

## ВИКОРИСТАННЯ НЕОБХІДНОЇ МОДИФІКАЦІЇ, ЯК ІНСТРУМЕНТУ ТРАНСФОРМАЦІЇ БАЗ ЗНАНЬ НАФТОГАЗОВОЇ ПРЕДМЕТНОЇ ОБЛАСТІ

Шекета В.І.

Івано-Франківський національний технічний університет нафти і газу  
кафедра прикладної математики  
sheketa@mail.ru

### Abstract

*Sheketa V.I. The using of necessary modification, as an tool for transformation of knowledge bases of oil and gas subject domain. In given paper the notion of necessary modification is introduced, as an tool for transformation of knowledge bases and preparing them for applying the formal-logical apparatus of predicate queries modifications, which are considered as a multifunctional mechanism for supporting dialog with an user in the information systems on the basis of databases and knowledge bases of oil and gas subject domain.*

Одним із ключових питань з точки зору математичного моделювання процесу побудови інформаційних інтелектуальних систем для нафтогазової предметної області [1] є спосіб представлення знань, на основі якого система повинна приймати рішення в певній ситуації. Представлення знань повинно бути задано способом, який дозволяє перехід до представлення фрагментів інформації про нафтогазовий об'єкт в термінах структур баз даних і знань.

Теорія модифікаційних предикатних запитів [2] є формально-логічним апаратом опису і вивчення процесів оновлення і модифікації баз даних і знань, логічного висновку на основі баз даних і знань [3,4,5,6,7]. 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 [8,9].

Розглядатимемо таку базу даних з точки зору фактів і процесів, що призводять до їх зміни, тобто з погляду семантики і синтаксису [9,11] такого представлення. Під синтаксисом будемо розуміти набір правил для поєднання символів в логічно коректні вирази. Під семантикою будемо розуміти спосіб інтерпретації виразів, що одержуються в результаті конкретних реалізацій синтаксичних правил.

В роботі [14] автором введено поняття модифікаційних предикатних запитів, як інструменту підтримки діалогу з користувачем інформаційних інтелектуальних систем в процесі побудови ним запитів, що розглядаються в ході динаміки процесу оновлення і модифікації баз даних і знань нафтогазової предметної області.

Проте виконані дослідження показали, що рівень структуризації інформації, що наповнює реальні бази знань нафтогазової предметної області не відповідає рівню, на якому можна було би виконувати безпосереднє застосування концепції модифікаційних предикатних запитів.

Тому ціллю даної статті є побудова ефективного механізму трансформації вихідних баз знань предметної області, з метою досягнення задовільного рівня логічної зв'язаності інформації, що є необхідною передумовою застосування формально-логічного апарату модифікаційних предикатних запитів для інформаційних систем на основі баз даних і баз знань нафтогазової предметної області.

Ми будемо розвивати власний формально-логічний апарат, на основі основних принципів абстрактного логічного програмування [11], і теоретичних аспектів програмування в мові Prolog [13].

Розглядатимемо базу знань як набір інформаційних сутностей атомарних предикатів з деякого скінченного інформаційного простору  $O$ . Всі зміни, що відбуватимуться в базі знань, будемо розглядати, як наслідок модифікаційних предикатних запитів, що генеруються ІС відповідно до вказівок користувача. Основою самих запитів є набір модифікаційних предикатних правил. Будемо розглядати 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  $o, o_i, p_i \in O$ ,  $\ll$  - дескриптор модифікації. Основна ідея такого запису правил в тому, що  $K_{B+}(o)$  означає, що атомарний предикат  $o$  повинен бути включений в базу знань  $K_B$ , а  $K_{B-}$  означає, що  $o$  - повинен бути виключений з бази знань.

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

Введемо поняття необхідної модифікації, яка буде описувати всі операції  $K_{B+}()$  і  $K_{B-}()$ , що спричиняються модифікаційним запитом для вихідної бази знань  $K_B$ .

**Означення 1.** Нехай  $Q_m$  – модифікаційний предикатний запит. Позначимо через  $\lambda_{nm}(Q_M)$  необхідні модифікації для  $K_B$ . В [13] показано, що  $\lambda_{nm}(Q_M)$  фактично є найменшою моделлю для  $Q_m$ , якщо  $\lambda_{nm}(Q_M)$  розглядати, як Горн-програму  $Q_m$  побудовану з незалежних атомарних предикатів виду  $K_{B+}(o)$  і  $K_{B-}(p)$ .

