УДК 517.9

КЛЮЧЕВЫЕ АНТИЦЕПИ РЕШЕТКИ ОПИСАНИЙ ИНТЕРВАЛОВ ПРИЗНАКОВОГО ПРОСТРАНСТВА

© **Ильченко А.В.**

Таврический национальный университет им. В.И. Вернадского
факультет математики и информатики
пр-т Вернадского, 4, г. Симферополь, 95007, Украина

Abstract. The notion of the key antichain for the lattice of characteristic space interval descriptions, key antichains family property, the algorithm of the key antichains construction are conside the paper.

Введение

Алгоритмы и методы кластерного анализа во многом составляют основу приложений анализа данных. Рост объема данных, подлежащих обработке, предъявляет ряд требований к используемым алгоритмам кластеризации: возможность находить кластеры в пространстве большой размерности, наглядность, легкость интерпретации полученных результатов, отсутствие необходимости приведения исходных данных к какому-либо каноническому виду и так далее.

В некоторых случаях выполнение многих из этих требований могут обеспечить алгоритмы и методы анализа формальных понятий [1], [2].

Целью статьи является рассмотрение понятия ключевой антицепи решетки описаний интервалов признакового пространства, изучение свойств семейства ключевых антицепей, построение и обоснование алгоритма поиска ключевых антицепей, основанного на использовании некоторых конструкций теории решеток, классического и обощенного анализа формальных понятий.

1. Признаковое пространство и его области

S – признаковое пространство, каждое измерение (признак, атрибут) которого является конечным линейно упорядоченным множеством.

G – множество объектов. Каждому объекту g из множества G соответствует некоторая точка s признакового пространства S – проекция объекта g в признаковое пространство S. Значениями координат такой точки являются значения, принимаемые атрибутами на объекте g.

Точки признакового пространства, которые являются проекциями объектов из множества G, удобно называть «занятыми» точками ($\bar{0}$-точками) пространства S.

$H_{\bar{0}}$ – проекция множества G в признаковое пространство S («занятая» область, $\bar{0}$-область пространства S).

$H_0 = S \setminus H_{\bar{0}}$ – дополнение проекции H_0 до пространства S («свободная» область, $\bar{0}$-область пространства S).

$S = H_0 \cup H_{\bar{0}}$ – представление признакового пространства в виде объединения «свободной» и «занятой» областей.

2. Сокращенная интервальная структура признакового пространства [4]

Интервал I, содержащийся в области H, называется максимальным интервалом области H, если не существует его собственного надинтервала, содержащегося в области H.

Множество всех максимальных интервалов области H называется сокращенной интервальной структурой этой области.

$I_m(H)$ – обозначение множества всех максимальных интервалов области H.

Каждую область H признакового пространства S можно представить в виде объединения интервалов ее сокращенной интервальной структуры.

Множество всех максимальных интервалов, определяемых «свободной» и «занятой» областями пространства S, называется сокращенной интервальной структурой признакового пространства, порождаемой проекцией множества G в пространство S.

Пространство S можно представить в виде объединения интервалов его сокращенной интервальной структуры.

3. Решетка интервалов признакового пространства

$P(S)$ – множество всех подмножеств признакового пространства S. Отношение включения «$\subseteq$» определяет отношение предшествования (порядка) на этом множестве. Известно [3], что множество $P(S)$, упорядоченное таким образом, является полной решеткой. Нижняя и верхняя грани такой решетки совпадают с теоретико-множественными операциями пересечения и объединения соответственно.

$I(S)$ – семейство всех интервалов признакового пространства. Это семейство обладает свойством замыкания [3]. Поэтому $I(S)$, рассматриваемое вместе с унаследованным отношением предшествования «$\subseteq$», является полной решеткой. В этой решетке нижняя грань совпадает с теоретико-множественным пересечением интервалов: $I \sqcap J = I \cap J$, а верхняя грань $I \sqcup J$ – это наименьший интервал, содержащий интервалы I и J (интервальная оболочка, интервальное замыкание объединения интервалов I и J). Наименьший элемент этой решетки – пустое множество, наибольший элемент – признаковое пространство S, рассматриваемое как интервал.

$I^{\cap}(S) \subseteq I(S)$ – подмножество интервалов, неразложимых в пересечение ($\cap$-неразложимых).

$I^{\cap}(I) = \{J \in I^{\cap}(S) | I \subseteq J\}$ – подмножество неразложимых в пересечение интервалов, содержащих интервал I.

$I^{\cap}_{min}(I) = min\ I^{\cap}(I)$ – антицепь из минимальных неразложимых в пересечение интервалов, содержащих интервал I.

В силу того, что решетка $(I(S), \subseteq)$ является конечной, семейство $I^{\cap}(S)$ является плотным по пересечению подмножеством этой решетки [3]. Поэтому каждый интервал $I \in I(S)$ представим в виде пересечения (разложим в пересечение) интервалов некоторого подмножества семейства $I^{\cap}(S)$, например

$$I = \cap I^{\cap}(I) \tag{1}$$

В (1) для вычисления пересечения достаточно использовать только минимальные элементы подмножества $I^{\cap}(I)$:

$$I = \cap I^{\cap}_{min}(I) \tag{2}$$

То есть, справедливо

Утверждение 1. *Для каждого интервала семейства $I(S)$ существует антицепь из $\cap$-неразложимых интервалов, являющаяся $\cap$-разложением этого интервала:*

$$\forall I \in I(S) \exists A \subseteq I^{\cap}(S) : A - \text{антицепь и } I = \cap A. \tag{3}$$

Например, антицепь $I^{\cap}_{min}(I)$.

Примечание 1. Для $\cap$-неразложимого интервала в качестве такой антицепи рассматривается одноэлементное подмножество, элементом которого является сам этот интервал.

Утверждение 1 справедливо для элементов любой конечной решетки. Для тех интервалов решетки $(I(S), \subseteq)$, которые не являются пустым множеством, справедливо более сильное утверждение.

Утверждение 2. *Для каждого интервала семейства $I(S)$, не являющегося пустым множеством, существует единственная антицепь из $\cap$-неразложимых интервалов, являющаяся $\cap$-разложением этого интервала:*

$$\forall I \in I(S) : I \neq \emptyset \Rightarrow \exists! A \subseteq I^{\cap}(S) : A - \text{антицепь и } I = \cap A \tag{4}$$

Очевидно, что $I^{\cap}_{min}(I)$ является такой антицепью.

4. Решетка описаний интервалов признакового пространства

Фиксируется некоторый способ описания интервалов признакового пространства. При этом, $d(I)$ – это обозначение описания интервала I, $I(d)$ – обозначение интервала, определяемого описанием d. Описание интервала рассматривается как предикат, определенный на пространстве S. Если сокращенная интервальная структура области H известна, то дизъюнкция описаний интервалов семейства $I_m(H)$ определяет сокращенную дизъюнктивную нормальную форму области H:

$$D_c(H) = \underset{I \in I_m(H)}{\vee} d(I) \,.$$

Считается, что знание сокращенной ДНФ одной из областей признакового пространства достаточно для определения сокращенной ДНФ другой области.

$D(S)$-обозначение семейства описаний всех интервалов признакового пространства S. Способ описания интервалов выбирается так, что каждому интервалу соответствует единственное его описание и каждое описание определяет единственный интервал, то есть соответствие между семействами $D(S)$ и $I(S)$ является взаимно однозначным.

На множестве $D(S)$ определяется отношение предшествования «$\leq$», порождаемое отношением предшествования «$\subseteq$», определенным на множестве интервалов $I(S)$:

$$d(I) \leq d(J), \text{если} I \supseteq J \tag{5}$$

Отношение «$\leq$» определяет частичный порядок на элементах семейства $D(S)$.

В силу определения $D(S)$ и способа определения отношения порядка «$\leq$» на этом множестве, $D(S)$ и $I(S)$ – двойственные множества. Поэтому $(D(S), \leq)$ – полная решетка. В этой решетке нижняя грань $d(I) \sqcap d(J)$ является описанием наименьшего интервала, содержащего объединение интервалов I и J, верхняя грань $d(I) \sqcup d(J)$ – это описание интервала, являющегося пересечением интервалов I и J.

Наименьший элемент этой решетки – описание признакового пространства, рассматриваемого как интервал. Наибольший элемент этой решетки – описание пустого множества, рассматриваемого как интервал.

$D^{\sqcup}(S) \subseteq D(S)$ – подмножество $\sqcup$-неразложимых описаний семейства $D(S)$.

$D^{\sqcup}(d) = \{m \in D^{\sqcup}(S) | \ m \leq d\}$ – подмножество $\sqcup$-неразложимых описаний, предшествующих описанию d.

$D^{\sqcup}_{max}(d) = max \ D^{\sqcup}(d)$ – антицепь из максимальных $\sqcup$-неразложимых описаний, предшествующих описанию d.

В силу отмеченной двойственности, в решетке $(D(S), \leq)$ справедливо утверждение, двойственное утверждению 2:

Утверждение 3. *Для каждого описания семейства $D(S)$, не являющегося описанием пустого множества, существует единственная антицепь из $\sqcup$-неразложимых описаний, являющаяся $\sqcup$-разложением этого описания:*

$$\forall d \in D(S) : I(d) \neq \varnothing \Rightarrow \exists! A \subseteq D^{\sqcup}(S) : A - \text{антицепь и } d = \sqcup \ A \tag{6}$$

Очевидно, что $D^{\sqcup}_{max}(d)$ является такой антицепью.

5. Соответствия Галуа и оператор замыкания

Объекту множества G соответствует некоторая точка пространства S. Точка признакового пространства может рассматриваться как одноточечный интервал и, следовательно, ей соответствует некоторое описание из множества $D(S)$. Поэтому, можно считать, что каждому объекту g множества G соответствует описание $d(g) \in D(S)$ одноточечного интервала, порождаемого проекцией объекта в признаковое пространство.

$P(G)$ – множество всех подмножеств множества G. Рассматриваются отображения $\varphi : P(G) \to D(S)$ и $\psi : D(S) \to P(G)$, которые определены следующим образом.

$\varphi(A)$-наибольшее (по отношению предшествования «$\leq$») интервальное описание, предшествующее описаниям элементов подмножества A:

$$\varphi(A) = \sqcap_{g \in A} d(g), \tag{7}$$

$\psi(A)$-подмножество тех элементов множества G, описания которых следуют за описанием d (проекции которых содержатся в интервале $I(d)$)

$$\psi(d) = \{g \in G | \ d \leq d(g)\}, \tag{8}$$

Пара отображений (φ, ψ) представляет собой соответствие Галуа [3] между решетками $(P(G), \subseteq)$ и $(D(S), \leq)$, а тройка $(P(G), D(S), \varphi)$ может рассматриваться как решеточный контекст [2].

Композиция отображений (φ, ψ) определяет отображение γна множестве $D(S)$:

$$\gamma = \varphi \circ \psi : D(S) \to D(S) \tag{9}$$

Отображение γ действует следующим образом: $\gamma(d)$ – это описание наименьшего (по включению подмножеств) интервала, содержащего те же точки «занятой» области пространства S, что и интервал, определяемый описанием d.

Так определенное отображение γ является оператором замыкания [3], то есть для выполняются следующие свойства:

$a)$	$d_1 \leq d_2 \Rightarrow \gamma(d_1) \leq \gamma(d_2)$	монотонность
$b)$	$d \preceq \gamma(d)$	экстенсивность
$c)$	$\gamma(\gamma(d)) = \gamma(d)$	идемпотентность

Элемент $d \in D(S)$ γ-замкнут, если $\gamma(d) = d$. Оператор γ порождает на элементах семейства $D(S)$ отношение γ-равносильности $R_\gamma : d_1 R_\gamma d_2 \Leftrightarrow \gamma(d_1) = \gamma(d_2)$. Отношение R_γ – это отношение эквивалентности. Оно разбивает множество интервальных описаний $D(S)$ на классы. В один класс попадают элементы, γ-замыкания которых совпадают: $[d] = \{e \in D(S) |\ \gamma(e) = \gamma(d)\}$. Такой класс, вместе с унаследованным отношением предшествования «$\leq$», является упорядоченным множеством, в котором существует единственный максимальный элемент. Это γ-замкнутый элемент класса. Минимальный элемент класса называется ключевым элементом класса. Ключевых элементов может быть несколько.

Понятие «ключевой элемент» используется в некоторых алгоритмах анализа формальных понятий [1]. Далее будет представлен один из алгоритмов поиска ключевых элементов, модифицированный для случай решетки интервальных описаний.

Ключевые элементы того класса, наибольший элемент которого является единицей решетки $(D(S), \leq)$, – это и есть описания максимальных интервалов «свободной» области. Можно воспользоваться алгоритмом, который отыскивает все ключевые элементы, порождаемые оператором γ, и, затем, среди найденных ключевых элементов отобрать те, для которых $\gamma(d) = 1$.

6. Некоторые свойства ключевых элементов конечной решетки

$K \subseteq D(S)$ – множество ключевых элементов решетки.

$K^{\sqcup} = D^{\sqcup}(S) \sqcap K$ – подмножество $\sqcup$-неразложимых ключевых элементов решетки.

$K(d) = \{k \in K |\ k < d\}$ – подмножество ключевых элементов, строго предшествующих элементу d.

$K_{max}(d) = max\ K(d)$ – антицепь из максимальных ключевых элементов, строго предшествующие элементу d.

$K^{\sqcup}_{max}(d) = max\ \{k \in K^{\sqcup} |\ k \leq d\}$-антицепь из максимальных $\sqcup$-неразложимых ключевых элементов, предшествующие элементу d.

Утверждение 4. *Элемент решетки является ключевым элементом тогда и только тогда, когда его γ-замыкание не совпадает с γ-замыканием ни одного из максимальных строго предшествующих ему ключевых элементов:*

$$d \in K \Leftrightarrow \gamma(d) \neq \gamma(k), \forall k \in K_{max}(d). \tag{10}$$

Это очевидное следствие определения ключевого элемента.

Теорема 1. *Каждый $\sqcup$-разложимый ключевой элемент конечной решетки представим в виде верхней грани строго предшествующих ему ключевых элементов:*

$$\forall d \in K \setminus K^{\sqcup} : d = \sqcup K(d). \tag{11}$$

Доказательство. Пусть $d \in K \setminus K^{\sqcup}$ – $\sqcup$-разложимый ключевой элемент решетки.

Верхнюю грань подмножества $K(d)$ ключевых элементов решетки, строго предшествующих элементу d, обозначим k:

$k = \sqcup K(d)$.

По свойству верхней грани, $k \leq d$ и, следовательно, по свойству монотонности оператора замыкания

$$\gamma(k) \leq \gamma(d), \tag{12}$$

и по свойству верхней грани для сравнимых элементов

$$k \sqcup d = d. \tag{13}$$

Допустим $k \neq d$. Тогда

$$k < d. \tag{14}$$

Подмножество $\sqcup$-неразложимых элементов конечной решетки является $\sqcup$-плотным подмножеством этой решетки [3] и, следовательно, каждый элемент d решетки представим в виде $d = d \sqcup D^{\sqcup}(d)$.

Отсюда, $d \sqcup k = \sqcup\ D^{\sqcup} \sqcup k$ и

$d = \sqcup\ D^{\sqcup}(d) \sqcup k$	учтено равенство 13
$= m_1 \sqcup ...m_n \sqcup k$	считаем, что $D^{\sqcup}(d) = \{m_1, ..., m_n\}$
$\leq \gamma(m_1) \sqcup ... \sqcup \gamma(m_n) \sqcup \gamma(k)$	экстенсивность оператора замыкания
$= \gamma(m_1) \sqcup ...\gamma(m_n) \sqcup \gamma(k)$	учтено, что $\gamma(k_j) = \gamma(m_j)$, где k_j-ключевой элемент того класса, в состав которого входит $m_j, j = 1, ..., n$.
$= \gamma(k)$.	учтено, что каждый $k_j \in K(d)$. Поэтому $k_j \leq k \Rightarrow \gamma(k_j) \leq \gamma(k) \Rightarrow \gamma(k_j) \sqcup \gamma(k) = \gamma(k)$.

Таким образом, $d \leq \gamma(k)$.

Из этого неравенства, в силу монотонности оператора следует, что

$$\gamma(d) \leq \gamma(\gamma(k)). \tag{15}$$

Идемпотентность оператора γ позволяет преобразовать неравенство (15) к виду

$$\gamma(k) \leq \gamma(d). \tag{16}$$

Совместное выполнение неравенств (12) и (16) влечет равенство

$$\gamma(k) = \gamma(d). \tag{17}$$

Совместное выполнение неравенства (14) и равенства (14) приводит к тому, что элемент d не является минимальным в своем классе, то есть не является ключевым элементом. Противоречие с условием. Следовательно $d = \sqcup K(d)$.

□

В (11) для вычисления верхней грани достаточно использовать только максимальные элементы подмножества $K(d)$. Поэтому справедливо

Следствие 1. Для каждого $\sqcup$-разложимого ключевого элемента конечной решетки антицепь из максимальных строго предшествующих ему ключевых элементов является $\sqcup$-разложением этого элемента:

$$\forall d \in K \setminus K^{\sqcup} : d = \sqcup\ K_{max}(d). \tag{18}$$

Следствие 2. Для каждого $\sqcup$-разложимого ключевого элемента конечной решетки антицепь из максимальных $\sqcup$-неразложимых строго предшествующих ему ключевых элементов является $\sqcup$-разложением этого элемента:

$$\forall d \in K \setminus K^{\sqcup} : d = \sqcup\ K^{\sqcup}_{max}(d). \tag{19}$$

Примечание 2. Для $\sqcup$-неразложимого ключевого элемента решетки в качестве такой антицепи рассматривается одноэлементное подмножество, элементом которого является сам этот ключевой элемент.

Для решетки интервальных описаний $(D(S), \leq)$ следствие 2, с учетом утверждения 3 и примечания 2, позволяет сформулировать следующее

Утверждение 5. *Для каждого ключевого описания решетки $(D(S), \leq)$, не являющегося описанием пустого множества, существует единственная антицепь из $\sqcup$-неразложимых ключевых описаний, являющаяся $\sqcup$-разложением этого описания:*

$$\forall d \in K : I(d) \neq \emptyset \Rightarrow \exists!\ A \subseteq K^{\sqcup} : A - \text{антицепь и } d = \sqcup\ A. \tag{20}$$

Очевидно, что $K^{\sqcup}_{max}(d)$ является такой антицепью.

7. Ключевые антицепи

Подмножество B непосредственно предшествует (по включению) множеству A (обозначение «$\prec$»), если B является подмножеством множества A и мощность B на единицу меньше мощности A:

$B \prec A \Leftrightarrow B \subset A$ и $\mid B \mid = \mid A \mid -1$.

Определение 1. Состоящая из $\sqcup$-неразложимых ключевых элементов антицепь, верхняя грань которой является ключевым описанием, называется ключевой антицепью.

$A(K^{\sqcup})$-семейство всех антицепей из $\sqcup$-неразложимых ключевых элементов семейства $D(S)$.

$KA(K^{\sqcup}) = \{A \in A(K^{\sqcup}) | \sqcup A \in K\}$-семейство всех ключевых антицепей.

Утверждение 6. *Если d – $\sqcup$-разложимый ключевой элемент, k-максимальный ключевой элемент, строго предшествующий элементу d, то ключевая антицепь элемента k непосредственно предшествует ключевой антицепи элемента d:*
$d \in K \setminus K^{\sqcup}$ и $k \in K_{max}(d) \Rightarrow K^{\sqcup}_{max}(k) \prec K^{\sqcup}_{max}(d)$.

Утверждение 7. *Если антицепь X, состоящая из $\sqcup$-неразложимых ключевых элементов, не является ключевой антицепью, то существует ее строгое подмножество, которое является ключевой антицепью, порождающей ключевой элемент того же класса, которому принадлежит элемент $\sqcup X$:*

$$X \in A(K^{\sqcup}) \text{ и } \sqcup X \notin K \Rightarrow \exists Z \subset X : Z \in KA(K^{\sqcup}) \text{ и } \gamma(\sqcup Z) = \gamma(\sqcup X). \tag{21}$$

Доказательство. Пусть $X \in A(K^{\sqcup})$ и $\sqcup X \notin K$. Тогда существует ключевой элемент k, который строго предшествует $\sqcup X$ и находится в том же классе, что и $\sqcup X$.

$Z = k^{\triangledown} \cap X$-антицепь из всех максимальных $\sqcup$-неразложимых ключевых элементов, предшествующих ключевому элементу k. Следовательно, в силу утверждения 5, Z является ключевой антицепью для элемента k. □

8. Порядковый идеал ключевых антицепей

Теорема 2. *Семейство ключевых антицепей $KA(K^{\sqcup})$ является порядковым идеалом на $(P(K^{\sqcup}), \subseteq)$, то есть, если $Y \in KA(K^{\sqcup})$ и $X \subseteq Y$, то $X \in KA(K^{\sqcup})$.*

Доказательство. Пусть $Y \in KA(K^{\sqcup})$ и $X \subseteq Y$. Допустим, что $X \notin KA(K^{\sqcup})$. Тогда, в соответствии с утверждением 5,

$$\exists Z \subset X : Z \in KA(K^{\sqcup}) \text{ и } \gamma(\sqcup Z) = \gamma(\sqcup X). \tag{22}$$

Покажем, что в этом случае $\gamma(\sqcup Y) = \gamma(\sqcup(Y \setminus (X \setminus Z)))$. Для этого понадобятся несколько вспомогательных утверждений.

$$1. \qquad X \setminus Z \subseteq X \Rightarrow \gamma(\sqcup(X \setminus Z)) \leq \gamma(\sqcup Z). \tag{23}$$

Доказательство.

$X \setminus Z \subseteq X \Rightarrow$
$\Rightarrow \sqcup(X \setminus Z) \leq \sqcup X$ свойство верхней грани
$\Rightarrow \gamma(\sqcup(X \setminus Z)) \leq \gamma(\sqcup X)$ монотонность γ
$\Rightarrow \gamma(\sqcup(X \setminus Z)) \leq \gamma(\sqcup Z)$ учтено (22) □

$$2. \qquad Z \subseteq Y \setminus (X \setminus Z) \Rightarrow \gamma(\sqcup Z) \leq \gamma(\sqcup(Y \setminus (X \setminus Z))) \tag{24}$$

Доказательство.

$Z \subseteq Y \setminus (X \setminus Z) \Rightarrow$
$\Rightarrow \sqcup Z \leq \sqcup(Y \setminus (X \setminus Z))$ свойство верхней грани
$\Rightarrow \gamma(\sqcup Z) \leq \gamma(\sqcup(Y \setminus (X \setminus Z)))$ монотонность γ □

$$3. \qquad \sqcup Y \leq \gamma(\sqcup(Y \setminus (X \setminus Z))) \tag{25}$$

Доказательство.

$$\begin{aligned}
&Y = (Y \setminus (X \setminus Z)) \cup (X \setminus Z) \Rightarrow \\
&\sqcup Y = \sqcup((Y \setminus (X \setminus Z)) \cup (X \setminus Z)) \\
&\quad = \sqcup(Y \setminus (X \setminus Z)) \sqcup (\sqcup(X \setminus Z)) && \text{ассоциативность } \sqcup \\
&\quad \leq \gamma(\sqcup(Y \setminus (X \setminus Z))) \sqcup \gamma(\sqcup(X \setminus Z)) && \text{монотонность } \gamma \\
&\quad \leq \gamma(\sqcup(Y \setminus (X \setminus Z))) \sqcup \gamma(\sqcup Z) && \text{учтено (22)} \\
&\quad \leq \gamma(\sqcup(Y \setminus (X \setminus Z))) \sqcup \gamma(\sqcup(Y \setminus (X \setminus Z))) && \text{учтено (24)} \\
&\quad = \gamma(\sqcup(Y \setminus (X \setminus Z)))
\end{aligned}$$

□

$$4. \qquad \sqcup\, Y \leq \gamma(\sqcup(Y \setminus (X \setminus Z))) \Rightarrow \gamma(\sqcup Y) \leq \gamma(\sqcup(Y \setminus (X \setminus Z))) \tag{26}$$

Доказательство.

$$\begin{aligned}
&\sqcup Y \leq \gamma(\sqcup(Y \setminus (X \setminus Z))) \Rightarrow \\
&\quad \Rightarrow \gamma(\sqcup Y) \leq \gamma(\gamma(\sqcup(Y \setminus (X \setminus Z)))) && \text{монотонность } \gamma \\
&\quad \Rightarrow \gamma(\sqcup Y) \leq \gamma(\sqcup(Y \setminus (X \setminus Z))) && \text{идемпотентность } \gamma
\end{aligned}$$

□

$$5. \qquad Y \setminus (X \setminus Z) \subseteq Y \Rightarrow \gamma(\sqcup(Y \setminus (X \setminus Z))) \leq \gamma(\sqcup Y) \tag{27}$$

Доказательство.

$$\begin{aligned}
&Y \setminus (X \setminus Z) \subseteq Y \Rightarrow \\
&\quad \Rightarrow \sqcup(Y \setminus (X \setminus Z)) \leq \sqcup Y && \text{свойство верхней грани} \\
&\quad \Rightarrow \gamma(\sqcup(Y \setminus (X \setminus Z))) \leq \gamma(\sqcup Y) && \text{монотонность } \gamma
\end{aligned}$$

□

Совместное выполнение неравенств (26) и (27) влечет равенство

$$\gamma(\sqcup(Y \setminus (X \setminus Z))) = \gamma(\sqcup Y) \tag{28}$$

Так как $Y \setminus (X \setminus Z) \subset Y$, то равенство (28) означает, что Y не является ключевой антицепью. Противоречие с условием теоремы. □

Тот факт, что $KA(K^{\sqcup})$ является порядковым идеалом на $(P(K^{\sqcup}), \subseteq)$, влечет все нижеследующие определения и утверждения текущего раздела.

$A_{-}(A) = \{B \in A(K^{\sqcup}) | B \prec A\}$ – подмножество антицепей из $\cap$-неразложимых ключевых элементов, непосредственно предшествующих антицепи A.

Определение 2. Антицепь $A \in A(K^{\sqcup})$ называется кандидатом (кандидатом в ключевую антицепь), если все ее строгие подмножества являются ключевыми антицепями и ее верхняя грань не является описанием пустого множества.

Утверждение 8. *Антицепь $A \in A(K^{\sqcup})$ является кандидатом тогда и только тогда, когда каждое непосредственно предшествующее ей подмножество является ключевой антицепью и ее верхняя грань не является описанием пустого множества:*

$A \in A(K^{\sqcup})$ – кандидат $\Leftrightarrow A_{-}(A) \subseteq KA(K^{\sqcup})$ и $\sqcup\, A \neq d(\varnothing)$.

Утверждение 9. *Антицепь-кандидат является ключевой антицепью тогда и только тогда, когда γ-замыкание ее верхней грани не совпадает с γ-замыканием верхней грани ни одной из непосредственно предшествующей ей антицепей.*

Определение 3. Подмножество $A \in P(K^{\sqcup})$ называется подмножеством-потенциальным кандидатом, если хотя бы одно из непосредственно предшествующих ему подмножеств является ключевой антицепью:

$A \in P(K^{\sqcup})$ – подмножество-потенциальный кандидат $\Leftrightarrow A_{-}(A) \cap KA(K^{\sqcup}) \neq \emptyset$.

Определение 4. Антицепь $A \in A(K^{\sqcup})$ называется антицепью-потенциальным кандидатом, если хотя бы одно из непосредственно предшествующих ему подмножеств является ключевой антицепью:

$A \in A(K^{\sqcup})$ – антицепь-потенциальный кандидат $\Leftrightarrow A_{-}(A) \cap KA(K^{\sqcup}) \neq \emptyset$.

Примечание 3. Требование $\sqcup A \neq d(\emptyset)$ включается в определения 2, 3 или 4 в зависимости от того, где удобнее или проще вычислить $\sqcup A$. Кроме этого, выражение $\sqcup A \neq d(\emptyset)$ часто удобнее записывать в виде $I(\sqcup A) \neq \emptyset$.

9. Алгоритм поиска ключевых антицепей

Обозначения алгоритма

Обозначение	Пояснение
i	Номер уровня, номер итерации
SC_i	Семейство подмножеств – потенциальных кандидатов i-го уровня
AC_i	Семейство антицепей – потенциальных кандидатов i-го уровня
C_i	Семейство антицепей – кандидатов i-го уровня
K_i	Семейство ключевых антицепей i-го уровня

Псевдокод алгоритма

Вход.	S – признаковое пространство;
	$H_{\bar{0}} \subseteq S$ – проекция множества объектов G в пространство S.
Выход.	$KA(K^{\sqcup})$ – семейство всех ключевых антицепей.
Метод.	
1)	$K_0 := \{d(S)\}$
2)	$C_1 := \{A = \{d\} \mid d \in D^{\sqcup}(S)$ и $I(\sqcup A) \neq \emptyset\}$
3)	$K_1 := \{A \in C_1 \mid \gamma(\sqcup A) \neq \gamma(\sqcup B), \forall B \in K_{max}(A)\}$
4)	$i := 1$
5)	Цикл пока $K_i \neq \emptyset$
6)	$i++$
7)	$SC_i := \{A = \{d_1, ..., d_i\} \subseteq K^{\sqcup} \mid A_{-}(A) \cap K_{i-1} \neq \emptyset\}$
8)	$C_i := \{A \in AC_i \mid A_{-}(A) \subseteq K_{i-1}$ и $I(\sqcup A) \neq \emptyset\}$
9)	$K_i := \{A \in C_i \mid \gamma(\sqcup B) \neq \gamma(\sqcup A), \forall B \in A_{-}(A)\}$
10)	Конец цикла
11)	Возврат $\bigcup_{j=0}^{i-1} K_j$

10. Пояснения к псевдокоду алгоритма

Шаг 1. Описание признакового пространства рассматривается как одноэлементная ключевая антицепь.

Шаг 2. Все $\sqcup$-неразложимые описания, не являющиеся описаниями пустого множества, рассматривают как кандидаты (определение 2).

Шаг 3. Те кандидаты, γ-замыкания верхних граней которых не совпадает с γ-замыканием верхней грани ни одного из строго предшествующих им максимальных ключевых элементов, являются ключевыми элементами (утверждение 4).

Шаг 4. Инициализируется счетчик итераций (счетчик уровней).

Шаг 5. Условие $K_i \neq \emptyset$ определяет выполнение условия очередной итерации алгоритма, реализующей построение необходимых конструкций следующего уровня из ключевых антицепей предшествующего уровня.

Шаг 6 определяет номер следующей итерации (номер следующего уровня).

Шаг 7 формирует семейство подмножеств – потенциальных кандидатов, непосредственно следующих за ключевыми антицепями семейства K_{i-1} (определение 3).

Шаг 8 отбирает из подмножеств – потенциальных кандидатов те, которые являются антицепями, формируя семейство антицепей – потенциальных кандидатов i-го уровня (определение 4).

Шаг 9 отбирает из антицепей – потенциальных кандидатов те, которые являются кандидатами (утверждение 8).

Шаг 10 отбирает из кандидатов тех, которые являются ключевыми антицепями (утверждение 9).

Шаг 12 объединяет результаты, полученные на разных уровнях, и возвращает это объединение как результат работы.

Примечание 4. На шаге 3 требуется знать множество $K_{max}(A)$. Здесь антицепь $A = \{d\}$ является одноэлементным подмножеством, состоящим из $\sqcup$-неразложимого интервального описания. Поэтому, для этой антицепи, множество $K_{max}(A)$ состоит из одного максимального ключевого элемента, строго предшествующего элементу d.

Для проверки условия этого шага удобно выписать все максимальные цепи, порождаемые $\sqcup$-неразложимыми описаниями $D^{\sqcup}(S)$ семейства $D(S)$. Тогда очевидно, что $K_{max}(A)$ состоит из максимального ключевого описания, строго предшествующего элементу d в цепи, содержащей d.

После шага 3 полученное множество K_1 – это фактически $K^{\sqcup}$ – множество $\sqcup$-неразложимых ключевых описаний.

Примечание 5. Процедура генерации элементов семейства SC_i (шаг 7) здесь не фиксируется. Можно применять любой вариант, обеспечивающий получение семейства SC_i, например, может быть применена процедура, использующая тот или иной линейный порядок на элементах множества $\sqcup$-неразложимых ключевых описаний [1].

Заключение

В статье рассмотрена решетка описаний интервалов признакового пространства. Доказана теорема о представимости каждого $\sqcup$-разложимого ключевого элемента этой решетки в виде верхней грани строго предшествующих ему ключевых элементов. Следствием этой теоремы является утверждение о единственности $\sqcup$–разложения ключевых элементов, не являющихся описанием пустого множества, в антицепь из $\sqcup$-неразложимых ключевых элементов.

Приведено определение понятия ключевой антицепи и исследованы свойства семейства ключевых антицепей. Вчастности показано, что семейство ключевых антицепей является порядковым идеалом на множестве-степени $\sqcup$-неразложимых ключевых элементов, упорядоченных отношением включения.

Предложен алгоритм поиска всех ключевых антицепей решетки интервальных описаний.

Список литературы

1. *Ganter, B., Wille, R.* Formal Concept Analysis – Mathematical Foundations. Springer-Verlag, Berlin., 1999.
2. *Gugisch, R.* Lattice Contexts – a Generalization in Formal Concept Analysis. Handuot to ICCS 2000, Darmstadt(2000) http://www.mathe2.uni-bayreuth.de/ralfg/papers/diplom.ps.gz.
3. *Биркгоф Г.* Теория решеток: Пер. с англ. – М.: Наука, 1984. – 568 с.
4. *Ильченко А.В.* Компактная компонентная и сокращенная интервальная структуры признакового пространства, порождаемые эмпирическими данными // Сб. Таврический вестник информатики и математики. – 2005. – №2. – С. 126-142.

Статья поступила в редакцию 30.04.2008