

## АЛГОРИТМ ТАКСОНОМИИ ВЕРШИН ВЗВЕШЕННОГО ГРАФА

Л.И. Макаров

### В в е д е н и е

Во многих областях научных исследований применяются графовые модели, в которых вершины графа соответствуют исследуемым объектам, а каждое ребро - отношению (связи) между парой объектов, задаваемому количественно в виде действительной величины - веса (длины) ребра. При этом естественно возникает задача таксономии вершин взвешенного графа, т.е. разбиения множества его вершин на таксоны (классы) - подмножества близких (в смысле длин соединяющих ребер) вершин графа [1,2].

В органической химии, например, исследование зависимости "структура-свойство" химических соединений основывается на гипотезе о том, что свойства соединений определяются соответствующими структурными фрагментами их молекул и структурно подобные химические соединения обладают сходными свойствами [3]. В качестве моделей, описывающих строение молекул химических соединений, используются молекулярные графы. Для определения меры подобия графов можно ввести величину расстояния между парой графов  $G$  и  $H$ , например, одну из следующих величин:

$$\sigma(G, H) = P_G + P_H - 2P_{GH}$$

$$\rho(G, H) = 1 - \frac{2P_{GH}}{P_G + P_H}$$

$$\mu(G, H) = \frac{P_G + P_H}{2P_{GH}} - 1$$

где  $P_G, P_H$  - порядки графов  $G$  и  $H$ , а  $P_{GH}$  - порядок их наибольшего общего подграфа. При этом для величины  $\sigma$  всегда выполняется неравенство треугольника, а для величин  $\rho$  и  $\mu$  это условие может и не выполняться. Тогда семейству химических соединений можно поставить в соответствие взвешенный граф семейства, вершины которого соответствуют молекулярным графам соединений, а длины ребер - расстояниям между ними. В этом случае вершины (молекулы) одного таксона графа семейства имеют достаточно большие общие подграфы (фрагменты) и соответствующие химические соединения имеют близкие свойства. Таким образом, решение задачи таксономии вершин графа семейства химических соединений может способствовать поиску фрагментов молекул, определяющих свойства соединений, синтезу новых химических соединений и прогнозу их свойств.

К задаче таксономии вершин взвешенного графа можно свести задачу группировки параметров [2], в которой требуется во множестве коррелируемых параметров выделить группы параметров, сильно связанных внутри группы и слабо связанных с параметрами других групп. В этом случае множеству коррелируемых параметров можно сопоставить граф, вершины которого соответствуют параметрам, ребра - связям между парами параметров, а их длины - значениям величины, обратной коэффициенту корреляции, например, одной

из величин:  $1 - |r_{ij}|$  или  $\frac{1}{|r_{ij}|}$ , где  $r_{ij}$  - коэф -  
 фициент корреляций между  $i$ -м и  $j$ -м параметрами.

Для таксономии вершин взвешенного графа могут быть ис -  
 пользованы известные алгоритмы [1,2,4], которые, однако, имеют  
 недостатки, затрудняющие их практическое применение, например,  
 необходимость предварительного задания количества или объема  
 таксонов, использование функций качества таксономии, монотонно  
 зависящих от числа одновершинных таксонов, и т.д.

Предлагаемый итеративный алгоритм таксономии вершин взве -  
 шенного графа ВЕГ основан на принципе "ближайшего соседа" [1]  
 и укрупняет таксоны за счет присоединения к ним ближайших вер -  
 шин или объединения ближайших таксонов. Укрупнение таксонов  
 происходит, если при этом не нарушаются ограничения на задан -  
 ные размеры таксонов и условия отделимости таксонов - отноше -  
 ния удаления таксонов друг от друга к близости вершин внутри  
 каждого таксона. При формировании очередного разбиения множе -  
 ства вершин графа на таксоны производится оценка качества так -  
 сономии по критерию, учитывающему суммарную отделимость таксо -  
 нов и равномерность распределения вершин по таксонам. Из набо -  
 ра полученных таксономий выбирается таксономия, лучшая по ка -  
 честву.

