УДК 519.7

ИССЛЕДОВАНИЕ ОСОБЕННОСТЕЙ НЕЧЕТКИХ ПРОСТРАНСТВ ТОЛЕРАНТНОСТИ ДЛЯ МОДЕЛИРОВАНИЯ ПОНЯТИЙНЫХ ЗНАНИЙ НА ОСНОВЕ ГИПЕРТЕКСТОВЫХ СТРУКТУР

ПАВЛОВ П.Ф., СОЛОВЬЕВА Е.А.

В целях моделирования понятийных знаний на основе гипертекстовых структур анализируются особенности нечетких пространств толерантности, вводятся понятия, позволяющие описывать такие пространства.

При моделировании системы понятийных знаний на основе гипертекстовых структур для учета ее динамики, нечетко проявленных связей и т.д. используется аппарат нечетких множеств (подмножеств), предложенных Л.А. Заде. Гипертекстовые структуры давно применяются как основа представления знаний в справочниках и энциклопедиях. Одним из наиболее признанных определений гипертекста является определение Субботина [4]: "Гипертекст – это форма организации текстового материала, при которой его единицы представлены не в линейной последовательности, а как система явно указанных возможных переходов, связей между ними". Основываясь на этом определении, введем понятие гипертекстовой системы как тройки (тривиальная гипертекстовая система):

$$ГС = < V, R, Br >,$$

где V – множество вершин; R – множество направленных дуг; Br – механизм навигации (браузер). Элементами множества вершин ($v \in V$) являются линейно-организованные тексты. Элементы множества направленных дуг соответствуют тем связям, которые установил явным образом разработчик (создатель, генератор) гипертекста. Дуга $r_{i,j} \in R$ вида $r_{i,j} = \langle v_i, v_j \rangle$ означает, что вершина $v_i \in V$ ссылается на вершину $v_j \in V$. Механизм навигации позволяет перемещаться от вершины к вершине по существующим связям.

Для моделирования нечеткости дуг, которая характерна для реальных гипертекстовых систем, исследуем особенности нечетких пространств толерантности.

Пусть A и B – произвольные непустые множества; $\underline{R}$ – нечеткое бинарное отношение между элементами множеств A и B с множеством принадлежностей $M = [0, 1]$ и функцией принадлежности

$$\mu_{\underline{r}}: A, B \to M.$$

Элементы множества M будем называть степенями принадлежности. Принадлежность со степенью α пары (x, y) нечеткому бинарному отношению $\underline{\mathrm{R}}$ будем обозначать:

$$(x,y) \underset{\alpha}{\in} \underline{R},\ \alpha = \mu_{\underline{r}}(x,y).$$

Нечеткое бинарное отношение $\underline{\mathrm{R}}$, определенное на множестве А, называется рефлексивным, если

$$\forall (x,x) \in A \mid A\left[\mu_{\underline{r}}(x,x)=1\right].$$

Нечеткое бинарное отношение $\underline{\mathrm{R}}$, определенное на множестве А, называется симметричным, если

$$\forall (x,y) \in A \mid A\left[\mu_{\underline{r}}(x,y)=\alpha \to \mu_{\underline{r}}(y,x)=\alpha\right].$$

Нечетким отношением толерантности, или сходства [1], называется рефлексивное и симметричное нечеткое бинарное отношение.

Нечеткое бинарное отношение $\underline{\mathrm{R}}$, определенное на множестве А, называется транзитивным, если

$$\forall\, (x,y),(y,z),(x,z) \in A \times A,$$

$$\left[\mu_{\underline{r}}(x,z) >= \max_{y}\left[\min\left(\mu_{\underline{r}}(x,y), \mu_{\underline{r}}(y,z)\right)\right]\right].$$

Нечетким отношением эквивалентности [2, с.110] или отношением подобия [1, с.107] называется нечеткое бинарное отношение, обладающее свойствами: рефлексивности, симметричности и транзитивности.

По аналогии с четкими отношениями [7], пару $\langle A, \underline{T} \rangle$ (где $\underline{T}$ – нечеткое отношение толерантности, определенное на множестве А) будем называть нечетким пространством толерантности.

Как и в случае четких отношений, всякая нечеткая эквивалентность является, очевидно, одновременно и нечеткой толерантностью.

