УДК 519.816

ФОРМАЛИЗАЦИЯ НЕЧЕТКОЙ ИНФОРМАЦИИ ПРИ ОЦЕНИВАНИИ КАЧЕСТВЕННЫХ ПРИЗНАКОВ В ВЕРБАЛЬНЫХ ШКАЛАХ

Н.М. Кораблёв

ХНУРЭ, г. Харьков, Украина, korablev@kture.kharkov.ua

Рассматривается построение моделей экспертного оценивания качественных признаков на основе семантических пространств в вербальных шкалах, значениями которых являются слова, выражающие степень интенсивности проявления признаков. Эти модели могут применяться для формализации информации в рамках как качественных, так и количественных признаков.

ФУНКЦИЯ ПРИНАДЛЕЖНОСТИ, ВЕРБАЛЬНАЯ ШКАЛА, ТЕРМ-МНОЖЕСТВО, СЕМАНТИЧЕСКОЕ ПРОСТРАНСТВО, ПРИЗНАК

Введение

Важным этапом обработки информации является этап ее формализации, то есть этап представления этой информации в виде, позволяющем на следующих этапах ее обработки применять аппараты известных математических теорий. Известно, что нечеткая экспертная информация трудно формализуема при использовании традиционных подходов [1, 2]. Модельный подход на основе аппарата теории нечетких множеств позволил устранить недостатки традиционных формализаций нечеткой экспертной информации [1–4]. С точки зрения этого подхода моделями экспертного оценивания признаков служат семантические пространства (СП), термы которых соответствуют уровням вербальных шкал, используемых для оценивания признаков [5, 6]. Однако не все модели, построенные на основе СП, обладают свойствами, обеспечивающими успешность решения практических задач на основе этих моделей. Одним из таких свойств является свойство полноты, которое состоит в возможности описания каждого из элементов универсального множества в лингвистических термах этого пространства [4, 7, 8].

В [9] рассмотрены вопросы формализации нечеткой экспертной информации на основе СП путем опроса одного эксперта о типичных значениях термов и о разбиении универсального множества признака. Вместе с тем для оценивания качественных признаков и для описания количественных признаков используют вербальные шкалы, значениями которых являются слова, выражающие степень интенсивности проявления признаков [10]. При формализации значений количественных признаков с помощью вербальных шкал имеет место недостаток, состоящий в том, что при описании объектов с пограничными значениями показателя эксперт испытывает трудности в связи со скачкообразным переходом от одного значения к другому. Устранить этот недостаток позволяет аппарат теории нечетких множеств, с позиций которого вербальным уровням количественного признака в соответствие ставятся не четкие интервалы значений, а нечеткие множества.

При построении лингвистической шкалы для качественных признаков нельзя однозначно определить для них универсальные множества, как для количественных признаков. Поэтому разработка методов формализации нечеткой экспертной информации на основе СП в рамках вербальных шкал является актуальной задачей.

1. Постановка задачи

Пусть X – некоторое множество x с функцией принадлежности (ФП) $\mu_{\tilde{A}}(x)$ к некоторому множеству $\tilde{A}$ с терм-множеством $T(x)=\{X_1, X_2, ..., X_m\}$, где X_1 – терм, соответствующий минимальной интенсивности проявления признака, X_m – терм, соответствующий максимальной интенсивности проявления признака на универсальном множестве $U=[a,b]$. Предположим, что экспертом определены типичные для термов X_l, $l=\overline{1,m}$ интервалы (x_l^1, x_l^2), $l=\overline{1,m}$, то есть интервалы, для всех точек которых ФП соответствующих термов равны единице.

Необходимо построить ФП терм-множеств СП на основе апостериорной информации, полученной в результате оценивания экспертом качественного признака X у совокупности объектов, которые удовлетворяли бы требованиям к ФП $\mu_l(x)$, $l=\overline{1,m}$ их терм-множеств [4]:

1. Существует $\hat{U}_l \neq \emptyset$, где $\hat{U}_l = \{x \in U : \mu_l(x)=1\}$ есть точка или отрезок.

2. Если $\hat{U}_l = \{x \in U : \mu_l(x)=1\}$, тогда $\mu_l(x), l=\overline{1,m}$ не убывает слева от $\hat{U}_l$ и не возрастает справа от $\hat{U}_l$.

3. ФП $\mu_l(x), l=\overline{1,m}$ имеют не более двух точек разрыва первого рода.

4. Для каждого $x \in U$ существует $l, l=\overline{1,m} : \mu_l(x) \neq 0$.