### 1. Алгоритм ВЕГ таксономии вершин взвешенного графа

Пусть задан связный взвешенный граф  $G(V, X)$ , имеющий  
 множество вершин  $V = \{v_i\}$ ,  $i = \overline{1, p}$ , и множество ребер  
 $X = \{x_j\}$ ,  $j = \overline{1, q}$ , с приписанным каждому ребру  $x_j$   
 положительным действительным числом  $l_j$  - длиной (весом) реб -  
 ра.

Для связаного подграфа  $G_t$  графа  $G$ ,  $G_t(V_t, X_t) \subseteq G$ ,  
 $V_t \subseteq V$ ,  $X_t \subseteq X$ ,  $|V_t| = p_t$ ,  $|X_t| = q_t$ , определим  
 его длину как величину  $l(G_t) = \sum l_j, x_j \in X_t$ . Удалением

вершины  $v$  от вершины  $u$  в связном подграфе  $G_t \subseteq G$  назовем величину  $r_{vu}(G_t) = \min l(P(v,u))$  по всем простым цепям  $P(v,u)$ , соединяющим  $v$  и  $u$  в  $G_t$ , а диаметром связного подграфа  $G_t$  - величину  $d(G_t) = \max_v \max_u r_{vu}$ , где в качестве  $r_{vu}$  можно выбрать одну из величин:  $r_{vu}(G_t)$  или  $r_{vu}(G)$ , причем  $r_{vu}(G_t) \geq r_{vu}(G)$ .

Для произвольного множества вершин  $V_t \subseteq V$  графа  $G$  назовем подграф  $G_t(V_t, X_t) \subseteq G$  собственным, если он связан. При этом по определению считаем, что одновершинный подграф  $G_t$  с  $p_t = 1$  является связным. Выделим следующие типы собственных подграфов  $G_t$  множества вершин  $V_t$ :  $G_t = B_t(V_t, X_t^B)$  - подграф, порожденный в  $G$  множеством вершин  $V_t$ , т.е.  $X_t^B$  содержит все ребра  $x = (v,u) \in X$ ,  $v \in V_t, u \in V_t$ ;  $G_t = T_t(V_t, X_t^T)$  - остов подграфа  $B_t$ , т.е. связанное дерево  $T_t$ , содержащее все вершины из  $V_t$ ,  $X_t^T \subseteq X_t^B$ ;  $G_t = G'_t(V_t, X'_t)$  - связный подграф  $G_t$ , для которого  $X_t^T \subseteq X'_t \subseteq X_t^B$ , например, такой, что  $X'_t$  содержит все ребра  $x = (v,u) \in X_t^B$ , длина которых  $l(v,u) \leq c \cdot \min_u l(v,u)$ ,  $c \geq 1$ , для всех  $v \in V_t$ .

Любое разбиение  $\hat{V} = \{V_t\}$ ,  $t = \overline{1, k}$ , множества  $V$  вершин графа  $G$  на подмножества  $V_t$ , для каждого из которых существует собственный подграф  $G_t(V_t, X_t) \subseteq G$ , назовем таксономией  $\hat{V}$ , а подмножества  $V_t$  - таксонами вершин графа  $G$ . Таксоны  $V_t$  и  $V_s$  назовем смежными, если в  $G$  существует непустое множество ребер  $X_{t,s} = \{x_{t,s}\} \subseteq X$ , соединяющих вершины из разных таксонов  $V_t$  и  $V_s$ . Отсюда следует, что таксономия  $\hat{V}$  вершин графа  $G$  порождает взвешенный граф  $\hat{G} = \hat{G}(\hat{V}, \hat{X})$ , вершины которого соответствуют таксонам

$V_t$ ,  $t = \overline{1, k}$ , каждое ребро  $\hat{x}_{ts} \in \hat{X}$  соответствует множеству ребер  $X_{ts}$ , соединяющих вершины смежных таксонов  $V_t$  и  $V_s$ , а его длина  $l(\hat{x}_{ts})$  - функция длин ребер  $x_{ts} \in X_{ts}$ .