Рассмотрим произвольное нечеткое пространство толерантности $\langle A, \underline{T} \rangle$. Выделим в его носителе множества $S \subseteq A$ такие, что

$$\forall x, y \in S\left[\mu_{\underline{t}}(x, y) > 0\right].$$

По аналогии с четким случаем (см., например, [5, с.93]) такие множества будем называть предклассами нечеткого пространства толерантности $\langle A, \underline{T} \rangle$.

Пусть С – некоторый предкласс в $\langle A, \underline{T} \rangle$. Тогда для любых $x, y \in C$

$$\mu_{\underline{t}}(x, y) = \alpha \in M,$$

т.е. степень принадлежности α является функцией двух переменных.

Предкласс C в $\langle A, \underline{T} \rangle$ такой, что

$$\forall\, x,y \in C\lfloor \mu_{\underline{t}}(x,y)=\alpha \rfloor,$$

где α – некоторое фиксированное число из M=[0,1], будем называть предклассом постоянной степени. Максимальный по включению предкласс $\widehat{c}$ постоянной степени такой, что $\langle A,\underline{T}\rangle$

$$\forall x,y \quad \widehat{c}\lfloor \mu_{\underline{t}}(x,y)=\alpha \rfloor,$$

будем называть предклассом степени α.

Как и в случае четких отношений [5, с.93], максимальные по включению предклассы будем называть классами, т.е. предкласс K является классом тогда и только тогда, когда

$$\forall y \in A \setminus K\,,\ \forall x \in K\lfloor \mu_{\underline{t}}(x,y)=0 \rfloor.$$

Рассмотрим множества классов $\{K_1,K_2,...,K_j\}$ в некотором нечетком пространстве толерантности $\langle A,\underline{T}\rangle$. Ядром, порожденным нечеткими классами $K_1,K_2,...,K_j$, будем называть множество, которое состоит из всех элементов x, входящих одновременно во все эти классы и только в них. Это ядро в дальнейшем будем обозначать $R(K_1,K_2,...,K_j)$. (Аналогичное четкое понятие вводится, например, в книге [5, с.101]).

Очевидно следующее включение:

$$R(K_1,K_2,...,K_j) \subseteq \bigcap_{i=1}^{j} K_i\,.$$

Подмножества ядер будем называть предъядрами. По аналогии с предклассами будем рассматривать предъядра степени α, $\alpha \in M$.

Введем в рассмотрение нечеткое бинарное отношение $\underline{\theta}$, определенное на множестве A следующим образом:

$$\mu_{\underline{\theta}}(x,y)>0,\ \ (x,y\in A),$$

если для всякого $z \in A$ из $\mu_{\underline{t}}(z,x)>0$ следует, что $\mu_{\underline{t}}(z,y)>0$, и наоборот, если $\mu_{\underline{t}}(z,y)>0$, то $\mu_{\underline{t}}(z,x)>0$. Очевидно, $\underline{\theta}$ рефлексивно, симметрично и транзитивно, т.е. является нечетким отношением эквивалентности на A.

Пусть $x \in A$. Ядром $R(x)$ элемента x назовем совокупность всех $y \in A$ таких, что $(x,y)\in\underline{\theta}$. Таким образом, ядра элементов являются классами эквивалентности $\underline{\theta}$ и образуют покрытие множества A.

Очевидна следующая лемма.

Лемма 1. Если $\mu_{\underline{\theta}}(x,y)>0$, то $\mu_{\underline{t}}(x,y)>0$.

Лемма 2. Пусть $\sum=\{\delta\}$ – вполне упорядоченное множество, элементам которого сопоставлены предклассы E_δ некоторого нечеткого пространства толерантности так, что из $\delta<\delta_1$ следует $E_\delta \subseteq E_{\delta_1}$. Тогда объединение $\bigcup\limits_{\Sigma} E_\delta = E$ является предклассом в том же нечетком пространстве толерантности.

Доказательство. Пусть $x \in E$ и $y \in E$. Обозначим через $\sum(x)$ подмножество в е такое, что при $\delta \in \sum(x)$ имеет место $x \in E_\delta$. Обозначим через $\delta(x)$ первый элемент множества $\Sigma(x)$ (ибо по теореме Цермело любое множество можно вполне упорядочить). Пусть для определенности $\delta(e)\le\delta(x)$. Поскольку $y \in E_{\delta(y)}$, а $E_{\delta(y)} \in E_{\delta(x)}$, то $y \in E_{\delta(x)}$. Следовательно, x и y входят в общий предкласс и значит, лемма 2 доказана.

