

УДК 681. 3:622. 276

В.Шекета, канд.техн.наук

Івано-Франківський національний технічний університет нафти і газу

МОДИФІКАЦІЙНІ ПРЕДИКАТНІ ЗАПИТИ ЯК ІНСТРУМЕНТ ПІДТРИМКИ ДІАЛОГУ З КОРИСТУВАЧЕМ В ІНФОРМАЦІЙНИХ СИСТЕМАХ НА ОСНОВІ БАЗ ДАНИХ І ЗНАНЬ

Виконано побудову формально-логічного апарату для модифікаційних предикатних запитів, як універсального інструменту підтримки діалогу з користувачем в інформаційних системах на основі баз даних і знань нафтогазової предметної області.

Умовні позначення

O	– скінченний інформаційний простір;
K_B	– база знань;
Q_M	– модифікаційний предикатний запит;
$K_{B+}(), K_{B-}()$	– модифікаційні літерали;
λ_{mt}	– абстрактний оператор модифікації;
\ll	– дескриптор модифікації;
O_I	– множина інерції;
Δ_L	– логічна програма;
E_H	– екземпляр Гербранда.

Одним із ключових питань з точки зору математичного моделювання процесу побудови інформаційних інтелектуальних систем для нафтогазової предметної області є спосіб представлення знань, на основі якого система повинна приймати рішення в

певній ситуації [1]. Таким чином, представлення знань повинно бути задано способом, який дозволяє перехід до представлення фрагментів інформації про нафтогазовий об'єкт в термінах структур баз даних (БД) і знань (БЗ). Розглядатимемо таку базу даних з точки зору фактів і процесів, що призводять до їх зміни, тобто з погляду семантики і синтаксису такого представлення [2,3]. Під синтаксисом будемо розуміти набір правил для поєднання символів в логічно коректні вирази, а під семантикою – спосіб інтерпретації виразів, що одержуються в результаті конкретних реалізацій синтаксичних правил.

Теорія модифікаційних предикатних запитів [4] є формально-логічним апаратом опису і вивчення процесів оновлення і модифікації баз даних і знань, логічного висновку на основі баз даних і знань [5-9]. В першому наближенні під базою даних будемо розуміти набір фактів. Основні ідеї такого підходу розглядаються в рамках конкретних реалізацій SQL або реалізацій для WWW і WEB[2,10].

На відміну від розглянутого підходу, де функцією логічного обґрунтування запитів перекладено на користувача бази даних, і від Prolog-програм, де передбачається побудова бази знань інженером когнітологом і постійний супровід користувача під час сеансу логічного висновку, в пропонованому нами підході інформаційна інтелектуальна система(ІС) на основі БД і БЗ сама формує та поповнює базу даних і знань і проводить логічний висновок.

Тому ціллю даного дослідження є розробка формально-логічного апарату підтримки діалогу користувача з ІС на основі БД і БЗ для нафтогазової предметної області в процесі побудови ним запитів, що розглядаються в ході динаміки процесу оновлення і модифікації баз даних і знань. У даній роботі ми будемо розвивати власний формально-логічний апарат, на основі основних принципів абстрактного логічного програмування [3] і теоретичних аспектів програмування в мові Prolog [11].

Розглядатимемо БЗ як набір інформаційних сутностей атомарних предикатів з деякого скінченного інформаційного простору O . Всі зміни, що відбуватимуться в БЗ, будемо розглядати як наслідок модифікаційних предикатних запитів, що генеруються ІС відповідно до вказівок користувача. Основою самих запитів є набір модифікаційних предикатних правил. Розглядатимемо два типи правил:

$$K_{B+}(o) \ll K_{B+}(o_1), \dots, K_{B+}(o_l), K_{B-}(p_1), \dots, K_{B-}(p_m), \quad (1)$$

$$K_{B-}(o) \ll K_{B+}(o_1), \dots, K_{B+}(o_l), K_{B-}(p_1), \dots, K_{B-}(p_m), \quad (2)$$

де $o, o_i, p_j \in O$. Основна ідея такого запису правил полягає в тому, що $K_{B+}(o)$ означає, що атомарний предикат o повинен бути включений в БЗ K_B , а K_{B-} означає, що o – повинен бути виключений з бази знань.

