного" виразу числення доменів першого порядку.

Введемо поняття "безпечного" виразу реляційного числення доменів вищих порядків (для ненормалізованих відношень) $\{ x_1\, x_2 \dots x_n\, P_1\, P_2 \dots P_k \mid \Psi(x_1, x_2, \dots, x_n, P_1, P_2, \dots, P_k) \}$, де x_i (i=1,…,n) – атомарні атрибути, P_j (j=1,…,k) – вкладені відношення. Для безпечних виразів елементи x_i, P_j (i=1,…,n, j=1,…,k) з лівої частини виразу (до символу "|") повинні бути елементами DOM(Ψ), де DOM(Ψ) – це множина символів, які або входять у Ψ, або є елементами кортежів деякого ненормалізованого відношення R, або є елементами кортежів вкладених відношень $P_1, P_2, \dots, P_k$ у ненормалізоване відношення R, яке входить у формулу Ψ. Тобто DOM(Ψ) визначається не залежно від вигляду Ψ, а як функція фактичних відношень, які підставляються замість змінних відношень у Ψ. Оскільки розглядаємо скінченні відношення, то і DOM(Ψ) є скінченною множиною. Змінні $x_1, x_2, \dots, x_n, P_1, P_2, \dots, P_k$ – вільні змінні Ψ.

<u>Означення</u> 1. Доменом формули $\Psi(R(x_1, \dots, x_n, P_1, \dots, P_k), a_1, \dots, a_m)$ називається множина всіх значень, яких можуть набувати зміні та константи формули Ψ: $\mathrm{DOM}(\Psi) = \{\pi_1(R) \cup \dots \cup \pi_n(R) \cup \pi_{n+1}(R) \cup \dots \cup \pi_{n+k}(R) \cup \{a_1, \dots, a_m\}\}$, де $a_1, \dots, a_m$ – константи, які входять до Ψ, $x_1, \dots, x_n$ – атомарні атрибути ненормалізованого відношення $R(x_1, \dots, x_n, P_1, \dots, P_k)$, яке входить у Ψ, $P_1, \dots, P_k$ – вкладені відношення ненормалізованого відношення R.

Введемо означення безпечного виразу реляційного числення зі змінними на доменах:

<u>Означення</u> 2. Безпечним виразом числення зі змінними на доменах називається вираз типу $\{ x_1\, x_2 \dots x_n\, P_1\, P_2 \dots P_k \mid \Psi(x_1, x_2, \dots, x_n, P_1, P_2, \dots, P_k) \}$ за умов:

1. якщо $\Psi(x_1, \dots, x_n, P_1, \dots, P_k)$ – істинне, то $x_i \in \mathrm{DOM}(\Psi)$, $P_j \in \mathrm{DOM}(\Psi)$, де i=1,…,n, j=1,…,k;
2. якщо $(\exists u)\,(\Phi(u))$ – підформула Ψ, то з істинності $\Phi(u)$ випливає, що $u \in \mathrm{DOM}(\Phi)$;
3. якщо $(\forall u)\,(\Phi(u))$ – підформула Ψ, то з істинності $\neg\Phi(u)$ випливає, що $u \in \mathrm{DOM}(\Phi)$.

<u>Означення</u> 3. Формула числення зі змінними на доменах $\Psi(x_1, x_2, \dots, x_n, P_1, P_2, \dots, P_k)$ називається безпечною, якщо вираз числення зі змінними на доменах $\{ x_1\, x_2 \dots x_n\, P_1\, P_2 \dots P_k \mid \Psi(x_1, x_2, \dots, x_n, P_1, P_2, \dots, P_k) \}$ є безпечним.

Для пояснення третьої умови в означенні 2 зауважимо, що формула $(\forall u)\,(\Phi(u))$ еквівалентна формулі $\neg(\exists u)\,(\neg\Phi(u))$, яка не є безпечною лише тоді, коли існує деяке u_0, для якого істинне $\neg\Phi(u_0)$ і u_0 не належить домену формули ¬Φ. Оскільки домени формул ¬Φ і Φ однакові, то третя умова означення 2 встановлює, що формула $(\forall u)\,(\Phi(u))$ безпечна, коли безпечною є формула $\neg(\exists u)\,(\neg\Phi(u))$.

Редукція числення доменів для ненормалізованих відношень до розширеної реляційної алгебри

Доведемо, що для довільної безпечної формули числення доменів для ненормалізованих відношень $\Psi(x_1, \ldots, x_n, P_{n+1}, \ldots, P_k)$, де $x_1, \ldots, x_n, P_{n+1}, \ldots, P_k$ – вільні змінні, можна сконструювати алгебраїчний вираз, значення якого $\{ x_1 \ldots x_n P_{n+1} \ldots P_k \mid \Psi(x_1, \ldots, x_n, P_{n+1}, \ldots, P_k) \}$.