Теорема 1. *Пусть* $<A,\underline{T}>$ *– нечеткое пространство толерантности. Тогда для всякого* $x \in A$ *существует класс нечеткой толерантности, содержащий* x.

Доказательство. Проведем трансфинитную индукцию. Первый шаг. Пусть $C_1=\{x\}$. Очевидно, C_1 является предклассом. Если не существует $y \ne x, y \in A$ таких, что $\mu_t(x,y)>0$, то C_1 – класс. Если такой $y \in A$ существует, то перейдем к следующему шагу.

Последующие шаги индукции определяем следующим образом. Если α является трансфинитом первого рода и $C_{\alpha-1}$ определено, то либо $C_{\alpha-1}$ уже есть класс, либо существует $y \notin C_{\alpha-1}$ такой, что

$$\forall x \in C_{\alpha-1}[\mu_t(x,y)>0].$$

Тогда будем полагать $C_\alpha = C_{\alpha-1}\cup\{y\}$.

Если α – трансфинит второго рода и $C_{\alpha'}$ определены при всех $\alpha'<\alpha$, то полагаем $C_\alpha=\bigcup\limits_{\alpha'<\alpha} C_{\alpha'}$.

В силу леммы 2 C_α является предклассом в $<A,\underline{T}>$. Для завершения доказательства достаточно положить $K=\bigcup\limits_{\alpha} C_\alpha$, где объединение берется по всем трансфинитам α, для которых C_α определено. Очевидно, K является максимальным предклассом, содержащим x, т.е. классом. Теорема 1 доказана.

Следствие 1. Для того чтобы выполнялось $\mu_{\underline{t}}(x,y)>0$, $x,y\in A$, необходимо и достаточно существование класса нечеткой толерантности K, одновременно содержащего x и y.

Следствие 2. Для всякого предкласса некоторого нечеткого пространства толерантности существует содержащий его класс.

Замечание. Аналог теоремы 1 для нечетких пространств толерантности доказан для случая конечного носителя в работе [7], для случая носителя произвольной мощности – в [6].

Сопоставим всякому $x\in A$ множество $H(x)$ содержащих его классов нечеткой толерантности.

Теорема 2. *Для совпадения множеств* $H(x)$ *и* $H(y)$*, где* $x,y\in A$*, необходимо и достаточно, чтобы* $\mu_{\underline{\theta}}(x,y)>0$.

Доказательство. *Необходимость.* Пусть $H(x)=H(y)$, и для некоторого $z\in A$ $\mu_{\underline{t}}(x,z)>0$. Тогда, согласно следствию 1, существует класс нечеткой толерантности K, содержащий одновременно x и z. Ясно, что $K\in H(x)$. Но тогда $K\in H(y)$, следовательно, $\mu_{\underline{t}}(z,y)>0$. Аналогично доказывается, что из $\mu_{\underline{t}}(z,y)>0$ следует $\mu_{\underline{t}}(z,x)>0$. Таким образом, $\mu_{\underline{\theta}}(x,y)>0$.

Достаточность. Пусть $\mu_{\underline{\theta}}(x,y)>0$ и K – класс нечеткой толерантности, содержащий x. Ввиду того, что

$$\forall z\in K\left[\mu_{\underline{t}}(z,x)>0\right],$$

получаем, что

$$\forall z\in K\left[\mu_{\underline{t}}(z,y)>0\right].$$

Поэтому $y\in K$. Таким образом, любой класс $K\in H(x)$ одновременно входит в $H(y)$, и наоборот. Следовательно, $H(x)=H(y)$. Теорема 2 доказана.

Следствие 3. Для любого $x\in A$ имеет место равенство $R(x)=R(H(x))$.