Таким чином, модифікаційні предикатні правила задають логічні обмеження щодо БЗ. Зокрема, правило (1) накладає на БЗ наступну умову: o належить БЗ, або принаймні одне з o_i $1 \leq i \leq l$ не належить БЗ, або принаймні одне з p_j $1 \leq j \leq m$ належить БЗ. Правило (2) стверджує, що для випадку, коли o і всі o_i $1 \leq i \leq l$ належать БЗ, а всі p_j $1 \leq j \leq m$ не належать БЗ, то можемо виключити o з БЗ, або вилучити одне з o_i , або додати одне з p_j .

Розглядатимемо тепер O як деяку скінчену множину, елементи якої називатимемо атомами. Розглянемо також множину $\{K_{B_1}, \dots, K_{B_l}\}$ БЗ, таких, що $K_{B_i} \subseteq O$, де $1 \leq i \leq l$. Вирази виду $K_{B+}(o)$, $K_{B-}(o)$, де $o \in O$ – деякий атом, назвемо модифікаційними літералами. Модифікаційні літерали $K_{B+}()$, $K_{B-}()$ розглядатимемо як взаємообернені і зарезервуємо для їх позначення в загальному випадку послідовність грецьких букв, починаючи з τ : τ, υ, φ . Обернений модифікаційний літерал для τ позначатимемо через τ^{-1} .

Нехай Q_M – модифікаційний предикатний запит і K_B – база знань. Введення семантики задасть специфікацію для множин БЗ, кожна з яких може бути вибрана як

модифікація вихідної бази знань K_B^{noch} після виконання модифікаційного предикатного запиту Q_M . Таку множину $\{K_{B_1}^m, \dots, K_{B_l}^m\}$ називатимемо множиною Q_M -модифікацій

вихідної БЗ K_B^{noch} . Кожна із Q_M -модифікацій буде в свою чергу теж задовольняти всі обмеження, накладені на K_B^{noch} .

Спочатку розглянемо випадок порожньої вихідної БЗ. Кожний модифікаційний предикатний запит можна розділити на дві частини: K_{B^+} - правила, тобто такі, що:

$$\{\Delta \in Q_M : head(\Delta) = K_{B^+}(o), \text{ для деякого } o \in O\}$$

і K_{B^-} - правила, такі, що:

$$\{\Delta \in Q_M : head(\Delta) = K_{B^-}(o), \text{ для деякого } o \in O\}.$$

Причому

$$Q_M = K_{B^+}^+(Q_M) \cup K_{B^-}^-(Q_M) \cup Q_M^S,$$

де Q_M^S - додаткові специфікації модифікаційного запиту, які не описують жодних модифікацій стосовно вихідної бази знань. Кожна із таких модифікацій повинна бути описана в $K_{B^+}^+(Q_M)$.

Почнемо розгляд процесу виконання модифікаційних запитів для операції вилучення або додавання правил до БЗ. Тоді в термінах теорії множини одержимо, що:

$$K_B^{noch} \cup \{o\} \Leftrightarrow \{K_{B^+}(o) \ll\},$$

$$K_B^{noch} \setminus \{o\} \Leftrightarrow \{K_{B^-}(o) \ll\}.$$

Нехай модифікаційний запит Q_M для вихідної БЗ K_B^{noch} задано у вигляді:

$$K_{B^+}(o_1) \ll, \dots, K_{B^-}(o_l) \ll, \dots, K_{B^-}(p_1) \ll, \dots, K_{B^-}(p_m) \ll,$$

Якщо Q_M не є логічно суперечливим, тобто $\{o_1, \dots, o_l\} \cap \{p_1, \dots, p_m\} = \emptyset$, тоді множина Q_M -модифікацій для K_B^{noch} матиме вигляд

$$Q_M \Big|_{K_B^{noch}} \Leftrightarrow (K_B^{noch} \cup \{o_1, \dots, o_l\}) \setminus \{p_1, \dots, p_m\}.$$