В качестве оценки близости вершин графа  $G$  в таксоне  $V_t$  примем величину  $l_t$ , зависящую от длин ребер его собственного подграфа  $G_t \subseteq G$ , а в качестве оценки удаления таксонов  $V_t$  и  $V_s$  примем длину  $l(\hat{x}_{ts}) = l_{ts}$  ребра  $\hat{x}_{ts}$ , соединяющего  $V_t$  и  $V_s$  в графе  $\hat{G}$ . При естественных предположениях, что  $\min l(x) \leq l_t \leq \max l(x)$ ,  $x \in X_t$ , и  $\min l(x_{ts}) \leq l_{ts} \leq \max l(x_{ts})$ ,  $x_{ts} \in X_{ts}$ , для величин  $l_t$  и  $l_{ts}$  могут быть выбраны разные функции, например, для

$l_t$ : средняя длина ребра  $G_t$ , т.е.  $l_t = \frac{l(G_t)}{q_t}$ ; средняя длина ребер  $x \in X_t$  граничных вершин  $G_t$ ; наименьшая длина ребра  $l_t = \min l(x)$ ,  $x \in X_t$ ; а для величины  $l_{ts}$ : длина кратчайшего ребра из множества  $X_{ts}$  т.е.  $l_{ts} = \min l(x_{ts})$ ,  $x_{ts} \in X_{ts}$ , средняя длина ребра из  $X_{ts}$  и т.д.

Определим отделимость таксона  $V_t$  от таксона  $V_s$ , принадлежащего множеству  $V(t)$  таксонов, смежных ему в  $G$ , как величину  $f_{ts} = \frac{l_{ts}}{l_t}$  отношения удаления этих таксонов к близости вершин графа в таксоне  $V_t$ . Для таксона  $V_t$  отклонение распределения по таксонам вершин графа  $G$  от равномерного зададим как величину  $g_t = p(k) \left| 1 - \frac{p_t}{p(k)} \right|$ , где  $p(k) = \frac{p}{k}$ . Тогда качество таксономии  $\hat{V}$  можно оценивать с помощью критерия

$$F = 1 - (\alpha F_0 + (1-\alpha)F_p),$$

где  $\alpha$  - весовой параметр,  $0 \leq \alpha \leq 1$ ,

$$F_0 = \frac{1}{n \cdot k} \sum_{t=1}^k \frac{1}{m_t} \sum_{s=1}^{m_t} \frac{1}{f_{ts}}, \quad 2 \leq k \leq p,$$

$n$  - параметр нормирования,  $1 \leq n \leq \frac{\max l_j}{\min l_j}$ ,  $j = \overline{1, q}$ ,

$m_t$  - число таксонов  $V_s \in V(t)$ , смежных  $V_t$  в  $\hat{G}$ ,

$$F_p = \frac{1}{2(k-1)p(k)} \sum_{t=1}^k \epsilon_t,$$

$F_0, F_p$  - нормированные средние величины по всем таксонам  $V_t$  заданной таксономии  $\hat{V}$ .

Из выражения для  $F_0$  видно, что если для одновершинных таксонов принято  $l_t = 0$ , то при  $\alpha = 1$  наилучшим качеством  $F = 1$  обладает таксономия, в которой каждая вершина графа является отдельным таксоном, и качество уменьшается с уменьшением количества одновершинных таксонов. Чтобы избежать этого, в каждый одновершинный таксон  $V_t = \{v\}$  можно ввести фиктивную вершину  $v'$ , соединенную с  $v$  ребром длины  $l'_t$ , не большей длины кратчайшего ребра из ребер, соединяющих в  $G$  вершину  $v$  со смежными вершинами  $u$ , т.е.  $0 \leq l'_t \leq \min l(v, u)$ ,  $u \in V$ .

Поскольку при этом алгоритм, основанный на принципе "ближайшего соседа", объединяет вершины  $v$  и  $v'$  в один таксон, не изменяя таксономии других вершин графа, то для одновершинного таксона  $V_t \in \hat{V}$  можно принять  $l_t = l'_t$ . В качестве  $l'_t$  можно использовать, например, одну из величин:  $l_t^* = \min_u l(v, u)$ ,  $l^* = \min_j l_j$ ,  $j = \overline{1, q}$ ;  $\frac{1}{2}(l_t^* + l^*)$ ,  $(l_t^* - l^*)$  и т.д. Из  $F_0$  следует, что при

$\alpha = 1$  и  $l_t = l_t^*$  качество таксономии  $F$  возрастает с уменьшением количества одновершинных таксонов.

