

7. Воеводин В. В. Численные методы алгебры. Теория и алгоритмы / В. В. Воеводин. — М. : Наука, 1966.
8. Воеводин В. В. Линейная алгебра / В. В. Воеводин. — М. : Наука, 1980.

In this paper the methods and algorithms of signal processing subject to informational redundancy by the solution of redefine systems of linear algebraic equations are regarded.

Key words: *previous and a posteriors of signal processing; location; adaptive processing; informational redundancy; redefine systems of linear algebraic equations; methods of solution.*

Отримано: 02.04.2009

УДК 519.766.23, 519.767.6

О. В. Нечипоренко^{*}, канд. техн. наук,

А. А. Верлань^{**}, канд. техн. наук,

Ю. О. Фуртат^{***}, аспірант

^{*} Східноєвропейський інститут економіки і менеджменту, м. Черкаси,

^{**} НТУУ “КПІ”, м. Київ,

^{***} ІПМЕ ім. Г.С. Пухова НАН України, м. Київ

СЕМАНТИЧНА ІНФОРМАЦІЯ ТА ЛІНГВІСТИЧНІ ЗМІНИ В ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЯХ

Розглядаються проблеми представлення неформалізованих процедур в інформаційних технологіях. Неформалізовані процедури визначаються як процедури, які використовують змістовну інформацію, що задається змістом і значенням понять і суджень предметної області. Визначається логіко-лінгвістичний опис предметної області і відповідне йому логіко-семантичне представлення нечіткої системи. Розглядається задача класифікації та структуризації знань.

Ключові слова: *неформалізована процедура, змістовна інформація, логіко-лінгвістичний опис, логіко-семантичне представлення, лінгвістична змінна.*

Вступ. Концептуальну основу традиційних інформаційних технологій складає алгоритм, тобто формалізоване знання, що має форму строгих суджень — формальних правил. Однак людина оперує в основному не формальними правилами, а змістовно-неформальними, які враховують зміст понять і суджень предметної області. Зміст інформації включає в себе семантику і значення, в якому використову-

ються ці поняття та судження [1]. Обробка за допомогою ЕОМ смислової інформації передбачає моделювання наближених міркувань людини, а значення, що приписуються поняттям і судженням предметної області, можуть відображати іншу важливу властивість людського мислення — здатність до структуризації та класифікації знань. Задачі, що розв'язуються з використанням змістовних знань, відносяться до неформалізованих. Змістовні знання формулюються користувачем у вигляді евристичних правил і утворюють логіко-лінгвістичний опис предметної області.

Таким чином, проблема, що розглядається, полягає в представленні смислової інформації в ЕОМ і в структуризації і класифікації знань з її допомогою.

Відмітимо, що в багатьох галузях практичної діяльності людини неформалізовані задачі утворюють, можливо, значно більший клас задач, ніж формалізовані. Тому непридатність ЕОМ, яка є алгоритмічною системою, для розв'язування неформалізованих задач стримує розвиток інформаційних технологій, зокрема, застосування ЕОМ в описових [2].

Логіко-лінгвістичний опис предметної області L в загальному випадку являє собою сукупність логіко-лінгвістичних описів L_Q зв'язків вхідних та вихідних параметрів (або просто параметрів) в термінах лінгвістичних змінних, $L_Q \in L$. Під логіко-лінгвістичним описом L_Q зв'язків параметрів ми розуміємо сукупність логіко-лінгвістичного правила Q та його конкретизації

$$q_i^Q \in q^Q, i = 1, \dots, n_Q, L_Q = (Q, q^Q),$$

де $Q \stackrel{\Delta}{=} (\text{Якщо } X, \text{ то } Y)$, $q_i^Q \stackrel{\Delta}{=} (\text{Якщо } X^{(i)}, \text{ то } Y^{(i)})$, q^Q — схема нечітких міркувань. Тут X, Y — назви лінгвістичних змінних, які в загальному випадку можуть бути складеними, $X^{(i)}, Y^{(i)}$ — значення лінгвістичних змінних. Якщо цим значенням надається деяка семантика (смысл), то $X^{(i)}$ и $Y^{(i)}$ — назви нечітких змінних. Логіко-лінгвістичні правила утворюють множину

$$P = \{Q_1, \dots, Q_i, \dots, Q_l\}, i = 1, \dots, l.$$

Семантика значень $X^{(i)}, Y^{(i)}$ формалізується за допомогою нечітких множин $X^{(i)}, Y^{(i)}$ універсальних множин U и V відповідно. Нагадаємо, що нечіткою множиною X на множині U називається сукупність пар:

$$X = \{ \langle \mu_X(u), u \rangle \}, \quad (1)$$

де $\mu_X : U \rightarrow [0, 1]$ — відображення множини U в одиничний відрізок $[0, 1]$, що називається функцією належності нечіткої множини X . Зна-