Таким чином, будь-яке оновлення вихідної БЗ K_B^{noch} можна розглядати як наслідок дії модифікаційного предикатного запиту. Але це є випадком самого простого запиту даного типу. В загальному випадку виражатимемо комплексні логічні умови. Для БЗ нафтогазової предметної області розглядатимемо запити типу:

$$Q_M = \{K_{B^+}(\text{насиченість_породи} > 1) \ll, K_{B^+}(\text{порода_колектор}) \ll, \\ K_{B^+}(\text{насиченість_породи} > 1), K_{B^-}(\text{пористість_породи} > 1)\}.$$

Дане означення модифікаційного предикатного запиту будується на принципі інерції БЗ[3]. В нашому випадку даний принцип говорить, що статус кожного елемента БЗ залишається сталим, доки він не буде змінений в результаті модифікації. Таким чином, кожна зміна, внесена в БЗ в процесі виконання модифікаційного запиту, відповідає деякому модифікаційному предикатному правилу із загальної множини модифікаційних правил, на основі яких власне і виконується запит. Даний принцип інерції БЗ тісно пов'язаний із принципом теорії абстрактного логічного програмування[11], згідно з яким кожний атомарний предикат, що може бути хибним, слід розглядати як хибний.

Нехай $K_B^{нов}$ - вихідна БЗ, яка ініціалізує всі атомарні предикати значенням “false”. Надалі, покажемо, що введене нами поняття модифікаційного предикатного запиту включає в себе, як частковий випадок згаданий принцип .

Основні ідеї теорії модифікаційних запитів тісно пов’язані із принципами логічного програмування[3], згідно з якими знання розглядається як множина логічних аксіом (правил), на якій вводиться поняття інтерпретатора, що дозволяє здійснювати логічний висновок на основі заданих початкових атомарних фактів. При такому підході правила БЗ інтерпретатора представлені у вигляді $T_d \leftarrow T_1, \dots, T_n$, де T_i – логічні твердження. Якщо T_1, \dots, T_n є істинними, тоді T_d – теж вважається істинним.

Побудуємо друге означення модифікаційних предикатних запитів, виходячи із альтернативного означення залишку для Q_M .

Означення 1. Нехай Q_M -модифікаційний запит, а K_{B_1} і K_{B_2} - дві БЗ. Означимо залишок для Q_M стосовно (K_{B_1}, K_{B_2}) як $Q_M^{K_{B_2}} | K_{B_1}$, який одержується в результаті:

1) видалення з Q_M кожного модифікаційного правила, body-частина якого не співставляється з K_{B_2} (в результаті одержимо $Q_M^{K_{B_2}}$);

2) видалення кожного модифікаційного літералу з body-частин правил в $Q_M^{K_{B_2}}$, що співставляється з K_{B_1} .

Означення 2. Нехай Q_M -модифікаційний запит і K_{B_1} , K_{B_2} – дві БЗ. Тоді справедливо:

$$\lambda_{nm}(Q_M^{K_{B_1}, K_{B_2}}) \in \text{когерентним і } K_{B_2} = K_{B_1} \circ \lambda_{nm}(Q_M^{K_{B_1}, K_{B_2}}),$$

$$\lambda_{nm}(Q_M^{K_{B_2}} | K_{B_1}) \in \text{когерентним і } K_{B_2} = K_{B_1} \circ \lambda_{nm}(Q_M^{K_{B_2}} | K_{B_1}).$$

Означення 3. Нехай Q_M -модифікаційний запит і $K_{B_2} \in Q_M$ -модифікацією для K_{B_1} . Тоді

$$\lambda_{nm}(Q_M^{K_{B_1}, K_{B_2}}) = \lambda_{nm}(Q_M^{K_{B_2}} | K_{B_1}) = \text{head}(Q_M^{K_{B_2}}).$$

Означення 4. Виходячи із вищесказаного, справедливим буде наступне :

1) БЗ $K_{B_2} \in Q_M$ -модифікацією БЗ K_{B_1} ;

2) $\lambda_{nm}(Q_M \cup \{o \ll o \in O : (K_{B_1}, K_{B_2})\}) \subset K_{B_2}$;

3) $\lambda_{nm}(Q_M^{K_{B_1}, K_{B_2}}) \cup O(K_{B_1}, K_{B_2}) \subset K_{B_2}$.

