

УДК 004.8(075.8)

ИСПОЛЬЗОВАНИЕ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ ПРИ ПОСТРОЕНИИ ИНФОРМАЦИОННЫХ СИСТЕМ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

С.В. Данилкин, О.Г. Иванова, А.В. Лагутин, Н.Г. Шахов

Кафедра «Информационные системы», ТГТУ

Ключевые слова и фразы: автоматическая классификация; виды и формы нечеткости; информационные системы; монотетические; неопределенность; политетические; синтез, стохастическая.

Аннотация: Рассмотрены различные варианты разрешения проблемы неопределенности в задачах автоматической классификации. Проанализированы основные свойства отношений эквивалентности, а также отношения частичной упорядоченности на целевом пространстве информационной системы.

Взаимосвязь понятий «однородность» и «классификация» представляется очевидной даже на интуитивном уровне. С.И. Ожегов в «Словаре русского языка» дает следующие трактовки этих понятий: «Однородный – относящийся к тому же роду, разряду; одинаковый» [1, с. 382]; «классифицировать – распределить по группам, разрядам, классам» [1, с. 238]. Если обе этих трактовки объединить в одну, то классификацию можно понимать как разбиение множества объектов на однородные группы или классы. В таком случае под однородностью подразумевается наличие у объектов одного класса общих свойств или признаков, определяющих некоторое сходство данных объектов и служащих основанием для отнесения этих объектов к одному классу.

Вместе с тем, во многих областях математики, к примеру, в прикладной статистике, понятие однородности оказывается основополагающим, так как обработка статистических данных производится только в однородных группах [2, с. 11]. Требование однородности исследуемого множества объектов не ограничивается

только лишь определением наблюдаемого объекта, так как любое реальное множество объектов являет собой систему дифференцированных, различающихся между собой элементов, что делает задачу разбиения исходного множества исследуемых объектов на однородные подмножества приоритетной при анализе систем любой природы: технических, биологических, социально-экономических, информационных.

Одной из основных особенностей задачи классификации является наличие как качественных, так и количественных признаков в описании объектов исходного множества, в силу чего при выделении однородных групп различают такие виды группировки исходных данных, как структурная и типологическая. *Структурной группировкой* именуется разбиение качественно однородного исходного множества объектов на классы, которые характеризуют общее строение исходного множества объектов [3, с. 96]. *Типологической группировкой* называется разбиение исходного множества объектов на классы определенного качества. Таким образом, структурная группировка представляет собой способ выделения количественно однородных групп объектов, а типологическая – способ выделения качественно однородных групп. При сопоставлении этих определений происходит своеобразное противопоставление категорий качества и количества, сильно упрощающее понимание этих категорий и являющееся, вообще говоря, неправомерным [4, с. 7].

Если в основе типологической группировки находится некоторый качественный признак, причем единственный, то задача классификации, как правило, решается элементарно, однако в подавляющем большинстве случаев ее необходимо проводить по количественным признакам, что в значительной степени усложняет задачу. В таком случае, на начальных этапах исследования рассуждения о качественной однородности исходных данных лишены всякого смысла, поскольку качественная однородность данных может быть установлена только в результате проведения анализа, основой которого, как справедливо указывали И.И. Елисеева и В.О. Рукавишников, должен быть синтез «теоретических концепций и опыта прошлых исследований» [2, с. 15].

Таким образом, представляется нецелесообразным различать методы выделения качественно и количественно однородных групп, однако имеет смысл, как отмечал И.Д. Мандель, «говорить только о непрерывном синтезе этих категорий в процессе классификации» [4, с. 10]. Методы выделения однородных групп объектов, в связи с вышеизложенным замечанием, условно объединяются в следующие основные подходы [4].

