

В.І. Шекета, к.т.н., доц.  
Івано-Франківський національний технічний університет нафти і газу

## ІНФОРМАЦІЙНО-ПОШУКОВІ ЗАДАЧІ НА ОСНОВІ ОБМЕЖЕНЬ ДЛЯ НАФТОГАЗОВОЇ ПРЕДМЕТНОЇ ОБЛАСТІ

*Основною ідеєю запропонованої процедури трансформації є перехід графа бази даних в підходящі домену і введення змінних для елементів інформаційної предикатної схеми. Для представлення семантики співпадань введено окрему множину обмежень, які можуть бути категоризовані як з точки зору множини міток, так і з точки зору представлення структурної частини семантики співпадань.*

Інформаційно-пошукові задачі на основі обмежень [1], [2], [3], [4], [6], [7] формують окремих клас пошукових проблем. Вони використовуються при вирішенні прикладних задач через встановлення обмежень або властивостей, яким повинен задовольняти кожен знайдений розв'язок. У загальному випадку інформаційно-пошукова задача формулюється у вигляді набору змінних із встановленням для кожної змінної окремого домену. Множина накладених обмежень визначає набори значень, які можуть одночасно приймати введені змінні. В роботі [8] виконано побудову формально-логічного апарата інформаційної предикатної схеми як середовища виконання трансформації запитів користувача за слабоструктурованою інформацією нафтогазової предметної області. Слідуючи логіці роботи [5], в якій проаналізовано проблему оптимізації абстрактних кон'юнктивних запитів через введення формалізації засобами формального апарата пошукових задач на основі обмежень, виникає питання про можливість застосування такого підходу до інформаційної предикатної схеми, побудованої у [8].

Тому ціллію даної статті є дослідження питання побудови формально-логічного апарата для процесу трансформації проблеми встановлення співпадань інформаційної предикатної схеми та об'єктів бази даних нафтогазової предметної області, представлених у вигляді орієнтованих графів в інформаційно-пошукову задачу на основі обмежень.

Під інформаційно-пошуковою задачею на основі обмежень будемо розуміти кортеж  $(Y, O, L)$ , де:

- $Y$  – множина змінних  $(y_1, \dots, y_m)$ ;
- $O$  – множина скінчених областей, що відповідають кожній змінній  $y_i \in Y$ ;
- $L$  – множина обмежень  $\{L_{K_1}, \dots, L_{K_m}\}$ , що обмежують граничні значення, які змінні

можуть приймати одночасно.  $K_i = (y_{K_{i1}}, \dots, y_{K_{im}})$  є кортежами змінних з  $Y$ ; і кожний  $L_{K_i}$  задається як  $(L_{K_i} \subseteq O_{K_{i1}} \times \dots \times O_{K_{im}})$ .

Оскільки змінні та їх множини значень зв'язані, ми будемо називати кожен кортеж  $\langle y_i, O_i \rangle$ , де  $y_i \in Y$ , і  $O_i \in O$  – доменна змінна. Відображення змінних для даної множини доменних змінних є співставленням  $\mu: Y \rightarrow O$ , таке, що  $\mu(y_i) \in O_i$  справджується для всіх  $y_i \in Y$ . Інтерпретація змінних  $\mu$  задовольняє обмеженням  $L_{(y_{K_{i1}}, \dots, y_{K_{im}})}$  якщо  $(\mu(y_{K_{i1}}), \dots, \mu(y_{K_{im}})) \in L_{(y_{K_{i1}}, \dots, y_{K_{im}})}$ . Розв'язком для пошукової задачі на основі обмежень вважатимемо інтерпретацію змінних таку, що одночасно задовольняє всі обмеження.

**Означення 1.** Упорядковану четвірку  $\Gamma = (B, R, f_1, f_2)$  вважатимемо напрямленим інформаційним графом [8], якщо  $B$  – множина вершин,  $R$  – множина дуг, а  $f_1$  і  $f_2$  – функції, які для кожної дуги визначають її початок та кінець. В якості синоніма для терміну вершина будемо використовувати термін вузол. Але будемо вважати обов'язковим використання терміну дуга, щоб підкреслити напрямленість інформаційних графів. В нашій моделі два вузли можуть бути з'єднані за допомогою більше, ніж з одної дуги. Крім того, ми припускаємо використання петель. Введемо наступні означення:

**Означення 2.** Нехай  $N$  – множина написів, тоді впорядкована п'ятірка  $\Gamma = (B, R, f_1, f_2, n)$  є  $(N-)$  описаним напрямленим графом, якщо  $(B, R, f_1, f_2, n)$  є напрямленим графом і  $n: B \cup R \rightarrow N$  є функцією, яка присвоює кожній вершині і дузі надпис з  $N$ .

