УДК 519.711.3:519.68

СВЯЗЬ ОТНОШЕНИЙ ТОЛЕРАНТНОСТИ И ЭКВИВАЛЕНТНОСТИ В МЕТАТЕКСТОВЫХ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИХ СИСТЕМАХ

ЛУГАНСКИЙ А.М., МАШТАЛИР В.П., ШЛЯХОВ В.В.

Рассматриваются вопросы моделирования связи метатекстовых конструкций в интерактивных информационных системах. Внимание фокусируется на трансформации композиций отношений толерантности в эквивалентность. Формулируются и доказываются необходимые и достаточные условия продуцирования отношения эквивалентности пересечением произвольного числа отношений толерантности.

Введение

В концептуальном и прикладном аспектах парадигма метатекстового анализа и синтеза информационных ресурсов играет достаточно заметную роль в развитии интерактивного и коммуникативного сервиса. Трактуя метатекст как ориентированную на пользователя устойчивую систему адаптации содержательных аспектов текстовой информации в контексте его онтологического, социального, индивидуального опыта, профессиональных навыков, приходим к необходимости некоторой унификации стратифицированного анализа. На текущем уровне понимания основными компонентами метатекста являются: оглавление, аннотация (реферат), рецензия, дискуссия, цитирование, индексированные списки по ключевым словам, контекстные связи с другими текстами, знание-ориентированная семантическая метамодель содержания, объектно-агентная ориентировочная схема, спецификация контрактов с другими метатекстами, позиционирование в классификационных схемах [1, 2]. Таким образом, метатекст представляется некоторым множеством разнородных объектов – признаков, в том числе в номинальных, дихотомических, порядковых, интервальных и пропорциональных шкалах.

Учитывая потенциальную контралатеральность локальных и глобальных условий принятия решений, следует выделить основные требования к характеристикам метатекстов. С одной стороны, на множестве всех допустимых признаков необходимо искать семейство достаточных признаков, а на нем – класс необходимых, т.е. множество характеристических свойств метатекстов, имеющих минимальную (в широком смысле) сложность и адекватно представляющих декларативные данные. С другой стороны, в критериях информативности признаков целесообразно учитывать не только величину потерь информации, но и информационную эффективность – некоторый функционал от дисбаланса получаемой «полезности» и требуемых «затрат».

Исходя из теоретических предпосылок, практических возможностей и конечных целей, можно прелиминарно определить признаковые пространства множеств метатекстов. Вместе с тем реляционные свойства признаков трансформируются не только от задачи к задаче, но и при использовании всевозможных процедур деривационного типа. В связи с этим важное значение имеет автоматический динамический анализ свойств бинарных отношений и их семейств.

Состояние вопроса исследований и цель работы

В силу простоты и удобства задания в качестве базового свойства взаимосвязи элементов метатекстов часто выступает отношение толерантности [3,4]. Традиционной и естественной экспликацией математического понятия «толерантность» является интерпретация интуитивного понятия сходства. Действительно, из свойства рефлексивности следует, что любой объект неразличим сам с собой, из симметричности – объекты различимы или нет независимо от порядка сравнения. Однако возникающие покрытия множеств признаков малоконструктивны с практической точки зрения. Более того, даже при выделении базисных классов толерантности не исключены ситуации поэлементной классификационной обработки всего множества признаков. Очевидно, что переход к разбиениям, т.е. отношениям эквивалентности, трактуемым как «обобщенное равенство», принципиально сокращает комбинаторную емкость решения задач установления соответствия некоторого запроса и представителя эталонных декларативных данных [5]. Но сложность или даже невозможность задания *a priori* хотя бы достаточного семейства отношений эквивалентности, имеющих к тому же способность изменяться в процессе последовательного анализа, приводит к принципиальным сложностям формализации схем принятия решения. Возникающая при этом потребность перехода к классам эквивалентности, вообще говоря, определяется различными факторами: влиянием целей и путей тематической интерпретации результатов анализа, вариациями способов задания метатекстов, наличием значимого количества потенциально приемлемых подходов, методов и алгоритмов их обработки [6].