Вероятностный подход основан на предположении о том, что объекты, принадлежащие одному из выделяемых классов, описываются одинаково распределенными случайными векторами, а для различных классов характерны различные распределения вероятностей. В специальной литературе этот подход традиционно именуется расщеплением смесей распределений, где каждый класс понимается как некоторая параметрически заданная одномодальная совокупность, а наблюдения над объектами, подлежащими классификации, трактуются как выборка из смеси таких совокупностей, так что задача заключается в разделении этих совокупностей, исходя из значений параметра, определяющего совокупность, и некоторых предположений, к примеру, о числе классов.

Вариативный подход состоит в разбиении множества объектов по выбранному исследователем признаку на интервалы группирования, в результате чего исходное множество объектов разбивается на группы таким образом, что объекты одной группы находятся на относительно небольшом расстоянии друг от друга. В многомерном же случае, при наличии нескольких признаков, данный подход представляет собой комбинационную группировку, для которой характерно поочередное использование признаков для выделения групп. Такой подход, когда единственный признак используется для разбиения всего множества объектов на группы, а также в случае поочередного использования различных признаков, когда каждый из них применяется для выделения одной группы, называется *монотетическим* [2, с. 7].

Структурный подход базируется на представлении об объектах как точках в многомерном пространстве. В этом случае задача состоит в выделении из исходного множества многомерных точек однородных подмножеств таким образом,

чтобы элементы каждого подмножества были в определенном смысле сходны между собой, а сами подмножества – классы объектов – отличались бы друг от друга, так что отыскивается своего рода «естественное» расслоение исходного множества на классы. Данный подход иногда именуется геометрическим, поскольку, используя понятия расстояния между объектами и расстояния между классами, выделяет геометрически удаленные группы. Наиболее последовательно геометрический подход реализован в методах кластерного анализа, которые в специальной литературе называются также методами автоматической классификации, численной таксономией или распознаванием образов с самообучением. В отличие от монотетического подхода к проблеме классификации объектов, кластерный анализ использует *одновременно все признаки* и называется *политетическим*.

Подробное исследование взаимосвязей между вышеизложенными подходами к решению проблемы выделения однородных групп объектов проведено И.Д. Манделем [4]. Вместе с тем, анализируя соотношение вероятностного и структурного подходов, в первую очередь необходимо отметить то обстоятельство, что многие зарубежные исследователи, такие как Дж. Хартиган [5], К. Фукунага [6], М. Вонг [7], рассматривают кластер-анализ чрезмерно широко, включая в него и задачи расщепления смесей, то есть задачи классификации в условиях отсутствия обучающих выборок, когда исходные данные об исследуемых объектах имеют вероятностную природу и каждый класс интерпретируется как одномодальная генеральная совокупность при неизвестном значении определяющего ее параметра, а классифицируемые объекты рассматриваются как выборки из смеси таких генеральных совокупностей. Как способ представления исходных данных понятие смеси использует также известный польский исследователь Я.В. Овсиньски [8, с. 392] при рассмотрении общей постановки задачи кластер анализа. В отечественной литературе подобное рассмотрение автоматической классификации прослеживается в работах М.И. Шлезингера [9] и А.В. Миленького [10]. Е.Е. Жук и Ю.С. Харин [11, с. 17–19] также указывают на существование в кластер-анализе вероятностного и геометрического подходов, отдавая предпочтение первому. Необходимо указать, что применимость методов расщепления

смесей вероятностных распределений к решению задач классификации зависит от обоснованности предположений о вероятностной природе исходных данных и корректности выдвигаемой гипотезы о распределении вероятностей, описывающих классы объектов, тогда как успешное применение геометрических методов классификации зависит только от адекватности выбранной меры близости объектов. Отнесение же группы вероятностно-статистических методов классификации в условиях отсутствия обучающих выборок к кластерному анализу в силу причин методологического характера представляется спорным, так что следует говорить не о вероятностном подходе к решению задачи автоматической классификации, а о теоретико-вероятностной модификации задачи кластер-анализа, как это было предложено С.А. Айвазяном [12, с. 146].

