

*Запропоновано підхід до побудови діагностичної системи на основі нечітких знань з реалізацією можливості навчання шляхом поповнення знань про предметну область.*

© В.Д. Саввакін, О.І. Провотар,  
2019

УДК 681.3

В.Д. САВВАКІН, О.І. ПРОВОТАР

## **ДІАГНОСТИЧНА СИСТЕМА НА ОСНОВІ НЕЧІТКИХ ЗНАНЬ**

**Вступ.** Найбільш вражаючою властивістю людського інтелекту є здатність приймати правильні рішення за умов неповної і нечіткої інформації. Побудова моделей наближених роздумів людини і використання їх у комп'ютерних системах представляє сьогодні одну з найважливіших проблем науки.

На практиці виникає багато задач, розв'язання яких точними методами або неможливе в принципі, або дуже громіздке і тому неефективне. Тому для розв'язання таких задач розроблялись і розробляються методи, що ґрунтуються на різних способах подання нечіткої інформації.

Математична теорія нечітких множин (fuzzy sets) і нечітка логіка (fuzzy logic) [1 – 3] є узагальненнями класичної теорії множин і класичної формальної логіки. Ці теорії були вперше запропоновані американським ученим Лотфі Заде (Lotfi Zadeh) в 1965 р. Основною причиною їх появи є необхідність використання (з подальшою формалізацією) нечітких і наближених міркувань при описанні процесів, систем, об'єктів тощо.

Період розвитку нечіткої логіки супроводжувався побудовою перших експертних систем на її основі та розробці нечітких контролерів. Нечіткі експертні системи для підтримки прийняття рішень знаходять широке застосування в медицині та економіці. Нарешті, з кінця 80-х років і до сьогодні, з'являються пакети програм для побудови нечітких експертних систем, а області застосування нечіткої логіки помітно розширюються.

Вона застосовується в автомобільній, аерокосмічній і транспортній промисловості, в області виробів побутової техніки, у сфері фінансів, аналізу і прийняття управлінських рішень та багатьох інших областях.

**Нечіткі системи логічного виведення.** Під нечіткою системою логічного виведення (системою специфікацій) [1, 4] розуміють упорядковану множину нечітких інструкцій, які при виконанні дають наближене (нечітке) вирішення проблеми.

Нехай  $x$  та  $y$  – вхідна та вихідна лінгвістичні змінні,  $A$  та  $B$  – деякі нечіткі множини, що задають значення елементів терм-множин змінних  $x$  та  $y$  відповідно. Найпростішою нечіткою системою логічного виведення може бути така конструкція:

*вхід* ( $x$ );  
**якщо**  $x \in A$ , **то**  $y \in B$ ;  
*вихід* ( $y$ ).

Інструкція «**якщо**  $x \in A$ , **то**  $y \in B$ » інтерпретується як нечітка імплікація  $A \rightarrow B$ , отже, задається нечітким відношенням на декартовому добутку областей визначення (чітких множинах)  $X$  вхідної змінної і  $Y$  вихідної змінної. Вихідне значення системи визначається за допомогою композиційного правила. А саме, якщо на вхід подається нечітка множина  $A'$ , то на виході отримуємо нечітку множину, яка визначається за формулою

$$B'(y) = \max_{x \in X} (\min(A'(x), \min(A(x), B(y)))) , y \in Y.$$

Більш складну нечітку систему утворює конструкція вигляду:

*вхід* ( $x$ );  
**якщо**  $x \in A_1$ , **то**  $y \in B_1$ ;  
**якщо**  $x \in A_2$ , **то**  $y \in B_2$ ;  
 .....  
**якщо**  $x \in A_m$ , **то**  $y \in B_m$ ;  
*вихід* ( $y$ ),

де  $A_i$  і  $B_i$  – нечіткі множини.