Отношения толерантности могут задаваться различными способами. Пусть $\mathfrak{M}=\{x_i\}_{i=1}^{n}$ – множество метатекстов, каждый элемент которого, в свою очередь, представляется семействами характеристических свойств $x_i=\{x_i^k\}_{k=1}^{m}$. Наконец, x_i^k – также может быть некоторым множеством, в частности, списком $x_i^k=(\alpha_1, \alpha_2, \ldots, \alpha_{l_k})$. Наиболее простой вид задания отношения $\mathfrak{T} \subset x_i \times x_j$ толерантности двух метатекстов связан с существованием у них общих элементов, т.е. $x_i \cap x_j \neq \varnothing$. Естественным обобщением является условие наличия нескольких совпадающих признаков: $\mathrm{card}\,(x_i \cap x_j) \geq p,\ p \leq m$. Далее, если ввести двоичные кортежи $x_i \leftrightarrow \langle \xi_1^i, \xi_2^i, \ldots, \xi_m^i \rangle$, в которых ξ_k^i, $k=\overline{1, m}$ принимает значение «1», если соответствующий признак имеется в метатексте и «0» в противном случае, то множество толерантно x_j, когда $\exists l \in L \subset \{1, 2, \ldots, m\}$: $\xi_l^i = \xi_l^j$, т.е. совпадают признаки из заданного подмножества. При анализе троичных кортежей $\langle \xi_1^i, \xi_2^i, \ldots, \xi_m^i \rangle$, в которых $\xi_k^i = -1$, когда о k-м признаке метатекста ничего не известно*), можно использовать метрические свойства: два метатекста толерантны, если

$$\rho(f(\langle \xi_1^i, \xi_2^i, \ldots, \xi_m^i \rangle), f(\langle \xi_1^j, \xi_2^j, \ldots, \xi_m^j \rangle)) \leq \varepsilon,$$

где $\rho(\cdot,\cdot)$ — некоторая метрика, а в общем случае — мера близости; ε — наперед заданный порог; f — инъекция m-мерного признакового пространства в $\mathbb{R}^1$. Так, если ϑ_1 — число случаев принадлежности (возможно с точностью до заданной величины) метатекстам x_i и x_j одинаковых признаков; ϑ_2 — количество случаев, когда метатексты x_i и x_j не имеют одинаковых признаков; ϑ_3 — число ситуаций, в которых метатекст x_i не имеет признаков, присущих x_j; ϑ_4 — количество случаев, когда вектор x_j не имеет признаков, свойственных x_i, то можно указать модифицированные (не обязательно используются одни и те же признаки) метрики или коэффициенты сходства [7–10]:

Кульжинского

$$\rho(x_i, x_j) = (\vartheta_1 + \vartheta_2)/m,$$

$$\rho(x_i, x_j) = (\vartheta_1/(\vartheta_1 + \vartheta_3) + \vartheta_1/(\vartheta_1 + \vartheta_4))/2,$$

$$\rho(x_i, x_j) = \vartheta_1 \big/ \sqrt{(\vartheta_1 + \vartheta_3)(\vartheta_1 + \vartheta_4)};$$

Рассела и Рао

$$\rho(x_i, x_j) = \vartheta_1/m;$$

Хаммана

$$\rho(x_i, x_j) = (\vartheta_1 + \vartheta_2 - \vartheta_3 - \vartheta_4)/m;$$

Джакарта

$$\rho(x_i, x_j) = \vartheta_1/(\vartheta_1 - \vartheta_3 - \vartheta_4);$$

Дейка

$$\rho(x_i, x_j) = 2\vartheta_1/(2\vartheta_1 - \vartheta_3 - \vartheta_4);$$

Сокаля и Снифа

$$\rho(x_i, x_j) = \vartheta_1/(\vartheta_1 - 2\vartheta_3 - \vartheta_4);$$

Жакара и Нидмана

$$\rho(x_i, x_j) = \vartheta_1/m;$$

Юла

$$\rho(x_i, x_j) = (\vartheta_1\vartheta_2 - \vartheta_3\vartheta_4)/(\vartheta_1\vartheta_2 - \vartheta_3\vartheta_4).$$

Применение различныых метрик обусловлено необходимостью задания при принятии решений различных весов данным. Например, коэффициент Дейка удваивает значение совпадающих признаков, первая метрика Кульжинского указывает на равнозначность совпадения и несовпадения признаков и т.д. Следует указать, что в самом общем виде любое всюду определенное соответствие $\varphi : X \to Y$ продуцирует отношение толерантности: два элемента из X толерантны, если пересечение их образов не пустое множество [11].