5. Для каждого $x \in U$ $\sum_{l-1}^{m} \mu_l(x)=1$.

ФП, которые удовлетворяют этим требованиям, обладают свойством полноты [4].

Математической основой построения методов обработки и анализа нечеткой информации с использованием нечетких множеств является алгебра нечетких чисел. Поэтому рассмотрим сначала построение совокупности нечетких чисел, используемых для формализации лингвистических значений признаков.

2. Построение совокупности нечетких чисел для формализации лингвистических значений признаков

Известно [5], что нечеткое число $\tilde{A}$ с ФП $\mu_{\tilde{A}}(x)$ называется нормальным, если $\max\limits_{x}\mu_{\tilde{A}}(x)=1, x\in R$. Нечеткое число $\tilde{A}$ с ФП $\mu_{\tilde{A}}(x)$ называется унимодальным, если существует единственная точка $x\in R$, для которой выполняется равенство $\mu_{\tilde{A}}(x)=1$. Нечеткое число $\tilde{A}$ с ФП $\mu_{\tilde{A}}(x)$ называется многомодальным, если точка $x\in R$: $\mu_{\tilde{A}}(x)=1$ не является единственной, и толерантным, если существуют интервалы, для всех точек котороых выполняется равенство $\mu_{\tilde{A}}(x)=1$. Этот интервал называется интервалом толерантности нечеткого числа $\tilde{A}$.

Рассмотрим толерантные и унимодальные $(L-R)$-числа с ФП [10]

$$\mu_{\tilde{A}}(x)=\begin{cases} L\left(\dfrac{a_1-x}{a_L}\right), 0\le\dfrac{a_1-x}{a_L}\le 1, a_L>0;\\ R\left(\dfrac{x-a_2}{a_R}\right), 0\le\dfrac{x-a_2}{a_R}\le 1, a_R>0;\\ 1,\ \dfrac{a_1-x}{a_L}<0\cap\dfrac{x-a_2}{a_L}<0;\\ 0,\ \dfrac{a_1-x}{a_L}>1\cup\dfrac{x-a_2}{a_R}>1 \end{cases} \tag{1}$$

и следующими условиями на функции L и R:

1) $L(0)=R(0)=1$, $L(1)=R(1)=0$,

2) $L(x)$ и $R(x)$ – монотонно убывающие функции на множестве $[0,1]$.

Нечеткое число $\tilde{A}$ записывается в виде $\tilde{A}\equiv(a_1,a_2,a_L,a_R)$ (или $\mu_{\tilde{A}}(x)\equiv(a_1,a_2,a_L,a_R)$), где a_1,a_2,a_L,a_R являются параметрами толерантного (L-R)-числа $\tilde{A}$. Отрезок $[a_1,a_2]$ называется интервалом толерантности, а a_L и a_R – соответственно левым и правым коэффициентами нечеткости. Функция $L\left(\dfrac{a_1-x}{a_L}\right)$ является левой границей ФП толерантного (L-R)-числа, а функция $R\left(\dfrac{x-a_2}{a_R}\right)$ является правой границей ФП толерантного (L-R)-числа. При $a_L=0$ предполагается, что $L\left(\dfrac{a_1-x}{a_L}\right)=0$, при $a_R=0$ предполагается, что $R\left(\dfrac{x-a_2}{a_R}\right)=0$.

Унимодальное (L-R)-число $\tilde{A}$ имеет ФП толерантного (L-R)-числа при условии $a_1=a_2$. Символически унимодальное (L-R)-число $\tilde{A}$ записывается в виде $\tilde{A}\equiv(a_1,a_L,a_R)$.

Обозначим через Λ совокупность всех толерантных и унимодальных чисел с условиями на функции L и R. Элементы совокупности Λ будем называть Λ- числами, которые используем для построения ФП качественных признаков в вербальных шкалах.

3. Построение функций принадлежности качественных признаков в вербальных шкалах

Построим ФП терм-множеств СП на основе апостериорной информации, полученной в результате оценивания экспертом качественного признака X у совокупности объектов. Предполагается, что оценивание признака осуществлялось в рамках вербальной шкалы с уровнями $X_l, l=\overline{1,m}$, $m\ge 2$, упорядоченными по возрастанию интенсивности проявления признака. Уровни используемой вербальной шкалы однозначно задают терм-множество – $T(x)=\{X_1,X_2,...,X_m\}$. В качестве универсального множества СП признака X выбрано множество $U=[0,1]$. Точка $x=0$ соответствует полному отсутствию проявления признака X и поэтому считается типичной точкой терма X_1, точка $x=1$ соответствует полному присутствию проявления признака X и поэтому считается типичной точкой терма X_m.

