5. Миллер Г.Б., Панков А.Р. Фильтрация случайного процесса в статистически неопределенной линейной стохастической дифференциальной системе // Автомат. и телемех. 2005. № 1. С. 59–71.

6. Пугачев В.С., Синицын И.Н. Стохастические дифференциальные системы. Анализ и фильтрация. М.: Наука, 1990.

7. Богданович В.А., Вострецов А.Г. Теория устойчивого обнаружения, различения и оценивания сигналов. М.: Физматлит, 2004.

8. Шмелев А.Б. Основы марковской теории нелинейной обработки случайных полей. М.: Изд-во МФТИ, 1998.

9. Тихонов В.И., Кульман Н.К. Нелинейная фильтрация и квазико-герентный прием сигналов. М.: Сов. радио, 1975.

10. Граничин О.Н. Введение в методы стохастической оптимизации и оценивания. СПб.: Изд-во С.-Петербургского ун-та, 2003.

11. Фирсов А.Н. Аналитическое решение нестационарного уравнения Колмогорова-Феллера с нелинейным коэффициентом сноса // Сборник трудов XXIII Международ. научно-практической конференции «Системный анализ в проектировании и управлении». Ч. 2. СПб.: Изд-во Политех-Пресс, 2019. С. 9–73.

12. Firsov A.N., Zhilenkov A.A. Mathematical analysis of transport systems modeled by the stationary Kolmogorov-Feller equation with a nonlinear drift coefficient // Vibroengineering Procedia (ISSN Print 2345-0533; ISSN online 2538-8479). June 2019. Vol. 25. Pp. 166–170.

13. Козлов В.Н. Управление энергетическими системами. СПб: Издательство Политехн. ун-та, 2008.

14. Snyder D.L. Filtering and Detection for Doubly Stochastic Poisson Processes // IEEE Transactions on Information Theory. IT-19. Jan. 1972. Pp. 91–103.

УДК 004.827
doi:10.18720/SPBPU/2/id20-158

Кулик Борис Александрович,
д-р физ.-мат. наук, ст. науч. сотр., вед. науч. сотр.

ЛОГИКО-ИНТЕРВАЛЬНЫЙ АНАЛИЗ СИСТЕМ НА ОСНОВЕ АЛГЕБРЫ КОРТЕЖЕЙ

Институт проблем машиноведения РАН,
Санкт-Петербург, Россия,
ba-kulik@yandex.ru

Аннотация. Предлагается новый подход к логической составляющей интервального анализа, в котором вместо минимаксных операций и сравнений интервалов по величине используются операции и сравнения алгебры множеств. Полученная модель позволяет моделировать и анализировать объекты, у которых имеются не только измеримые, но и не имеющие меры упорядоченные атрибуты, при этом для них можно помимо логического анализа объектов применять методы упорядочения и кластеризации.

Ключевые слова: интервальный анализ, темпоральная логика, интервальная логика, алгебра кортежей, квантование интервалов, мера, кластеризация.

Boris A. Kulik,
Dr. Sci. (Phys.-Math.), Leading Sci. Researcher

LOGICAL-INTERVAL ANALYSIS OF SYSTEMS BASED ON N-TUPLE ALGEBRA

Institute for Problems in Mechanical Engineering of RAS,
St. Petersburg, Russia,
ba-kulik@yandex.ru

Abstract. A new approach to the logical component of interval analysis is proposed, in which set algebra operations and comparisons are used instead of minimax operations and comparisons of intervals by value. The resulting model allows you to model and analyze objects that have not only measurable, but also non-measurable ordered attributes, and you can use ordering and clustering methods for them in addition to logical analysis of objects.

Keywords: interval analysis, temporal logic, interval logic, n-tuple algebra, interval quantization, measure, clustering.

Введение

Интервальный анализ представляет собой интенсивно развивающееся направление математики. Первая обстоятельная монография на эту тему была опубликована в 1966 г. [1], а на русском языке – в 1981 г. [2]. Основная идея интервального анализа состоит в том, чтобы в аналитических функциях использовать вместо вещественных чисел интервалы.

В монографии [3] отмечается, что алгебраические свойства классической интервальной арифметики (IR) существенно более бедны по сравнению с привычными числовыми системами, такими как кольца целых чисел, поля рациональных, вещественных и комплексных чисел, поскольку

• все интервалы с ненулевой шириной, т. е. большинство элементов IR, не имеют обратных элементов по отношению к операциям,

• арифметические операции над интервалами связаны друг с другом весьма слабыми соотношениями (вроде субдистрибутивности), а полноценная дистрибутивность умножения (и деления) относительно сложения и вычитания не имеет места.