Якщо навіть заданий вираз є безпечним, деякі підформули його формули можуть такими не бути [2, с.110]. Тому, використовуючи методологію [2], доведемо індукцією за розміром (кількістю операторів) підформули Φ формули Ψ з вільними змінними $y_1, \ldots, y_l, S_{l+1}, \ldots, S_m$, що існує вираз реляційної алгебри для

$$(\mathrm{DOM}\ (\Psi))^m \cap \{ y_1 \ldots y_l S_{l+1} \ldots S_m \mid \Phi(y_1, \ldots, y_l, S_{l+1}, \ldots, S_m)\},$$

де $D^m = D \times D \times \ldots \times D$ (m разів). Спочатку доведемо потрібні леми.

Лема 1. Для будь-якої формули Ψ числення зі змінними на доменах для ненормалізованих відношень існує вираз реляційної алгебри, який визначає унарне відношення (множину) DOM (Ψ).

Доведення. Якщо R – ненормалізоване відношення степеня (арності) k, то нехай $E (R) = \pi_1(R) \cup \pi_2(R) \cup \ \ldots \cup \pi_k(R)$. Тоді шуканий вираз є об’єднанням $E (R)$ по всіх змінних-відношеннях, які містить формула Ψ, і сталого відношення $\{a_1, a_2, \ldots, a_n\}$, де всі a_i – символи констант, які входять у Ψ.■

Лема 2. Для довільної формули Ψ числення зі змінними на доменах для ненормалізованих відношень існує еквівалентна їй формула Ψ' цього числення, яка не містить $\wedge$ і $\forall$. Якщо Ψ – безпечна формула, то Ψ' теж безпечна формула.

Доведення. Використовуючи закон де Моргана, замінимо кожну підформулу $\Psi_1 \wedge \Psi_2$ формули Ψ на $\neg (\neg\Psi_1 \vee \neg\Psi_2)$. Також замінимо кожну підформулу $(\forall u)\ (\Psi_1(u))$ на $\neg(\exists u)\ (\neg\Psi_1(u))$. Це перетворення є еквівалентним, оскільки якщо не існує u, для якого Ψ_1 хибне, то Ψ_1 істинне для всіх u.

В результаті отримаємо формулу Ψ', еквівалентну Ψ. Якщо Ψ безпечна, то згідно з означенням 2, якщо $(\forall u)\,(\Psi_1\ (u))$ – підформула Ψ, то $\neg(\exists u)\,(\neg\Psi_1(u))$ також задовольняє умову безпеки, а отже, і Ψ' – безпечна формула.■

Теорема 1. Для кожного безпечного виразу реляційного числення зі змінними на доменах для ненормалізованих відношень існує еквівалентний вираз розширеної реляційної алгебри.

Доведення. Нехай $\{ x_1 \ldots x_n P_{n+1} \ldots P_k \mid \Psi(x_1, \ldots, x_n, P_{n+1}, \ldots, P_k) \}$ – безпечний вираз числення зі змінними на доменах для ненормалізованих відношень. Згідно з лемою 2 припустимо, що Ψ містить тільки оператори $\vee$, $\neg$ та квантор $\exists$. Відповідно до леми 1 позначимо E – вираз реляційної алгебри для множини DOM (Ψ). Нехай E^k означає $E \times E \times \ldots \times E$ (k разів). Доведемо індукцією за кількістю операторів в підформулі Φ формули Ψ, що якщо Φ містить вільні доменні змінні $y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m$, то для $\mathrm{DOM}\ (\Psi)^m \cap \{ y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m \mid \Phi\ (y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m) \}$ існує еквівалентний вираз реляційної алгебри.

У частковому випадку, якщо Φ є самою формулою Ψ, отримаємо алгебраїчний вираз для $\mathrm{DOM}\ (\Psi)^k \cap \{ x_1 \ldots x_n P_{n+1} \ldots P_k \mid \ \Psi(x_1, \ldots, x_n, P_{n+1}, \ldots, P_k)\}$. Оскільки Ψ – безпечна формула, то перетин з $\mathrm{DOM}\ (\Psi)^k$ не змінить відношення, яке ця формула визначає. Отже, теорему доведено.

Для доведення теореми використаємо метод індукції.

