

Долинніна О.М., к.т.н., доцент
Саратовський державний технічний університет ім. Гагаріна Ю.О.
Соколов Д.І., аспірант
Саратовський державний технічний університет ім. Гагаріна Ю.О.

МЕТОД УПРАВЛІННЯ БАЗОЮ ЗНАНЬ ДЛЯ ПРИЙНЯТТЯ РІШЕНЬ У ПРОДУКЦІЙНИХ ЕКСПЕРТНИХ СИСТЕМАХ

Одним з найбільш поширених способів подання знань в інтелектуальних системах вважаються продукції. Це обумовлюється їх перевагами. По-перше, продукції універсальні. З їх допомогою можна представити найрізноманітніші людські знання практично в будь-якій предметній області. По-друге, вони незалежні. Додавання, зміна або видалення однієї продукції не представляє небезпеки для працездатності системи в цілому. І нарешті, будучи комбінацією логічних і мережевих моделей, вони мають здатність реалізовувати різноманітні алгоритми, маючи при цьому високу наочність. Проблеми починаються при зростанні кількості продукції в базі знань. Чим більше згенеровано продукцій, тим складніше стежити за тим, щоб вони не суперечили один одному. А їх незалежність ускладнює завдання пошуку цих протиріч. Така експертна система в цілому не втратить працездатність, але, в деяких випадках, буде давати некоректні результати. У забезпеченні коректної роботи експертної системи велику роль відіграє метод управління продукціями. Від його (методу) вибору залежить також швидкість і ефективність її (системи) роботи. Завдання вибору правильного методу часто є визначальним в процесі створення експертних систем.

Існує кілька класичних методів управління продукційною базою знань: принцип пріоритетного вибору, принцип «стопки книжок», принцип «класної дошки» тощо. Кожен з них добре підходить для застосування, якщо виконуються певні умови. Однак реальні експертні системи досить складні і застосування конкретного методу до них не представляється можливим. Тому частіше розробляються оригінальні методи, або комбінують класичні методи керування, або додатково використовують спеціальні алгоритми. Крім цього, для роботи з кількістю продукцій зазвичай більше 1000 базу знань розбивають на рівні, а для продукцій виділяють сфери застосування. Подібна структура дозволяє значно скоротити фронт готових продукцій, а, отже, і час для прийняття рішення. Однак, у випадку, коли ми стикаємося з даними, що мають високий ступінь ненадійності, такий підхід буде недоцільним. Якщо знання не детерміновані – застосовують імовірнісний підхід. Проблемою імовірнісного підходу є відсутність статистичних даних, на основі яких він будується.

Проаналізувавши різні методи управління, ми прийшли до висновку, що найбільший позитивний ефект при управлінні продукціями може бути досягнутий за допомогою об'єднання структурного та імовірнісного підходів. Був розроблений метод, заснований на двох класичних принципах: «стопки книг» і «принципу пріоритетного вибору». Крім того, він може враховувати сфери застосування продукцій, якщо вони задані в базі знань. Експерти задають пріоритети для продукцій, тим самим визначаючи порядок їх активації, а в спеціальній таблиці («стопці книг») ведеться підрахунок успішно активованих продукцій. «Стопки книжок» складає N продукцій, які найчастіше зустрічаються. При цьому N визначається експертом. У міру накопичення статистики використань вибір продукцій може змінюватися і залежить не тільки від пріоритету, але і від частоти їх використання. Метод управління продукціями був успішно апробований при побудові експертної системи PharmExS з підбору безрецептурних препаратів для лікування симптомів простудних захворювань. Вона ґрунтується на продукційній моделі подання знань і складається з двох рівнів. На першому рівні враховуються фактори, пов'язані з вибраним симптомом і визначаються групи препаратів, з яких в подальшому буде проводитися вибір. Правила на першому рівні мають наступний вигляд:

$$id_i, pr_i, field_i: IF s_condition_1 \& s_condition_2 \& \dots \& s_condition_n \text{ then } lek_1 \& \dots \& lek_n \quad (1)$$

де $id_i \in ID$, де $ID = \{ id_1, id_2, \dots, id_n \}$, унікальний серійний номер правила;

$pr_i \in Pr$, де $Pr = \{ pr_1, pr_2, \dots, pr_n \}$, пріоритет правила, що показує його важливість з точки зору експерта;

$field_i \in Field$, де $Field = \{ field_1, field_2, \dots, field_6 \}$ основний симптом.

$s_condition_n \in Symptom_Condition$, де $Symptom_Condition = \{ s_condition_1, s_condition_2, \dots,$

$s_condition_n \}$, фактори, які визначають основний симптом;

$lek_i \in Lek$, где $Lek = \{ lek_1, lek_2, \dots, lek_n \}$ група препаратів, із яких в подальшому буде виконаний вибір.

Другий рівень призначений для вибору конкретного препарату. Тому враховуються такі характеристики як ціна, виробник і дозування ліків, фізичні особливості покупця, а також інші суб'єктивні чинники. Правила на другому рівні мають наступний вигляд:

$$Id_i, lek_i: IF c_condition_1 \& c_condition_2 \& \dots \& c_condition_n \text{ then } drug_i \quad (2)$$

де $id_i \in ID$, де $ID = \{ id_1, id_2, \dots id_n \}$, унікальний серійний номер правила;
 $lek_i \in Lek$, де $Lek = \{ lek_1, lek_2, \dots lek_n \}$ група препаратів, із переліку яких буде виконуватись вибір;
 $c_condition_n \in Common_Condition$, де $Common_Condition = \{ c_condition_1, c_condition_2, \dots c_condition_n \}$, додаткові умови (ціна, виробник і дозування ліків, фізичні особливості покупця, а також інші суб'єктивні фактори);
 $drug_i \in Drug$, де $Drug = \{ drug_1, drug_2, \dots drug_m \}$, один з доступних для вибору препаратів $m=1, \dots, 579$.

В якості сфери застосування продукції використовуються симптоми, до яких вони відносяться. Пріоритети визначаються експертами, і лежать в діапазоні від 1 (найнезначніший) до 100 (найбільш значущий). На основі отриманих даних було виділено 4 групи пріоритетів: терміновий (продукції з пріоритетом 100), високий (продукції з пріоритетом від 60 до 99), звичайний (продукції з пріоритетом від 30 до 59) і низький (продукції з пріоритетом, меншим 30). Кожна продукція має унікальний ідентифікатор і, якщо вона використана, то ідентифікатор записується в «стопку книг». Для автоматизації процесу консультації кожної продукції був поставлений у відповідність список питань. В результаті роботи запропонованого методу система на кожному кроці обирає найбільш підходящу для продовження консультації продукцію і задає питання із списку питань, пов'язаного з нею.

Розроблений метод добре показав себе для управління базою знань експертної системи PharmExS. По-перше, він однаково працював на обох рівнях бази знань, хоча продукції мали різні структури. По-друге, спочатку, не маючи накопиченої статистики, система часто задавала зайві питання, затягуючи процес консультації. По ходу роботи і накопичення статистики в «стопці книг» відсоток консультацій, в яких задавалися зайві питання, значно зменшився. Таким чином, застосування принципів, сформованих у ході розробки методу, представляється корисним для подальших досліджень в галузі управління продукційними базами знань.