Нехай, наприклад, модифікаційний предикатний запит  $Q_m$  має вигляд  $\{K_{B+}(o) \ll, K_{B-}(p) \ll K_{B+}(o), K_{B-}(r) \ll K_{B-}(o)\}$  тоді  $\lambda_{nm}(Q_M) = \{K_{B+}(o), K_{B-}(p)\}$ .

В результаті виконання модифікаційних предикатних запитів база знань змінюється, але статус деяких елементів залишається незмінним. Тобто згідно розглянутого принципу інерції [14] при створенні набору модифікаційних правил ми власне вказуємо тільки опис того, чи потрібно змінити певні елементи. Якщо для певного елемента немає правила модифікації, то відповідно, згідно принципу інерції, він не буде змінюватися в ході виконання модифікацій.

Таким чином, множину всіх модифікаційних літералів, що описують елементи бази знань, статус яких не змінюється в процесі переходу від бази знань  $K_{B_1}$  до бази знань  $K_{B_2}$ ,  $Q_m$  і  $K_{B_1} \gg K_{B_2}$  будемо називати множиною інерції для  $K_{B_1}$ ,  $K_{B_2}$  і означати, як:

$$O_i(K_{B_1}, K_{B_2}) = \{K_{B+}(o) : o \in K_{B_1} \cap K_{B_2}\} \cup \{K_{B-}(o) : o \notin K_{B_1} \cup K_{B_2}\}.$$

**Означення 2.** Залишком для  $Q_m$  стосовно баз знань  $(K_{B_1}, K_{B_2})$   $Q_M^{K_{B_1}, K_{B_2}}$  будемо називати новий модифікаційний запит, який одержується із  $Q_m$  шляхом видалення із тіла кожного модифікаційного правила в  $Q_m$  всіх модифікаційних літералів, які належать також і  $O_i(K_{B_1}, K_{B_2})$ .

**Означення 3.** Базу знань  $K_{B_2}$  будемо називати  $Q_m$ -модифікацією для  $K_{B_1}$ , якщо множина необхідних модифікацій для  $Q_M^{K_{B_1}, K_{B_2}}$  є когерентною і якщо  $K_{B_2} = K_{B_1} \circ \lambda_{nm}(Q_M^{K_{B_1}, K_{B_2}})$

Для ілюстрації наведених вище означень розглянемо наступний приклад.  
**Приклад 1.** Припустимо, нам потрібно будувати процедуру логічного висновку для інтелектуальної системи на основі баз даних і знань нафтогазової предметної області Plast Researcher виходячи із чотирьох характеристик нафтогазоносної породи:

*{пористість, насиченість, проникність, коефіцієнт\_опору}*.

Причому, виходячи із досвіду експертів по опрацюванню даних характеристик породи, ми будемо мати справу із певним набором обмежень. А саме:

1. Характеристики “пористості” і “насиченості” є найбільш продуктивним для виконання процедури логічного висновку тому одна із них повинна обов’язково бути присутньою в вихідній базі знань  $K_B^{noc}$ .

2. Використання характеристики “проникність” є неефективним без одночасного опрацювання характеристики “коефіцієнт\_опору”. Це так звані парні характеристики і якщо в  $K_B^{noc}$  не буде включено характеристику “коефіцієнт\_опору”, то не потрібно включати і характеристику “проникність”.

3. Спільне використання характеристик “коефіцієнт\_опору” і “проникність” не є ефективним, оскільки вони мають різну фізичну природу і методики їх вимірювання базуються на різних принципах.

4. Неефективним є також спільне використання характеристик “насиченість”, “коефіцієнт\_опору” відповідно до причин, викладених в пункті 3.

Згідно перелічених обмежень в якості початкового складу бази знань прийmemo  $K_B^{noc} = \{пористість, проникність\}$ .