Базис. Нуль операторів у Φ. Тоді Φ є атомарним виразом. Цей атомарний вираз може бути одного з видів: $x_1\ \Theta\ x_2$, $x_1\ \Theta\ x_1$, $x_1\ \Theta\ c$, $R(x_{i_1}, x_{i_2}, \ldots, x_{i_m}, Q_{i_{m+1}}, \ldots, Q_{i_l})$, де Θ – операція порівняння, а c – константа. Якщо атомарний вираз має вигляд $x_1\ \Theta\ x_2$, то шуканий еквівалентний вираз реляційної алгебри: $\sigma_{1\ \Theta\ 2}\ (E \times E)$. Аналогічно записуються еквівалентні алгебраїчні вирази для $x_1\ \Theta\ x_1$ та $x_1\ \Theta\ c$.

Якщо атомарний вираз реляційного числення доменів для ненормалізованих відношень має вигляд $R(x_{i_1}, x_{i_2}, \ldots, x_{i_m}, Q_{i_{m+1}}, \ldots, Q_{i_l})$, то шуканий алгебраїчний вираз запишемо так: $\pi_{j_1, j_2, \ldots, j_k}(\sigma_F(R))$, де F – формула, яка містить терм $u = v$ щоразу, коли x_{i_u} та x_{i_v} або Q_{i_u} та Q_{i_v} є однаковими змінними і $u < v$. Всі терми зв'язані операторами $\wedge$. Список $j_1, j_2, \ldots, j_k$ повинен бути таким, що $x_{i_{j_1}} = x_1, \ldots, x_{i_{j_k}} = x_k$, $Q_{i_{j_1}} = Q_1, \ldots, Q_{i_{j_k}} = Q_k$. Наприклад, якщо $\Psi = R(x_2, x_1, x_2, x_3, Q_5)$, то шуканий вираз реляційної алгебри запишемо так: $\pi_{2,1,4,5}(\sigma_{1=3}(R))$.

Індукція. Припустимо, що Φ містить хоча б один оператор і що гіпотеза індукції справедлива для всіх підформул формули Ψ, які мають менше операторів, ніж Φ.

Випадок 1. $\Phi(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m) = \Phi_1(u_1, \ldots, u_t, R_{t+1}, \ldots, R_n) \vee \Phi_2(v_1, \ldots, v_z, S_{z+1}, \ldots, S_p)$, де всі u_i є різними y_j, а всі v_i – різними y_j (але деякі u та v можуть бути однаковими y_j); і всі R_i є різними Q_j, а всі S_i – різними Q_j (але деякі R та S можуть бути однаковими Q_j). Нехай E_1 – алгебраїчний вираз для

$$\mathrm{DOM}(\Psi)^n \cap \{ u_1 \ldots u_t R_{t+1} \ldots R_n \mid \Phi_1(u_1, \ldots, u_t, R_{t+1}, \ldots, R_n) \},$$

а E_2 – алгебраїчний вираз для

$$\mathrm{DOM}(\Psi)^p \cap \{ v_1 \ldots v_z S_{z+1} \ldots S_p \mid \Phi_2(v_1, \ldots, v_z, S_{z+1}, \ldots, S_p) \}.$$

Визначимо E_1' так: $E_1' = \pi_{i_1, \ldots, i_m}(E_1 \times E^{m-n})$, де i_k – це q таке, що $u_q = y_k$ або $R_q = Q_k$, якщо таке u_q або R_q існує, і i_k – унікальне ціле число між $n+1$ і m в протилежному випадку.

Аналогічно визначимо $E_2' = \pi_{j_1, \ldots, j_m}(E_2 \times E^{m-p})$, де j_k – це таке q, що $v_q = y_k$ або $S_q = Q_k$, якщо таке v_q або S_q існує, і j_k – унікальне ціле число між $p+1$ і m в протилежному випадку.

Тоді шуканий вираз реляційної алгебри – $E_1' \cup E_2'$.

Наприклад, якщо $\Phi(y_1, y_2, y_3, Q_4)$ має вигляд $\Phi_1(y_1, y_3, Q_4) \vee \Phi_2(y_2, Q_4)$, то $E_1' = \pi_{1,4,2,3}(E_1 \times E)$ і $E_2' = \pi_{3,1,4,2}(E_2 \times E \times E)$.