Таким образом, цель работы — изучение связей эквивалентностей и толерантностей.

Постановка задачи

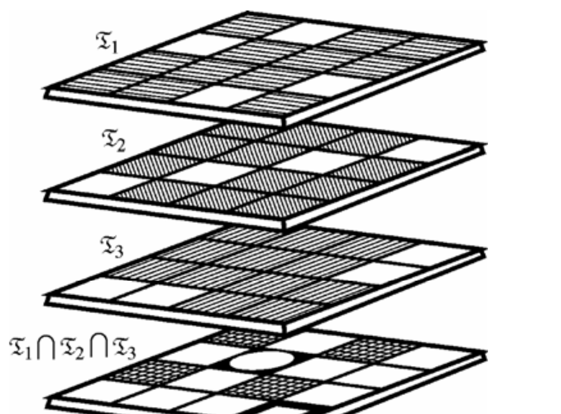
Рассмотрим пример, когда для композиции отношений толерантности выполняется свойство транзитивности. Пусть на множестве $\mathfrak{M} = \{x, y, z, u\}$ заданы следующие отношения толерантности:

$$\mathfrak{T}_1 = \{(x, z); (x, u); (y, z); (y, u)\};$$

$$\mathfrak{T}_2 = \{(x, y); (x, z); (y, u)\};$$

$$\mathfrak{T}_3 = \{(x, y, z); (x, z); (y, u)\}.$$

Как легко видеть, объединение этих отношений — полное, транзитивное замыкание [8] каждого из них продуцирует тривиальную (несущественную) эквивалентность. Если бы произведение толерантностей являлось эквивалентностью, то это означало бы, что компоненты метатекстов заданы с существенной избыточностью, т.е. имеются признаки, «поглощающие» другие характеристики. Вместе с тем на рисунке показано пересечение толерантностей, в результате которого получаем три класса эквиалентности $\{x, z\}$, $\{y\}$, $\{u\}$. Следовательно, одна из задач заключается в поиске условий, при которых пересечение толерантностей трансформируется в отношение эквивалентности.



Пример трансформации пересечения толерантностей в эквивалентность

*) Такая ситуация весьма типична для частично определенных запросов в задачах классификационного анализа и при неполностью определенных прототипах при синтезе информационных структур.

Условия трансформации толерантностей в эквивалентность

Рассмотрим множество произвольной природы $\mathfrak{M}$, на декартовом квадрате $\mathfrak{M}^2$ которого задан набор $\{\mathfrak{T}_k\}_{k=1}^n$ бинарных отношений толерантности. Применительно к произвольному отношению $\mathfrak{T}$ будем говорить: $\forall x, y \in \mathfrak{M}$ справедливо $x t y$ или $x\mathfrak{T}y$, если отношение $t \in \mathfrak{T}$ выполняется, и $\langle x, y\rangle \notin \mathfrak{T}$, если не выполняется. Под пересечением двух отношений $\mathfrak{R}, \mathfrak{P}$ будем понимать: $x\mathfrak{R}\bigcap\mathfrak{P}y \Leftrightarrow x\mathfrak{R}y, x\mathfrak{P}y$.