Касательно соотношения вариативного и структурного подходов, здесь лишь укажем, что при использовании вариативного подхода, являющегося разновидностью типологической группировки, классы имеют субъективный характер, а сама группировка является полностью управляемой, так что «естественное» расслоение, отыскиваемое методами кластерного анализа, в случае применения вариативных методов не имеет места. Главным же отличием двух подходов является то, что понятия «близость» и «сходство» объектов в типологической группировке неформализованы, в отличие от структурного подхода, где они формализованы и выражаются рядом соотношений. Достаточно полный обзор метрик и мер близости, формализующих понятие сходства и используемых в задачах классификации, содержится в работах [13, 14].

Собственно процесс решения задачи классификации, независимо от природы исходных данных, в общем, состоит из восьми этапов [12, с. 42–43]: установочного, на котором формулируется постановка задачи на содержательном уровне; постановочного, в ходе которого определяется тип прикладной задачи в терминах теории классификации; информационного, состоящего в выработке плана сбора исходной информации, ее предварительном анализе и редактировании; априорного математико-постановочного, заключающегося в выборе на основании выводов, полученных в результате реализации предыдущих этапов, базовых математических моделей для математической постановки конкретной задачи клас-

сификации; разведочного, предусматривающего применение специальных методов предварительного анализа исходных данных с целью выявления их вероятностной и геометрической природы; апостериорного математико-постановочного, в процессе которого уточняется выбор базовой математической модели с учетом результатов реализации разведочного этапа процесса решения задачи классификации; вычислительного, целью которого является программная реализация выбранного математического аппарата для решения конкретной задачи; итогового, на котором производится анализ и интерпретация результатов проведенного исследования. Таким образом, вид задачи классификации определяется в результате реализации первых трех этапов процесса исследования; к примеру, если предварительная выборочная информация отсутствует, а априорные сведения о классах объектов являются лишь некоторыми предположениями самого общего характера, то задача относится к классу задач распознавания образов с самообучением.

Вместе с тем, на практике зачастую оказывается, что задаче свойственна нечеткость [15], значительно затрудняющая или вообще делающая невозможным получение решения, так что на первый план выходит проблема устранения нечеткости, присущей задаче классификации. Понятие нечеткости является общенаучным и может быть определено как внешнее выражение качества внутренней основы явлений, специфика которого заключается в непрерывности перехода от отсутствия проявления к полному выявлению качества предметов, свойств и отношений реального мира, что находит свое отражение в познавательной и мыслительной деятельности индивида. Содержание понятия нечеткости включает в себя последовательный ряд абстракций более низкого уровня. По отношению к человеческому сознанию выделяются такие категориальные виды нечеткости, как объективная и субъективная нечеткость; в свою очередь, объективная нечеткость может характеризоваться как стохастической, так и нестохастической детерминированностью. Объективная стохастическая нечеткость имеет такие формы проявления, как неопределенность и случайность. В данном случае неопределенность выступает в качестве нечеткой закономерности проявления свойств предмета, а случайность может быть определена как событие, имеющее нечеткое основание. Формами проявления объективной нестохастической нечеткости являются неде-

терминированность, рассматриваемая как нечеткость связи между предметами, свойствами или отношениями; размытость, характеризующая границы явлений, процессов, предметов, а также их классов и, кроме того, имена и область применимости предиката в логике; неоднозначность, определяемая как нечеткость значения признака объекта; неполнота, представляющая собой отсутствие всей возможной информации о рассматриваемом предмете или явлении, частными случаями которой выступают недостаточность – как отсутствие необходимой информации – и неадекватность как описание предмета по аналогии с рассмотренными ранее; неточность, являющаяся нечеткостью измерения или вычисления; неопределенность, трактуемая как нечеткость предела проявления характеристики предмета; случайность, определяемая как нечеткая реализация одной из нескольких существующих возможностей. Субъективная нечеткость имеет такие формы проявления, как неясность, под которой подразумевается нечеткость восприятия; размытость, которая в данном случае являет собой характеристику представления индивида о явлениях, процессах, предметах, свойствах, отношениях; недетерминированность, определяемая как свойство процесса логического вывода, производимого индивидом, в нечетких условиях; неоднозначность, понимаемая как нечеткость результата процесса интерпретации информации; неточность, которая в данном случае трактуется как мера соответствия знаний индивида о предмете объективным характеристикам рассматриваемого предмета; неопределенность, понимаемая как нечеткость отношения между объектом реального мира и представлением о нем в сознании индивида, а также как нечеткость смыслового значения имени, выражающего некоторое понятие.