Коректність формули $E_1' \cup E_2'$ випливає з того, що E_1' означає $\mathrm{DOM}(\Psi)^m \cap \{ y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m \mid \Phi_1(u_1, \ldots, u_t, R_{t+1}, \ldots, R_n) \}$, а E_2' – $\mathrm{DOM}(\Psi)^m \cap \{ y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m \mid \Phi_2(v_1, \ldots, v_z, S_{z+1}, \ldots, S_p) \}$, де кожне u та кожне v є одним із y, а кожне R та S є одним з Q. Звідси випливає, що $E_1' \cup E_2'$ означає $\mathrm{DOM}(\Psi)^m \cap \{ y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m \mid \Phi(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m) \}$.

Випадок 2. $\Phi(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m) = \neg\Phi_1(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m)$. Нехай E_1 – алгебраїчний вираз для

$$\mathrm{DOM}(\Psi)^m \cap \{ y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m \mid \Phi_1(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m) \}.$$

Тоді $E^m \setminus E_1$ – алгебраїчний вираз для

$$\mathrm{DOM}(\Psi)^m \setminus \{ y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m \mid \Phi_1(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m) \},$$

яке, своєю чергою, еквівалентне

$$\mathrm{DOM}(\Psi)^m \cap \{ y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m \mid \neg\Phi_1(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m) \}.$$

Випадок 3. $\Phi(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m) = (\exists y_{m+1})(\exists Q_{m+2})(\Phi_1(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m, y_{m+1}, Q_{m+2}))$. Нехай E_1 – алгебраїчний вираз для $\mathrm{DOM}(\Psi)^{m+2} \cap \{ y_1 \ldots y_{m+1} \mid \Phi_1(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m, y_{m+1}, Q_{m+2}) \}$. Оскільки Ψ – безпечна формула, то за означенням 2 $\Phi_1(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m, y_{m+1}, Q_{m+2})$ не може бути істинною, якщо y_{m+1} і Q_{m+2} не належить множині $\mathrm{DOM}(\Phi_1)$, яка є підмножиною $\mathrm{DOM}(\Psi)$. Отже, $\pi_{1,2,\ldots,m}(E_1)$ означає відношення $\mathrm{DOM}(\Psi)^m \cap \{ y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m \mid (\exists y_{m+1})(\exists Q_{m+2})(\Phi_1(y_1, \ldots, y_l, Q_{l+1}, \ldots, Q_m, y_{m+1}, Q_{m+2})) \}$, що завершує індукцію і доводить теорему. ■

Приклад 1. Нехай $R(w, x, Q)$ та $S(y, z, P)$ – ненормалізовані відношення, визначені на спільному домені $dom(R) = dom(S)$. Вкладені відношення Q та P мають однаковий степінь і спільний домен $dom(Q) = dom(P)$. Вираз числення зі змінними на доменах для ненормалізованих відношень

$$\{ w, x, Q \mid R(w, x, Q) \wedge (\forall y)(\forall P)(\neg S(w, y, P) \wedge (\neg S(x, y, P))) \} \qquad (1)$$

визначає множину таких кортежів у R, в яких жоден з атомарних компонентів не є першим атомарним компонентом якогось кортежу в S.

Побудуємо вираз розширеної реляційної алгебри, еквівалентний виразу (1).

Вираз (1) є безпечним, оскільки:

1. Формула не задовольняється для w, x та Q, якщо w, x, Q не є кортежем R.

2. Щоразу, коли y або P не є символом, який входить в деякий кортеж S, очевидно, істинна формула $\neg S(w, y, P) \wedge \neg S(x, y, P)$.

Нехай E означає $\pi_1(R) \cup \pi_2(R) \cup \pi_3(R) \cup \pi_1(S) \cup \pi_2(S) \cup \pi_3(S)$. Використовуючи доведення леми 2, виключимо $\wedge$ та $\forall$. Після взаємного знищення пар заперечень $\neg$ отримаємо вираз

$$\{ w, x, Q \mid \neg(\neg R(w, x, Q) \vee (\exists y)\ (\exists P)\ (S(w, y, P) \vee S(x, y, P)\ )\ )\ \}.$$

Застосуємо доведення теореми 1. Вираз R означає $E^3 \cap \{ w, x, Q \mid R(w, x, Q) \}$, вираз S означає і $E^3 \cap \{ w, y, P \mid S(w, y, P) \}$ і $E^3 \cap \{ x, y, P \mid S(x, y, P) \}$.

Відповідно до випадку 1 теореми 1, для $E^4 \cap \{ w, x, y, P \mid S(w, y, P\ ) \vee S(x, y, P\ ) \}$ одержимо такий вираз реляційної алгебри: $E_1 = \pi_{1,4,2,3}(S \times E) \cup \pi_{4,1,2,3}(S \times E)$.