**Означення 3.** Інформаційною предикатною схемою на множині унарних предикатів  $\Pi$  будемо називати об'єкт  $\eta = (B^{(n)}, R^{(n)}, f_1^{(n)}, f_2^{(n)}, n^{(n)})$ , де кожному елементу приписаний відповідний предикат  $(n: B^{(n)} \cup R^{(n)} \rightarrow \Pi)$ .

**Базові принципи трансформації**

Ми будемо ілюструвати процес трансформації проблеми встановлення співпадання інформаційної предикатної схеми в пошукову задачу на основі обмежень. Основною ідеєю трансформації буде перехід графа бази даних в підходящі домени і ввід змінних для елементів інформаційної предикатної схеми. Більше того, будуть введені обмеження, що представлятимуть семантику співпадань. Вони можуть бути категоризовані з точки зору множини міток і з точки зору представлення структурної частини семантики співпадань.

У загальному випадку припустимо, що ми маємо об'єкт бази даних  $\Omega = (B^{(\Omega)}, R^{(\Omega)}, f_1^{(\Omega)}, f_2^{(\Omega)}, n^{(\Omega)})$  та інформаційну предикатну схему  $\eta = (B^{(n)}, R^{(n)}, f_1^{(n)}, f_2^{(n)}, n^{(n)})$ . Наступна трансформація буде виконана:

1. З графа бази даних ми виводимо домени  $O_B = B^{(\Omega)}$  і  $O_R = R^{(\Omega)}$ . У нашому випадку ми одержимо:

$$O_B = \{b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8, b_9, b_{10}, b_{11}\},$$

$$O_R = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8, r_9, r_{10}, r_{11}, r_{12}\}.$$

2. Змінна  $y_i$  вводиться для кожного елемента в  $B^{(n)} \cup R^{(n)}$ . Зауважимо, що ми використовуємо змінні як синоніми для відповідних елементів і кажемо, що " $y_i$  є дуга".  $Y$  є множиною цих змінних.

3. Для кожної змінної  $y_i \in Y$  ми означимо її домен  $O_i$  як  $O_i := O_B$ , якщо  $y_i$  є вузлом і  $O_i := O_R$ , якщо  $y_i$  є дугою.  $O$  є множиною  $O_i$ .

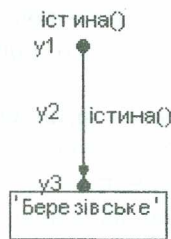


Рис. 1. Інформаційна предикатна схема

Згідно з рисунком 1 одержуємо:

$$Y = \{y_1, y_2, y_3\}, O_{y_1} = O_2 = O_B, O_{y_2} = O_R. \tag{1}$$

Для кожної змінної  $y_1$  ми введемо унарний обмежувач для мітки  $(L_{y_1}^{np} = \{(o_i) \in O_i \mid n^{(n)}(y_1)(n^{(\Omega)}(o_i)) = \text{істина}\})$ .

На рисунку 1  $y_1$  відповідає предикат *істина* (), а надпису "родовище" відповідає предикат  $Y = \text{родовище}$ , який також відповідає вузлам  $b_5, b_7, b_8$ . Тому

$$L_{(y_1)}^{np} = \{(o_i) \in O_1 \mid \text{істина}(n^{(\Omega)}(o_i))\} = \{(b_1), (b_2), (b_3), (b_4), (b_5), (b_6), (b_7), (b_8), (b_9), (b_{10}), (b_{11})\};$$

$$L_{(y_2)}^{np} = \{(o_2) \in O_2 \mid \text{істина}(n^{(\Omega)}(o_2))\} = \{(r_1), (r_2), (r_3), (r_4), (r_5), (r_6), (r_7), (r_8), (r_9), (r_{10}), (r_{11}), (r_{12})\}; \tag{2}$$

$$L_{(y_3)}^{np} = \{(o_3) \in O_3 \mid \text{Родовище} = n^{(\Omega)}(o_3)\} = \{(b_5), (b_7), (b_8)\}.$$