#### Итеративный алгоритм таксономии ВЕГ

Пусть заданы взвешенный граф  $G(V, X)$ ,  $|V| = p$ ,  $|X| = q$ , список длин  $l(v, u)$  его ребер  $x = (v, u) \in X$  и набор значений управляющих параметров  $\alpha, c, d, k_0, p_0$ , где  $\alpha$  - весовой параметр критерия качества,  $0 \leq \alpha \leq 1$ ;  $c$  - порог отделимости пары таксонов,  $1 \leq c \leq \frac{\max l(x)}{\min l(x)}$ ,  $x \in X$ ;  $d$  - допустимый диаметр таксона,  $\min l(x) \leq d \leq d(G)$ ;  $k_0$  - допустимое число таксонов;  $2 \leq k_0 \leq p$ ;  $p_0$  - допустимое число вершин в таксоне,  $1 \leq p_0 \leq p-1$ .

1. Подготовка исходных данных. Начальной таксономией  $\hat{V}^1$  является множество одновершинных таксонов  $\hat{V}^1 = \{v_t^1\}$ ,  $p_t = 1$ ,  $t = \overline{1, p}$ , для которых задаются величины  $l_t = l_t^*$  и вычисляется качество таксономии  $F(\hat{V}^1)$ . Граф  $\hat{G}^1(\hat{V}^1, \hat{X}^1)$  совпадает с графом  $G(V, X)$ . Формируется список  $\bar{X}^1$  ребер  $x \in X = \hat{X}^1$ , упорядоченный по неубыванию их длин.

2. Объединение ближайших таксонов. Для  $i$ -й итерации алгоритма исходными являются данные, полученные в  $(i-1)$ -й итерации: таксономия  $\hat{V}^i = \{v_t^i\}$ ,  $|V_t^i| = p_t^i$ , множество собственных подграфов таксонов  $\{G_t^i\}$ ,  $t = \overline{1, k^i}$ , граф  $\hat{G}^i(\hat{V}^i, \hat{X}^i)$ ,  $\hat{X}^i = \{x_j^i\}$ ,  $j = \overline{1, q^i}$ , упорядоченный по неубыванию длин список ребер  $\bar{X}^i$ , качество таксономии  $F(\hat{V}^i)$ ,  $i \leq p - k_0$ .

Объединение таксонов производится в порядке, задаваемом списком ребер  $\bar{X}^i$ . Для очередного ребра  $x_j^i = (v_t^i, v_s^i) \in \bar{X}^i$ , соединяющего таксоны  $V_t^i$  и  $V_s^i$ , имеющие собствен-

ные подграфы  $G_t^i$  и  $G_s^i$ , производится объединение этих таксонов в один новый таксон  $\tilde{V}_r^i$  и построение его собственного подграфа  $\tilde{G}_r^i$  при выполнении ограничений

$$d(\tilde{G}_r^i) \leq d, \quad \tilde{p}_r^i \leq p_0, \quad k_0 < k_j^i \quad (1)$$

и условий отделимости

$$f_{ts} = \frac{l_{ts}}{l_t} \leq 0, \quad f_{st} = \frac{l_{ts}}{l_s} \leq 0, \quad (2)$$

где  $d(\tilde{G}_r^i)$  - диаметр собственного подграфа  $\tilde{G}_r^i$  множества вершин  $\tilde{V}_r^i = V_t^i \cup V_s^i$ ,  $\tilde{p}_r^i = |\tilde{V}_r^i|$ ,  $k_j^i$  - количество таксонов, сформированных в  $i$ -й итерации перед выбором очередного ребра  $x_j^i$ , имеющего длину  $l_{ts}$ ,  $l_t$  и  $l_s$  - оценки близости вершин в таксонах  $V_t^i$  и  $V_s^i$  соответственно.

Если для очередного ребра хотя бы одно неравенство из (1), (2) не выполняется, то формируются выходные данные  $i$ -й итерации:  $\hat{V}^{i+1}$ ,  $\{G_b^{i+1}\}$ ,  $\hat{G}^{i+1}$ ,  $\bar{X}^{i+1}$ ,  $F(\hat{V}^{i+1})$  и производится переход к  $(i+1)$ -й итерации.

Если для всех ребер  $x_j^i \in \bar{X}^i$  выполняются неравенства (1), (2) или для каждого из ребер  $x_j^i$  не выполняется хотя бы одно неравенство из (1), (2), то переход к п.3, а таксономия  $\hat{V}^i$  при этом является конечной таксономией  $\hat{V}_0$  работы алгоритма.

3. Результат таксономии. В качестве результирующей таксономии выбирается лучшая (при заданном наборе значений параметров) таксономия  $\hat{V}^*$ , для которой оценка качества  $F(\hat{V}^*) = \max_i F(\hat{V}^i)$ ,  $1 \leq i \leq p-k_0$ .



## ПРИМЕЧАНИЯ.

1. Завершение очередной  $i$ -й итерации и формирование новой таксономии  $\hat{V}^{i+1}$  можно производить и по другим условиям, например, после каждого объединения пары таксонов или при последовательном укрупнении таксонов после завершения формирования очередного таксона.

2. Алгоритмы построения собственных подграфов таксонов  $V_t$  достаточно просты и зависят от выбранного типа подграфа:  $B_t$ ,  $T_t$  или  $G'_t$ . Заданный граф  $G(V, X)$  также может быть некоторым подграфом исходного взвешенного графа  $H(V, X_H)$ , например,  $G = B(V, X_H^B)$ ,  $G = T(V, X_H^T)$  или  $G = G'(V, X_H')$ .

3. Для нахождения наилучшей по качеству таксономии можно организовать работу алгоритма в циклическом режиме при разных значениях порога отделимости  $c$ . При этом для  $c < 1$  конечная таксономия  $\hat{V}_0$  совпадает с начальной одновершинной таксономией  $\hat{V}^1$ , а для  $c \geq \frac{\max l(x)}{\min l(x)}$ ,  $x \in X$ , конечная таксономия  $\hat{V}_0$  состоит из одного таксона, содержащего все вершины графа. Для конечной таксономии  $\hat{V}_0$  отделимость таксонов друг от друга превышает заданное значение  $c$ .

Объем памяти, требуемый для работы алгоритма, не превосходит  $a \cdot q$ , а вычислительная сложность -  $b \cdot q \log q$  (где  $a, b$  - константы) в случае, если заданный граф  $G$  является кратчайшим остовом исходного взвешенного графа  $H(V, X_H)$  и нет ограничений на диаметр таксонов.

## 2. Эталоны таксонов вершин взвешенного графа

При известном разбиении взаимосвязанных объектов на таксоны возникает задача нахождения в каждом таксоне наиболее типичных представителей - эталонов и ядер таксонов - подмножеств

объектов, близких к эталонам. Типичность объекта данного таксона можно определять в зависимости от его положения (удаления) относительно объектов этого таксона или относительно объектов других таксонов.

Пусть множеству  $V$  объектов соответствует взвешенный граф  $G(V, X)$  и известна таксономия  $\hat{V} = \{V_t\}$ ,  $t = \overline{1, k}$ ,  $|V_t| = p_t$ , множества его вершин на таксоны  $V_t$  с собственными подграфами  $G_t(V_t, X_t)$ . Рассмотрим следующие критерии типичности  $\Phi(v)$  вершин  $v$  графа  $G$  в таксономии  $\hat{V}$ .

Типичность вершины  $v \in V_t$  зависит только от ее положения относительно других вершин  $G_t$ . Тогда типичность  $v$  можно характеризовать величиной среднего удаления  $v$  от вершин таксона  $\Phi_1(v) = \frac{1}{p_t} \sum r_{vu}$ ,  $u \in V_t$ , а за эталоны  $v(t)$  таксона  $V_t$  принять "центральные" вершины  $G_t \subseteq G$ , для которых  $\Phi_1(v(t)) = \min \Phi_1(v)$ ,  $v \in V_t$ .

Типичность вершины  $v \in V_t$  зависит только от ее положения в графе  $G$  относительно вершин других таксонов. В качестве критерия типичности  $v$  можно принять величину среднего удаления в  $G$  вершины  $v$  от вершин других таксонов  $\Phi_2(v) = \frac{1}{p-p_t} \sum r_{vu}(G)$ ,  $u \in V \setminus V_t$ , а за эталоны таксона  $V_t$  принять или вершины  $v(t)$ , в среднем наиболее удаленные от вершин других таксонов, для которых  $\Phi_2(v(t)) = \max \Phi_2(v)$ ,  $v \in V_t$ , или вершины  $v'(t)$ , в среднем наименее удаленные от них, для которых  $\Phi_2(v'(t)) = \min \Phi_2(v)$ ,  $v \in V_t$ .