Відповідно до випадку 3 теореми 1, для $E^3 \cap \{ w, x, P \mid (\exists y)\ (S(w, y, P\ ) \vee S(x, y, P\ )\ )\ \}$ отримаємо такий еквівалентний вираз реляційної алгебри: $\pi_{1,2,3}(E_1)$.

Побудувавши каскадні проекції, отримаємо: $E_2 = \pi_{1,4,2}(S \times E) \cup \pi_{4,1,2}(S \times E)$.

Цей вираз означає множину кортежів, один з компонентів яких є першим атомарним компонентом кортежу з S, а другий атомарний компонент входить у деякий кортеж R або S.

Відповідно до випадку 2, для $E^3 \cap \{ w, x, Q \mid \neg R(w, x, Q) \}$ матимемо такий еквівалентний вираз реляційної алгебри: $E^3 \setminus R$.

Відповідно до випадку 1, для $E^3 \cap \{ w, x, Q \mid \neg R(w, x, Q) \vee (\exists y)\ (\exists P)\ (S(w, y, P) \vee S(x, y, P)\ )\ \}$ отримаємо такий еквівалентний вираз реляційної алгебри: $(E^3 \setminus R) \cup E_2$ (тут проекції не записуємо, оскільки $\pi_{1,2,3}$ і для $E^3 \setminus R$ і для E^3 залишає всі кортежі без змін).

Повний вираз числення зі змінними на доменах (1) еквівалентний алгебраїчному виразу $E^3 \setminus (\ (E^3 \setminus R) \cup E_2\ )$. Оскільки R і E_2 є підмножинами E^3, то цей вираз є еквівалентним $R \setminus E_2$, тобто $R \setminus \pi_{1,4,2}(S \times E) \cup \pi_{4,1,2}(S \times E)$.

Редукція розширеної реляційної алгебри до числення доменів для ненормалізованих відношень

Доведемо, що для довільного виразу реляційної алгебри існує еквівалентний йому безпечний вираз реляційного числення доменів для ненормалізованих відношень.

Теорема 2. Якщо E – вираз розширеної реляційної алгебри, то існує еквівалентний йому безпечний вираз реляційного числення доменів для ненормалізованих відношень.

Доведення. Для доведення використаємо метод індукції за кількістю операторів в $E(R)$, де $R(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n)$ – ненормалізоване відношення, яке містить атомарні атрибути $x_1, \ldots, x_t$ та вкладені відношення $Q_{t+1}, \ldots, Q_n$.

Базис. Нуль операторів. Тоді E може бути:

1) стале відношення $E = \{\{x_{1_1}, \ldots, x_{1_t}, Q_{1_{t+1}}, \ldots, Q_{1_n}\}, \ldots, \{x_{i_1}, \ldots, x_{i_t}, Q_{i_{t+1}}, \ldots, Q_{i_n}\}, \ldots, \{x_{k_1}, \ldots, x_{k_t}, Q_{k_{t+1}}, \ldots, Q_{k_n}\}\}$, де $\{x_{i_1}, \ldots, x_{i_t}, Q_{i_{t+1}}, \ldots, Q_{i_n}\}$ – i-й кортеж сталого відношення, k – кількість кортежів сталого відношення. В такому випадку вираз реляційного числення доменів для ненормалізованих відношень, еквівалентний E, запишемо у вигляді

$$\{ u_1, \ldots, u_t, S_{t+1}, \ldots, S_n \mid (u_1 = x_{1_1} \wedge \ldots \wedge u_t = x_{1_t}, S_{t+1} = Q_{1_{t+1}}, \ldots, S_n = Q_{1_n}) \vee \ldots \vee (u_1 = x_{k_1} \wedge \ldots \wedge u_t = x_{k_t}, S_{t+1} = Q_{k_{t+1}}, \ldots, S_n = Q_{k_n})\};$$

2) змінна R, яка означає відношення $E = R(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n)$. Вираз реляційного числення доменів для ненормалізованих відношень, еквівалентний E, запишемо у вигляді $\{ x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n \mid R(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n)\ \}$.

Оскільки символи, які стоять в лівій частині виразу (до символа “|”), є атомарними елементами кортежів ненормалізованого відношення R, або елементами кортежів відношень $Q_{t+1}, \ldots, Q_n$, вкладених у ненормалізоване відношення R, то цей вираз – безпечний.

Індукція. Припустимо, що E має хоча б один оператор і що теорема істинна для виразів з меншою кількістю входжень операторів в E.