4. Для кожної змінної  $y_i \in Y$  і для початкової вершини  $y_{f_1}$  та кінцевої  $y_{f_2}$  ми введемо дві структури, обмежені  $L_{(y_i, y_{f_1})}^{npo}$  і  $L_{(y_i, y_{f_2})}^{kin}$ . Вони будуть забезпечувати збереження графової структури для інформаційної предикатної схеми (тобто в даному випадку властивості морфізму

$$L_{(y_i, y_{f_1})}^{npo} = \{(o_i, o_{f_{2m}}) \in O_i \times O_{f_1} \mid n^{(\Omega)}(o_i) = o_{f_{2m}}\}; \tag{3}$$

$$L_{(y_i, y_{f_1})}^{kin} = \{(o_i, o_{f_{2m}}) \in O_i \times O_{f_2} \mid f_2^{(\Omega)}(o_i) = o_{f_{2m}}\}.$$

На рисунку власне є тільки одна дуга  $y_2$ . Її початком є вершина  $y_1$ , кінцем –  $y_3$ . З точки зору введених означень, ми читаємо це, як «якщо  $y_2 \rightarrow r_1$ , тоді  $b_1$  є дійсним значенням для  $y_1$ ». А саме:

$$L_{(y_2, y_1)}^{noc} = \{(o_2, o_{1_m}) \in O_2 \times O_1 \mid \eta^{(o)}(o_2) = o_{1_m}\} =$$

$$= \{(r_1, b_1), (r_2, b_1), (r_3, b_1), (r_4, b_2), (r_5, b_4), (r_6, b_2), (r_7, b_2), (r_8, b_2), (r_9, b_3), (r_{10}, b_4), (r_{11}, b_4), (r_{12}, b_4)\},$$

$$L_{(y_2, y_3)}^{kii} = \{(o_2, o_{3_m}) \in O_2 \times O_3 \mid f_2^{(o)}(o_2) = o_{3_m}\} =$$

$$= \{(r_1, b_2), (r_2, b_3), (r_3, b_4), (r_4, b_3), (r_5, b_3), (r_6, b_5), (r_7, b_6), (r_8, b_7), (r_9, b_8), (r_{10}, b_9), (r_{11}, b_{10}), (r_{12}, b_{11})\}.$$

5. Для кожної пари змінних, що є вузлами, і для кожної пари змінних, що є ребрами, ми введемо поняття ін'єктивних обмежень. Ці обмеження будуть гарантувати, що ніякі дві вершини і ніяких два ребра в інформаційній предикатній схемі не будуть відображені в одну вершину, або ребро графа бази даних  $(L_{(y, y_j)}^{in} = \{(o_i, o_{j_m}) \in O_i \times O_j \mid o_i \neq o_{j_m}\})$ .

У нашому випадку ми маємо лише дві вершини і одне ребро. Тому ми потребуємо тільки одне обмеження для вершин  $y_1, y_3$ :

$$L_{(y_1, y_3)}^{in} = \{(o_1, o_{3_m}) \in O_1 \times O_3 \mid o_1 \neq o_{3_m}\} = \{(b_1, b_2), (b_1, b_3), (b_1, b_4), \dots, (b_1, b_{11}),$$

$$(b_2, b_1), (b_2, b_3), (b_2, b_4), \dots, (b_2, b_{11}), (b_3, b_1), (b_3, b_2), (b_3, b_4), \dots, (b_3, b_{11}),$$

$$\dots, (b_{11}, b_1), (b_{11}, b_2), (b_{11}, b_3), \dots, (b_{11}, b_{10})\}. \tag{4}$$

6. Позначимо через  $O$  множину всіх введених обмежень.

7. Розв'язком для введеної нами пошукової задачі на основі обмежень буде  $(b_2, r_6, b_5)$ ,  $(b_2, r_8, b_7)$  і  $(b_3, r_9, b_8)$  для змінних  $(y_1, y_2, b_3)$ .