чення функції належності $\mu_X(u)$ для $u \in U$ називається ступенем належності. Вираз (1) іноді записують в еквівалентній формі

$$X = \mu_X(u_1)/u_1 + \mu_X(u_2)/u_2 + \dots + \mu_X(u_m)/u_m.$$

Кажуть, що елементарне висловлювання $p = X$ транслюється в розподілі можливостей [3]

$$\pi(X) = \tilde{X} : p = X \rightarrow \pi(X) \rightarrow \tilde{X}.$$

Приклад 1. Нехай задано логіко-лінгвістичний опис зв'язку параметру структурної вкладеності програми X її складністю.

$Q = \overset{\Delta}{\Delta}$ — якщо вкладеність програми X , то складність програми Y ;

$q_1 = \overset{\Delta}{\Delta}$ — якщо вкладеність програми НЕВЕЛИКА, то складність програми НЕЗНАЧНА;

$q_2 = \overset{\Delta}{\Delta}$ — якщо вкладеність програми СЕРЕДНЯ, то складність програми СЕРЕДНЯ;

$q_3 = \overset{\Delta}{\Delta}$ — якщо вкладеність програми ГЛИБОКА, то складність програми ДУЖЕ ВИСОКА.

В цьому описі $X =$ “вкладеність програми”, $Y =$ “складність програми”. Семантика використаних значень лінгвістичних змінних може бути, наприклад, визначена за п'ятибальною шкалою і представлена нечіткими множинами таким чином:

1) лінгвістична змінна “вкладеність програми”

$X^{(1)} =$ НЕВЕЛИКА; $X^{(1)} = [\text{НЕВЕЛИКА}] = 1/1-0,8/2+0,4/3-0,1/4-0/5$;

$X^{(2)} =$ СЕРЕДНЯ; $X^{(2)} = [\text{СЕРЕДНЯ}] = 0/1-0,4/2+1/3-0,4/4-0/5$;

$X^{(3)} =$ ГЛИБОКА; $X^{(3)} = [\text{ГЛИБОКА}] = 0/1+0/2-0,3/3-0,6/4+1/5$;

2) лінгвістична змінна “складність програми”

$Y^{(1)} =$ НЕЗНАЧНА; $Y^{(1)} = [\text{НЕЗНАЧНА}] = 1/1-0,9/2-0,3/3+0,1/4-0/5$;

$Y^{(2)} =$ СЕРЕДНЯ; $Y^{(2)} = [\text{СЕРЕДНЯ}] = 0/1+0,6/2+1/3-0,6/4-0/5$;

$Y^{(3)} =$ ДУЖЕ ВИСОКА; $Y^{(3)} = [\text{ДУЖЕ ВИСОКА}] = 0/1-0/2-0,3/3-0,6/4+1/5$.

Представлення смислової інформації в інформаційних технологіях можна здійснити за допомогою теорії нечітких множин. Однак кожне нечітке умовне висловлювання виду “Якщо X , то Y , інакше Z ” (X, Y, Z — нечіткі підмножини універсальних множин U, V та V відповідно) наділяється логічними структурами індуктивного та дедуктивного типів [4, 5]. Виникає проблема логіко-семантичного представлення нечіткої системи [5].

Під логіко-семантичним представленням нечіткої системи (що відповідає логіко-лінгвістичному опису зв'язків параметрів) розуміється сукупність моделі нечіткої системи та операторів перетворення множин вхідних параметрів в вихідні з урахування логічної структури системи. Множини вхідних та вихідних параметрів системи описуються лінгвістичними змінними, значення яких наділені певною семантикою. Вибір конкретної логічної структури залежить від типу причинно-наслідкових зв'язків, що розглядається.

Нечіткі умовні висловлювання можуть бути представлені двома моделями: 1) індуктивною моделлю (RI), в якій причина обумовлена виводом, а наслідок — посиланнями; модель наділена логічною структурою індуктивного типу; 2) дедуктивною моделлю (RD), в якій причина обумовлена посиланнями, а наслідок — виводом; модель наділена логічною структурою дедуктивного типу.