Таким образом, виды и формы нечеткости могут быть упорядочены в соответствии с иерархией, представленной на рис. 1

Каждой из рассмотренных форм неопределенности в задачах автоматической классификации соответствует определенная методология ее формализации или устранения.

Касательно неопределенности целей исследования следует указать, что поскольку формы ее проявления носят субъективный характер, необходимым условием для ее устранения является полное описание предмета исследования, сбор всей возможной информации.

Введем в рассмотрение понятие целевого пространства информационных систем. При анализе и проектировании информационных систем (ИС) всегда предполагается наличие некоторого исходного множества ИС и возможность осуществить сравнение его элементов для выбора предпочтительного варианта среди имеющихся. Выделение исходного множества информационных систем $X = \{x\}$ из множества X_{Σ} всех ИС осуществляется по признакам, характеризующим рассматриваемую конкретную задачу. Несомненно, что важнейшим из признаков, служащих для характеристики информационных систем, является цель функционирования. При проектировании из всего множества информационных систем, прежде всего, отбираются те, которые отвечают одной и той же (главной) цели функционирования или одному и тому же множеству (главных) целей функционирования.

Ввиду особой роли, которую играет главная цель функционирования, введем для нее специальное обозначение $\text{aim}x$. Отношение A «иметь общую главную цель функционирования или общее множество главных целей», т.е. xAy , если $\text{aim}x = \text{aim}y$, $x, y \in X_{\Sigma}$, является отношением эквивалентности. Отсюда следует, что в множестве $X \subset X_{\Sigma}$ информационных систем, имеющих общую цель ξ ($\xi = \text{aim}x$, $x \in X$), все элементы по данному признаку эквивалентны. Разумеется, на множестве X могут существовать (и, как правило, имеются) и другие отношения эквивалентности R .

Для возможности сопоставления информационных систем $x \in X$, имеющих общую цель, предположим, что для некоторых элементов X задано также отношение xPy частичного порядка.

Подмножество X множества X_{Σ} всех информационных систем с заданными на нем отношениями эквивалентности и частичной упорядоченности, имеющее в числе признаков информационных систем общую главную цель (или цели), будем называть *целевым пространством*. Таким образом, ЦПИС есть тройка (кортеж из 3 элементов)