Нетрудно заметить, что свойство толерантности инвариантно относительно операции пересечения отношений. Действительно, непосредственно из определения вытекает рефлексивность: если $\forall x \in \mathfrak{M}$ выполнены соотношения $x\mathfrak{R}x$ и $x\mathfrak{P}x$, то выполнено и $x\mathfrak{R}\bigcap\mathfrak{P}x$. Далее, пусть отношения $\mathfrak{R}, \mathfrak{P}$ симметричны. Учитывая, что произвольное отношение $\mathfrak{R}$ симметрично тогда и только тогда, когда $\mathfrak{R}=\mathfrak{R}^{-1}$, где $\mathfrak{R}^{-1}$ – обратное отношение, т.е. $x\mathfrak{R}^{-1}y$ равносильно $y\mathfrak{R}x$ [12], предположим, что выполнено отношение $x(\mathfrak{R}\bigcap\mathfrak{P})^{-1}y$. Очевидно, тогда выполнено отношение $y(\mathfrak{R}\bigcap\mathfrak{P})x$, т.е. одновременно выполняются отношения $y\mathfrak{R}x$ и $y\mathfrak{P}x$ или, что равносильно, $x\mathfrak{R}^{-1}y$ и $x\mathfrak{P}^{-1}y$. Следовательно, $(\mathfrak{R}\bigcap\mathfrak{P})^{-1}=\mathfrak{R}^{-1}\bigcap\mathfrak{P}^{-1}$, но тогда в силу симметричности отношений $\mathfrak{R}$ и $\mathfrak{P}$ имеем $(\mathfrak{R}\bigcap\mathfrak{P})^{-1} = \mathfrak{R}\bigcap\mathfrak{P}$, что и означает инвариантность симметричности отношений относительно пересечения. Таким образом, необходимо ответить на вопрос: какие условия обеспечивают выполнение отношения транзитивности на пересечении отношений толерантности?

Обозначим множество отношений эквивалентности через $\mathfrak{E}$. Предположим, что на множестве $\mathfrak{M}$ задано N отношений толерантности $\mathfrak{T}_k$. Зафиксируем произвольное значение $k \in \{1, 2, \dots, N\}$ и введем в рассмотрение следующие множества:

$$\left.\begin{aligned} \mathfrak{L}_k&=\{(x, y, z) \in \mathfrak{M}^3 \ : \ x\mathfrak{T}_k z, y\mathfrak{T}_k z, \langle x, y\rangle \notin \mathfrak{T}_k\}, \\ \mathfrak{N}_k&=\{(x, y, z) \in \mathfrak{M}^3 \ : \ x\mathfrak{T}_k z, y\mathfrak{T}_k z\}, \end{aligned}\right\} \quad (1)$$

где $\mathfrak{M}^3$ – декартов куб множества $\mathfrak{M}$. Обозначим

$$\left.\begin{aligned} \mathfrak{L}&=\bigcup\nolimits_{k=1}^N \mathfrak{L}_k, \ \mathfrak{N}=\bigcap\nolimits_{k=1}^N \mathfrak{N}_k, \\ \mathfrak{S}&=\mathfrak{L}\bigcap\mathfrak{N}, \ \ \mathfrak{T}=\bigcap\nolimits_{k=1}^N \mathfrak{T}_k. \end{aligned}\right\} \quad (2)$$

Теорема 1. *Необходимым и достаточным условием эквивалентности пересечения произвольного числа отношений толерантности является пустота множества* $\mathfrak{S}$*, т.е.* $\mathfrak{T} \subset \mathfrak{E} \Leftrightarrow \mathfrak{S}=\varnothing$.

Доказательство. Рассмотрим необходимость. Пусть $\mathfrak{T}$ – отношение эквивалентности. Допустим, что $\mathfrak{S}=\mathfrak{L}\bigcap\mathfrak{N} \neq \varnothing$, следовательно, найдется тройка $(x, y, z) \in \mathfrak{M}^3$ такая, что $(x, y, z) \in \mathfrak{L}$ и $(x, y, z) \in \mathfrak{N}$. Тогда из (2) вытекает справедливость импликаций:

$$\forall k \ : \ k \in \{1, 2, \dots, N\} \ \Rightarrow \ (x, y, z) \in \mathfrak{N}_k, \quad (3)$$

$$\exists l \ : \ l \in \{1, 2, \dots, N\} \ \Rightarrow \ (x, y, z) \in \mathfrak{L}_l. \quad (4)$$

С учетом определений (1) из (3) получаем выполнение отношений $x\mathfrak{T}_k z, y\mathfrak{T}_k z$ при любом значении $k \in \{1, 2, \dots, N\}$, а из (4) находим, что $\langle x, y\rangle \notin \mathfrak{T}_l$. Но тогда из определения пересечения отношений следует $x\mathfrak{T}z, y\mathfrak{T}z, \langle x, y\rangle \notin \mathfrak{T}$, что противоречит транзитивности отношения $\mathfrak{T}$ и нарушает посылку об эквивалентности отношения $\mathfrak{T}$. Таким образом, $\mathfrak{S}=\varnothing$, что и требовалось.