Пусть a – некоторое действительное число. Тогда через {a} будем обозначать дробную часть a. Введем в рассмотрение нечеткое пространство толерантности с множеством принадлежностей [0,1]:

$$\underline{T}_{p,q,r}=<M_{p,q,r};\underline{\tau}>;\quad p,q,r\in N;$$

носителем которого является множество $M_{p,q,r}$, состоящее из всевозможных наборов вида

$$x=\{v_0,v_1,...,v_k\},$$

где для всех i=1,..,k все v_i попарно-различны и

$$v_i[1,q+1],\ k\in[1,r],\ v_0=1,..,(p+1).$$

Нечеткое отношение толерантности $\underline{\tau}$ вводится следующим образом:

Пусть

$$x=\{v_0(x);\ v_1(x),...,v_k(x)\};$$

$$y=\{v_0(y);\ v_1(y),...,v_k(y)\}.$$

Тогда $\mu_{\underline{t}}(x,y)>0$ в том и только в том случае, если существуют i=1,…,k; j=1,…,l такие, что $v_i(x)=v_j(y)$, причем

$$\mu_{\underline{t}}(x,y)=\begin{cases}\{v_i(x)\}, & если\ \ v_i(x)\notin z,\\ v_i(x)-1, & если\ \ v_i(x)\in z.\end{cases}$$

Пусть A – нечеткое конечное множество. Имеет место теорема о каноническом представлении нечетких пространств толерантности, являющаяся аналогом четкого случая (см. [7]).

Теорема 3. *Для любого нечеткого пространства толерантности* $<A,\underline{T}>$ *существует отображение*

$$\varphi:\ <A,\underline{T}>\to\underline{T}_{p,q,r},$$

являющееся изоморфным вложением для некоторых p *и* q.

Доказательство. Рассмотрим множество всех классов нечеткой толерантности $<A,\underline{T}>$. Пусть количество этих классов равно q. Тогда данные классы можно перенумеровать некоторым числом от 1 до q. Всякому $x\in A$ сопоставим следующий набор:

$$x\to\{v_0(x);v_1(x),...,v_k(x)\},$$

где для всякого i=1,...,k

$$v_i(x)=n_s(x)+w_t(x),$$

здесь $\{n_s(x)\}$ – номера классов, в которые входит x; $\{w_t(x)\}$ – степени принадлежности элемента x во всех предклассах постоянной степени соответствующих классов.

Теперь рассмотрим ядро $R(K_{n_1},...,K_{n_k})$, а число $v_0(x)$ однозначно выделяет x среди остальных элементов множества. Следовательно, отображение φ взаимно-однозначно.

Пусть

$$x=\{v_0(x);\ v_1(x),...,v_k(x)\};$$

$$y=\{v_0(y);\ v_1(y),...,v_k(y)\}.$$

Тогда если для некоторых i,j имеет место $v_i(x)=v_j(y)$, то существует общий класс нечеткой толерантности, содержащий оба данных элемента и, следовательно, $\mu_{\underline{t}}(x,y)>0$, причем

$$\mu_{\underline{t}}(x,y)=\begin{cases}\{v_i(x)\}, & если\ \ v_i(x)\notin z,\\ v_i(x)-1, & если\ \ v_i(x)\in z.\end{cases}$$

Наоборот, если $\mu_{\underline{t}}(x,y) > 0$, то существует общий класс нечеткой толерантности, который их содержит, поэтому существуют $v_i(x)$, $v_j(y)$ такие, что

$$v_i(x) = v_j(y).$$

Таким образом, если наборы

$$\{v_0(x); v_1(x),..., v_k(x)\}$$

рассматривать как элементы $\underline{T}_{p,q,r}$, то из толерантности элементов в $< A, \underline{T} >$ следует толерантность их образов в $\underline{T}_{p,q,r}$, и наоборот. Теорема 3 доказана.

По аналогии с четким случаем (см. [Вагнер, 1965]), нечеткое бинарное отношение $\underline{r}$, определенное на паре непустых множеств A и B, будем называть вполне эффективным тогда и только тогда, если

1) $\forall x \in A \;\; \exists y \in B \;\; [\mu_{\underline{r}}(x,y) > 0]$;

2) $\forall y \in B \;\; \exists x \in A \;\; [\mu_{\underline{r}}(x,y) > 0]$.

В противном случае нечеткое бинарное отношение будем называть частичным.

Пусть $< A, r_a >, < B, r_b >$ – модели типа <2> [3, с.46]. Нечеткое отображение

$$f: < A, r_a > \; < B, r_b >$$

будем называть нечеткой корреспонденцией, если

$$\forall x, y \in B\,[(x,y) \in R_b \supset [f^{-1}(x), f^{-1}(y)] \in R_a].$$

Нечеткий гоморфизм, являющийся одновременно корреспонденцией, будем называть нечетким k-гоморфизмом.