Випадок 1. $E = E_1 \cup E_2$. Тоді вирази E_1 і E_2 кожний мають меншу кількість входжень операторів порівняно з E. За гіпотезою індукції ми можемо знайти безпечні вирази реляційного числення доменів для ненормалізованих відношень $\{ x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n \mid \Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \}$ – еквівалентний E_1 і $\{ x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n \mid \Psi_2(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \}$ – еквівалентний E_2. В такому випадку вираз реляційного числення доменів для ненормалізованих відношень $\{ x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n \mid \Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \vee \Psi_2(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \}\}$ є еквівалентним E. Якщо $x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n$ задовольняє $\Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \vee \Psi_2(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n)$, то кожний елемент $x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n$ належить $\mathrm{DOM}(\Psi_1)$ або $\mathrm{DOM}(\Psi_2)$. Оскільки $\mathrm{DOM}(\Psi_1 \vee \Psi_2) = \mathrm{DOM}(\Psi_1) \cup \mathrm{DOM}(\Psi_2)$, то E еквівалентний деякому безпечному виразу. Отже, повна формула $\Psi(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) = \Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \vee \Psi_2(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n)$ істинна тільки тоді, коли $x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n \in \mathrm{DOM}(\Psi)$ і будь-яка підформула $(\exists x_i)\, \Phi(x_i)$ або $(\forall x_i)\, \Phi(x_i)$ чи $(\exists Q_i)\, \Phi(Q_i)$ або $(\forall Q_i)\, \Phi(Q_i)$ із Ψ повинна входити в Ψ_1 або Ψ_2. Тому, відповідно до гіпотези індукції, ці підформули не порушують безпечності виразу.

Випадок 2. $E = E_1 \setminus E_2$. Тоді, як і у випадку 1, для E_1 та E_2 існують безпечні вирази. Очевидно, що вираз розширеної реляційної алгебри E еквівалентний такому виразу реляційного числення доменів для ненормалізованих відношень:

$\{ x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n \mid \Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \wedge \neg\Psi_2(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \}\}$. Оскільки

$\mathrm{DOM}(\Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \wedge \neg\Psi_2(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n)) = \mathrm{DOM}(\Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \cup \Psi_2(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n))$, то наведений вище вираз є безпечним.

Випадок 3. $E = E_1 \times E_2$. Нехай E_1 та E_2, як і у випадку 1, еквівалентні деяким безпечним виразам, і нехай E_1 позначає ненормалізоване відношення $\Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n)$, а E_2 позначає ненормалізоване відношення $\Psi_2(y_1, \ldots, y_l, S_{l+1}, \ldots, S_m)$, де $x_1, \ldots, x_t$ і $y_1, \ldots, y_l$ – атомарні атрибути, $Q_{t+1}, \ldots, Q_n$ і $S_{l+1}, \ldots, S_m$ – вкладені відношення. Тоді еквівалентний до E вираз реляційного числення доменів для ненормалізованих відношень запишемо у вигляді:

$\{w_1, \ldots w_t, F_{t+1}, \ldots, F_n, w_{n+1}, \ldots, w_{n+l}, F_{n+l+1}, \ldots, F_{n+m} \mid \exists x_1 \ldots \exists x_t\, \exists Q_{t+1} \ldots \exists Q_n\, \exists y_1 \ldots \exists y_l\, \exists S_{l+1} \ldots \exists S_m\, (\Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \wedge \Psi_2(y_1, \ldots, y_l, S_{l+1}, \ldots, S_m) \wedge w_1{=}x_1 \wedge \ldots \wedge w_n{=}x_n \wedge F_{t+1}{=}Q_{t+1} \wedge \ldots \wedge F_n{=}Q_n \wedge w_{n+1}{=}y_1 \wedge \ldots \wedge w_{n+l}{=}y_l \wedge F_{n+l+1}{=}S_{l+1} \wedge \ldots \wedge F_{n+m}{=}S_m)\}$.

Наведений вираз безпечний, оскільки w_i обмежується значеннями, яких може набувати x_i при $i \le t$, та значеннями, яких може набувати y_{i-n} при $n < i \le n+l$; а F_i обмежується значеннями, яких може набувати Q_i при $t < i \le n$, та значеннями, яких може набувати S_{i-n}, якщо $n+l < i \le n+m$.

Випадок 4. $E = \pi_{i_1, i_2, \ldots, i_l, i_{t+1}, \ldots, i_m}(E_1)$.