$$\text{ЦПИС} = \langle X, \{R\}, \{P\} \rangle, \quad (1)$$

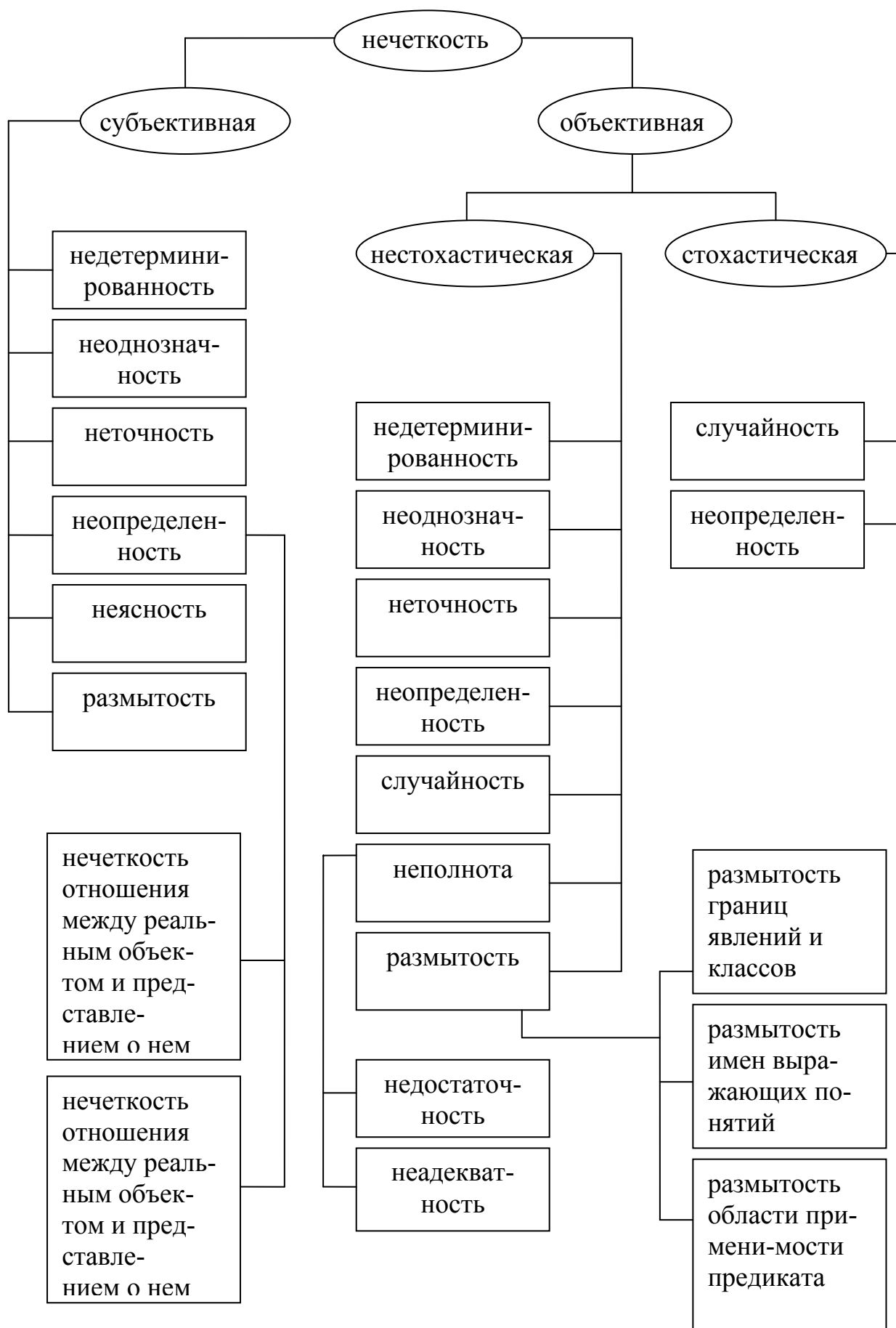


Рис. 1. Виды и формы нечеткости

при условии общности главной цели (целей) (из $x \in X$ вытекает, что $A = \text{aim}x \in \{R\}$) и замкнутости $\{R\}, \{P\}: R', P'$, логически вытекающие из R, P , также входят в ЦПИС.

Далее без дополнительных пояснений будем использовать обычные обозначения отношений: \sim для эквивалентности; \geq, \leq для нестрогого предшествования; $>, <$ для строгого предшествования.

В соответствии с определением ЦПИС является моделью Тарского.

Структура отношений на ЦПИС. Для дальнейшего изложения будем использовать алгебраические понятия теории структур [16], [17]. Рассмотрим множество $\mathbf{R} = \{R\}$ всех эквивалентностей, определенных на ЦПИС; оно по определению не пусто. Если $R_1, R_2 \in \mathbf{R}$, то примем, что

$$R_1 \leq R_2 \stackrel{\Delta}{=} (\text{из } xR_1y \text{ вытекает } xR_2y)$$

$$\text{и из } R \in \mathbf{R}, R \leq R' \text{ следует, что } R' \in \mathbf{R}. \quad (2)$$