Існує два основних способи визначення виходу  $B'$ . В обох використовується так зване поняття агрегації правил, тобто врахування сумарного ефекту від роботи всіх правил. Оператор агрегації **Agg** діє як  $s$ -норма, але дозволяється використання довільної  $t$ -норми.

Перший спосіб визначення виходу полягає у попередній агрегації нечітких відношень  $R = \text{Agg}(R_1, R_2, \dots, R_m)$ . Результат  $B'$  при заданому вході  $A'$  визначається за допомогою композиційного правила:  $B' = A' \circ R$ . Якщо оператор агрегації є операцією знаходження максимуму, то  $B'$  визначається за формулою

$$B' = A' \circ \bigcup_{i=1}^m R_i.$$

Другий спосіб полягає у визначенні виходів для кожного правила за допомогою використання композиції  $B'_i = A' \circ R_i$ ,  $i = 1, \dots, m$ . Далі здійснюється агрегація отриманих виходів за правилом  $B' = \text{Agg}(B'_1, B'_2, \dots, B'_m)$ , тобто

$$B' = \bigcup_{i=1}^m (A' \circ R_i).$$

**Твердження 1.** При використанні max-min композицій спільно з операцією максимуму в ролі оператора агрегації результати, отримані обома механізмами логічного виведення, будуть еквівалентними, тобто справедливе співвідношення

$$A' \circ \bigcup_{i=1}^m R_i = \bigcup_{i=1}^m (A' \circ R_i).$$

Більш цікавою видається ситуація, коли система має не один, а кілька входів:

**вхід**  $(x_1, x_2, \dots, x_n)$ ;  
**якщо**  $x_1 \in A_{11} \wedge x_2 \in A_{12} \wedge \dots \wedge x_n \in A_{1n}$  **то**  $y \in B_1$ ;  
**якщо**  $x_1 \in A_{21} \wedge x_2 \in A_{22} \wedge \dots \wedge x_n \in A_{2n}$  **то**  $y \in B_2$ ;  
 ...  
**якщо**  $x_1 \in A_{m1} \wedge x_2 \in A_{m2} \wedge \dots \wedge x_n \in A_{mn}$  **то**  $y \in B_m$ ;  
**вихід**  $(y)$ ,

де  $x_j$ ,  $j = 1, \dots, n$  – вхідні лінгвістичні змінні,  $y$  – вихідна лінгвістична змінна;  $A_{ij}$  та  $B_i$  – нечіткі множини. Логічна зв'язка “ $\wedge$ ” інтерпретується як  $t$ -норма нечітких множин [1].

На відміну від випадку з однієї вхідної змінної подання імплікації у вигляді відношення в алгоритмах з багатьма вхідними параметрами неможливо. У зв'язку з цим використовується інша процедура знаходження виходу, яка використовує так звані рівні істинності правил типу

**якщо**  $x_1 \in A_{i1} \wedge x_2 \in A_{i2} \wedge \dots \wedge x_n \in A_{in}$  **то**  $y \in B_i$ .

У випадку двох входів  $x_1$  і  $x_2$ , процедура виконання алгоритму буде складатися з наступних кроків:

- 1) для кожного правила  $R$ ,  $i = 1, \dots, m$  обчислюємо рівень істинності правила

$$\alpha_i = \min \left[ \max_{X_1} (A'_1(x_1) \wedge A_{i1}(x_1)), \max_{X_2} (A'_2(x_2) \wedge A_{i2}(x_2)) \right];$$

- 2) для кожного правила обчислюємо індивідуальні виходи

$$B'_i(y) = \min(\alpha_i, B_i(y));$$

- 3) обчислюємо агрегатний вихід

$$B'(y) = \max(B'_1, B'_2, \dots, B'_m).$$

Ця процедура називається *max-min* процедурою або процедурою логічного висновку Мамдані (імплікація інтерпретується як операція мінімум, агрегація виходів правил – як операція максимум).