Теорема про уявність. Логіко-семантичне представлення нечіткої системи є представленням вигляду $X = Y \circ RID$ і $Y = X \partial RID$, де

$$\mu_{RID}(u, v) = \begin{cases} I, \text{ якщо } \mu_X(u) \geq \mu_Y(v) \begin{pmatrix} \text{логічна структура} \\ \text{індуктивного типу} \end{pmatrix} \\ \mu_X(u), \text{ якщо } \mu_X(u) < \mu_Y(v) \text{ (будь-яка логічна структура)} \\ \mu_Y(v), \text{ якщо } \mu_X(u) \geq \mu_Y(v) \begin{pmatrix} \text{логічна структура} \\ \text{дедуктивного типу} \end{pmatrix} \end{cases}$$

\circ і ∂ — оператори перетворення (композиції нечітких відношень) множини вхідних параметрів в вихідні

$$\mu_X(u) = \max_v \{ \mu_Y(v) \wedge \mu_{RI}(u, v) \}$$

і

$$\mu_{Y(v)} = \inf_{u'} \in \{ \mu_X(u) > \mu_{RD}(u, v) \} \{ \mu_{RD}(u', v) \}.$$

Логіко-семантичне представлення нечіткої системи, що відповідає логіко-лінгвістичному опису “Якщо X , то Y , інакше Z ” (X, Y, Z — назви лінгвістичних змінних) буде несуперечливим, якщо існує модель RID нечіткої системи, а також оператори \circ (max–min-композиція) та ∂ (∂ -композиція) перетворення вхідних параметрів у вихідні, що забезпечують узгоджену взаємодію індуктивного і дедуктивного механізмів нечіткого логічного: $X = (X \partial RID) \circ RID$ і $Y = (Y \circ RID) \partial RID$. Нагадаємо, що X і Y — нечіткі множини, що визначають семантику значень лінгвістичних змінних X та Y відповідно.

Приклад 2. Розглянемо логіко-лінгвістичний опис зв'язку параметра структурної вкладеності програми з її складністю з прикладу 1.

Індуктивна модель нечіткої системи у відповідності до теореми про уявності має наступний вигляд

$$RI = \begin{pmatrix} 1,0 & 0,8 & 0,4 & 0,1 & 0 \\ 0 & 0,4 & 0,4 & 0,1 & 0 \\ 0 & 0 & 1,0 & 0,1 & 0 \\ 0 & 0 & 0,3 & 0,4 & 0 \\ 0 & 0 & 0,3 & 0,6 & 0 \end{pmatrix}.$$

Дедуктивна модель тієї ж самої нечіткої системи запишеться як

$$RD = \begin{pmatrix} 1,0 & 0,9 & 0,3 & 0,1 & 0 \\ 0,8 & 0,8 & 0,4 & 0,4 & 0 \\ 0,4 & 0,6 & 1,0 & 0,6 & 0,3 \\ 0,1 & 0,4 & 0,4 & 0,6 & 0,6 \\ 0 & 0 & 0,3 & 0,6 & 1,0 \end{pmatrix}.$$

Модель нечіткої системи RID включає в себе обидві ці моделі. Вибір конкретної моделі залежить від типу причинно-наслідкових зв'язків, що розглядаються.

Припустимо, що визначено значення ГЛИБОКА лінгвістичної змінної “вкладеність програми”. Тоді, застосовуючи ∂ -композицію, отримаємо значення лінгвістичної змінної “складність програми”

$$Y^{(i)} = X^{(i)} \partial RD = [ГЛИБОКА] \partial RD = [ДУЖЕ ВИСОКА].$$

В системі враховується дедуктивний причинно-наслідковий зв'язок.

Нехай тепер нам відоме значення ДУЖЕ ВИСОКА лінгвістичної змінної “складність програми”. Потрібно визначити значення лінгвістичної змінної “вкладеність програми”. В цьому випадку необхідно використати (max-min)-композицію, в нечіткій системі враховується індуктивний причинно-наслідковий зв'язок. Тоді

$$X^{(i)} = Y^{(i)} \circ RI = [ДУЖЕ ВИСОКА] \circ RI = [ГЛИБОКА].$$

Структуризація і класифікація знань — це одна з основних задач проблеми представлення неформалізованих процедур в інформаційних технологіях. Якщо поняття і судження предметної області можуть розумітись в значенні “загальне правило — окреме правило”, “загальне правило — виняток з правил”, то для розв'язання задачі, що розглядається, можна застосувати теоретичний апарат формальних систем з винятками [6]. Основні положення цього апарату необхідно узагальнити на нечіткий випадок.