Вполне эффективное нечеткое бинарное отношение $g\underline{p}$ на паре непустых множеств A и B будем называть нечеткой обобщенной дифункциональностью, если существуют пространства нечетких толерантностей $< A, \underline{T_a} >$ и $< B, \underline{T_b} >$, а также нечеткий многозначный k-гоморфизм

$$f: < A, \underline{T_a} > \to < B, \underline{T_b} >$$

такие, что $\mu_{g\underline{r}}(x,y) > 0$ тогда и только тогда, когда $\mu_{\underline{f}}(x,y) > 0$.

Вполне эффективное нечеткое бинарное отношение $g\underline{p}$ на паре непустых множеств A и B будем называть нечетким дифункциональным отношением, если существуют пространтва нечетких эквивалентностей $< A, \underline{E_a} >$ и $< B, \underline{E_b} >$, а также нечеткий многозначный k-гоморфизм

$$f: A \setminus \underline{E_a} \to B \setminus \underline{E_b},$$

индуцирующий нечеткую биекцию фактормножеств

$$\underline{\hat{f}}: A \setminus \underline{E_a} \to B \setminus \underline{E_b},$$

(о фактормножествах по нечетким эквивалентностям см. [2, с.113]) такие, что

$$\forall x \in A, y \in B \left\lfloor \mu_{\underline{d}}(x,y) > 0 \leftrightarrow \mu_{\underline{f}}(x,y) > 0 \right\rfloor.$$

Из теоремы 3 с учетом приведенных выше определений получаем

Следствие 4. Пусть $g\underline{p}$ – нечеткое обобщенное дифункциональное отношение на паре непустых множеств A и B. Тогда существуют числа $p, q, p', q' \in N$ и гоморфизм

$$\varphi: < AB, g\underline{p} > \qquad < M_{p,q,q} M_{p',q',q'}; \underline{f} >$$

(где $\underline{f}$ – нечеткий многозначный k-гоморфизм из $\underline{T}_{p,q,q}$ в $\underline{T}_{p',q',q'}$) является изоморфным вложением моделей типа <1>.

Замечание. Так как всякое нечеткое обобщенное дифункциональное отношение является нечетким дифункциональным отношением, то описанное в следствии 4 представление имеет место и для нечетких дифункциональных отношений.

Таким образом, доказана теорема о каноническом представлении нечетких пространств толерантности. Данная теорема позволяет описать с точностью до изоморфизма модели нечетких дифункциональных отношений, которые используются при моделировании нечетких отношений между узлами гипертекста, используемого как метод представления знаний в естественноязыковой форме.

Литература: 1. *Кофман А.* Введение в теорию нечетких множеств. М.: Радио и связь, 1982. 432 с. 2. *Кузьмин В.Б.* Построение групповых решений в пространствах четких и нечетких бинарных отношений. М.: Наука, 1982. 168 с. 3. *Мальцев А.И.* Алгебраические системы. М.: Наука, 1970. 392 с. 4. *Субботин М.М.* Новая информационная технология: создание и обработка гипертекстов // НТИ. Сер. 2. 1988. N 5. С. 2-7. 5. *Шрейдер Ю.А.* Равенство, сходство, порядок. М.: Наука, 1971. 256 с. 6. *Шрейдер Ю.А.* Пространства толерантности // Кибернетика. 1976. №2. С. 42-48. 7. *Якубович С.М.* Аксиоматическая теория сходства // НТИ.Сер.2. М.: ВИНИТИ, 1968. №10. С. 15-19.

Поступила в редколлегию 27.04.2001

Рецензент: д-р техн. наук, проф. Шабанов-Кушнаренко С.Ю.

Павлов Павел Федорович, старший преподаватель кафедры программного обеспечения ЭВМ, старший научный сотрудник научно-учебной лаборатории приобретения знаний ХНУРЭ. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 40-95-91,

pavlo_pavlov@mail.ru.

Соловьева Екатерина Александровна, д-р техн. наук, профессор кафедры программного обеспечения ЭВМ, заведующая научно-учебной лабораторией приобретения знаний ХНУРЭ. Научные интересы: все, что связано с познанием сущности мира и человека – системология, моделирование знаний, когнитология, теория классификации, искусственный интеллект. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 40-95-91, 47-71-85.