Ми прагнемо сформуванати базу знань  $K_B$ , яка би задовольняла всім чотирьом накладеним обмеженням. Таким чином в термінах введених нами вище означень ми одержуємо  $K_{B_1} = \{пористість, проникність\}$ . Побудуємо модифікаційний предикатний запит  $Q_M$ :

$$\{K_{B_+}(насиченість) \ll K_{B_-}(пористість), K_{B_+}(пористість) \ll K_{B_-}(насиченість), K_{B_+}(коефіцієнт_опору) \ll K_{B_+}(проникність), K_{B_-}(проникність), \ll K_{B_-}(коефіцієнт_опору), K_{B_-}(пористість) \ll K_{B_+}(коефіцієнт_опору), K_{B_-}(коефіцієнт_опору) \ll K_{B_+}(насиченість)\}$$

Покажемо, що  $K_{B_2} = \{пористість\}$  є  $Q_M$ -модифікацією для  $K_{B_1}$ .

Виходимо з того, що

$$O = \{пористість, насиченість, проникність, коефіцієнт_опору\}$$

Тоді

$$O_i(K_{B_1}, K_{B_2}) = \{K_{B_+}(пористість), K_{B_-}(насиченість),$$

$$K_{B_-}(коефіцієнт_опору)\}$$

$$Q_M^{K_{B_1}, K_{B_2}} = \{K_{B_+}(\text{насиченість}) \ll K_{B_-}(\text{пористість}), K_{B_+}(\text{пористість}) \\ \ll K_{B_+}(\text{коефіцієнт_опору}) \ll K_{B_+}(\text{проникність}), K_{B_-}(\text{проникність}) \\ \ll K_{B_-}(\text{пористість}) \ll K_{B_+}(\text{коефіцієнт_опору}), \\ K_{B_-}(\text{коефіцієнт_опору}) \ll K_{B_+}(\text{насиченість})\}$$

Звідки

$$\lambda_{nt}(Q_M^{K_{B_1}, K_{B_2}}) = \{K_{B_+}(\text{пористість}), K_{B_-}(\text{проникність})\}.$$

Збережена властивість когерентності

$$K_{B_2} = K_{B_1} \circ \lambda_{nt}(Q_M^{K_{B_1}, K_{B_2}}).$$

Таким чином ми показали, що  $K_{B_2}$  є  $Q_M$ -модифікація для  $K_{B_1}$ .

Тепер ми покажемо, що  $K_B^{Q_M}$  є  $Q_M$ -модифікацією для  $K_B^{noch}$ .

Припустимо, що правила з модифікаційного запиту  $Q_M$  є такими, що не є всі одночасно хибними. Тоді  $E_M$  є найменшою моделлю для  $(\Delta_R^{noch})^{E_M} \cup \langle (\Delta_R^{in})^{E_M} \rangle \cup Q_M$ . Для того, щоби обчислити залишок

$Q_M \Big|_{K_B^{noch}, K_B^{Q_M}}$  для  $Q_M$  ми повинні видалити з тіла кожного модифікаційного

правила всі атомарні предикати  $K_{B_+}(o)$ , такі, що  $o \in K_B^{noch} \cap K_B^{Q_M}$ , і всі

атомарні предикати  $K_{B_-}(o)$ , такі, що  $o \in K_B^{in} \cap K_B^{Q_M}$ . Це і будуть ті

атомарні предикати, що є істинними в  $E_M$ , завдяки множині  $\langle (\Delta_R^{in})^{E_M} \rangle$ .

Тому  $E_M$  залишається найменшою моделлю модифікованого запиту

$$(\Delta_R^{noch})^{E_M} \cup \langle (\Delta_R^{in})^{E_M} \rangle \cup (Q_M \Big|_{K_B^{noch}, K_B^{Q_M}}).$$

Тоді необхідна модифікація  $\lambda_{nt}(Q_M \Big|_{K_B^{noch}, K_B^{Q_M}})$  є найменшою моделлю

для  $Q_M \Big|_{K_B^{noch}, K_B^{Q_M}}$ . Розглянемо тепер дві множини

$$O_+ = \{o : K_{B_+}(o) \in \lambda_{nt}(Q_M \Big|_{K_B^{noch}, K_B^{Q_M}})\} \text{ і}$$

$$O_- = \{o : K_{B_-}(o) \in \lambda_{nt}(Q_M \Big|_{K_B^{noch}, K_B^{Q_M}})\}.$$