Это отношение логического следования будем, записывать в виде

$$R_1 \leq R_2 \stackrel{\Delta}{=} xR_1y \rightarrow xR_2y. \quad (2')$$

Определенное таким образом на \mathbf{R} отношение \leq рефлексивно, антисимметрично, транзитивно, т.е. устанавливает на \mathbf{R} частичный порядок. Нетрудно установить, что пара $\langle \mathbf{R}, \leq \rangle$ является *полной структурой*. Действительно, пусть $\mathbf{T} \subset \mathbf{R}$ – произвольное непустое подмножество \mathbf{R} . Определим отношение R_* следующим образом

$$xR_*y \stackrel{\Delta}{=} (xR_1y, \forall R \in \mathbf{T}). \quad (3)$$

Поскольку определенное в (3) отношение R_* (его называют *пересечением отношений*) есть эквивалентность, то $R_* \in \mathbf{R}$.

Покажем, что R_* – точная нижняя грань \mathbf{T} . Пусть $R \in \mathbf{T}$ – произвольный элемент \mathbf{T} , по определению R_*

$$xR_*y \rightarrow xRy,$$

т.е. $R_* \leq R$, значит, R_* – нижняя грань \mathbf{T} . Пусть R_{**} – еще одна нижняя грань \mathbf{T} , тогда $R_{**} \leq R$ для любого $R \in \mathbf{T}$, а это значит, что

$$xR_{**}y \rightarrow xRy \quad \forall R \in \mathbf{T},$$

а следовательно, $xR_{**}y \rightarrow xR_*y$, что по определению означает, что $R_{**} \leq R_*$. Значит, R_* есть точная нижняя грань \mathbf{T} .

Определим отношение R^* следующим образом

$$xR^*y = \overset{\Delta}{\{ \exists (R_1, R_2, \dots, R_n) \in \mathbf{T}, \\ \exists (x = z_1, z_2, \dots, z_n, z_{n+1} = y) \in X, z_i R_i z_{i+1} (i = \overline{1, n}) \}}. \quad (4)$$

Отношение (4) является транзитивным замыканием.

Отношение R^* является эквивалентностью. Если для некоторых $x, y \in X$ и $R \in \mathbf{T}$ имеет место xRy , то согласно (4) xR^*y , т.е. $R \leq R^*$, откуда следует, что R^* – верхняя грань для \mathbf{T} . Пусть R^{**} – еще одна верхняя грань для \mathbf{T} . Предположим, что для $z_1, z_2, \dots, z_n \in X$ и $R_1, R_2, \dots, R_n \in \mathbf{T}$ имеют место отношения $z_i R_i z_{i+1}, (i = \overline{1, n})$. Поскольку по определению $R^{**} \geq R_i$, то $z_i R^{**} z_{i+1}$, следовательно, $z_1 R^{**} z_{n+1}$, а это означает $xR^*y \rightarrow xR^{**}y$ или, по определению, $R^* \leq R^{**}$. Таким образом, R^* – точная верхняя грань $R, R \in \mathbf{T}$.

В целом приведенными рассуждениями доказано, что \mathbf{R} есть полная структура; роль единицы \mathbf{R} , очевидно, играет отношение $A \in \mathbf{R}$ «иметь одну и ту же главную цель (цели)», связывающее все элементы ЦПИС. Итак, R – полная структура с единицей (или *полная моноструктура*).

В дальнейшем, наряду с понятием структуры, часто используется также понятие *полугруппы*.

Из приведенных определений следует, что множество \mathbf{R} отношений эквивалентности на ЦПИС является коммутативной полугруппой с единицей (*коммутативным моноидом*) относительно операции пересечения отношений $R \in \mathbf{R}$, определенной согласно (3)

$$(R = R_1 \ R_2) = \overset{\Delta}{(xRy \rightarrow xR_1y \wedge xR_2y)}, \quad (5)$$

где \wedge – логический символ «И». Действительно, R является эквивалентностью и выполняются как ассоциативный закон: $R_1(R_2R_3) = (R_1R_2)R_3$, так и коммутативный закон

$$R_1R_2 = R_2R_1.$$

Роль единицы этой подгруппы, очевидно, играет отношение $A \in \mathbf{R}$: поскольку оно справедливо для всех без исключения элементов ЦПИС, то для произвольного $R \in R$ справедливо $AR = RA = R$.