**Робота із змінними та маршрутами**

Введемо змінні  $\{y_1, y_2, y_3, y_4, y_5\}$  з їх доменами. Більше того, ми введемо обмеження  $L_{(y_1)}^{np}$ ,  $L_{(y_2)}^{np}$ ,  $L_{(y_3)}^{np}$ ,  $L_{(y_4)}^{np}$  і  $L_{(y_5)}^{np}$ , що одержуються з множини міток інформаційної предикатної схеми, обмеження  $L_{(y_2, y_1)}^{noc}$ ,  $L_{(y_2, y_3)}^{kii}$ ,  $L_{(y_4, y_1)}^{noc}$  і  $L_{(y_4, y_3)}^{noc}$ , що виводяться із структури інформаційної предикатної схеми, та ін'єктивні обмеження  $L_{(y_1, y_3)}^{in}$ ,  $L_{(y_3, y_4)}^{in}$ ,  $L_{(y_1, y_5)}^{in}$  і  $L_{(y_2, y_4)}^{in}$ . Для представлення змінних, розміщених на  $y_3$  і  $y_5$ , зробимо наступне. Для кожної пари  $y_i$  і  $y_j$  змінних з  $b(y_i) = b(y_j)$  ми введемо обмеження, яке гарантуватиме, що мітки співставлення  $y_i$  з  $y_j$  задовольняють:

$$L_{(y_i, y_j)}^{zm} = \{(o_i, o_{j_m}) \in O_i \times O_j \mid \eta^{(o)}(o_i) = \eta^{(o)}(o_{j_m})\}. \tag{5}$$

У нашому випадку ми одержимо:

$$L_{(y_3, y_5)}^{zm} = \{(o_3, o_{5_j}) \in O_3 \times O_5 \mid \eta^{(o)}(o_3) = \eta^{(o)}(o_{5_j})\} = \{(b_1, b_1), (b_2, b_2), (b_3, b_3), (b_4, b_4), (b_5, b_5),$$

$$(b_5, b_7), (b_5, b_8), (b_6, b_6), (b_7, b_5), (b_7, b_7), (b_7, b_8), (b_8, b_5), (b_8, b_7), (b_8, b_8),$$

$$(b_9, b_9), (b_{10}, b_{10}), (b_{11}, b_{11})\}.$$

Тоді дана інформаційно-пошукова задача на основі обмежень матиме своїм рішенням  $(b_2, r_6, b_5, r_8, b_7)$ , що відповідає рисунку 2.

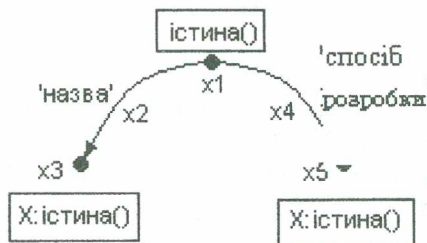


Рис. 2. Змінні в інформаційній предикатній схемі

Для обробки обходів графа ми потребуємо додаткового домену. Ми вкладемо маршрут графа бази даних в новий домен. З графа бази даних ми введемо домен  $O_T = T^{+(O)}$ .

Для розглянутої вище бази даних ми одержимо:

$$L_{(y_2, y_1)}^{noc} = \{(o_2, o_{1_m}) \in O_2 \times O_1 \mid \eta^{(o_2)} = o_{1_m}\} =$$

$$= \{(r_1, b_1), (r_2, b_1), (r_3, b_1), (r_4, b_2), (r_5, b_4), (r_6, b_2), (r_7, b_2), (r_8, b_2), (r_9, b_3), (r_{10}, b_4), (r_{11}, b_4), (r_{12}, b_4)\},$$

$$L_{(y_2, y_3)}^{kin} = \{(o_2, o_{3_m}) \in O_2 \times O_3 \mid f_2^{(o_2)} = o_{3_m}\} =$$

$$= \{(r_1, b_2), (r_2, b_3), (r_3, b_4), (r_4, b_3), (r_5, b_3), (r_6, b_5), (r_7, b_6), (r_8, b_7), (r_9, b_8), (r_{10}, b_9), (r_{11}, b_{10}), (r_{12}, b_{11})\}.$$

5. Для кожної пари змінних, що є вузлами, і для кожної пари змінних, що є ребрами, ми введемо поняття ін'єктивних обмежень. Ці обмеження будуть гарантувати, що ніякі дві вершини і ніяких два ребра в інформаційній предикатній схемі не будуть відображені в одну вершину, або ребро графа бази даних  $(L_{(y, y_j)}^{in} = \{(o_i, o_{j_m}) \in O_i \times O_j \mid o_i \neq o_{j_m}\})$ .