Обозначим относительные частоты появления объектов, у которых интенсивность признака X оценена уровнями $X_l, l=\overline{1,m}$, соответственно через $a_l,\ l=\overline{1,m},\ \sum\limits_{l=1}^{m}a_l=1$. Будем предполагать, что нечеткие числа, соответствующие термам $X_l, l=\overline{1,m}$ с ФП $\mu_l(x), l=\overline{1,m}$, принадлежат совокупности Λ и удовлетворяют дополнительному условию: если $L(x)$, $R(x)$ – нелинейные, то они имеют центральную симметрию относительно точки перегиба.

Построение ФП терм-множеств СП будет осуществляться таким образом, чтобы обеспечивалось выполнение требований, предъявляемых к этим функциям в рамках определения СП, и чтобы площади фигур, ограниченных графиками функций $\mu_l(x), l=\overline{1,m}$ и осью абсцисс, равнялись $a_l, l=\overline{1,m}$. Построение необходимо ограничить логичными требованиями на области нечеткости между соседними термами (или параметры нечеткости нечетких чисел, соответствующих термам). С одной стороны, эту область желательно сделать как можно меньше, тогда соответственно будет меньше нечеткость построенной в виде СП модели. С другой стороны, необходимо опираться на содержательный смысл области нечеткости, поэтому предлагается мощность (длину) этой области для крайних термов вычислять как $\min(a_1,a_2)$ или соответственно $\min(a_{m-1},a_m)$, а для средних термов вычислять, исходя из соотношений между числами

a_{l-1}, a_l, a_{l+1}, $l=\overline{3, m-2}$. Графики построенных ФП будут представлять собой криволинейные трапеции со средними линиями, равными a_l, $l=\overline{1, m}$.

Построим функцию принадлежности терма X_m:

1. Если $a_m \le a_{m-1}$, то

$$\mu_m(x)=\begin{cases} 0, & 0 \le x \le 1-\dfrac{3a_m}{2}; \\ L\left(\dfrac{1-\dfrac{a_m}{2}-x}{a_m}\right), & 1-\dfrac{3a_m}{2} < x \le 1-\dfrac{a_m}{2}; \\ 1, & 1-\dfrac{a_m}{2} < x \le 1. \end{cases} \tag{2}$$

На рис. 1 и 2 изображены функции принадлежности терма X_m для частного случая: $L(x)=1-x,\ 0\le x\le 1$ – функции принадлежности T-чисел.

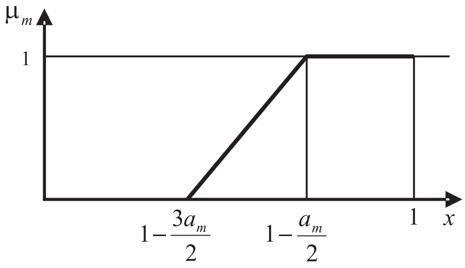

Рис. 1. Функция принадлежности (2)

2. Если $a_m > a_{m-1}$, то

$$\mu_m(x)=\begin{cases} 0, & 0 \le x \le 1-a_m-\dfrac{a_{m-1}}{2}; \\ L\left(\dfrac{1-a_m+\dfrac{a_{m-1}}{2}-x}{a_{m-1}}\right), & \\ & 1-a_m-\dfrac{a_{m-1}}{2} < x \le 1-a_m+\dfrac{a_{m-1}}{2}; \\ 1, & 1-a_m+\dfrac{a_{m-1}}{2} < x \le 1. \end{cases} \tag{3}$$

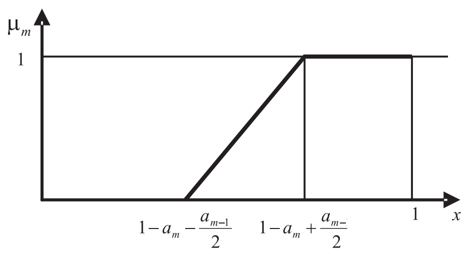

Рис. 2. Функция принадлежности (3)

Построим функцию принадлежности терма X_{m-1}:

1. Если $a_{m-1} \ge \max(a_m, a_{m-2})$, то

$$\mu_{m-1}(x)=\begin{cases} 0, & 0 \le x \le \sum\limits_{l=1}^{m-3} a_l+\dfrac{a_{m-2}}{2}; \\ R\left(\dfrac{\sum\limits_{l=1}^{m-2} a_l+\dfrac{a_{m-2}}{2}-x}{a_{m-2}}\right), & \\ & \sum\limits_{l=1}^{m-3} a_l+\dfrac{a_{m-2}}{2} < x \le \sum\limits_{l=1}^{m-2} a_l+\dfrac{a_{m-2}}{2}; \\ 1, & \sum\limits_{l=1}^{m-2} a_l+\dfrac{a_{m-2}}{2} < x \le 1-\dfrac{3a_m}{2}; \\ L\left(\dfrac{x-1+\dfrac{3a_m}{2}}{a_m}\right), & 1-\dfrac{3a_m}{2} < x \le 1-\dfrac{a_m}{2}; \\ 0, & 1-\dfrac{a_m}{2} < x \le 1. \end{cases} \tag{4}$$

На рис. 3–6 изображены функции принадлежности терма X_{m-1} для частного случая: $L(x)=1-x,\ R(x)=1-x,\ 0\le x\le 1$.

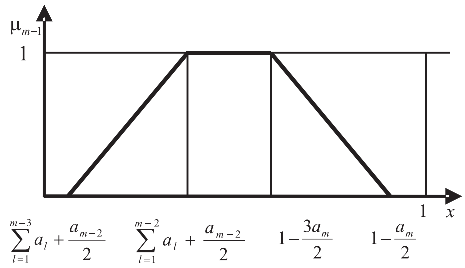

Рис. 3. Функция принадлежности (4)

2. Если $a_m < a_{m-1} < a_{m-2}$, то

$$\mu_{m-1}(x)=\begin{cases} 0, & 0 \le x \le 1-a_m-\dfrac{3a_{m-1}}{2}; \\ R\left(\dfrac{1-a_m-\dfrac{a_{m-1}}{2}-x}{a_{m-1}}\right), & \\ & 1-a_m-\dfrac{3a_{m-1}}{2} < x \le 1-a_m-\dfrac{a_{m-1}}{2}; \\ 1, & 1-a_m-\dfrac{a_{m-1}}{2} < x \le 1-\dfrac{3a_m}{2}; \\ L\left(\dfrac{x-1+\dfrac{3a_m}{2}}{a_m}\right), & 1-\dfrac{3a_m}{2} < x \le 1-\dfrac{a_m}{2}; \\ 0, & 1-\dfrac{a_m}{2} < x \le 1. \end{cases} \tag{5}$$

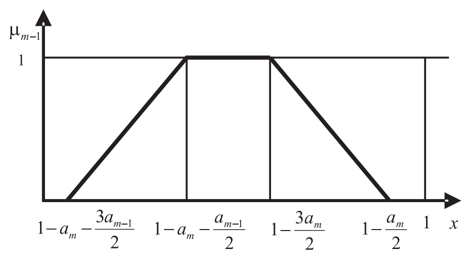

Рис. 4. Функция принадлежности (5)

3. Если $a_{m-2} < a_{m-1} < a_m$, то

$$\mu_{m-1}(x)=\begin{cases} 0, & 0 \le x \le \sum\limits_{l=1}^{m-3} a_l + \dfrac{a_{m-2}}{2}; \\ R\left(\dfrac{\sum\limits_{l=1}^{m-2} a_l + \dfrac{a_{m-2}}{2} - x}{a_{m-2}}\right), & \sum\limits_{l=1}^{m-3} a_l + \dfrac{a_{m-2}}{2} < x \le \sum\limits_{l=1}^{m-2} a_l + \dfrac{a_{m-2}}{2}; \\ 1, & \sum\limits_{l=1}^{m-2} a_l + \dfrac{a_{m-2}}{2} < x \le 1 - a_m - \dfrac{a_{m-1}}{2}; \\ L\left(\dfrac{x - 1 + a_m + \dfrac{a_{m-1}}{2}}{a_{m-1}}\right), & 1 - a_m - \dfrac{a_{m-1}}{2} < x \le 1 - a_m + \dfrac{a_{m-1}}{2}; \\ 0, & 1 - a_m + \dfrac{a_{m-1}}{2} < x \le 1. \end{cases} \qquad (6)$$