Далее, непосредственно из определения полугрупповой операции следует, что для любого $R \in R - RR = R$, т.е. каждый элемент полугруппы является идемпотентом. Итак, $S = \langle \{R\}, \cdot \rangle$ – коммутативный моноид идемпотентов или, в терминологии теории полугрупп, коммутативная связка, причем структура R и полугруппа S сопряжены.

Заметим, что в полугруппе S формально иным образом может быть введено частичное упорядочение, а именно $R_1 \leq R_2$, тогда и только тогда, когда R_1 «делится» на R_2 . Иначе говоря, существует отношение $R \in S$ такое, что $R_1 = R_2R$; рефлексивность, транзитивность и антисимметричность этого отношения проверяются без труда. Относительно указанного частичного порядка полугруппа S является полной структурой \mathbf{R}' , причем структура \mathbf{R}' и полугруппа S также являются сопряженными.

Отметим также, что и полугрупповая операция в структуре R может быть определена формально иным образом, а именно

$$R_1R_2 = R_2R_1 = \inf \{R_1, R_2\},$$

причем относительно этой операции R является коммутативной связкой S' , сопряженной с R . Из единственности коммутативной связки, сопряженной с данной структурой, вытекает изоморфность S и S' , R и R' .

Для иллюстрации приведем примеры некоторых эквивалентностей, используемых на практике при формировании ЦПИС:

– отношение $R_{\text{оп}}$ «иметь один и тот же прототип (эталон)»: $xR_{\text{оп}}y$ тогда и только тогда, когда $x, y \in X$ имеют общий прототип x_0 ;

- отношение $R_{пр}$ «иметь одну и ту же программную реализацию», например, одинаковые структуру и язык реализации и базы;
- отношение $R_{а0}$ «использовать одно и то же информационное (программное, аналитическое, математическое) обеспечение»;
- отношение R_d динамической эквивалентности: $xR_d y, x, y \in \text{ЦПИС}$, тогда и только тогда, когда x, y обеспечивают одинаковое значение функционала $Q: Q(x) = Q(y)$, или функционалов $Q_q: Q_q(x) = Q_q(y)$, определяющих качество выполнения главной цели (целей) в одном и том же для x и y диапазоне условий функционирования ($q = \overline{1, n}$);
- отношение $R_{оу}$ «иметь один и тот же закон (способ, метод) управления»;
- отношение $R_{вз}$ взаимозаменяемости: $xR_{вз} y$ означает, что x взаимозаменяем с y .

Таковы основные свойства отношений эквивалентности на ЦПИС.

Что касается отношений частичной упорядоченности на ЦПИС, то они проистекают из наличия общей цели (целей) $\text{aim}x$ и обычно состоят в сравнении систем между собою по характеристикам P_A ее выполнения (таким, например, как качество, точность, надежность, достоверность, быстродействие и т.д.), по объему P_V аппаратуры, обеспечивающей выполнение цели, объему и составу P_H требуемой информации и другим техническим условиям и требованиям.

Пусть $P \in \mathbf{P}$, где \mathbf{P} – множество отношений частичного порядка на ЦПИС, а $x_1 > x_2 > \dots > x_n > \dots$ – убывающая цепочка элементов $x_i \in X$ относительно P . Для приложений целесообразно считать, что выполнено следующее свойство: любая убывающая цепочка систем из ЦПИС $x_1 > x_2 > \dots > x_n > \dots$ имеет минимальный элемент, т.е. такую систему x_{n+p} , что $x_i \geq x_{n+p}$ для любого i (иначе говоря, указанная цепочка где-то стабилизируется: $x_{n+p} = x_{n+p+1} = \dots$). Данное предположение для ЦПИС оправдано тем, что для реализации цели всегда необходимо иметь какой-то минимум средств, устройств, информации и т.д., а характеристики выполнения цели ограничены и с какого-то значения практически неразличимы.