Аналогичной «бедностью» отличается и логическая составляющая IR, в которой используются порядковые свойства интервалов. Важную роль в частично упорядоченных множествах играет вычисление верхней ($\vee$) и нижней ($\wedge$) грани заданного подмножества. Пусть заданы интервалы $A = [a_1, a_2]$, $B = [b_1, b_2]$. Тогда

$$A\vee B = [\min\{a_1, b_1\},\max\{a_2, b_2\}];\ A\wedge B = [\max\{a_1, b_1\},\min\{a_2, b_2\}]. \quad (1)$$

Данное определение операций $\vee$ и $\wedge$ предложено также для логико-интервальных математических моделей, рассмотренных в [4, 5].

Стоит отметить, что предложенная алгебра не является полной, так как операция $\wedge$ не всегда выполнима. Например, $[2, 4] \wedge [6, 8]$ не определено (получается неправильный интервал [6, 4]).

Данную проблему можно решить за счет расширения классической IR. В предложенной Каухером [6] полной IR можно использовать не только *правильные* интервалы типа $[a_1, a_2]$, где $a_1 \leq a_2$, но и *неправильные*, у которых $a_1 > a_2$

Не определены в IR также и дополнения интервалов.

Кроме рассмотренных были предложены другие методы логико-интервального анализа, к ним, в частности относится интервальная логика В.И. Левина [7] и темпоральная логика Аллена [8].

Интервальная логика Левина разработана в рамках непрерывной (НЛ) или бесконечнозначной логики [9, 10], в которой значения истинности переменных содержатся в непрерывном интервале действительных чисел, а операции $\vee$ и $\wedge$ для пары переменных определяются как максимум и минимум соответственно их значений. Операция отрицания в них определена, но в разных версиях НЛ по-разному. В [7, 10] дизъюнкция $\vee$, конъюнкция $\wedge$ и отрицание $\neg$ определяются следующим образом. Пусть $C = [A, B]$ – замкнутый интервал вещественных чисел с центром $M=(A+B)/2$. Тогда для любых переменных $a,b \in C$

$a \vee b = \max(a,b)$; $a \wedge b = \min(a,b)$; $\neg a = 2M - a$.

В НЛ не соблюдаются некоторые законы булевой алгебры, в частности, закон исключенного третьего и закон ортогональности

$a \vee b = a \vee (\neg a \wedge b)$.

Интервальные операции в [7] определены так. Пусть заданы интервалы $A = [a_1, a_2]$, $B = [b_1, b_2]$. Тогда

$A \vee B = [\max\{a_1, b_1\}, \max\{a_2, b_2\}]$; $A \wedge B = [\min\{a_1, b_1\}, \min\{a_2, b_2\}]$;

$\neg A = [\neg a_2, \neg a_1]$.

При сравнении с (1) ясно, что в интервальной логике Левина алгоритм вычисления верхних и нижних граней существенно отличается от алгоритмов в классическом интервальном анализе. Например, если в интервальном анализе в результате операции $A \vee B$ образуется интервал, содержащий в себе интервалы A и B, то в интервальной логике это не соблюдается. В интервальной логике также не соблюдаются некоторые законы булевой алгебры, в частности, законы дистрибутивности и противоречия.

В темпоральной логике Аллена [8, 11] определены 7 базисных интервальных отношений. Эти отношения обозначаются так: ***b*** (before), ***m*** (meets), ***o*** (overlaps), ***f*** (finishes), ***s*** (starts), ***d*** (during), ***e*** (equates). Еще 6 отношений образуются за счет инверсий первых шести отношений. Инверсия отношения обозначается символом «*». Например, отношение A ***m**** B верно тогда и только тогда, когда верно B ***m*** A (см. таблицу 1).

Пусть A и B – интервалы, а Ω – множество всех 13-ти интервальных отношений, т. е. $\Omega = \{\boldsymbol{b}, \boldsymbol{m}, \boldsymbol{o}, \boldsymbol{d}, \boldsymbol{s}, \boldsymbol{f}, \boldsymbol{e}, \boldsymbol{b}*, \boldsymbol{m}*, \boldsymbol{o}*, \boldsymbol{d}*, \boldsymbol{s}*, \boldsymbol{f}*\}$. Тогда *элементарным предложением* в логике Аллена называется выражение $A\ \alpha\ B$, где $\Omega \in \alpha$. *Предложением* называется дизъюнкция $S = (A\ \alpha_i\ B) \vee (A\ \alpha_j\ B) \vee \ldots$ элементарных предложений, где $\alpha_i, \alpha_j, \ldots\ \Omega \in$. Сокращенно предложение S можно записать как $S = A\ \omega\ B$, где $\alpha\{ = \omega_i, \alpha_j, \ldots\}$. Например, $A\ \{\boldsymbol{m}, \boldsymbol{o}, \boldsymbol{d}*\}\ B = (A\ \boldsymbol{m}\ B) \vee (A\ \boldsymbol{o}\ B) \vee (A\ \boldsymbol{d}*B)$.