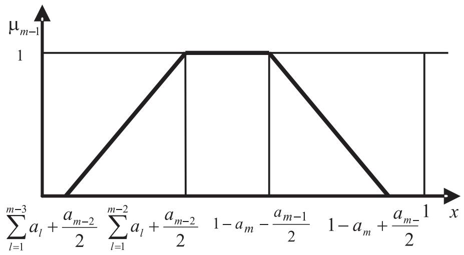

Рис. 5. Функция принадлежности (6)

4. Если $a_{m-1} \le \min(a_m, a_{m-2})$, то

$$\mu_{m-1}(x)=\begin{cases} 0, & 0 \le x \le 1 - a_m - \dfrac{3a_{m-1}}{2}; \\ R\left(\dfrac{1 - a_m - \dfrac{a_{m-1}}{2} - x}{a_{m-1}}\right), & 1 - a_m - \dfrac{3a_{m-1}}{2} < x \le 1 - a_m - \dfrac{a_{m-1}}{2}; \\ L\left(\dfrac{x - 1 + a_m + \dfrac{a_{m-1}}{2}}{a_{m-1}}\right), & 1 - a_m - \dfrac{a_{m-1}}{2} < x \le 1 - a_m + \dfrac{a_{m-1}}{2}; \\ 0, & 1 - a_m + \dfrac{a_{m-1}}{2} < x \le 1. \end{cases} \qquad (7)$$

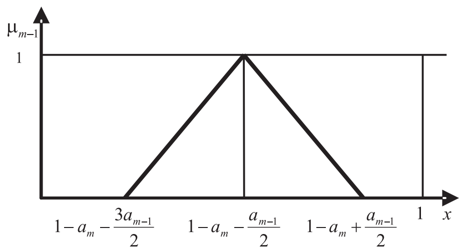

Рис. 6. Функция принадлежности (7)

Аналогично ФП $\mu_{m-1}(x)$ строятся ФП $\mu_l(x), l = \overline{2, m-2}$.

Построим ФП для терма X_l при условии четного числа термов:

1. Если $a_1 \le a_2$, то

$$\mu_1(x)=\begin{cases} 1, & 0 \le x \le \dfrac{a_1}{2}; \\ L\left(\dfrac{x - \dfrac{a_1}{2}}{a_1}\right), & \dfrac{a_1}{2} < x \le \dfrac{3a_1}{2}; \\ 0, & \dfrac{3a_1}{2} < x \le 1. \end{cases} \qquad (8)$$

На рис. 7 и 8 изображены ФП для частного случая: $L(x) = 1 - x,\ 0 \le x \le 1$.

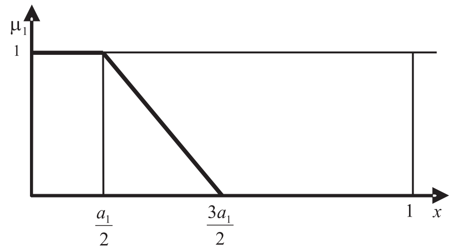

Рис. 7. Функция принадлежности (8)

2. Если $a_1 > a_2$, то

$$\mu_1(x)=\begin{cases} 1, & 0 \le x \le a_1 - \dfrac{a_2}{2}; \\ L\left(\dfrac{x - a_1 + \dfrac{a_2}{2}}{a_2}\right), & a_1 - \dfrac{a_2}{2} < x \le a_1 + \dfrac{a_2}{2}; \\ 0, & a_1 + \dfrac{a_2}{2} < x \le 1. \end{cases} \qquad (9)$$

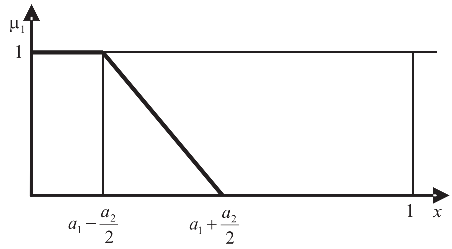

Рис. 8. Функция принадлежности (9)

При нечетком числе термов получаем:

1. Если $a_1 \le a_2$, то

$$\mu_1(x)=\begin{cases} 1, & 0 \le x \le \dfrac{a_1}{2}; \\ R\left(\dfrac{x - \dfrac{a_1}{2}}{a_1}\right), & \dfrac{a_1}{2} < x \le \dfrac{3a_1}{2}; \\ 0, & \dfrac{3a_1}{2} < x \le 1. \end{cases} \qquad (10)$$

2. Если $a_1 > a_2$, то

$$\mu_1(x)=\begin{cases} 1, & 0 \le x \le a_1 - \dfrac{a_1}{2}; \\ R\left(\dfrac{x-a_1+\dfrac{a_2}{2}}{a_2}\right), & a_1 - \dfrac{a_2}{2} < x \le a_1 + \dfrac{a_2}{2}; \\ 0, & a_1 + \dfrac{a_2}{2} < x \le 1. \end{cases} \quad (11)$$

Вид ФП (10) и (11) совпадает соответственно с видом ФП (8) и (9), изображенных на рис. 7 и 8.