**Твердження 2.** При використанні *max-min* композицій і логічного виведення Мамдані результати будуть еквівалентними, тобто справедливе співвідношення

$$B'(y) = \max_{x \in X} (A'(x) \wedge (R(x, y))) = \max_{i=1}^m (\alpha_i \wedge B_i(y)).$$

Грунтуючись на фундаментальному результаті Фунахаші про те, що за допомогою нечітких систем можна апроксимувати з будь-якою заданою точністю будь-яку неперервну на компактній функцію, з'являються можливості використання нечітких специфікацій для розв'язання задач діагностики в різних постановках і ймовірно-можливісних підходів для отримання оцінок виходу нечітких систем логічного виведення. Використовуючи ці оцінки (зокрема, для гіпотез), можна будувати оптимальні в деякому сенсі гіпотези і досліджувати питання про їх оптимальну кількість.

Застосування нечітких систем логічного виведення для розв'язання задач діагностики з відповідними методами навчання може виявитися більш ефективним при реалізації у порівнянні з нейронними мережами.

**Приклад застосування нечітких систем логічного виведення.** Нехай  $X_1 = \{5, 10, 15, 20\}$ ,  $X_2 = \{5, 10, 15, 20\}$  та  $X_3 = \{35, 36, 37, 38, 39, 40\}$  – простори для визначення значень елементів терм-множин:

- “Кашель” = {“слабкий”, “помірний”, “сильний”};
- “Нежить” = {“слабка”, “помірна”, “сильна”};
- “Температура” = {“нормальна”, “підвищена”, “висока”, “дуже висока”} відповідно.

Визначимо елементи цих терм-множин наступним чином:

$$\begin{aligned} \text{“Кашель”}: \quad \text{“слабкий”} &= 1/5 + 0.5/10; \\ &\text{“помірний”} &= 0.5/5 + 0.7/10 + 1/15; \\ &\text{“сильний”} &= 0.5/10 + 0.7/15 + 1/20. \end{aligned}$$

$$\begin{aligned} \text{“Нежить”}: \quad \text{“слабка”} &= 1/5 + 0.5/10; \\ &\text{“помірна”} &= 0.5/10 + 1/15; \\ &\text{“сильна”} &= 0.7/15 + 1/20. \end{aligned}$$

$$\begin{aligned} \text{“Температура”}: \quad \text{“нормальна”} &= 0.5/35 + 0.8/36 + 0.9/37 + 0.5/38; \\ &\text{“підвищена”} &= 0.5/37 + 1/38; \\ &\text{“висока”} &= 0.5/38 + 1/39; \\ &\text{“дуже висока”} &= 0.8/39 + 1/40. \end{aligned}$$

Нехай  $Y = \{\text{Грип, ГРЗ, Ангіна, Запалення легенів}\}$  – простір для визначення значень лінгвістичної змінної  $y$ .

Тоді залежність хвороби пацієнта від симптомів може бути описана, наприклад, наступною системою специфікацій:

**вхід**  $(x_1, x_2, x_3)$ ;

**якщо**  $x_1 \in \text{«слабкий»} \wedge x_2 \in \text{«слабка»} \wedge x_3 \in \text{«підвищена»}$ , **то**  $y \in \text{«0.5/Грип} + 0.5/ГРЗ + 0.4/Ангіна + 0.8/Запалення легенів»}$ ;

**якщо**  $x_1 \in \text{«слабкий»} \wedge x_2 \in \text{«помірна»} \wedge x_3 \in \text{«висока»}$ , **то**  $y \in \text{«0.8/Грип} + 0.7/ГРЗ + 0.8/Ангіна + 0.3/Запалення легенів»}$ ;

**якщо**  $x_1 \in \text{«слабкий»} \wedge x_2 \in \text{«помірна»} \wedge x_3 \in \text{«дуже висока»}$ , **то**  $y \in \text{«0.9/Грип} + 0.7/ГРЗ + 0.8/Ангіна + 0.2/Запалення легенів»}$ ;