Необходимо иметь в виду, что при введении ЦПИС проектирование рассматривается на уровне систем и подсистем; построение систем из элементов при этом не рассматривается, а получаемые результаты носят обобщенный характер.

Список литературы

1. Ожегов, С.И. Словарь русского языка : ок. 57000 слов / под ред. Н.Ю. Шведовой. – М. : Русский язык, 1985. – 797 с.
2. Елисеева, И.И. Группировка, корреляция, распознавание образов (Статистические методы классификации и измерения связей) / И.И. Елисеева, В.О. Рукавишников. – М. : Статистика, 1977. – 144 с.
3. Смоляк, С.А. Устойчивые методы оценивания / С.А. Смоляк, Б.Н. Титаренко. – М. : Статистика, 1982. – 208 с.
4. Мандель, И.Д. Кластерный анализ / И.Д. Мандель. – М. : Финансы и статистика, 1988. – 176 с.
5. Hartigan, J.A. Clustering Algorithms / J.A. Hartigan. – New York: John Wiley & Sons, 1975.
6. Фукунага, К. Введение в статистическую теорию распознавания образов : пер. с англ. / К. Фукунага. – М. : Наука, 1979. – 368 с.
7. Wong, M.A. A Hybrid Clustering Method for Identifying High-Density Clusters / M.A. Wong // Journal of American Statistical Association. – 1982. – Vol. 77. – Pp. 841 – 847.
8. Owsinski, J.W. Clustering – Modelling, Capacities, Limits, Applications / J.W. Owsinski // Control and Cybernetics. – 1995. – Vol. 24. – Pp. 391 – 397.
9. Шлезингер, М.И. О самопроизвольном различении образов / М.И. Шлезингер // Читающие автоматы. – Киев, 1965. – С. 38 – 65.
10. Миленский, А.В. Классификация сигналов в условиях неопределенности / А.В. Миленский. – М. : Сов. радио, 1975. – 328 с.
11. Жук, Е.Е. Устойчивость в кластер-анализе многомерных наблюдений / Е.Е. Жук, Ю.С. Харин. – Мн. : Белгосуниверситет, 1998. – 240 с.

12. Прикладная статистика: Классификация и снижение размерности : справ. изд. / С.А. Айвазян и др. ; под ред. С.А. Айвазяна. – М. : Финансы и статистика, 1989. – 607 с.
13. Лейбкинд, А.Р. Математические методы и модели формирования организационных структур управления / А.Р. Лейбкинд, Б.Л. Рудник, А.А. Тихомиров. – М. : МГУ, 1982. – 232 с.
14. Walesiak, M. Ugolniona Miara Odleglosci w Statystycznej Analizie Wielowymiarowej / M. Walesiak. – Wroclaw: Wydawnictwo AE; 2002.
15. Вятчинин, Д.А. Формы проявления нечеткости / Д.А. Вятчинин // Гуманитарно-экономический вестник. – 1998. – № 1. – С. 66 – 69.
16. Батыршин, И.З. Иерархические алгоритмы кластеризации на базе классов толерантности / И.З. Батыршин, Р.Г. Халитов // Повышение эффективности технологических процессов химических, нефтехимических и биотехнологических производств : тезисы докладов Республиканской науч.-практич. конф. – Казань, 1986. – С. 109.
17. Нгуен, М.Х. Применение нечетких отношений в классификации / М.Х. Нгуен // Нечеткие системы поддержки принятия решений : сб. науч. тр. – Калинин, 1989. – С. 99 – 107.

USE OF AUTOMATIC CLASSIFICATION FOR INFORMATION SYSTEM DEVELOPMENT IN UNCERTAINTY CONDITIONS

S.V. Danilkin, O.G. Ivanova, A.V. Lagutin, N.G. Shakhov

Key words and phrases: automatic classification; fuzziness types and forms; information systems; monotetic; uncertainty; polytetic; synthesis, stochastic.

Abstract: Different variants of uncertainty problem solution in automatic classification tasks are considered. The main properties of equivalence ratio are analysed as well as partial order ratio on target space of information system.