В логіко-лінгвістичному описі предметної області $L=\{L_Q\}$ Q_i представляють собою логіко-лінгвістичні правила з множини P , $i=1, \dots, I$. Множину P можна розбити на M класів, що не перетинаються, $P_1, \dots, P_m, \dots, P_M$ функціонально однотипних правил (правил, у яких в правій частині знаходиться одна й та сама змінна). Деякі пари цих правил з різних класів пов'язані відношенням “загальне правило — окреме правило” (ЗО-відношення), для позначення якого будемо використовувати знак \ll — якщо $(Q \in P_i) \& (P \in P_j) \& (Q \ll P)$, то Q — опис окремих зв'язків, а P — опис загальних зв'язків для $Q, i \neq j$.

Кожний з класів P_m містить логіко-лінгвістичні правила $Q_1^m, \dots, Q_j^m, \dots, Q_{m_i}^m, m=1, \dots, m_i$. Деякі пари цих логіко-лінгвістичних правил пов'язані відношенням “множина посилянь — підмножина посилянь” (МП-відношення), для позначення якого будемо використовувати знак \prec — якщо $(Q^m \in P_m) \& (P^m \in P_m) \& (Q^m \prec P^m)$, то Q^m — виняток з правила P^m , а P^m — загальне правило для Q^m .

Окрім того, кожна множина нечітких висловлювань $q^Q = \{q_1^Q, \dots, q_i^Q, \dots, q_{n_Q}^Q\}$ може бути розбита на підмножини $q_1^Q, \dots, q_i^Q, \dots, q_{n_Q}^Q$, що не перетинаються, для яких побудовані моделі нечітких систем $\tilde{R}_1^Q, \dots, \tilde{R}_k^Q, \dots, \tilde{R}_K^Q$ відповідно, $k=1, \dots, K$. Моделі нечітких систем пов'язані відношенням “загальне правило — виключення з правил” (ПВ-відношення), точніше різновидом ПВ-відношення — відношенням нечіткої спеціалізації. Для позначення цього відношення будемо використовувати знак \triangleleft — якщо $\tilde{R}_1^Q \triangleleft \tilde{R}_2^Q$, то \tilde{R}_1^Q — виняток (нечітка спеціалізація) моделі \tilde{R}_2^Q , а \tilde{R}_2^Q — загальна модель для \tilde{R}_1^Q .

Розглянуті типи відношень на множині нечітких умовних висловлювань використовуються для класифікації (виділення класів функціонально однотипних висловлювань і встановлення зв'язків між ними) знань.

Приклад 3. Розглянемо БЗ, що складається з чотирьох логіко-лінгвістичних правил, пронумерованих в порядку їх надходження:

1. Якщо $X_1 \& X_3$, то X_2
2. Якщо X_2 , то Y
3. Якщо X_3 , то X_2
4. Якщо $X_1 \& X_2$, то Y

Для цих правил можна побудувати фрейм-вказівник, значення слотів якого відображають ЗО-, ПП-, ПВ-відношення (рис. 1), де в рядках “вивід” і “посилання” вказуються номери правил, в яких використовуються відповідні базові поняття.

Назва X	Вивід	Посилання
X_1	[1, 3]	1, 4
X_2	[4, 2]	2, 4
V	[4, 2]	1, 3

Рис. 1. Фрейм-вказівник для прикладу

Розглянуті правила за допомогою фрейма-вказівника об'єднані в семантичну мережу, як показано на рис. 2. В процесі логічного виводу правилу 1 надається пріоритет по відношенню до правила 3, а правилу 4 — по відношенню до правила 2. Ця ситуація і відображена у фреймі-вказівнику за допомогою впорядкованого розміщення вказівників у стовпчику “вивід”: першим виконується правило, вказівник на яке розташований зліва у відповідному рядку.

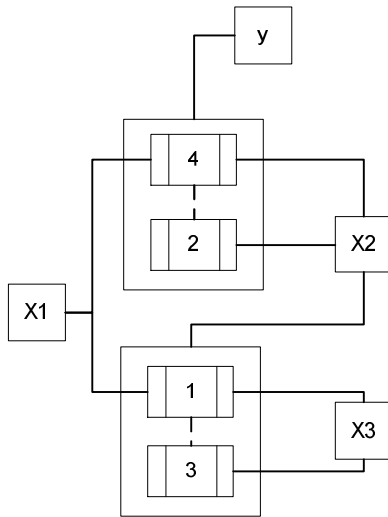


Рис. 2. Семантична мережа

В режимі повернень при реалізації неформалізованих процедур необхідно забезпечити заборони логіко-лінгвістичних правил у випадку використання винятків. У відповідності до цієї концепції необхідно встановити порядок перебору логіко-лінгвістичних правил, що гарантуватиме неможливість вибору якогось правила до всіх його винятків, і встановити порядок перебору моделей нечітких систем

для обраного логіко-лінгвістичного, що гарантуватиме неможливість вибору якоїсь моделі нечіткої системи до її спеціалізації.