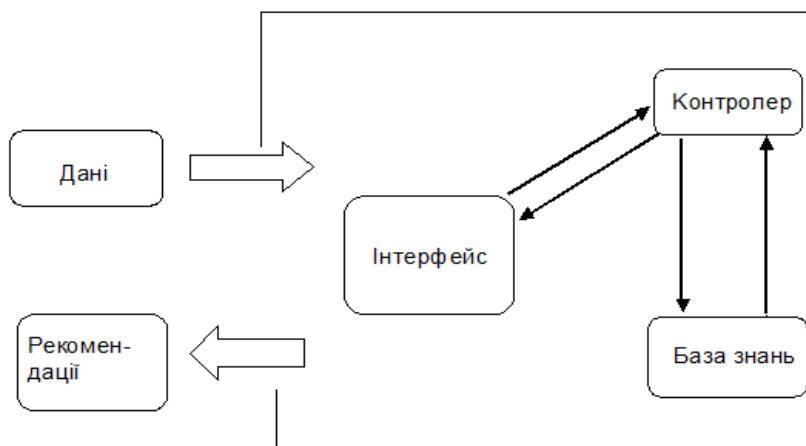
**вихід**  $(y)$ ,

де  $x_1, x_2, x_3$  – вхідні лінгвістичні змінні, що приймають значення з терм-множин «Кашель», «Нежить» і «Температура» відповідно,  $y$  – вихідна лінгвістична змінна.

Якщо на вхід  $x_1$  цього алгоритму подати величину  $A'_1 = 1/5 + 0.7/10$ , на вхід  $x_2$  – величину  $A'_2 = 1/5 + 0.5/10$ , на вхід  $x_3$  – величину  $A'_3 = 1/36 + 0.9/37$ , то згідно з процедурою виконання цього алгоритму отримуємо:

$$V = 0.5/\text{Грип} + 0.5/\text{ГРЗ} + 0.4/\text{Ангіна} + 0.5/\text{Запалення легенів}.$$

**Експертно-діагностична система на основі нечітких знань.** Одним із застосувань нечіткої логіки, зокрема, нечітких систем логічного виведення, є експертно-діагностичні системи для лікарів, основним призначенням яких є встановлення захворювання або захворювань пацієнта за результатами обстеження та суб'єктивних оцінок пацієнта. Архітектура запропонованої системи показана на рисунку.



РИСУНОК

Автоматизація визначення попереднього діагнозу пацієнтів могла б суттєво знизити рівень залежності коректності обраного лікування від професійної компетенції лікарів, до яких звернувся пацієнт. Окрім того, використання автоматизованих систем дозволило б прибрати фактор людської помилки при встановленні діагнозу. Позбавлення від вищеперерахованих недоліків збільшить якість медичного обслуговування та зменшить час встановлення попереднього діагнозу.

Прийняття рішення системою після введення даних, що характеризують стан пацієнта, буде здійснюватися на основі бази знань, в якій будуть сформовані нечіткі правила відповідності того чи іншого симптому (чи групи симптомів) певному захворюванню. Наприклад, відповідність симптомів і агентів (захворювань) наведено в таблиці. Тоді така залежність може бути описана нечіткими правилами вигляду:

**вихід** ( $x_1, x_2, x_3$ );

**якщо**  $x_1 \in$  “слабкий”  $\wedge x_2 \in$  “слабкий”  $\wedge x_3 \in$  “підвищена”, **то**  $y \in$  “0.3/squil + 0.3/ squil + 0.5/APIS”;

**якщо**  $x_1 \in$  “слабкий”  $\wedge x_2 \in$  “помірна”  $\wedge x_3 \in$  “висока”, **то**  $y \in$  “0.3/squil + 0.5/ IGN + 0.6/APIS”;

**якщо**  $x_1 \in$  “слабкий”  $\wedge x_2 \in$  “помірна”  $\wedge x_3 \in$  “дуже висока”, **то**  $y \in$  “0.3/squil + 0.5/ IGN + 0.9/ ACON”;

**вихід** ( $y$ ),

де елементи терм-множин визначаються, як і раніше.