Примеры некоторых нелинейных функций $L(x)$ и $R(x)$, которые могут использоваться при построении ФП СП рассмотренным методом, имеют вид:

$$1.\ L(x)=\begin{cases} 2(x-\frac{1}{2})^2+\frac{1}{2}, & 0 \le x \le \frac{1}{2}; \\ -2(x-\frac{1}{2})^2+\frac{1}{2}, & \frac{1}{2} < x \le 1. \end{cases}$$

$$2.\ R(x)=\begin{cases} \sqrt{\frac{1}{4}-\frac{x}{2}}+\frac{1}{2}, & 0 \le x \le \frac{1}{2}; \\ \sqrt{\frac{x}{2}-\frac{1}{4}}+\frac{1}{2}, & \frac{1}{2} < x \le 1. \end{cases}$$

Предложенный метод построения ФП терм-множеств СП можно применять не только в условиях апостериорной информации, представленной к обработке в настоящее время. Эксперт может строить СП в отсутствие такой информации в настоящий момент, пользуясь информацией, которой он обладал раньше на основании своего опыта. Этот метод может применяться для построения ФП терм-множеств СП, формализующего информацию по оцениванию некоторого объекта несколькими экспертами. Задачу в такой постановке можно считать двойственной задачей по отношению к задаче оценивания одним экспертом некоторой совокупности объектов. ФП термов СП по данному методу могут быть получены даже в том случае, когда невозможно привлечение экспертов для проведения стандартных процедур парных сравнений результатов [2] с целью получения значений этих функций.

Выводы

Разработан метод формализации нечеткой информации, полученной в результате оценивания качественных признаков в вербальных шкалах. Метод имеет существенное достоинство, состоящее в том, что построение ФП формализованной информации может осуществляться в условиях большого числа оцениваемых объектов, и для этого построения не требуется никакой дополнительной информации (полученной от экспертов, проводящих стандартные парные сравнения объектов друг с другом). Разработанный метод может применяться для построения ФП моделей экспертного оценивания не только в рамках информации, полученной непосредственно после проведения оценочных процедур, но также на основе информации из предыдущего опыта их проведения.

Список литературы: 1. *Андрейчиков А.В., Андрейчикова О.Н.* Интеллектуальные информационные системы. – М.: Финансы и статистика, 2004. – 424 с. 2. *Борисов А.Н, Крумберг О.А., Федоров И.П.* Принятие решений на основе нечетких моделей: Примеры использования. – Рига.: Зинатне, 1990. – 184 с. 3. *Прикладные нечеткие системы / Под ред. Т. Тэрано, К. Асни, М. Сугэно:* Пер. с япон. – М.: Мир, 1993. – 368 с. 4. *Рыжов А.П.* Элементы теории нечетких множеств и измерения нечеткости – М.: Диалог-МГУ, 1998. – 116 с. 5. *Аверкин А.Н., Батыршин И.З., Блишун А.Ф.и др.* Нечеткие множества в моделях управления и искусственного интеллекта. – М.: Наука. Гл. ред. физ-мат. лит., 1986. – 312 с. 6. *Ежкова И.В.* Семантически – инвариантная формализация лингвистических оценок // В кн. Семиотические аспекты формализации интеллектуальной деятельности. – М.: МДНТП, 1983. – С.48-51. 7. *Полещук О.М.* Методы представления экспертной информации в виде совокупности терм-множеств полных ортогональных семантических пространств // Вестник МГУЛ. – 2002. – № 5 (25). – С. 198–216. 8. *Полещук О.М.* О развитии систем обработки нечеткой информации на базе полных ортогональных семантических пространств» // Вестник МГУЛ. – 2003. №1 (26). – С. 112–117. 9. *Кораблев Н.М.* Формализация нечеткой информации на основе опроса одного эксперта // Системи обробки інформації. – 2007. – Вип. 9 (67). – С. 20–23. 10. *Литвак Б.Г.* Экспертные оценки и принятие решений. – М.: Патент, 1996. – 271 с.

Поступила в редколлегию 10.04.2008