Тоді запит  $(\Delta_R^{noch})^{E_M} \cup \langle (\Delta_R^{in})^{E_M} \rangle \cup (Q_M \Big|_{K_B^{noch}, K_B^{QM}})$  складається із трьох незалежних частин, причому залишок  $Q_M \Big|_{K_B^{noch}, K_B^{QM}}$  більше не містить атомарних предикатів, які би входили в інші дві частини. Тоді множина  $E_M$  буде складатися з:

$$\begin{aligned} & \{K_{B+}^{\in}(o) : o \in K_B^{noch}\} \cup \{K_{B-}^{\in}(o) : o \in K_B^{'noch}\}, \\ & \{K_{B+}(o) : o \in K_B^{QM} \cap K_B^{noch}\} \cup \{K_{B-}(o) : o \in K_B^{'QM} \cap K_B^{'noch}\}, \\ & \{K_{B+}(o) : o \in O_+\} \cup \{K_{B-}(o) : o \in O_-\} \end{aligned}$$

на основі чого, робимо висновок, що:

$$\begin{aligned} K_B^{QM} &= \{o : K_{B+}(o) \in E_M\} = (K_B^{QM} \cap K_B^{noch}) \cup O_+, \\ K_B^{'QM} &= O \setminus K_B^{QM} = \{o : K_{B-}(o) \in E_M\} = (K_B^{'QM} \cap K_B^{'noch}) \cup O_-. \end{aligned}$$

Таким чином множина необхідних модифікацій  $\lambda_{nm}(Q_M \Big|_{K_B^{noch}, K_B^{QM}})$  є когерентною. Також  $K_B^{QM} = (K_B^{noch} \cup O_+) \setminus O_-$ . Звідки слідує, що  $K_B^{QM}$  є  $Q_M$  модифікацією для  $K_B^{noch}$ .

Введені означення є обгрунтованими, оскільки при їх побудові ми не виходили за рамки процедури обчислення обгрунтованих семантик прийнятих в теорії абстрактного логічного програмування і стабільних семантик для абстрактних логічних програм.

## Висновки

В даній роботі введено поняття необхідної модифікації, як інструменту трансформації баз знань нафтогазової предметної області. Подальші дослідження даного напрямку будуть напрямлені на розширення формально-логічного апарату необхідних модифікацій в рамках побудови загальної теорії модифікаційних предикатних запитів для інформаційних інтелектуальних систем на основі баз даних і знань нафтогазової предметної області.

## Література

1. Шекета В.І. Інформаційна система для прогнозування нафтогазоносних покладів. Дисертація на здобуття наукового ступеня кандидата технічних наук. Херсон, 1999.-130с.
2. Hemerly, A. S., Casanova, M. A., Furtado, A. L.: Cooperative behavior through Request Modification. In: Proc. 10th Conference on the Entity-Relationship Approach, 607–621, 1991.
3. 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. 324-331. Philadelphia. 1997.
4. Howe A. E. and Dreilinger D. SavvySearch: A meta-search engine that learns which search engines to query. *AI Magazine*, 18(2), 1997.
5. Sormunen E, Laaksonen J, Keskustalo H, Kemppainen H, Laitinen H, Pirkola, (1998) 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, p. 22-32.
6. 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), pages 210–216, Kansas City, MO, November 1999. ACM Press.
7. Sormunen E (2000b) 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 34:25-32.
8. 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.
9. Flake G. W. and Lawrence S. Efficient SVM regression training with SMO. *Machine Learning*, 46(1/3):271–290, 2002.
10. Niemel I. a 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), pages 289–303. MIT Press, 1996.
11. Lifschitz V. Foundations of logic programming. In Principles of Knowledge Representation, pages 69–127. CSLI Publications, 1996.
12. 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, pages 113–125, Berlin, 1995. Springer.

13. 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, pages 110–131, Berlin, 1997. Springer.

14. Шекета В.І. Модифікаційні предикатні запити, як інструмент підтримки діалогу з користувачем в інформаційних системах на основі баз даних і знань. Вісник Тернопільського державного технічного університету. Серія Математичне моделювання.-2003. Здано до друку.

---

Дата надходження до редколегії: 02.12.2003 р.