Для доказательства достаточности предположим, что $\mathfrak{S}=\varnothing$, но отношение $\mathfrak{T}$ – не эквивалентность. Это означает существование тройки $(x, y, z) \in \mathfrak{M}^3$, на которой нарушается транзитивность пересечения отношений толерантности. Покажем, что в этом случае данная тройка принадлежит $\mathfrak{S}$.

Действительно, если отношение $\mathfrak{T}$ не транзитивно, то $\exists\,(x, y, z) \in \mathfrak{M}^3$, для которой

$$x\mathfrak{T}z, \ y\mathfrak{T}z, \quad (5)$$

$$\langle x, y\rangle \notin \mathfrak{T}. \quad (6)$$

Но из определения пересечения отношений и (5) следует, что $x\mathfrak{T}_k z, y\mathfrak{T}_k z$ при любом $k \in \{1, 2, \dots, N\}$, т.е. $(x, y, z) \in \bigcap\nolimits_{k=1}^N \mathfrak{N}_k=\mathfrak{N}$. С другой стороны, (6) означает, что для некоторого номера $l \in \{1, 2, \dots, N\}$ имеет место $\langle x, y\rangle \notin \mathfrak{T}_l$, т.е. $(x, y, z) \in \mathfrak{L}_l \subset \mathfrak{L}=\bigcup\nolimits_{k=1}^N \mathfrak{L}_k$. В итоге получаем $(x, y, z) \in \mathfrak{L}\bigcap\mathfrak{N}=\mathfrak{S}$, следовательно, $\mathfrak{S} \neq \varnothing$, что противоречит исходной посылке. Тем самым, отношение $\mathfrak{T}$ транзитивно, а окончательно – оно является эквивалентностью, что и требовалось доказать.

Замечание. Из условий формирования множеств $\mathfrak{L}_k$ и $\mathfrak{N}_k$ непосредственно следует, что для проверки необходимого и достаточного условий эквивалентности выбирается не более половины всех комбинаций троек, так как в рассмотрение принимаются только элементы, имеющие «предпосылки» существования транзитивности. Более того, множества $\mathfrak{L}_k$ состоят из троек с различными элементами. Действительно, если предположить, что два элемента множества $\mathfrak{L}_k$ совпадают, то $x \neq y$, поскольку $\langle x, y\rangle \notin \mathfrak{T}_k$. Но тогда $x = z$ и, следовательно, $\langle y, z\rangle \notin \mathfrak{T}_k$, что противоречит (1). Таким образом, все множество $\mathfrak{L}$ состоит из различных троек. Ясно, что и в множествах $\mathfrak{N}_k$ можно также учитывать только тройки из различных элементов, поскольку в теореме 1 фигурирует условие $\mathfrak{L}\bigcap\mathfrak{N}=\varnothing$, а $\mathfrak{N}=\bigcap\nolimits_{k=1}^N \mathfrak{N}_k$. Наконец, в качестве практически важного довода в пользу анализа несовпадающих элементов троек выступает поиск существенной (не тривиальной – одноэлементной) эквивалентности.

Следует подчеркнуть, что существуют более простые (но не совпадающие) условия необходимости и достаточности эквивалентности отношения, представленного пересечением произвольного числа отношений толерантности. Например, если $\mathfrak{N}=\varnothing$, то этого достаточно, чтобы $\mathfrak{T} \subset \mathfrak{E}$, поскольку $\mathfrak{S} \subset \mathfrak{N}$. В свою очередь, пустота множества $\mathfrak{N}=\bigcap\nolimits_{k=1}^N \mathfrak{N}_k$ вытекает из пустоты множества $\mathfrak{N}_l$ для какого-то номера $l \in \{1, 2, \dots, N\}$.

С другой стороны, нетрудно указать более простое по сравнению с утверждением теоремы 1 необходимое условие. Рассмотрим множество $\mathfrak{L}^*=\bigcap_{k=1}^{N}\mathfrak{L}_k$. Если $\mathfrak{T}\subset\mathfrak{C}$, т.е. свойство транзитивности выполняется, то $\mathfrak{L}^*=\varnothing$. В противном случае, если найдется тройка $(x, y, z)\in\mathfrak{L}^*$, то из (1) будет следовать $x\mathfrak{T}_k z$, $y\mathfrak{T}_k z$, $\langle x, y\rangle\notin\mathfrak{T}_k$, что противоречит транзитивности. Приведенные рассуждения позволяют сформулировать теорему.