У нашому випадку ми маємо лише дві вершини і одне ребро. Тому ми потребуємо тільки одне обмеження для вершин  $y_1, y_3$ :

$$L_{(y_1, y_3)}^{in} = \{(o_1, o_{3_m}) \in O_1 \times O_3 \mid o_1 \neq o_{3_m}\} = \{(b_1, b_2), (b_1, b_3), (b_1, b_4), \dots, (b_1, b_{11}),$$

$$(b_2, b_1), (b_2, b_3), (b_2, b_4), \dots, (b_2, b_{11}), (b_3, b_1), (b_3, b_2), (b_3, b_4), \dots, (b_3, b_{11}),$$

$$\dots, (b_{11}, b_1), (b_{11}, b_2), (b_{11}, b_3), \dots, (b_{11}, b_{10})\}. \tag{4}$$

6. Позначимо через  $O$  множину всіх введених обмежень.

7. Розв'язком для введеної нами пошукової задачі на основі обмежень буде  $(b_2, r_6, b_5)$ ,  $(b_2, r_8, b_7)$  і  $(b_3, r_9, b_8)$  для змінних  $(y_1, y_2, b_3)$ .

### Робота із змінними та маршрутами

Введемо змінні  $\{y_1, y_2, y_3, y_4, y_5\}$  з їх доменами. Більше того, ми введемо обмеження  $L_{(y_1)}^{np}$ ,  $L_{(y_2)}^{np}$ ,  $L_{(y_3)}^{np}$ ,  $L_{(y_4)}^{np}$  і  $L_{(y_5)}^{np}$ , що одержуються з множини міток інформаційної предикатної схеми, обмеження  $L_{(y_2, y_1)}^{noc}$ ,  $L_{(y_2, y_3)}^{kin}$ ,  $L_{(y_4, y_1)}^{noc}$  і  $L_{(y_4, y_5)}^{noc}$ , що виводяться із структури інформаційної предикатної схеми, та ін'єктивні обмеження  $L_{(y_1, y_3)}^{in}$ ,  $L_{(y_3, y_4)}^{in}$ ,  $L_{(y_1, y_5)}^{in}$  і  $L_{(y_2, y_4)}^{in}$ . Для представлення змінних, розміщених на  $y_3$  і  $y_5$ , зробимо наступне. Для кожної пари  $y_i$  і  $y_j$  змінних з  $b(y_i) = b(y_j)$  ми введемо обмеження, яке гарантуватиме, що мітки співставлення  $y_i$  з  $y_j$  задовольняють:

$$L_{(y_i, y_j)}^{zm} = \{(o_i, o_{j_m}) \in O_i \times O_j \mid \eta^{(o_i)} = \eta^{(o_{j_m})}\}. \tag{5}$$

У нашому випадку ми одержимо:

$$L_{(y_3, y_5)}^{zm} = \{(o_3, o_{5_m}) \in O_3 \times O_5 \mid \eta^{(o_3)} = \eta^{(o_{5_m})}\} = \{(b_1, b_1), (b_2, b_2), (b_3, b_3), (b_4, b_4), (b_5, b_5),$$

$$(b_5, b_7), (b_5, b_8), (b_6, b_6), (b_7, b_5), (b_7, b_7), (b_7, b_8), (b_8, b_5), (b_8, b_7), (b_8, b_8),$$

$$(b_9, b_9), (b_{10}, b_{10}), (b_{11}, b_{11})\}.$$

Тоді дана інформаційно-пошукова задача на основі обмежень матиме своїм рішенням  $(b_2, r_6, b_5, r_8, b_7)$ , що відповідає рисунку 2.

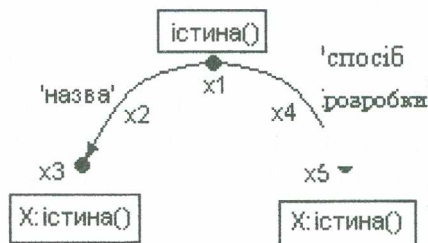


Рис. 2. Змінні в інформаційній предикатній схемі

Для обробки обходів графа ми потребуємо додаткового домену. Ми вкладемо маршрут графа бази даних в новий домен. З графа бази даних ми введемо домен  $O_T = T^{+(S)}$ .