В целях сокращения $A\ \{\boldsymbol{m}, \boldsymbol{o}, \boldsymbol{d}*\}\ B$ можно записать как $A\ \boldsymbol{mod}*\ B$. Отрицанием предложения $S = A\ \omega\ B$ является предложение $\neg S = A\ \omega\neg\ B$, где $\omega\backslash\Omega = \omega\neg$.

Таблица 1

Базисные интервальные отношения

Интервальные отношения	Схемы	Отношения между концами интервалов
A ***b*** B B ***b**** A	|==A==| |==B==|	$a_2 < b_1$
A ***m*** B B ***m**** A	|==A==| |===B===|	$a_2 = b_1$
A ***o*** B B ***o**** A	|===A===| |====B====|	$a_1 < b_1$; $b_1 < a_2$; $a_2 < b_2$
A ***d*** B B ***d**** A	|==A==| |====B=====|	$b_1 < a_1$; $a_2 < b_2$
A ***s*** B B ***s**** A	|==A==| |====B=====|	$a_1 = b_1$; $a_2 < b_2$
A ***f*** B B ***f**** A	|==A==| |===B=====|	$b_1 < a_1$; $a_2 = b_2$
A ***e*** B	|====A====| |====B====|	$a_1 = b_1$; $a_2 = b_2$

Основными задачами в логике Аллена является оценка совместимости предложений, выраженных в форме «посылка $\rightarrow$ следствие». В работе [11] показано, что темпоральная логика Аллена не полна и недостаточно выразительна. Там же предложены методы частичного устранения этих трудностей. Кроме того, в этой логике речь идет только о непрерывных интервалах, для которых не определено дополнение, поскольку дополнением интервала во многих случаях является разрывный интервал.

В данной работе расматривается основанная на алгебре кортежей [12, 13] математическая модель для логико-интервального анализа, отличающаяся от рассмотренных выше. В ней используются все выразительные средства логических исчислений, и в то же время имеются возможности для анализа метрических и порядковых свойств исследуемых объектов [14].

1. Краткие сведения об алгебре кортежей

Алгебра кортежей (АК) представляет собой алгебраическую систему для моделирования и анализа произвольных многоместных отношений. В ней отражены все выразительные средства исчисления высказываний и предикатов.

Отношения в АК выражаются с помощью четырех типов структур, называемых ***АК-объектами***. Каждый АК-объект погружен в определенное пространство ***атрибутов***. Имена АК-объектов содержат идентификатор, к нему добавляется заключенная в прямые скобки последовательность имен атрибутов, определяющих ***схему отношения***, в которой задан этот АК-объект. Например, имя $R[XYZ]$ означает, что данный АК-объект задан в пространстве $X \times Y \times Z$. Структуры АК матричные, причем в ячейках матриц записываются не элементы, а подмножества доменов соответствующих атрибутов, называемые ***компонентами***. Рассмотрим основные структуры АК – C-системы и D-системы.

C-система записывается в виде матрицы, ограниченной прямыми скобками. Так, $R[XYZ] = \begin{bmatrix} A_1 & * & A_3 \\ B_1 & B_2 & * \end{bmatrix}$ есть C-система, состоящая из компонент, при этом $A_1 \subseteq X$, $A_3 \subseteq Z$ и т. д. Компонента «*» – ***полная компонента***, ее значение есть множество, равное домену соответствующего (по месту расположения) атрибута (например, значение * в первой строке – домен атрибута Y). Данную C-систему можно, используя декартово произведение, преобразовать в обычное отношение следующим образом:

$R[XYZ] = (A_1 \times Y \times A_3) \cup (B_1 \times B_2 \times Z)$.

Образующиеся при этом преобразовании кортежи элементов называются в АК ***элементарными кортежами***.

C-система, состоящая из одной строки, называется ***C-кортежем***.

C-кортеж представляет многоместное отношение, равное декартову произведению его компонент.