Теорема 2. *Для того чтобы пересечение произвольного числа любых отношений толерантности было эквивалентностью, необходима пустота множества $\mathfrak{L}^*$ и достаточна пустота хотя бы одного из множеств $\mathfrak{N}_k$, т.е. $\mathfrak{T}\subset\mathfrak{C}\Rightarrow\mathfrak{L}^*=\varnothing$, если $\exists l\in\{1, 2, \ldots, N\}:\mathfrak{N}_l=\varnothing\Rightarrow\mathfrak{T}\subset\mathfrak{C}$.*

Условия теоремы 2 в ряде практических задач могут оказаться более удобными с точки зрения их программной реализации.

Обратимся теперь к примеру, рассмотренному выше. В таблице сведены все отношения толерантности.

Пример отношений толерантности

$\mathfrak{M}$	x	y	z	u
x	$\mathfrak{T}_1, \mathfrak{T}_2, \mathfrak{T}_3$	$\mathfrak{T}_2, \mathfrak{T}_3$	$\mathfrak{T}_1, \mathfrak{T}_2, \mathfrak{T}_3$	$\mathfrak{T}_1$
y	$\mathfrak{T}_2, \mathfrak{T}_3$	$\mathfrak{T}_1, \mathfrak{T}_2, \mathfrak{T}_3$	$\mathfrak{T}_1, \mathfrak{T}_3$	$\mathfrak{T}_1, \mathfrak{T}_2,$
z	$\mathfrak{T}_1, \mathfrak{T}_2, \mathfrak{T}_3$	$\mathfrak{T}_1, \mathfrak{T}_3$	$\mathfrak{T}_1, \mathfrak{T}_2, \mathfrak{T}_3$	$\mathfrak{T}_3$
u	$\mathfrak{T}_1$	$\mathfrak{T}_1, \mathfrak{T}_2,$	$\mathfrak{T}_3$	$\mathfrak{T}_1, \mathfrak{T}_2, \mathfrak{T}_3$

Сформируем множества $\mathfrak{L}_k, \mathfrak{N}_k,\ k=1,2,3$. Очевидно,

$$\mathfrak{L}_1=\{(z, u, x); (u, z, x); (x, y, z); (y, x, z); (x, y, u); (y, x, u); (z, u, y); (u, z, y)\},$$

$$\mathfrak{L}_2=\{(y, z, x); (z, y, x); (x, u, y); (u, x, y)\},$$

$$\mathfrak{L}_3=\{(x, u, z); (u, x, z); (y, u, z); (u, y, z)\}.$$

Для упрощения построения множеств $\mathfrak{N}_k$ введем в рассмотрение вспомогательные множества

$$\bar{\mathfrak{N}}_k=\{(x, y, z)\in\mathfrak{M}^3 : x\mathfrak{T}_k z, y\mathfrak{T}_k z, x\mathfrak{T}_k y\}.$$

Ясно, что $\bar{\mathfrak{N}}_k\cap\mathfrak{L}_k=\varnothing$, $\mathfrak{N}_k=\bar{\mathfrak{N}}_k\cup\mathfrak{L}_k$. Как следует из замечания к теореме 1, в $\mathfrak{N}_k$ или, что равносильно, $\bar{\mathfrak{N}}_k$ можно рассматривать только различные элементы, т.е. $\bar{\mathfrak{N}}_1=\bar{\mathfrak{N}}_2=\varnothing$ и $\mathfrak{N}_1=\mathfrak{L}_1$, $\mathfrak{N}_2=\mathfrak{L}_2$. Наконец,

$$\bar{\mathfrak{N}}_3=\{(x, y, z); (y, x, z); (x, z, y); (z, x, y); (y, z, x); (z, y, x)\}.$$

В данном примере $\mathfrak{N}_1\cap\mathfrak{N}_2=\mathfrak{L}_1\cap\mathfrak{L}_2=\varnothing$, следовательно, $\mathfrak{N}\cap\mathfrak{L}=\varnothing$, т.е. пересечение отношений толерантности представляет собой отношение эквивалентности.