Означення 5. Нехай Q_M -модифікаційний запит і K_{B_1} – БЗ. Якщо БЗ $K_{B_2} \in Q_M$ – модифікацією для K_{B_1} , тоді K_{B_2} є моделлю для Q_M .

Введене нами означення модифікаційного запиту задовольняє також принцип мінімальності [12]. Тобто, якщо виконати Q_M -модифікацією для K_{B_1} , то одержана БЗ K_{B_2} відрізнятиметься від вихідної мінімально. Для вимірювання цієї характеристики використаємо принцип симетричної різниці для БЗ [12]. А саме:

$$\text{dist}(K_{B_1}, K_{B_2}) = (K_{B_1} \setminus K_{B_2}) \cup (K_{B_2} \setminus K_{B_1}).$$

Означення 6. Нехай Q_M -модифікаційний запит, а K_{B_1} – БЗ. Якщо $K_{B_2} \in Q_M$ -модифікацією для K_{B_1} , тоді різниця $\text{dist}(K_{B_2}, K_{B_1})$ є мінімальною, тобто

$dist(K_{B_3}, K_{B_1}) \Rightarrow K_{B_3}$ є моделлю для Q_M . Таким чином, якщо додамо до модифікаційного запиту Q_M деякі правила, що задовольняються процедурою модифікації, то це не вплине на результат модифікації вихідної БЗ K_{B_1} .

Означення 7. Нехай K_{B_2} є Q_M -модифікацією для K_{B_1} . Припустимо, що Q_M є модифікаційним запитом, таким, що K_{B_2} є моделлю для Q_M . Тоді K_{B_2} є $(Q_M \cup Q_M)$ -модифікацією для K_{B_1} .

Означення 8. Якщо БЗ K_{B_1} є результатом дії модифікаційного запиту Q_M , тоді K_{B_1} є унікальною Q_M -модифікацією.

Нехай Q_M – деякий модифікаційний запит. Введемо означення відповідного йому оберненого модифікаційного запиту Q_M^{-1} , отриманого із Q_M внаслідок одночасної заміни всіх модифікаційних літералів на їм обернені.

Означення 9. Нехай Q_M -модифікаційний запит і нехай K_{B_1} – БЗ. Тоді K_{B_2} є Q_M -модифікацією для K_{B_1} тоді і тільки тоді, якщо $O \setminus K_{B_2}$ є Q_M^{-1} модифікацією для $O \setminus K_{B_1}$.

Розглянемо твердження t логічної програми виду $S_A \leftarrow R_{A_1}, \dots, R_{A_1}(\overline{V_{A_1}, \dots, V_{A_m}})$, для якого означимо модифікаційне предикатне правило $R_{Q_M}(t)$ як :

$$K_{B_+}(S_A) \ll K_{B_+}(R_{A_1}), \dots, K_{B_+}(R_{A_1}), K_{B_-}(V_{A_1}), \dots, K_{B_-}(V_{A_m}) .$$

Таким чином, для логічної програми Δ_L означаємо відповідний модифікаційний предикатний запит $R_{Q_M}(\Delta_L)$

$$R_{Q_M}(\Delta_L) = \{R_{Q_M}(t) : t \in \Delta_L\}.$$

Відповідно до введеного означення можна подати моделі логічних програм і стабільні моделі логічних програм [12,13], як прості моделі і Q_M -модифікації для самих модифікаційних предикатних запитів, що дослідимо надалі.