Дополнением C-кортежа является ***диагональная C-система***. Например, если $P = [A\ \ B\ \ C]$, то $\overline{P} = \begin{bmatrix} \overline{A} & * & * \\ * & \overline{B} & * \\ * & * & \overline{C} \end{bmatrix}$ (это соотношение выражает в структурах АК одно из свойств декартовых произведений).

D-кортеж, по сути, является сокращенной записью диагональной C-системы, в частности, $\overline{P} = \left]\overline{A} \quad \overline{B} \quad \overline{C}\right[$ (используются перевернутые скобки).

D-система – это матрица компонент, ограниченная перевернутыми прямыми скобками. С помощью D-систем легко вычислять дополнение C-систем. Для этого достаточно заменить все компоненты C-системы их дополнениями, а прямые скобки заменить перевернутыми.

Так, дополнение C-системы $T[XYZ] = \begin{bmatrix} A & * & C \\ D & E & * \end{bmatrix}$ – это D-система $\overline{T}[XYZ] = \left]\begin{matrix} \overline{A} & \varnothing & \overline{C} \\ \overline{D} & \overline{E} & \varnothing \end{matrix}\right[$. Получающаяся при выполнении этой операции компонента $\varnothing$ называется ***пустой компонентой***, она является дополнением компоненты $*$. Компоненты $\varnothing$ и $*$ названы ***фиктивными компонентами***.

D-систему $\overline{T}[XYZ]$ можно преобразовать в обычное отношение с помощью алгоритма:

$\overline{T}[XYZ] = ((\overline{A} \times Y \times Z) \cup (X \times Y \times \overline{C})) \cap ((\overline{D} \times Y \times Z) \cup (X \times \overline{E} \times Z))$.

D-система, содержащая одну строку, является D-кортежем.

Однотипные АК-объекты – это структуры, заданные в одном пространстве атрибутов.

Если компоненты АК объектов представить как одноместные предикаты с соответствующими переменными (например, компоненту A, отнесенную к атрибуту X, можно представить как предикат $A(x)$), то C-кортежу соответствует конъюнкция, а D-кортежу – дизъюнкция одноместных предикатов с разными переменными. Отсюда понятно, что C-системы представляют дизъюнктивную нормальную форму (ДНФ), а D-системы – конъюнктивную нормальную форму (КНФ).

Например, C-систему $T[XYZ] = \begin{bmatrix} A & * & C \\ D & E & * \end{bmatrix}$ можно представить как логическую формулу $(A(x) \wedge C(z)) \vee (D(x) \wedge E(y))$.

Правила выполнения операций объединения и пересечения для C- и D-структур имеют свою специфику. Также в АК обоснованы алгоритмы проверки включения одних типов АК-объектов в другие. Рассмотрим подробно операцию пересечения однотипных C-кортежей. Пусть даны два C-кортежа $P = [A_1\ A_2\ \ldots\ A_n]$ и $Q = [B_1\ B_2\ \ldots\ B_n]$. Тогда

$$P \cap Q = [A_1 \cap B_1\ A_2 \cap B_2 \ldots A_n \cap B_n].$$

Если хотя бы одно пересечение $A_i \cap B_i$ равно $\varnothing$, то $P \cap Q = \varnothing$. Пересечение C-систем вычисляется как объединение непустых пересечений каждого C-кортежа одной C-системы с каждым C-кортежем другой. Объединение C-кортежей за исключением некоторых определенных частных случаев нельзя представить как один C-кортеж. В общем случае справедливо:

$$[A_1 \quad A_2 \quad \ldots \quad A_n] \cup [B_1 \quad B_2 \quad \ldots \quad B_n] = \begin{bmatrix} A_1 & A_2 & \ldots & A_n \\ B_1 & B_2 & \ldots & B_n \end{bmatrix}.$$

Подробно алгоритмы и методы анализа интеллектуальных систем на основе АК, включая квантификацию, логико-вероятностное моделирование и т. д., рассмотрены в [13].

Основным структурообразующим элементом АК является C-кортеж. Более сложные АК-объекты образуются либо за счет объединения C-кортежей (C-системы), либо могут быть преобразованы в их объединение с помощью определенных алгоритмов. Если известны меры компонент C-кортежа, то мера самого C-кортежа вычисляется как произведение мер его компонент. Так, если C-кортеж $R = [A \ B \ C]$ задан в измеримых атрибутах, и меры его компонент равны соответственно $\mu(A)$, $\mu(B)$ и $\mu(C)$, то $\mu(R) = \mu(A) \cdot \mu(B) \cdot \mu(C)$. На этом свойстве основаны, в частности, методы логико-вероятностного анализа в АК [12].