Для розглянутої вище бази даних ми одержимо:

$$O_T = \{(r_1), (r_2), (r_3), (r_4), (r_5), (r_6), (r_7), (r_8), (r_9), (r_{10}), (r_{11}), (r_{12}), (r_1, r_4), (r_1, r_6), (r_1, r_7), (r_1, r_8), (r_2, r_9), (r_3, r_5), (r_3, r_{10}), (r_3, r_{11}), (r_3, r_{12}), (r_4, r_9), (r_5, r_9), (r_1, r_4, r_9), (r_3, r_5, r_9)\}.$$

Розглянемо рисунок 3.



Рис. 3. Маршрути в інформаційній предикатній схемі

Як ми бачимо, змінна  $y_2$  одержує  $O_T$  в якості свого домену, тоді як  $y_1, y_3$  і  $y_5$  одержують  $O_B$  і  $y_4$  одержує  $O_R$ , як і раніше. Обмеження  $L_{(y_1)}^{np}$ ,  $L_{(y_3)}^{np}$ ,  $L_{(y_4)}^{np}$  і  $L_{(y_5)}^{np}$ , і  $L_{(y_1, y_3)}^{поч}$ ,  $L_{(y_4, y_5)}^{поч}$ ,  $L_{(y_1, y_3)}^{кін}$ , і  $L_{(y_1, y_5)}^{кін}$  означаються так само, як і раніше. Тільки обмеження, що включає  $y_2$ , повинно бути адаптовано. А саме:

1. Для кожної дуги  $y_i$  з доменом  $O_T$  ми введемо унарний обмежувач мітки  $L_{(y_i)}^{np}$ , що представляє семантику відповідного предикату в інформаційній предикатній схемі. Маршрут задовольнятиме обмеження, якщо надписи всіх його ребер встановлюють для відповідного предиката  $y_i$  значення істини

$$L_{(y_i)}^{np} = \{(o_i) \in O_i \mid \forall r_j \in o_i : n^{(n)}(y_i)(n^{(o)}(r_j)) = \text{істина}\} \quad (6)$$

2. Для кожного ребра  $y_i \in R_{ij}$  з доменом  $O_T$  і вершинами  $y_{f_1}, y_{f_2}$  ми введемо два структурних обмежувачі  $L_{(y_i, y_{f_1})}^{поч}$  і  $L_{(y_i, y_{f_2})}^{кін}$ , що гарантуватиме збереження графової структури інформаційної предикатної схеми:

$$L_{(y_i, y_{f_1})}^{поч} = \{(o_i, o_{f_1}) \in O_i \times O_{f_1} \mid f_{1r}^{(o)}(o_i) = o_{f_1}\}, \quad (7)$$

$$L_{(y_i, y_{f_2})}^{кін} = \{(o_i, o_{f_2}) \in O_i \times O_{f_2} \mid f_{2r}^{(o)}(o_i) = o_{f_2}\}.$$

У нашому випадку це приводить до:

$$L_{(y_2)}^{np} = \{(o_2) \in O_2 \mid \forall r_j \in o_2 : \text{істина}(n^{(n)}(y_1)(n^{(o)}(r_j)))\}$$

$$= \{(r_1), (r_2), (r_3), (r_4), (r_5), (r_6), (r_7), (r_8), (r_9), (r_{10}), (r_{11}), (r_{12}), (r_1, r_4), (r_1, r_6), (r_1, r_7), (r_1, r_8), (r_2, r_9), (r_3, r_5), (r_3, r_{10}), (r_3, r_{11}), (r_3, r_{12}), (r_4, r_9), (r_5, r_9), (r_1, r_4, r_9), (r_3, r_5, r_9)\};$$

$$L_{(y_2, y_1)}^{поч} = \{(o_2, o_{1m}) \in O_2 \times O_1 \mid f_{1r}^{(o)}(o_2) = o_{1m}\} = \{(r_1, b_1), (r_2, b_1), (r_3, b_1), (r_4, b_2), (r_5, b_4), (r_6, b_2), (r_7, b_2), (r_8, b_2), (r_9, b_3), (r_{10}, b_4), (r_{11}, b_4), (r_{12}, b_4), (r_1, r_4, b_1), (r_1, r_6, b_1), (r_1, r_7, b_1), (r_1, r_8, b_1), (r_2, r_9, b_1), (r_3, r_5, b_1), (r_3, r_{10}, b_1), (r_3, r_{11}, b_1), (r_3, r_{12}, b_1), (r_4, r_9, b_2), (r_5, r_9, b_4), (r_1, r_4, r_9, b_1), (r_3, r_5, r_9, b_1)\};$$

$$L_{(y_2, y_3)}^{кін} = \{(o_2, o_3) \in O_2 \times O_3 \mid f_{2r}^{(o)}(o_2) = o_3\} = \{(r_1, b_2), (r_2, b_3), (r_3, b_4), (r_4, b_3), (r_5, b_3), (r_6, b_5), (r_7, b_6), (r_8, b_7), (r_9, b_8), (r_{10}, b_9), (r_{11}, b_{10}), (r_{12}, b_{11}), (r_1, r_4, b_3), (r_1, r_6, b_5), (r_1, r_7, b_6), (r_1, r_8, b_7), (r_2, r_9, b_8), (r_3, r_5, b_3), (r_3, r_{10}, b_9), (r_3, r_{11}, b_{10}), (r_3, r_{12}, b_{11}), (r_4, r_9, b_8), (r_5, r_9, b_8), (r_1, r_4, r_9, b_8), (r_3, r_5, r_9, b_8)\}.$$

Крім того, ми потребуємо двох додаткових обмежувачів, що представляють умови на довжину маршруту.

Для кожної дуги  $y_i \in R_{ij}$  з доменом  $O_T$  ми введемо два обмежувачі. Вони відображатимуть умови на довжину маршруту, які означені як:  $\varphi_{\min}^{(ij)}$  і  $\varphi_{\max}^{(ij)}$ :

$$(L_{y_i}^{\min} = \{(o_i) \in O_i \mid \varphi_{\min}^{(ij)}(y_i) \leq \text{stretch}(o_i)\}, \quad L_{y_i}^{\max} = \{(o_i) \in O_i \mid \varphi_{\max}^{(ij)}(y_i) \geq \text{stretch}(o_i)\}). \quad (8)$$

Оскільки  $\varphi_{\min}^{(ij)}(y_2) = 1$  і  $\varphi_{\max}^{(ij)}(y_2) = +\infty$ , то і всі обмеження можна описати як  $L_{(y_2)}^{pp}$ . Ми можемо помітити, що домен ребер  $O_R$  є особливим випадком домену обходів  $O_T$ , з уже існуючими в ньому обмеженнями, тому ми повинні задати більш точні обмеження на ін'єктивне відображення. Оскільки  $O_R$  є підмножиною  $O_T$ , ми введемо зв'язування, що поєднує змінні, що відповідають маршрутам, і змінні, що відповідають ребрам.

Для кожної пари змінних  $(y_i, y_j)$ , що є ребрами, ми введемо ін'єктивне обмеження. Ці обмеження гарантуватимуть, що ніякі ребра в інформаційній предикатній схемі не співставлятимуться з тим самим маршрутом в графі бази даних. Тобто  $(C_{(y_i, y_j)}''' = \{(o_i, o_{j_m}) \in O_i \times O_j \mid o_i \neq o_{j_m}\})$ . Ми вводимо це означення, виходячи з припущення, що ребро інформаційної предикатної схеми є рівним атомарному маршруту, що складається з цього ребра.

У нашому випадку це приводить до:

$$(L_{(y_2, y_4)}''' = \{(o_2, o_{4_m}) \in O_2 \times O_4 \mid o_2 \neq o_{4_m}\}) = \{(r_1, r_2), (r_1, r_3), (r_1, r_4), \dots, (r_1, r_{12}), (r_2, r_1), (r_2, r_3), (r_2, r_4), \dots, (r_2, r_{12}), \dots, (r_{12}, r_1), (r_{12}, r_2), (r_{12}, r_3), \dots, (r_{12}, r_{11}), (r_1, r_4, r_1), (r_1, r_4, r_2), (r_1, r_4, r_3), \dots, (r_1, r_4, r_{12}), \dots, (r_3, r_5, r_9, r_1), (r_3, r_5, r_9, r_2), (r_3, r_5, r_9, r_3), \dots, (r_3, r_5, r_9, r_{12})\}.$$

Цей крок і завершує трансформацію. Для нашого прикладу ми знайшли наступні рішення:  $(b_1, (r_1), b_2, r_6, b_5)$ ,  $(b_1, (r_1, r_4), b_3, r_9, b_8)$ ,  $(b_1, (r_2), b_3, r_9, b_8)$ ,  $(b_1, (r_3, r_5), b_3, r_9, b_8)$ , і  $(b_1, (r_3), b_4, r_{10}, b_9)$ , що відповідає рисунку 3.

Таким чином, основною ідеєю запропонованої процедури трансформації є перехід графа бази даних в підходящі домени і введення змінних для елементів інформаційної предикатної схеми. Для представлення семантики співпадань введено окрему множину обмежень, які можуть бути категоризовані як і з точки зору множини міток, так і з точки зору представлення структурної частини семантики співпадань.

Одержаний результат є обґрунтованим, оскільки для кожної пари змінних, що є ребрами, ми ввели ін'єктивні обмеження, які гарантують, що ніякі ребра в інформаційній предикатній схемі не співставлятимуться з тим самим маршрутом в графі бази даних.

**Висновки.** В даній роботі виконано побудову формально-логічного апарата для процедури трансформації проблеми встановлення співпадань інформаційної предикатної схеми та об'єктів бази даних нафтогазової предметної області, представлених у вигляді орієнтованих графів в інформаційно-пошукову задачу на основі обмежень, для якої одержано рішення для випадку набору із п'яти змінних.

Подальші розвідки в даному напрямі будуть спрямовані на дослідження питання коректності та повноти такої трансформації.

### ЛІТЕРАТУРА:

1. E. F. Codd. Data models in database management. In Proceedings of the Workshop on Data Abstraction, Databases, and Conceptual Modelling, pages 112–114, Pingree Park, CO, USA, June, 1980.
2. P. Buneman, S. Davidson, M. Fernandez, and D. Suciu. Adding structure to unstructured data. In Proceedings of the International Conference on Database Theory (ICDT), pages 336–350, Delphi, Greece, January 1997.
3. M. Deutsch, M. Fernandez, and D. Suciu. Storing semistructured data with STORED. In Proceedings of the ACM SIGMOD International Conference on Management of Data, pages 431–442, Philadelphia, PA, USA, June 1999.

4. *S. Chawathe, H. Garcia-Molina, J. Hammer, K. Ireland, Y. Papakonstantinou, J. Ullman, and J. Widom.* The TSIMMIS project: Integration of heterogenous information sources. In Proceedings of the Information Processing Society of Japan (IPSJ) Conference, pages 7–18, Tokyo, Japan, October 1994.
5. *K. Chandra and P. M. Merlin.* Optimal implementation of conjunctive queries in relational databases. In Proceedings of the ACM Symposium on Theory of Computing, pages 77–90, Boulder, CO, USA, May 1977.
6. *D. Florescu, A. Levy, and D. Suciu.* Query containment for conjunctive queries with regular expressions. In Proceedings of the Symposium on Principles of Database Systems (PODS), pages 139-148, Seattle, WA, USA, June 1998.
7. *Солдатов В.Н., Чудинов И.Л., Ямпольский В.З.* Банки данных в нефтяной промышленности. – Новосибирск: Наука, 1988. – 126 с.
8. *Шекета В.І.* Побудова інформаційної предикатної схеми як середовища виконання трансформації запитів користувача по напівструктурованій інформації нафтогазової предметної області // Науковий вісник ІФНТУНГ. – 2003.

ШЕКЕТА Василь Іванович – кандидат технічних наук, доцент кафедри прикладної математики Івано-Франківського національного технічного університету нафти і газу.

Наукові інтереси:

- бази даних та знань;
- експертні інформаційні системи;
- дослідження семантик абстрактного логічного програмування.

Тел. роб.: 8 (0342) 24-21-27

E-Mail: sheketa@rambler.ru

abba1973@mail.ru

Подано 10.10.2003