Выводы и перспективы

Сформулированы и доказаны необходимые и достаточные условия, при которых суперпозиция отношений толерантности трансформируется в эквивалентность. С практической точки зрения в информационно-поисковых системах эти свойства создают предпосылки для синтеза систем прелиминарной обработки больших объемов текстовой информации, более точно – для кластеризации данных в метатекстовых пространствах. При этом локально или глобально покрытия (слабая кластеризация), индуцируемые различными подмножествами толерантностей, могут трансформироваться в разбиения (строгая кластеризация). В результате достигается сокращение комбинаторной сложности решения аналитических и поисковых задач. Для дальнейшего развития обсуждаемого подхода необходимо изучение методов и алгоритмов построения упорядоченных по вложению классов толерантностей и эквивалентностей в конкретизированных метатекстовых пространствах.

В заключение отметим, что при иерархическом (монотетическом, политетическом или смешанном) анализе метатекстовых информационных структур отношения толерантности наряду с отношениями эквивалентности могут эффективно использоваться не только для обработки объектов на отдельных стратах, но и для учета связей между классификационными рубриками.

Литература: **1.** *Yang Y., Pedersen J.O.* A comparative study on feature selection in text categorization // Proc. of 14th Int. Conf. on Machine Learning (ICML-97) / D.H.Fisher (ed.).– San Francisco: Morgan Kaufmann Publishers, 1997. P.412–420. **2.** *Орлов П.И., Луганский А.М., Марков В.И.* Информационная система «Университет». Харьков: НУВД, 1999. 92с. **3.** *Skowron A., Stepaniuk J.* Tolerance approximation spaces // Fundamenta Informaticae. 1996. Vol.27, No.2,3. P.245–253. **4.** *Marcus S.* Tolerance rough sets, cech topologies, learning process // Bul.of the Polish Academy of Technical Sciences. 1994. Vol.42, No.3. P.471–487. **5.** *Sakaiand H., Okuma A.* An algorithm for finding equivalence relations from tables with non-deterministic information // Lecture Notes on Artificial Intelligence. 1999. Vol.1711. P.64–72. **6.** *Машталир В.П.* Точечно-множественные методы в задачах обработки информации. Харьков: Бизнес Информ, 2001. 199с. **7.** *Фор А.* Восприятие и распознавание образов. М.: Машиностроение, 1989. 272 с. **8.** *Korenjak-Иerne S.* Adapted methods for clustering large datasets of mixed units// Informatica: An International Journal of Computing and Informatics. 1999.Vol. 23, №4. P.507–511. **9.** *Baeza-Yates R., Navarro G.* Faster approximate string matching // Algorithmica. 1999. Vol.23, №2. P.127–158. **10.** *Bunke H.* Structural and syntactic pattern recognition // Handbook of Pattern Recognition and Computer Vision / ChenC.H., PauL.F. and WangP.S.P.(eds.). Singapore – New Jersey– London – Hong Kong: World Scientific Publishing Co. Pte. Ltd., 1995. P.163–209. **11.** *Мальцев А.И.* Алгебраические системы. М.: Наука, 1970. 392с. **12.** *Макаров И.М., Виноградская Т.М., Рубчинский А.А., Соколов В.Б.* Теория выбора и принятия решений. М.: Наука, 1982. 328 с.

Поступила в редколлегию 24.12.2003

Рецензент: д-р техн. наук, проф. Бодянский Е.В.

Луганский Александр Михайлович, начальник научно-исследовательского информационно-компьютерного центра Национального университета внутренних дел. Научные интересы: разработка информационно-управляющих систем. Адрес: Украина, 61080, Харьков, пр. 50-летия СССР, 27, тел. 50-31-43. E-mail: alex@adm.univd.kharkov.ua.

Машталир Владимир Петрович, д-р техн. наук, ст. науч. сотр., профессор кафедры информатики ХНУРЭ. Научные интересы: распознавание образов. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 7021-419, e-mail: mashtalir@kture.kharkov.ua.

Шляхов Владислав Викторович, канд. техн. наук, доцент, ведущий научный сотрудник кафедры ПО ЭВМ ХНУРЭ. Научные интересы: алгебраические структуры, искусственный интеллект. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 7021–446.