Нехай E_1 еквівалентний безпечному виразу $\{ x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n \mid \Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n)\}$. Тоді еквівалентний до E вираз реляційного числення доменів для ненормалізованих відношень запишемо у вигляді:

$\{w_1, \ldots w_l, F_{t+1}, \ldots, F_m \mid \exists x_1 \ldots \exists x_l\, \exists Q_{t+1} \ldots \exists Q_m\, (\Psi_1(x_1, \ldots, x_t, Q_{t+1}, \ldots, Q_n) \wedge w_1{=}x_1 \wedge \ldots \wedge w_l{=}{=}\, x_l \wedge F_{t+1}{=}Q_{t+1} \wedge \ldots \wedge T_m{=}Q_m)\}$.

Аналогічно випадку 3, наведений вираз реляційного числення доменів для ненормалізованих відношень є безпечним.

Розглянемо випадок, коли проекція виконується на деяку підмножину потужності l множини атомарних атрибутів ненормалізованого відношення Ψ_1 та підмножину потужності m множини атрибутів вкладеного відношення Q. Тобто E_1 еквівалентний безпечному виразу $\{ x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n) \mid \Psi_1(x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n))\}$. Тоді еквівалентний до E вираз реляційного числення доменів для ненормалізованих відношень запишемо у вигляді:

$\{w_1, \ldots w_l, F(u_{t+1}, \ldots, u_m) \mid \exists x_1 \ldots \exists x_l\, \exists y_{t+1} \ldots \exists y_m\, (\Psi_1(x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n) \wedge w_1{=}x_1 \wedge \ldots \wedge w_l{=}x_l \wedge u_{t+1}{=}y_{t+1} \wedge \ldots \wedge u_m{=}y_m\,)\}$

Аналогічно випадку 3, наведені вирази реляційного числення доменів для ненормалізованих відношень безпечні.

Випадок 5. $E = \sigma_F(E_1)$. Нехай E_1 еквівалентний безпечному виразу $\{ x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n) \mid \Psi_1(x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n))\}$. Тоді еквівалентний до E вираз реляційного числення доменів для ненормалізованих відношень запишемо у вигляді:

$$\{ x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n) \mid \Psi_1(x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n)) \wedge F'\},$$

де F' є F, у якому кожний операнд, який позначає i-й атрибут, замінюється на x_i при $1 \le i \le t$, та y_i при $t < i \le n$. Цей вираз є безпечним, оскільки кожний компонент $x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n)$ обмежується символами з $\Psi_1(x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n))$.

Випадок 6. $E = \mathrm{NEST}_F(E_1)$. Нехай E_1 еквівалентний безпечному виразу $\{ x_1, \ldots, x_t, x_{t+1}, \ldots, x_n \mid \Psi_1(x_1, \ldots, x_t, x_{t+1}, \ldots, x_n)\}$, а F: $y_i = x_i$, де $i = t+1, \ldots, n$ позначає правило групування – утворення вкладеного відношення $Q(y_{t+1}, \ldots, y_n)$ над однаковими значеннями даних у множині атрибутів $\{ x_1, \ldots, x_t \}$. Тобто оператор групування NEST_F у цьому випадку перетворює відношення $\Psi_1(x_1, \ldots, x_t, x_{t+1}, \ldots, x_n)$ на ненормалізоване відношення $R(x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n))$. Еквівалентний до E вираз реляційного числення доменів для ненормалізованих відношень запишемо у вигляді:

$$\{ x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n) \mid \exists x_{t+1} \ldots \exists x_n\, (\Psi_1(x_1, \ldots, x_t, x_{t+1}, \ldots, x_n) \wedge y_{t+1}{=}x_{t+1} \wedge \ldots \wedge y_n{=}x_n\,)\}.$$

Цей вираз є безпечним, оскільки x_i при $1 \le i \le t$ обмежується значеннями з безпечного виразу $\Psi_1(x_1, \ldots, x_t, x_{t+1}, \ldots, x_n)$, а y_i обмежується значеннями, яких може набувати x_i, якщо $t+1 < i \le n$, з безпечного виразу $\Psi_1(x_1, \ldots, x_t, x_{t+1}, \ldots, x_n)$.

Випадок 7. $E = \mathrm{UNNEST}_F(E_1)$. Нехай E_1 еквівалентний безпечному виразу $\{ x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n) \mid \Psi_1(x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n))\}$, а F: $x_i = y_i$, де $i = t+1, \ldots, n$ позначає правило розгрупування – перетворення ненормалізованого відношення $\Psi_1(x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n)$ на плоске відношення $R(x_1, \ldots, x_t, x_{t+1}, \ldots, x_n)$. Еквівалентний до E вираз реляційного числення доменів для ненормалізованих відношень запишемо у вигляді:

$$\{ x_1, \ldots, x_t, x_{t+1}, \ldots, x_n \mid \exists y_{t+1} \ldots \exists y_n\, (\Psi_1(x_1, \ldots, x_t, Q(y_{t+1}, \ldots, y_n)) \wedge x_{t+1}{=}y_{t+1} \wedge \ldots \wedge x_n{=}y_n)\}.$$

Аналогічно до випадку 6 цей вираз реляційного числення доменів для ненормалізованих відношень безпечний. ■

Отже, вище доведено редукцію числення доменів для ненормалізованих відношень до розширеної реляційної алгебри та редукцію розширеної реляційної алгебри до числення доменів для ненормалізованих відношень, тому вирази реляційного числення доменів для ненормалізованих відношень та вирази розширеної реляційної алгебри – еквівалентні.