ТАБЛИЦЯ

Симптоми	Агенти
Кашель	<b>APIS;</b> MEZ, NAT-MSEC, SPONG; squil
Температура	<b>ACON;</b> ANT-T, APIS, ARN, ARS, BELL, BRY, CACT, CHIN-S, CON, FERR-P, GELS, IP, KALI-I, LYC, MEZ, NAT-M, NUX-V, PHOS, PULS, RHUS-T, SEC, SPONG; squil, stram, tarax, verat
Нежить	<b>APIS, BELL;</b> GELS, IGN, PHOS, PULS; squil, stram

Відправною точкою для створення подібної системи може слугувати створення бази знань, що покриває переважну більшість поширених симптомів та захворювань одного чи декількох органів людини, наприклад, зору та слуху. Подальший розвиток системи може полягати у додаванні до бази знань нових інформаційних блоків про різні органи людського організму.

Діагностичні системи такого типу можуть відслідковувати стан пацієнта «у динаміці», враховувати попередню історію захворювань, за рахунок чого максимально підвищувати точність встановленого попереднього діагнозу.

Контроль над системою здійснює лікар-експерт, до функціональних обов'язків якого може входити поповнення бази знань (лікар може вносити дані про нові симптоми та взаємопов'язані із ними хвороби, тим самим поповнюючи базу знань), введення до системи результатів попередньої ітерації та їх обробка, аналіз результатів обстеження та призначення лікування інше.

Альтернативним варіантом застосування є використання експертно-діагностичних систем на базі державних чи приватних медичних лабораторій, останні з яких користуються високою популярністю серед громадян. Вже сьогодні лабораторії надають не просто чисельну інформацію про рівень того чи іншого компонента в організмі людини, а і проводять поверхневий аналіз результатів. Описана вище система могла б вдало доповнювати наявне у лабораторіях обладнання, а головне у випадку критичних результатів максимально швидко сповіщати клієнта електронною поштою про необхідність звернутися до лікаря.

**Висновки.** Таким чином, запропонований підхід дозволяє описувати нечіткі знання, які стосуються як діагностики, так і захворювання. Особливістю діагностичної системи є наявність інтелектуального компонента, який забезпечує реалізацію різних функціональностей виключно за рахунок власних можливостей без зміни програмного коду.

*В.Д. Саввакин, А.И. Проватар*

#### ДИАГНОСТИЧЕСКАЯ СИСТЕМА НА ОСНОВЕ НЕЧЕТКИХ ЗНАНИЙ

Предложен подход к построению диагностической системы на основе нечетких знаний с реализацией возможности обучения путем пополнения знаний о предметной области.

*V.D. Savvakina, A.I. Provotar*

#### FUZZY KNOWLEDGE BASED DIAGNOSTIC SYSTEM

An approach to the construction of a fuzzy knowledge based diagnostic system with the implementation of the possibility of learning by adding the knowledge about subject area is proposed.

**Список літератури**

1. Rutkowski L. Metody i Techniki Sztucznej Inteligencji (in Polish). Wydawnictwo Naukowe PWN, Warszawa, 2009. 452 s.
2. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. М.: Телеком, 2006. 382 с.
3. Zadeh L.A., Fuzzy Sets. Information and Control. 1965. Vol. 8. P. 338 – 353.
4. Провотар А.И., Лапко А.В. О некоторых подходах к вычислению неопределенностей. *Проблеми програмування*. 2010. № 2-3. С. 22 – 27.

Одержано 04.05.2019

***Про авторів:***

*Саввакін Володимир Дмитрович,*

магістр Національного технічного університету України  
«Київський політехнічний інститут імені Ігоря Сікорського»,

*Провотар Олександр Іванович,*

завідувач кафедри Київського національного університету імені Тараса Шевченка.

E-mail:aprowata1@bigmir.net