Твердження 1. Нехай Δ_L -логічна програма. Тоді

1) множина атомів A_M є моделлю для Δ_L тоді і тільки тоді, якщо A_M є моделлю для $R_{Q_M}(\Delta_L)$;

2) множина атомів для A_M є стабільною моделлю для Δ_L тоді і тільки тоді, якщо A_M є $R_{Q_M}(\Delta_L)$ -модифікацією для $\{0\}$.

Для того, щоб розширити коло застосування модифікаційних предикатних запитів і зробити їх більш гнучкими, введемо поняття змінних для модифікаційного запиту. Семантичні аспекти такого розширення базових означень не виходитимуть за рамки базової семантики модифікаційних предикатних запитів.

Будемо притримуватися тих самих означень для алфавіту, термінів, атомів, як і раніше, також використаємо ті самі означення модифікаційних літералів, модифікаційних правил і модифікаційних предикатних запитів, тільки тепер атомарні предикати будуть містити змінні.

Екземпляри атомів конструюватимемо через заміну змінних в атомі на базові терми. Аналогічно екземпляри модифікаційних літералів будуть конструюватися через заміну змінних в модифікаційних правилах на базові терми. Екземпляром Гербранда для модифікаційного запиту Q_M будемо називати множину всіх базових екземплярів модифікаційних правил запиту Q_M , що можуть бути сконструйовані з використанням базових термів універсуму Гербранда [12].

Введемо також змінні для початкової БЗ. Для цього розглядатимемо її, як множину атомарних предикатів, які можуть містити змінні. Тоді введемо означення екземпляра Гербранда для БЗ $K_B, E_H(K_B)$ як множину всіх базових екземплярів атомарних предикатів для БЗ K_B .

Тепер дамо нове означення для модифікаційного запиту:

Означення 10. Нехай Q_M -модифікаційний запит і K_{B_1} - вихідна БЗ, причому Q_M і $K_{B_1}^1$ можуть містити змінні. БЗ $K_{B_2} \in Q_M$ -модифікацією K_{B_1} , якщо $K_{B_2} \in E_H(Q_M)$ - модифікацією для $E_H(K_{B_1})$. Проілюструємо введені означення на прикладі.

Приклад 1. Нехай алфавіт O_A складається з двох констант {істина,хиба}, однієї змінної $\{Y\}$ і двох унарних предикатів {колектор(),нафтогазоносний()}. Нехай модифікаційний предикатний запит Q_M задано у вигляді

$$K_{B_+}(\text{нафтогазоносний}(Y)) \ll K_{B_+}(\text{колектор}(Y));$$

$$K_{B_-}(\text{колектор}(Y)) \ll K_{B_+}(\text{нафтогазоносний}(Y)).$$

Нехай $K_{B_1} = \{\text{колектор}(Y), \text{нафтогазоносний}(Y)\}$. Якщо {істина, хиба} приймемо в якості універсуму Гербранда, то тоді базисом Гербранда буде

$$\{\text{колектор}(істина), \text{колектор}(хиба), \\ \text{нафтогазоносний}(істина), \text{нафтогазоносний}(хиба)\}$$

Екземпляром Гербранда для Q_M буде:

$$\{K_{B_-}(\text{нафтогазоносний}(істина)) \ll K_{B_+}(\text{колектор}(істина)) \\ K_{B_-}(\text{нафтогазоносний}(хиба)) \ll K_{B_+}(\text{колектор}(хиба)) \\ K_{B_-}(\text{колектор}(істина)) \ll K_{B_+}(\text{нафтогазоносний}(істина)) \\ K_{B_-}(\text{колектор}(хиба)) \ll K_{B_+}(\text{нафтогазоносний}(хиба))\}.$$

А екземпляром Гербранда для K_{B_1} буде

$$\{\text{колектор}(істина), \text{колектор}(хиба), \\ \text{нафтогазоносний}(істина), \text{нафтогазоносний}(хиба)\}$$

Таким чином, Q_M - модифікацією для K_{B_1} будуть

$$K_{B_2}^1 = \{\text{колектор}(істина), \text{нафтогазоносність}(хиба)\}$$

$$K_{B_2}^2 = \{\text{нафтогазоносність}(істина), \text{колектор}(хиба)\}.$$

Даний приклад завершує введення базових означень для формально-логічного апарату модифікаційних предикатних запитів для інформаційних систем на основі баз даних і знань нафтогазової предметної області. Введені означення є обґрунтованими, оскільки при їх побудові ми не виходили за рамки процедури обчислення обґрунтованих семантик, прийнятих в теорії абстрактного логічного програмування і стабільних семантик для абстрактних логічних програм.