Висновки

Доведено еквівалентність виразів числення доменів для ненормалізованих відношень та виразів розширеної реляційної алгебри. З цією метою:

1. Введено означення безпечного виразу реляційного числення доменів для ненормалізованих відношень, що дало змогу унеможливити представлення за допомогою реляційних числень нескінченних відношень і виключити результати, що не мають змісту чи не можуть бути реалізовані.

2. Доведено теорему про редукцію числення доменів для ненормалізованих відношень до розширеної реляційної алгебри.

3. Доведено теорему про редукцію розширеної реляційної алгебри до числення доменів для ненормалізованих відношень.

1. E.F. Codd A Relational Model of Data for Large Shared Data Banks // Communications of the ACM: Volume 13, Number 6. – June, 1970. – P. 377–387. 2. Ульман Дж. Основы систем баз данных / пер. с англ. М.Р. Когаловского и В.В. Когутовского; под ред. М.Р. Когаловского. – М.: Финансы и *статистика, 1983. – 334 с. 3. Makinouchi A. A consideration on normal form of not-necessarily-normalized relation in the relational data model. In Proc. 3rd International Conference on Very Large Databases, pages 447-453, Tokyo, Oct. 1977. 4. Silberschatz A. Database System Concepts: 5th Edition / A. Silberschatz, Henry F. Korth, S. Sudarshan. - McGraw-Hill , August 9, 2005. Режим доступу: http://codex.cs.yale.edu/avi/db-book/db5/slide-dir/ch9.ppt 5. Грабовецький Ю.В., Пасічник В.В. Методичні вказівки до вивчення теми "Ненормалізовані реляційні моделі даних" курсу "Бази та банки даних і знань" для студентів спеціальності 2202 "Автоматизовані системи обробки інформації й управління". Ч.1,2. – Львів: ЛПІ, 1990. – 44 с. 6. Берко А.Ю. Системи баз даних та знань. Книга 1. Організація баз даних та знань / Берко А.Ю., Верес О.М., Пасічник В.В. – Львів: Магнолія 2006, 2008. – 456 с. 7. Такеути Г. Теория доказательств / пер. с англ. – М.: Мир, 1978. – 412 с. 8. Григорович А. Г., Григорович В. Г. Реляційне числення доменів для ненормалізованих відношень // Вісник Східноукраїнського національного університету імені Володимира Даля: науковий журнал. – № 8(179) Ч.2. – 2012 . – Луганськ. – С.24–30.*

УДК 004.89

В. Литвин
Національний університет "Львівська політехніка",
кафедра інформаційних систем та мереж

ОНТОЛОГІЧНИЙ ПІДХІД ДО ПОБУДОВИ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ УПРАВЛІННЯ СУХОПУТНИМИ ВІЙСЬКАМИ

Розглянуто побудову системи прийняття рішень управління Сухопутними військами з використанням онтологічного підходу. Описано побудову сховищ даних такої системи, ядром якої є онтологія.

Ключові слова: онтологія, інтелектуальна система прийняття рішень, база знань, Сухопутні війська.

In the paper considers the construction of the system of decision-making control of the Army Ontology-based. We describe the construction of repositories of such a system, the core of which is the ontology.

Key words: ontology, intelligent system, knowledge base, ground forces.

Постановка проблеми у загальному вигляді

Сухопутні війська (СВ) є системою, що являє собою множину складових частин – об'єктів, але меншого масштабу. Органи управління підрозділами відповідальні за їх розвиток та боєздатність. Тому формування необхідної інформаційної бази для прийняття управлінських рішень важливе як для СВ загалом, так і для окремого підрозділу нижчого рівня.

Під час математичного моделювання бойових дій [1] можна виділити ряд параметрів, які впливають на результат. До таких параметрів для моделювання бойових дій Сухопутних військ належать:

– відстань між військами;
– характеристики ходових властивостей механізованих військ;
– місцевість: