

УДК 519.767

*С.Г. Лелеков*Севастопольский национальный технический университет, г. Севастополь, Украина
Украина, 99053, г. Севастополь, ул. Университетская, 33

Метод учета влияния модификаторов на нечеткие значения лингвистических переменных

*S.G. Lelekov**Sevastopol National Technical University, c. Sevastopol, Ukraine
Ukraine, 99053, c. Sevastopol, Universitetskaya st., 33*

Method of Accounting for the Influence of Modifiers on the Fuzzy Values of Linguistic Variables

*С.Г. Лелеков*Севастопольський національний технічний університет, м. Севастополь, Україна
Україна, 99053, м. Севастополь, вул. Університетська, 33

Метод обліку впливу модифікаторів на нечіткі значення лінгвістичних змінних

В статье рассматривается проблема учета влияния слов-модификаторов типа: очень, слегка, немного, слишком и т.п. на функции принадлежности нечетких значений лингвистических переменных. Предлагается метод коррекции функций принадлежности нечетких переменных. Представлены экспериментальные данные по результатам использования предложенного метода. Полученные результаты могут быть использованы в системах нечеткого моделирования, в основе которых лежит нечеткая логика.

Ключевые слова: лингвистическая переменная, нечеткая переменная, диагностические экспертные системы

The article considers the problem of accounting for the impact of words-modifiers: very little, a little, too, etc. on the membership function of the fuzzy values of linguistic variables. It is offered the method of correction of the membership functions of fuzzy variables. The experimental data on the results of the use of the proposed method. The results obtained can be used in systems of fuzzy modeling, which are based on fuzzy logic.

Key words: linguistic variables, fuzzy variables, diagnostic expert system

У статті розглядається проблема обліку впливу модифікаторів типу: дуже, злегка, трохи і т.п. на функції належності нечітких значень лінгвістичних змінних. Пропонується метод корекції функцій належності нечітких змінних. Наведено експериментальні дані за результатами використання запропонованого методу. Отримані результати можуть бути використані в системах нечіткого моделювання, в основі яких лежить нечітка логіка.

Ключові слова: лінгвістична змінна, нечітка змінна, діагностичні експертні системи.

Введение

В настоящее время нечеткое моделирование является одним из наиболее перспективных направлений прикладных исследований в области управления и принятия решений. Диагностические экспертные системы являются неотъемлемой частью систем поддержки принятия решений в технике, медицине, биологических науках. Так, например, в гидробиологии идентификация гидробиологических объектов занимает важное место в системе мониторинга окружающей среды. Выполнять работу по определению видового состава проб могут только специалисты-эксперты, в подготовку которых входит

не только изучение соответствующих университетских курсов, но и длительная последующая стажировка по узко определенной тематике. Такая же ситуация наблюдается и в других дисциплинах.

Для представления знаний экспертов в компьютерных системах в настоящее время используются различные программы-оболочки [1]. В них имеющиеся знания представляются в виде трех основных компонент:

- описания объектов в выбранной системе признаков или правил;
- описания исходных систем классификации объектов, например, иерархическая система классификации живых существ в биологии;
- справочной информации самого разнообразного характера: текстовой, графической, табличной (в т.ч. базы данных), мультимедийной.

Одной из основных проблем при создании и использовании компьютерных баз знаний является неоднозначность естественного языка, применяемого как экспертами при описании различных объектов, так и пользователями при ответах на вопросы экспертной системы. Эксперт описывает объект в системе признаков, каждый из которых можно рассматривать как лингвистическую переменную.

Лингвистической называется переменная, значениями которой являются слова естественного языка. Например, «*Форма хвостового плавника рыбы*» будет лингвистической переменной, если она принимает значения: *вогнутый, прямой, выпуклый, почти прямой, сильно выпуклый, слегка вогнутый* и т.д.

Пользователь вынужден использовать терминологию эксперта, хотя их представления о сущности того или иного значения лингвистической переменной могут отличаться.

В этой ситуации вполне оправданным становится использование аппарата нечетких множеств для работы с лингвистическими переменными. Тогда каждое из значений лингвистической переменной может быть представлено нечеткими переменными. Однако это не решает полностью указанную выше проблему: эксперт не может задать все возможные значения лингвистической переменной в виду многообразия способов построения ее конструкций в естественном языке.

Решением здесь может быть перечисление экспертом некоторых базовых значений, в терминах которых и происходит описание объекта, а пользователю предоставляется возможность использовать смысловые модификации базовых терминов.

Многообразие значений лингвистических переменных, чаще всего, возникает при использовании слов-модификаторов, относящихся, обычно, к количественным наречиям, например, очень, почти, едва, слегка и др.

Целью данной работы является разработка метода учета влияния слов-модификаторов на функции принадлежности нечетких значений лингвистических переменных, что позволит повысить качество нечеткого вывода в диагностических экспертных системах.

Постановка задачи

Нечеткой переменной [2] назовем тройку $t=(\alpha, U, \Lambda(x, \mu_x))$, где α – название нечеткой переменной, U – область ее определения, $\Lambda(x, \mu_x)$ – функция принадлежности, которая задает μ_x – степень соответствия значения $x \in U$ нечеткой переменной α .

Лингвистической переменной назовем пятерку

$$L=(\chi, T, U, G, M),$$

где χ – название переменной;

$T=(t_1, t_2, \dots, t_n)$ – базовое множество нечетких значений переменной;

U – универсальное множество значений области определения переменных из T , при этом U является общим для всех переменных из T ;

G – грамматика, порождающая новые значения переменной из T ;

M – семантические правила, задающие функции принадлежности, определенные на множестве U каждого значения переменной из T , в т.ч. и для вновь образованных значений в соответствии с грамматикой G .

Например, лингвистическая переменная «Температура кофе» может быть представлена пятеркой:

$L_1 = (\langle \text{«Температура кофе»}, T_1 = \{\langle \text{«холодный»}, \langle \text{«комнатный»}, \langle \text{«теплый»}, \langle \text{«горячий»}\}, U_1 = \{10-95\}, G_1, M_1)$, где G_1 – набор правил порождения различных синтаксических конструкций из термов множества T_1 и набора модификаторов (обычно, количественных наречий), например: *очень, едва, немного, слишком* и т.п. Будем предполагать, что синтаксические правила G соответствуют общепринятым нормам построения подобных конструкций в русском языке.

Семантические правила M_1 удобно представить графически (рис. 1). Здесь нечеткие переменные из T_1 выражены значениями функции принадлежности:

$\langle \text{«горячий»} \rangle = \{0/1, 0.2/2, 0.5/3, 1/4\}$;

$\langle \text{«теплый»} \rangle = \{0.1/1, 0.5/2, 1/3, 0.4/4\}$;

$\langle \text{«комнатный»} \rangle = \{0.4/1, 0.9/2, 0.6/3, 0.2/4\}$;

$\langle \text{«холодный»} \rangle = \{1/1, 0.6/2, 0.2/3, 0/4\}$.

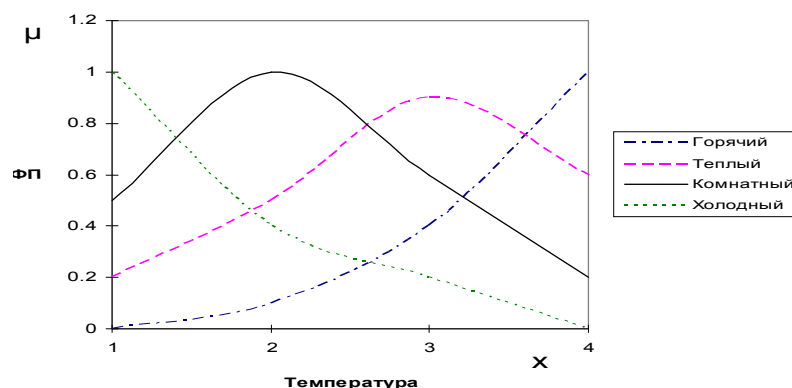


Рисунок 1 – Функции принадлежности нечетких значений лингвистической переменной «Температура кофе»

По оси X представлена температура в относительных единицах, покрывающих универсальное множество U_1 от 10 до 95. Ось Y соответствует значениям функции принадлежности μ нечетких значений из множества T_1 лингвистической переменной L_1 , лежащим в диапазоне $0 \leq \mu \leq 1$.

Функции принадлежности представлены для четырех объектов, которые в данном случае можно рассматривать, как диапазоны температур, соответствующих представлениям эксперта о значениях лингвистической переменной «Температура кофе». В целом при работе с лингвистическими переменными реальные числовые значения области определения U несущественны. Более важно соотношение значений между собой, что и выражают функции принадлежности.

Необходимым и достаточным условием возможности использования лингвистических переменных для описания объектов может служить требование о наличии линейного порядка на множестве определения U нечетких значений лингвистических переменных. Другими словами, должна существовать некоторая количественная мера, позволяющая задать этот линейный порядок. Например, в качестве такой меры может служить температура, радиус кривизны, скорость, количество и т.п.

При выполнении этого условия функции принадлежности можно записывать относительно множества базовых нечетких переменных из множества T . Поэтому далее,

при записи функций принадлежности будем использовать упрощенную запись, например, <горячий> = {0, 0.2, 0.5, 1}.

В диагностических экспертных системах при описании j -го объекта по признаку i , представляющему собой лингвистическую переменную, функции принадлежности значений этой переменной могут быть записаны в виде

$$D_i(j) = \{d_1, d_2, \dots, d_k, \dots, d_n\},$$

где n – количество базовых нечетких переменных лингвистической переменной, используемой для представления i -го признака; d_k – значение функции принадлежности объекта j значению k лингвистической переменной, представляющей признак i .

Например, в таксономической экспертной системе идентификации рыб признак «Форма хвостового плавника рыб» может быть представлен лингвистической переменной с 4 базовыми нечеткими переменными:

- 1) хвостовой плавник сильно закруглен;
- 2) хвостовой плавник округлый;
- 3) хвостовой плавник почти прямой;
- 4) хвостовой плавник вогнутый.

Тогда функции принадлежности могут быть заданы следующим образом:

$$D_1(\text{вид Labrys merula}) = \langle 0.2, 0.4, 1.0, 0.0 \rangle,$$

$D_1(\text{вид Thalassoma pavo}) = \langle 0.0, 0.0, 0.3, 1.0 \rangle$ и т.д. Другими словами *Labrys merula* имеет почти прямой хвостовой плавник, может несколько округлый. У *Thalassoma pavo* – вогнутый.

Как было отмечено выше, описание объектов экспертом в некоторой системе признаков, представляющих собой лингвистические переменные, предназначено для использования пользователями. Однако они могут указывать значения лингвистических переменных несколько иначе, чем предполагалось изначально экспертом, но в соответствии с правилами порождения новых значений лингвистических переменных, заданных грамматикой G . Особую значимость здесь приобретает использование слов-модификаторов, относящихся, обычно, к количественным наречиям. Допущение возможности вводить новые значения лингвистических переменных ставит задачу коррекции функций принадлежности базовых значений. Для решения этой задачи предлагается использование операций разряжения и размывания.

Операция разряжения

В русском языке модификаторы представлены в основном количественными наречиями, которые можно разбить на 2 группы.

В первую группу войдут те, которые «размывают» и так нечеткие значения лингвистической переменной, например, *более менее горячий, около пяти, где-то 12* и т.п.

Назовем эту группу группой типа «Равно» и включим в нее следующие модификаторы: *приблизительно, около, где-то, почти, примерно, вероятно, кажется, более или менее, скорее всего, точно*.

Во вторую группу типа «Больше-Меньше» войдут модификаторы, которые изменяют значение переменной в одну сторону: *немного, чуть, значительно, очень, слегка, несколько, много, не более, не менее, сильно, слишком, едва*. Например, *немного горячий, почти холодный, очень горячий* и т.п.

Использование слов-модификаторов «размывает» и так нечеткие значения лингвистических переменных. Так как эксперт задает только базовые значения лингвистических переменных, то учесть различные смысловые оттенки в пределах грамматики G не представляется возможным. Для решения этой проблемы воспользуемся принципом обобщения для нечетких множеств [3].

Принцип обобщения позволяет расширить область определения U , включив в нее наряду с имеющимися точками произвольные нечеткие подмножества множества U .

Предположим, что f – отображение $U \rightarrow V$, а A – нечеткое подмножество вида

$$A = \mu_1 u_1, \dots, \mu_n u_n,$$

где $\mu_i u_i$ означает, что элемент u_i имеет значение функции принадлежности μ_i .

Тогда принцип обобщения утверждает, что

$$f(A) = f(\mu_1 u_1, \dots, \mu_n u_n) \equiv \mu_1 f(u_1), \dots, \mu_n f(u_n).$$

Используя принцип обобщения, зададим операцию разряжения следующим образом:

$$f: u_i \rightarrow v_{pi},$$

где $i = 1 \dots |U|$, p – порядок разряжения.

Функции принадлежности для элементов с номерами от i до p_i можно получить с помощью линейной интерполяции.

Например, если исходное множество задано $A = \{ \mu_1 u_1, \mu_2 u_2 \}$, то разряжение 3-го порядка даст $A' = \{ \mu_3 u_3, \mu_6 u_6 \}$. Множество V будет определено, как $V = \{ u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8 \}$. Для удобства вычислений, смысл которых станет понятен ниже, добавим еще 2 элемента – u_0 и u_9 – со значениями функции принадлежности пренебрежительно малыми и близкими к нулю.

Таким образом, из 10 элементов множества V уже определены функции принадлежности для четырех: $\mu_0 = \mu_9 = 0$, μ_3 , μ_6 . Остальные значения функции принадлежности найдем, используя линейную интерполяцию.

Пусть $j = p_i$, $m = p(i+1)$.

Тогда из уравнения прямой, соединяющей точки j и m , получим:

$$\frac{\mu_r - \mu_j}{\mu_m - \mu_j} = \frac{r - j}{m - j}, \text{ или}$$

$$\mu_r = \frac{(\mu_m - \mu_m)(r - j)}{m - j} + \mu_j.$$

Так, если $\mu_6 = 1$, $\mu_3 = 0.5$, получим: $\mu_1 = 0.16$, $\mu_2 = 0.33$, $\mu_4 = 0.67$, $\mu_5 = 0.87$, $\mu_7 = 0.67$, $\mu_8 = 0.33$.

Применив операцию растяжения второго порядка к нечетким значениям лингвистической переменной «Температура кофе», получим следующие значения функции принадлежности нечетких переменных, определенных уже не 4 значениях, а на одиннадцати :

cool (холодный):

0,010	0,505	1,000	0,800	0,600	0,400	0,200
0,105	0,010	0,010	0,010			

room (комнатный):

0,010	0,205	0,400	0,650	0,900	0,750	0,600
0,400	0,200	0,105	0,010			

warm (теплый):

0,010	0,055	0,100	0,300	0,500	0,750	1,000
0,700	0,400	0,205	0,010			

hot (горячий):

0,010	0,010	0,010	0,105	0,200	0,350	0,500
0,750	1,000	0,505	0,010			

Здесь нулевые значения μ_0 , и μ_{10} приняты равными 0.01.

Основным достоинством этого подхода является увеличение множества U без конкретной привязки к возможным значениям температур! Любой человек никогда не станет связывать значение лингвистической переменной «горячий кофе» с точным обозначением его фактической температуры в градусах.

Операция размывания

Влияние модификаторов на функции принадлежности нечетких значений лингвистических переменных можно представить довольно простой рекурсивной зависимостью:

$\mu_x = \mu_x^{y(x)}$, где в левой части μ_x – новое значение функции принадлежности в x -й точке оси X , на которой представлены экземпляры множества U после выполнения операции разряжения, в правой части – μ_x – предыдущее значение функции принадлежности в x -й точке, $y(x)$ – функция преобразования, где $x=1...|V|$.

В [2] в качестве $y(x)$ предлагается использовать константу, например, равную 0,5. Однако такое решение не может охватить всех особенностей влияния модификаторов на поведение функции принадлежности.

Исходя из того, что множество модификаторов можно разбить на 2 группы, рассмотрим поведение функции $y(x)$ в двух случаях: для модификаторов 1-й группы и для модификаторов 2-й группы.

Использование модификаторов первой группы приводит к росту значений функции принадлежности в окрестностях нечеткой переменной. Например, *горячий кофе* или *более менее горячий кофе*, или *кажется, горячий кофе*. По мере отдаления от базового значения, значения функции принадлежности меняются в меньшей степени. Хорошей математической моделью здесь может служить функция, построенная на основе использования функции Гаусса

$$y = 1 - e^{-a_1 x^2 - a_2}$$

Графики этой функции при разных значениях коэффициентов a_1 и a_2 представлены на рис. 2 и 3.

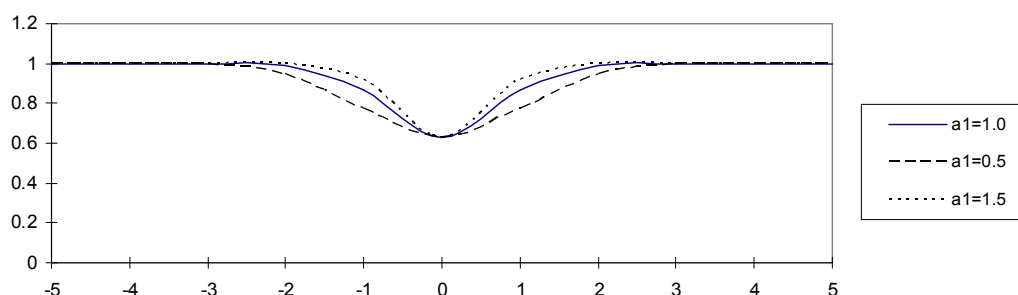


Рисунок 2 – Влияние коэффициента a_1 на функцию $y(x)$, при $a_2=1$

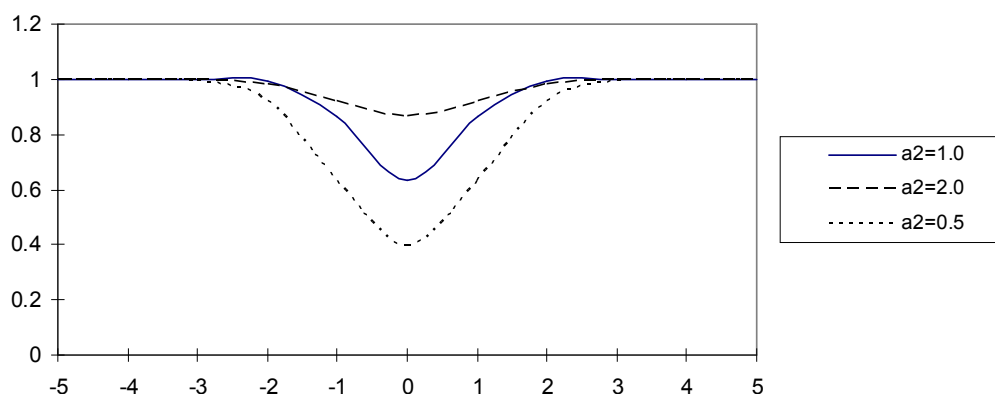


Рисунок 3 – Влияние коэффициента a_2 на функцию $y(x)$, при $a_1=0.5$
Коэффициент a_1 задает широту влияния модификатора на значения функции при-

надлежности. Коэффициент a_2 – степень этого влияния.

Например, нечеткая переменная «теплый кофе» с модификатором «более менее» будет иметь следующие значения функции принадлежности ($a_1=1$, $a_2=0,5$):

warm:

0,010	0,055	0,100	0,300	0,504	0,800	1,000
0,758	0,404	0,205.				

Модификаторы второй группы воздействуют на значения функций принадлежности более сложным образом. Так, например, модификатор «слегка», применительно к нечеткой переменной «теплый кофе», должен смещать значения функции принадлежности в сторону значений, соответствующих нечеткой переменной «комнатный кофе», тогда как модификатор «очень» – для этой же нечеткой переменной в сторону значений, соответствующих нечеткой переменной «горячий кофе».

В качестве функции $y(x)$ для модификаторов второй группы, сдвигающих значения функции принадлежности влево (типа «Меньше»), предлагается использовать сложную функцию следующего вида (рис. 4а):

$$y = \begin{cases} 1 - e^{a_1 \Delta x^2 - a_2}, & \text{если } x \leq x_0 \\ \sqrt{|\Delta x| + 1}, & \text{если } x > x_0 \end{cases}.$$

Здесь $x_0 = x_{\text{баз}} - (p-1)$, где $x_{\text{баз}}$ – крайняя левая точка функции принадлежности, в которой она принимает максимальное значение, p – порядок операции разряжения, $\Delta x = x - x_0$.

Кроме того, перед использованием функции $y(x)$, потребуется предварительно модифицировать значения функции принадлежности следующим образом:

$$\forall i, \mu_i = \begin{cases} \mu_i, & \text{если } x_i \leq x_0 \\ \mu_i - \delta_\mu, & \text{если } x_i > x_0 \end{cases}.$$

Здесь, $\delta_\mu = \mu_{x_{\text{баз}}} - \mu_{x_0}$

Аналогично, в качестве функции $y(x)$ для модификаторов второй группы, сдвигающих значения функции принадлежности вправо (типа «Больше»), предлагается использовать сложную функцию вида (рис. 4б):

$$y = \begin{cases} 1 - e^{a_1 \Delta x^2 - a_2}, & \text{если } x \geq x_0 \\ \sqrt{|\Delta x| + 1}, & \text{если } x < x_0 \end{cases}.$$

Здесь $x_0 = x_{\text{баз}} + (p-1)$, где $x_{\text{баз}}$ – крайняя правая точка функции принадлежности, в которой она принимает максимальное значение, p – порядок операции разряжения, $\Delta x = x - x_0$.

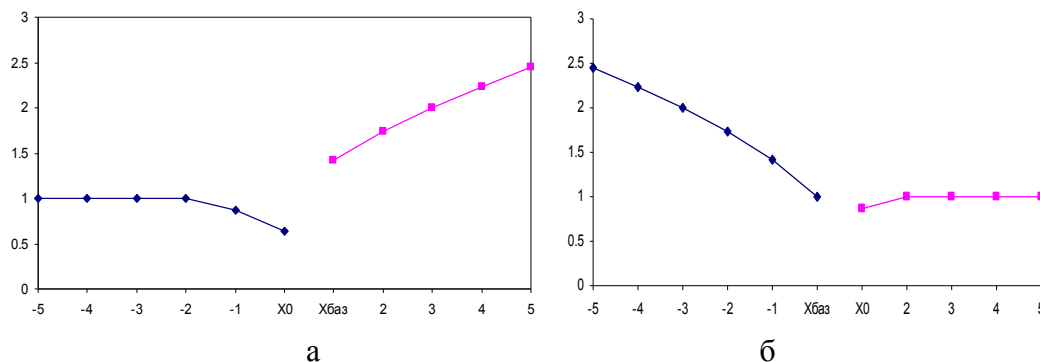


Рисунок 4 – Функция $y(x)$ для модификаторов 2-й группы:
а) «Меньше»; б) «Больше»

Кроме того, перед использованием функции $y(x)$, потребуется предварительно

модифицировать значения функции принадлежности следующим образом:

$$\forall i, \mu_i = \begin{cases} \mu_i, & \text{если } x_i \geq x_0 \\ \mu_i - \delta_\mu, & \text{если } x_i < x_0 \end{cases}$$

Здесь, $\delta_\mu = \mu_{x \text{ баз}} - \mu_{x0}$

Рассмотрим взаимодействие модификатора «слегка» и нечеткой переменной «теплый кофе». Получим:

warm:

0.010	0.055	0.101	0.319	0.584	0.834	0.666
0.251	0.023	0.000	0.000.			

Для описания примерно этой же ситуации, в которой отправной точкой является нечеткая переменная «комнатный кофе» с модификатором «очень», что можно интерпретировать как *слишком горячий для кофе с комнатной температурой*, получим:

room:

0.000	0.002	0.063	0.301	0.666	0.834	0.600
0.400	0.200	0.105	0.010.			

Коэффициент корреляции между двумя рядами значений $r=0.96$.

В табл. 1 показаны результаты расчетов по влиянию множественного применения модификатора «очень» к лингвистической переменной «горячий кофе».

Таблица 1 – Влияние множественного применения модификатора «очень» к лингвистической переменной «горячий кофе»

Нечеткая переменная	Функция принадлежности									
Горячий кофе	0.010	0.010	0.010	0.105	0.200	0.350	0.500	0.750	1.000	
Очень горячий кофе	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.094	0.381	
Очень очень горячий кофе	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
Очень очень очень горячий кофе	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
Очень очень очень очень горячий кофе	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	

Выводы

Использование слов-модификаторов в языке позволяет сделать речь, с одной стороны, более устойчивой к «помехам», с другой, менее однозначной. Это является значительной трудностью при переносе человеческих суждений в компьютерные базы знаний. Предложенный метод можно рассматривать как попытку формализации использования слов-модификаторов применительно к нечетким значениям лингвистических переменных в процессе создания и использования компьютерных баз знаний. В статье не рассматривается вопрос взаимодействия вновь полученных функций принадлежности с базовыми функциями в процессе нечеткого вывода, что является предметом отдельного исследования. Однако предложенный подход сделает нечеткий вывод более гибким, что повысит качество диагностических процедур.

Литература

1. Лелеков С.Г. Использование нечетких множеств в системах идентификации биологических объектов : [текст] / С.Г. Лелеков // Искусственный интеллект. Интеллектуальные системы : Материалы IX Международной н.-т. конф. – Донецк : ИПИИ «Наука і освіта» – 2008. – Т. 2. – С. 251-254.
2. Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH/ Леоненков А.В. – СПб. : БХВ – С. Петербург, 2003. – 736с.
3. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений / Заде Л.А. – М. : Мир, 1976. – 168 с.

Literatura

1. Lelekov S.G. Ispolzovanie nechetkih mnogestv v sistemah identifikacii biologicheskikh obektov [Tekst] / S.G. Lelekov // Iskusstvennyj intellekt. Intellektualnye sistemy: Materialy IX Megdunarodhoj n.-t. konf. – Donetsk: IPII «Nauka I osvita» – 2008. – T.2. – S. 251-254.
2. Leonenkov A.V. Nechetkoe modelirovanie v srede MATLAB i fuzzyTECH/ Leonenkov A.V. – SPb. : BHV – S. Peterburg, 2003. – 736 s.
3. Zadeh L.A. Ponyatie lingvisticheskoy peremennoj I ego primenenie k prinyatiyu pribligennih reshenii / L.A. Zadeh. – M. : Mir, 1976. –168 s.

RESUME

S.G. Lelekov

Method of Accounting for the Influence of Modifiers on the Fuzzy Values of Linguistic Variables

The article considers the problem of knowledge representation in the diagnostic expert systems. For example, in Hydrobiology are used taxonomic expert system for the determination of the species composition of the samples in the course of implementation of environmental monitoring. The knowledge of experts not always possible to provide precise quantitative values determinative characters. Therefore, there is a necessity of application of the linguistic variables, the values of which you can provide, using fuzzy sets. The rollover options of possible values of linguistic variables is practically impossible due to the rules recursiveness of their formation with the help of various kinds of words-type modifiers: very little, a little, too, etc. The article considers the methods automatic correction of the membership functions of fuzzy values of linguistic variables. For this purpose proposed the use 2 operations: discharge and Gaussian blur. The experimental data on the results of the use of the proposed method are presented. The results obtained can be used in diagnostic expert systems, based on the fuzzy logic.

Статья поступила в редакцию 02.04.2013.