2. Логико-интервальный анализ систем

Рассмотрим класс отношений, у которых значения всех атрибутов упорядочены. Такие отношения встречаются нередко, причем в них присутствуют не только количественные, но и качественные атрибуты, значения которых заданы именами. Кроме того, предусматривается еще один тип атрибутов, в котором значения заданы конечным числом интервалов. Домены таких атрибутов можно преобразовать в линейно упорядоченные конечные множества квантов с помощью изложенного в [13] метода квантования интервалов. Например, некий числовой атрибут в интервале [0,12] представлен в системе тремя интервалами:

$A = [1, 6]$, $B = [3, 10]$ и $C = [3, 8]$.

Для квантования расположим все границы интервалов в порядке возрастания (0, 1, 3, 6, 8, 10, 12) и все интервалы, ограниченные ближайшими точками, будем считать ***квантами***. Можно их обозначить соответствующими цифрами, например, 1 = [0, 1]; 2 = [1, 3]; 3 = [3, 6] и т. д. Тогда интервалы A, B и C можно выразить как множества:

$A = \{2, 3\}$, $B = \{3, 4, 5\}$, $C = \{3, 4\}$.

Для интервалов в таком виде выполнимы все операции и сравнения алгебры множеств. Например, легко проверить, что $C \subset B$ или, что дополнением интервала может оказаться разрывный интервал, например, $\overline{B} = \{1, 2, 6\}$. Для использования измеримых свойств системы можно для каждого кванта вычислить его меру, например для рассматриваемой системы $\mu(1) =1$; $\mu(3) = 3$; $\mu(4) = 2$.

Отношения с упорядоченными атрибутами можно описать с помощью алгебры кортежей. Исследования показали, что структуры АК с упорядоченными значениями атрибутов обладают новыми свойствами [14]. Одним из таких свойств является возможность кластеризации структур АК. В дальнейшем будем представлять в виде квантов не только «интервальные» атрибуты, но и любые упорядоченные атрибуты с конечным числом значений. В таких атрибутах значения (кванты) именуются числами 1, 2, 3, … Если не заданы длины квантов, то мера каждого их них принимается равной 1.

Если меры квантов известны и отличаются друг от друга (например, когда значения атрибута сформированы в результате квантования некоторой совокупности интервалов), то для таких атрибутов необходима ***нормализация интервалов,*** с помощью которой средняя мера всех квантов становится равной 1. Пусть заданы мера L интервала, содержащего все кванты, и меры квантов l_i ($i = 1, 2, \ldots, M$). Тогда при нормализации ***мера нормализованного интервала*** принимается равной M, а ***меры нормализованных квантов*** (μ_i) вычисляются по формуле

$$\mu_i = \frac{l_i \cdot M}{L}. \qquad (2)$$

В дальнейшем при определении степени связности рассматриваемых структур можно считать меры всех квантов одинаковыми и равными 1. При вычислении расстояний необходимо для нормализованных квантов использовать соответствующие меры (2).

Рассмотрим структуру, показанную на рис. 1, где домены атрибутов X и Y представлены системой квантов с именами 1, 2, 3,…, 9. Закрашенная область представляет собой АК-объект, который можно выразить с помощью C-системы $R[XY] = \begin{bmatrix} \{1,2,3,7\} & \{1,8,9\} \\ \{3,4,5,8,9\} & \{2,3,4,7,8\} \end{bmatrix}$.

Рассмотрим возможные соотношения для компонент одного атрибута. Компоненты могут быть непрерывными (например, {3, 4, 5}) либо разрывными (например, {3, 7, 8}). Непрерывная компонента называется ***интервалом***.

Определение 1. ***Расстояние*** $D(I_r, I_s)$ между интервалами I_r и I_s одного атрибута при условии $I_r < I_s$ определяется так:

$D(I_r, I_s) = \min(I_s) - \max(I_r) - 1$.

Например, расстояние между парой интервалов {2,3} и {5,6,7} равно 1, а между парой интервалов {1,2,3} и {4,5} расстояние равно 0.

Чтобы исследовать различные варианты упорядоченности в n-мерном пространстве, а также решать другие задачи в подобных структурах (например, вычисление степени связности объектов, задачи кластеризации и т. д.), необходимо более удобное представление АК-объектов. В данном случае неудобство представления структуры заключается в том, что часть C-кортежей, из которых составлен АК-объект, могут быть разрывными, поскольку их компоненты (например, $\{2, 3, 4, 7, 8\}$) содержат разрывы. Чтобы преобразовать исходную C-систему в эквивалентную C-систему, где каждый C-кортеж не имеет разрывов (например, C-кортеж $[\{1, 2, 3\}\ \ \{8, 9\}]$), нужно воспользоваться следующим соотношением.

Пусть дан C-кортеж $Q[X_1\ X_2\ \ldots\ X_n] = [C_1\ C_2\ \ldots\ C_n]$, в котором каждая компонента C_i ($i = 1, 2, \ldots, n$) разбита на k_i непересекающихся непустых множеств C_{ij} ($j = 1, 2, \ldots, k_i$). Тогда C-система, содержащая $k_1 \times k_2 \times \ldots \times k_n$ C-кортежей, записанных как элементы декартова произведения

$$\{C_{11},\ldots,C_{1k_1}\}\times\{C_{21},\ldots,C_{2k_2}\}\times\ldots\times\{C_{n1},\ldots,C_{nk_n}\}, \tag{3}$$

будет эквивалентна исходному C-кортежу Q.

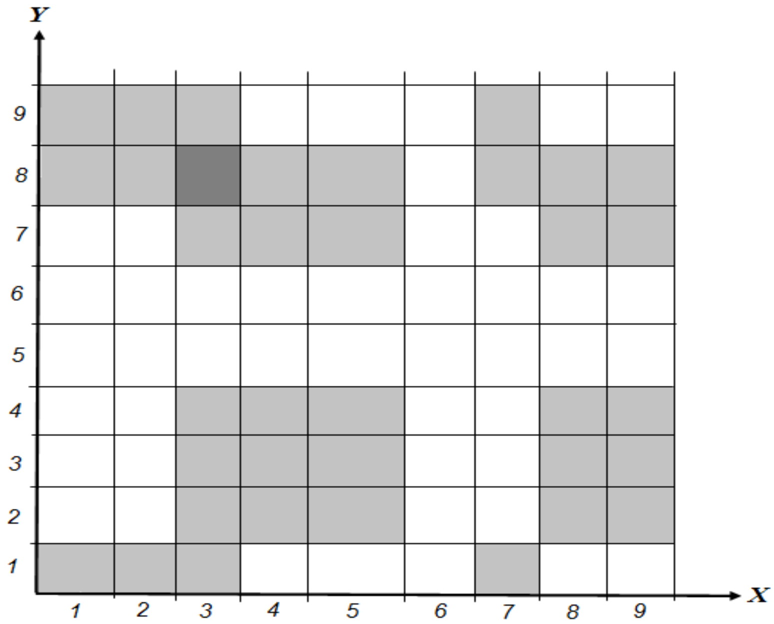

Рис. 1. Изображение C-системы R[XY]

Чтобы преобразовать каждый разрывный C-кортеж C-системы $R[XY]$ в совокупность неразрывных C-кортежей, разобьем каждую разрывную компоненту на совокупность интервалов и преобразуем C-кортежи с разрывными компонентами в C-системы в соответствии с (3). Тогда $R[XY]$ можно представить в «интервальном» виде

$$R[XY] = \begin{bmatrix} \{\{1,2,3\},\{7\}\} & \{\{1\},\{8,9\}\} \\ \{\{3,4,5\},\{8,9\}\} & \{\{2,3,4\},\{7,8\}\} \end{bmatrix},$$

после чего, используя (3), записать $R[XY]$ как C-систему:

$$R[XY]=\begin{bmatrix}\{1.2.3\} & \{1\}\\ \{1,2,3\} & \{8,9\}\\ \{7\} & \{1\}\\ \{7\} & \{8,9\}\\ \{3,4,5\} & \{2,3,4\}\\ \{3,4,5\} & \{7,8\}\\ \{8,9\} & \{2,3,4\}\\ \{8,9\} & \{7,8\}\end{bmatrix}$$

C-кортежи, у которых все компоненты являются интервалами, в *n*-мерном пространстве представляют собой полностью заполненный гиперпрямоугольник, который называется ***брусом***. Рассмотрим различные варианты связей между брусами. В нашем двумерном примере (рис. 1) содержатся три варианта связей между парами брусов:

1) *брусы имеют непустое пересечение*, к таким относится пара [{1,2,3} {8,9}] и [{3,4,5} {7,8}] – результатом их пересечения является элементарный брус [{3} {8}];

2) *брусы соприкасаются гранями* (в данном случае грани одномерные), к таким брусам относятся пары ([{7} {8,9}]; [{8,9} {7,8}]) и ([{1,2,3} {1}]; [{3,4,5} {2,3,4}]);

3) *брусы соприкасаются в одной точке* (пара [{7} {1}] и [{8,9} {2,3,4}]).

Очевидно, что, если расстояние между компонентами хотя бы для одной пары в сравниваемых брусах больше 0, то такая пара брусов не связана. Например, пара [{3,4,5} {7,8}] и [{7} {8,9}] не связана, т. к. расстояние между первыми компонентами равно 1.

Обобщим полученные результаты для брусов произвольной размерности.

Определение 2. ***Связными*** называются пары брусов, у которых каждая пара компонент имеет либо непустое пересечение, либо нулевое расстояние.

Определение 3. Пусть для пары связных брусов в *n*-мерном пространстве *k* – число пар компонент с непустым пересечением. Тогда ***степень связности*** этих брусов есть *k*.

Из Определения 3 ясно, что брусы со степенью связности 0 соприкасаются в одной точке, а пересечение брусов со степенью связности *n* непусто.

Определение 4. Назовем ***гранулами*** брусы или совокупности связанных брусов.

Произвольная совокупность гранул в гранулируемых отношениях является алгебраической системой, в которой определены операции, совпадающие с операциями над АК-объектами, а к отношениям АК (равенство и включение) добавлено отношение порядка, а также рассматриваемые ниже отношения связности и близости.

Рассмотрим метод построения графа связности брусов. Пусть C-система Q в n-мерном пространстве разложена в соответствии с (3) на m брусов B_i ($i = 1, 2, ..., m$). Построим для всех возможных пар этих брусов ***граф связности*** G с взвешенными ребрами, в котором пара брусов (B_i, B_j) соединена ребром веса k, если и только если эти брусы связны и степень их связности равна k. Ясно, что число требуемых проверок для формирования графа G равно $m(m–1)/2$.

Если в графе G имеется хотя бы одно ребро веса k, то ***графом k-й степени связности*** ($0 \le k \le n - 1$) называется невзвешенный граф G_k, в котором сохраняются все ребра графа G веса w ($w \ge k$), а все ребра весом менее k удаляются.

Определение 5. Если в АК-объекте граф G_k существует, то ***множеством гранул k-й степени связности*** ($0 \le k \le n - 1$) называется множество изолированных вершин и связных компонент графа G_k.

Например, для АК-объекта $R[XY]$ (рис. 1) число гранул 1-й степени связности равно 5, а число гранул 0-й степени – 4.

Пусть в структуре АГО решается задача кластеризации и предполагается, что связные компоненты графа G_k являются кластерами. Тогда число связных компонент этого графа равно числу кластеров. Если критерием качества кластеризации является заданный интервал числа возможных кластеров, то нужное решение в некоторых случаях можно получить за счет варьирования степени связности гранул.

Рассмотрим случай с разреженными структурами, когда связи всех порядков учтены, но число гранул в пространстве велико и требуется объединить в один кластер несвязанные близкие гранулы. Для этого можно определить расстояние между парами гранул и произвести кластеризацию объекта, используя известные методы кластеризации по расстоянию [15, 16]. Для определения расстояния между гранулами необходимо для каждой гранулы построить минимальные брусы, включающие эту гранулу (универсумы гранул). Тогда расстояние между парой (U_r, U_s) универсумов гранул вычисляется так.

Каждый универсум гранулы представлен C-кортежем размерности n. Для каждого атрибута X_i (i =1, 2,…, n) определим расстояние D_i между соответствующей парой компонент (если пара компонент имеет непустое пересечение, то расстояние между ними принимается равным 0). Тогда расстояние между парой гранул определяется с помощью формулы

$$D(U_r\ ,\ U_s) = \sqrt{D_1^2 + D_2^2 + \ldots + D_n^2}\ .$$

Величина $D(U_r\ ,\ U_s)$ по сути является геометрическим расстоянием между ближайшими гранями или точками сравниваемых брусов.

Рассмотрим методы упорядочивания в данном пространстве. Хотя значения атрибутов имеют линейный порядок, в пространстве нескольких таких атрибутов часто встречаются нарушения линейного порядка. Например, две «точки» с координатами (2, 4) и (3, 2) несравнимы по порядку. Ясно, что в пространстве с размерностью 2 и более порядок элементов устанавливается как доминирование.

Как известно, ***отношением доминирования*** называется соотношение между векторами $(a_1, a_2, \ldots, a_n) \le (b_1, b_2, \ldots, b_n)$, которое справедливо тогда и только тогда, когда $a_i \le b_i$ для всех $i = 1, 2, \ldots, n$. Для того чтобы анализировать отношение порядка в рассматриваемой модели, сформулируем еще 2 определения.

Определение 6. Отношение ***порядка*** $(I_r, < I_s)$ для пары интервалов (I_r, I_s) одного атрибута справедливо, если и только если $max(I_r) < min(I_s)$.

Если для пары интервалов (I_r, I_s) одного атрибута справедливо $I_r, \cap I_s = \varnothing$, то либо $I_r < I_s$, либо $I_s < I_r$. Отношение порядка можно ослабить, если определить его для пересекающихся интервалов.

Определение 7. Отношением ***квазипорядка*** $(\prec)$ для интервалов (I_r, I_s) при условии $I_r, \cap I_s \neq \varnothing$ называется соотношение $I_r \prec I_s$ такое, что $min(I_r) \le min(I_s)$ и $max(I_r) \le max(I_s)$.

С учетом определений 6 и 7 можно исследовать отношение доминирования для объектов в многомерном пространстве, при этом можно применять строгий порядок (Определение 6), либо более слабый вариант – квазипорядок (Определение 7).

Заключение

В работе описаны и проанализированы возможности применения алгебры кортежей в моделях с упорядоченными или интервальными атрибутами. Введены понятия гранул, определены степени связности между гранулами и предложены способы вычисления расстояний между несвязными гранулами. Полученные зависимости позволяют расширить область применения задач классификации.

Благодарности

Работа частично поддержана РФФИ (проекты №№ 18-07-00132, и 19-08-00079).

Список литературы

1. Moore R.E. Interval analysis. N.J.: Prentice-Hall, Englewood Cliffs, 1966. 159 p.
2. Шокин Ю.И. Интервальный анализ. Новосибирск: Наука, 1981. 112 с.

3. Шарый С.П. Конечномерный интервальный анализ. Новосибирск: Изд-во “XYZ”, 2019. 629 с.
4. Городецкий А.Е., Тарасова И.Л. Нечеткое математическое моделирование плохо формализуемых процессов и систем. СПб.: Изд-во Политехн. ун-та, 2010. 335 с.
5. Городецкий А.Е., Тарасова И. Л., Шкодырев В. П. Математическое моделирование интеллектуальных систем управления: Моделирование детерминированных интеллектуальных систем управления. СПб.: Изд-во Политехн. ун-та, 2016. 181 с.
6. Kaucher E. Interval analysis in the extended interval space IR // Fundamentals of numerical computation (Computer-oriented numerical analysis) / Alefeld G., Grigorieff R.D. (eds.) Wien: Springer, 1980. Pp. 33–49.
7. Левин В.И. Интервальная логика и некоторые ее применения // Логические исследования. 2004. № 1. С. 174–188.
8. Allen J.A. Maintaining knowledge about temporal intervals // Communications of the ACM. 1983. Vol. 20 (11). Pp. 832–843.
9. McNaughton R. A theorem about infinite-valued sentential logic // J. Symb. Logic, 1951. Vol. 16 (1). Pp. 1–13.
10. Гинзбург С.А. Математическая непрерывная логика и изображение функций. М.: Энергия, 1968. 136 с.
11. Плесневич Г.С., Нгуен Тхи Минь Ву. Алгоритмы дедукции для некоторых расширений интервальной логики Аллена // Искусственный интеллект и принятие решений. 2016. № 1. С. 75–88.
12. Кулик Б.А. Вероятностная логика на основе алгебры кортежей // Известия РАН. Теория и системы управления. 2007. № 1. С. 118–127.
13. Кулик Б.А., Зуенко А.А., Фридман А.Я. Алгебраический подход к интеллектуальной обработке данных и знаний. СПб.: Изд-во Политехн. ун-та, 2010. 235 с.
14. Kulik B., Fridman A. Algebra of clusterizable relations // Advances in Intelligent Systems and Computing. 2019. Vol. 836. Pp. 295–304.
15. Everitt B. Cluster analysis. Chichester, West Sussex, UK: Wiley, 2011. 330 p.
16. Журавлев Ю.И., Рязанов В.В., Сенько О.В. Распознавание. Математические методы. Программная система. Практические применения. М: Фазис, 2006. 176 с.

УДК 531.001.362
doi:10.18720/SPBPU/2/id20-159

Фирсов Андрей Николаевич[1],
доктор техн. наук, профессор СПбПУ;
Журавская Анжелика[2],
аспирант

О МЕТОДАХ ТЕОРИИ ПОДОБИЯ И РАЗМЕРНОСТИ

[1,2] Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Россия, [1] anfirs@yandex.ru, [2] hella94@mail.ru

Аннотация. Предлагаемая работа знакомит читателя с проблематикой теории подобия и размерности и подчеркивает её важность при решении прикладных инженерно-технических задач. Статья носит методический характер и, в первую