Проблема повернень при реалізації неформалізованих процедур існує в зв'язку з використанням в описі значень лінгвістичних змінних таких слів, як БІЛЬШІСТЬ, ДЕЯКІ, ДУЖЕ, БІЛЬШ-МЕНШ тощо.

Приклад 4. Нехай p = (зріст ДУЖЕ МАЛЕНЬКИЙ),

Q = (Якщо зріст X , то вага Y),

q_1 = (Якщо зріст МАЛЕНЬКИЙ, то вага ЛЕГКА),

q_2 = (Якщо зріст ДУЖЕ МАЛЕНЬКИЙ, то вага ДУЖЕ ЛЕГКА),

де МАЛЕНЬКИЙ і ЛЕГКИЙ формалізовані нечіткими множинами [МАЛЕНЬКИЙ] = $1/160 + 0,5/170 + 0/180$, [ЛЕГКИЙ] = $1/60 + 0,4/70 + 0/80$. Розподіл можливостей p , q_1 , q_2 і Q , має такий вигляд:

$$p \rightarrow \pi_{\text{зріст}} = [\text{ДУЖЕ МАЛЕНЬКИЙ}] = \pi_{(\text{маленький})}^2 = (1 \quad 0,25 \quad 0)$$

$$q_1 \rightarrow \pi_{(\text{маленький, легкий})} = \tilde{R}_1 = \begin{pmatrix} 1 & 0,4 & 0 \\ 0,5 & 0,4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$q_2 \rightarrow \pi_{(\text{дуже маленький, дуже легкий})} = \tilde{R}_2 = \begin{pmatrix} 1 & 0,16 & 0 \\ 0,25 & 0,16 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$Q \rightarrow \pi_{\text{зріст, вага}}^Q = \tilde{R}_1 \cup \tilde{R}_2 = \begin{pmatrix} 1 & 0,4 & 0 \\ 0,5 & 0,4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Застосовуючи композиційне правило виводу до заданих нечітких множин, отримуємо

$$(1 \quad 0,25 \quad 0) \circ \begin{pmatrix} 1 & 0,4 & 0 \\ 0,5 & 0,4 & 0 \\ 0 & 0 & 0 \end{pmatrix} = (1 \quad 0,4 \quad 0).$$

Отримана в результаті нечіткого виводу нечітка множина відповідає значенню ЛЕГКИЙ лінгвістичної змінної “вага” — r = (вага ЛЕГКА). Це значення не співпадає з інтуїтивно очікуваним результатом ДУЖЕ ЛЕГКИЙ. Таким чином, при нечіткому виводі моделі R_1 і R_2 повинні застосовуватись окремо.

Висновки. Представлення неформалізованих процедур в інформаційних технологіях зменшує семантичний розрив між людиною і ЕОМ, розширює область застосування обчислювальної техніки за рахунок оброки смислової інформації.

Список використаних джерел:

1. Петров Ю. А. Культура мышления: методологические проблемы научно-педагогической науки / Ю. А. Петров. — М. : Изд-во МГУ, 1990. — 118 с.
2. Дородницын А. А. Информатика: предмет и задачи / А. А. Дородницын // Природа, 1985, № 2. — С. 26—29.
3. Нечеткие множества в моделях управления и искусственного интеллекта / под ред. Д. А. Поспелова. — М. : Наука, 1986. — 312 с.
4. Танака Х. Модель нечеткой системы, основанная на логической структуре. — В кн.: Нечеткие множества и теория возможностей / Х. Танака, Т. Цукияма, К. Асаи ; пер. с англ. ; под ред. Р. Р. Ягера. — М. : Радио и связь, 1986. — С. 186—189.
5. Романов А. А. Логико-семантическая модель причинно-следственных связей в нечеткой системе / А. А. Романов, Ю. И. Шемакин // НТИ. Сер. 2. Информационные процессы и системы, 1990, № 6. — С. 26—32.
6. Кузнецов В. Б. Представление ЭВМ неформальных процедур: продукционные системы / В. Б. Кузнецов ; с послесл. Д. А. Поспелова. — М. : Наука, 1989. — 160 с.

The problems of presentation of the unformalized procedures in information technologies are considered. The unformalized procedures are determined as procedures, which use semantic information which is set by content and meaning of the subject domain's concepts and arguments. Logical-linguistic description of the subject domain is determined and corresponding logical-semantic conception of the fuzzy system. The task of knowledge classification and structuring is considered.

Key words: *unformalized procedure, semantic information, logical-linguistic description, logical-semantic conception, linguistic variable.*

Отримано: 28.11.2009