Висновки

В даній статті виконано побудову формально-логічного апарату для модифікаційних предикатних запитів як універсального інструменту підтримки діалогу з користувачем в інформаційних системах на основі баз даних і знань нафтогазової предметної області. Подальші розвідки даного напрямку будуть зосереджені на розширенні введених означень і дослідженні їх властивостей.

This paper is dealing with the construction of formal-logical apparatus for the predicate queries modifications, as an universal instrument for supporting dialog with an user in the information systems on the basis of databases and knowledgebases of oil and gas subject domain.

Література

1. Шекета В.І. Інформаційна система для прогнозування нафтогазоносних покладів. Дисертація на здобуття наукового ступеня кандидата технічних наук. - Херсон, 1999.-130 с.
2. Flake G.W. and Lawrence S. Efficient SVM regression training with SMO. Machine Learning, 46(1/3):pp.271–290, 2002.
3. Lifschitz V. Foundations of logic programming. In Principles of Knowledge Representation, pp. 69–127. CSLI Publications, 1996.
4. Hemerly A. S., Casanova M. A., Furtado A. L.: Cooperative behavior through Request Modification. In: Proc. 10th Conference on the Entity-Relationship Approach, pp.607–621, 1991.
5. Magennis M. and van Rijsbergen C. J. The potential and actual effectiveness of interactive query expansion. Proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.-pp. 324-331. Philadelphia. 1997
6. Howe A. E. and Dreilinger D. SavvySearch: A meta-search engine that learns which search engines to query. AI Magazine, 18(2), 1997.
7. Sormunen E., Laaksonen J., Keskustalo H., Kempainen H., Laitinen H. The IR Game - A Tool for Rapid Query Analysis in Cross-Language IR Experiments. PRICAI '98 Workshop on Cross Language Issues in Artificial Intelligence. Singapore, Nov 22-24, 1998, pp. 22-32.
8. Glover E. J. , Lawrence S. , Birmingham W. P. and Giles C. L. Architecture of a metasearch engine that supports user information needs. In Eighth International Conference on Information Knowledge Management (CIKM'99), pp. 210–216, Kansas City, MO, November 1999. ACM Press.
9. Sormunen E. A Novel Method for the Evaluation of Boolean Query Effectiveness across a Wide Operational Range. In: Belkin NJ, Ingwersen P and Leong M-K SIGIR 2000, Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Athens July 24-28, 2000. Special Issue of the SIGIR Forum 34:pp.25-32.
10. Glover E. , Flake G. , Lawrence S. , Birmingham W. P. , Kruger A. , Giles C. L. and Pennock D. Improving category specific web search by learning query modifications. In Symposium on Applications and the Internet, SAINT, San Diego, CA, January 8–12 , 2001.
11. Alferes J.J. and Pereira L.M. Update-programs can update programs. In Non-Monotonic Extensions of Logic Programming (Bad Honnef, 1996), volume 1216 of Lecture Notes in Computer Science, pp.110–131, Berlin, 1997. Springer.
12. Niemel I. and Simons P. Efficient implementation of the well-founded and stable model semantics. In M.J. Maher, editor, Proceedings of the 1996 Joint International Conference and Symposium on Logic Programming (JICSLP-96) (Bonn, Germany, September 2-6, 1996), pp.289–303. MIT Press, 1996.
13. Berman K. , Schlipf J. and Franco J. Computing the well-founded semantics faster. In Logic Programming and Nonmonotonic Reasoning (Lexington, KY, 1995), volume 928 of Lecture Notes in Computer Science, pp. 113–125, Berlin, 1995. Springer.

Одержано 14.09.2003 р.