Типичность вершины  $v \in V_t$  зависит как от ее положения относительно вершин своего таксона, так и относительно вершин других таксонов. Тогда в качестве критерия типичности  $v$  можно принять величину  $\Phi_3(v) = \frac{\Phi_1(v)}{\Phi_2(v)}$ , а за эталоны так-

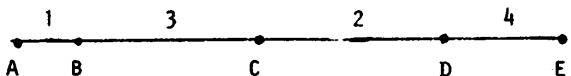
сона  $V_t$  принять вершины  $v(t)$ , для которых  $\Phi_3(v(t)) = \min \Phi_3(v), v \in V_t$ .

Ядро  $W_t$  таксона  $V_t$  также может быть определено разными способами. Рассмотрим два списка  $S'_t$  и  $S''_t$  вершин  $v \in V_t$ , упорядоченных по неубыванию значения величины  $f(v)$ , начиная с одной из вершин-эталонов  $v(t) = v_1 \in S_t$ . Величину  $f(v_i)$  определим для списка  $S'_t$  как  $f(v_i) = |\Phi(v_i) - \Phi(v(t))|$  для выбранного критерия типичности  $\Phi$ , а для списка  $S''_t$  как  $f(v_i) = \min_{v(t) \in V_t} r_{v_i, v(t)}$ , - величину наименьшего удаления  $v_i$  от эталонов  $v(t)$  таксона  $V_t$ . Тогда в качестве ядра  $W_t$  таксона  $V_t$  можно выбрать по каждому из списков или множество  $W_t = \{v_i \mid i = \overline{1, N_t}\}$ ,  $1 \leq N_t \leq p_t$ , первых  $N_t$  вершин в списке, или множество  $W_t = \{v_i \mid f(v_i) \leq \delta\}$  вершин  $v_i$ , отличающихся по значению  $f(v_i)$  от эталонов на величину не более  $\delta$ .

Различные понятия эталонов и ядер таксонов могут быть полезны при исследовании совокупностей объектов, в которых классы объектов имеют сложную структуру - состоят из подклассов, содержат ошибочно включенные объекты и т.д.

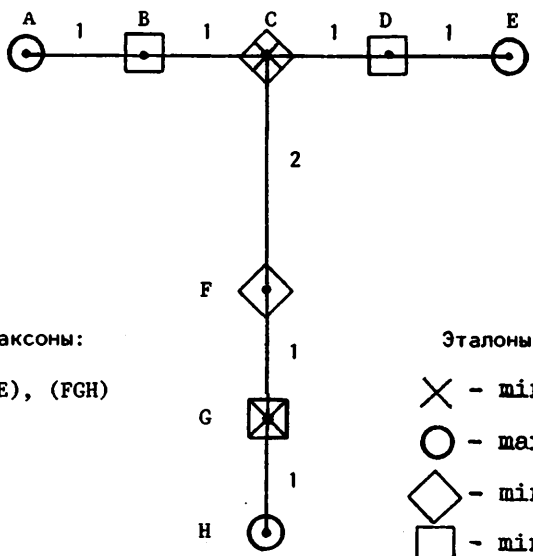
Например, при поиске общих фрагментов молекулярных структур химических соединений, обладающих общим свойством, может оказаться, что это свойство определяется структурно близкими, но разными фрагментами, задающими разбиение соединений на подклассы, которые соответствуют ядрам таксона этих соединений. В этом случае поиск общих фрагментов необходимо производить отдельно для соединений, входящих в каждое из ядер таксона, определенных по критериям  $\Phi_1(v)$  или  $\Phi_3(v)$ .

Если объекты - зависимые параметры, а расстояния между ними - величины обратные коэффициентам корреляции, то эталоны таксонов, определенные по значениям  $\max \Phi_2(v)$ , дают сово-



C	Таксоны	F	Примечание
$C < 1,5$	(A)(B)(C)(D)(E)	0,26	$\alpha = 1$ $l'_t = \frac{1}{2} (l_t^* + l_t)$ $l^* = 1$
	(AB)(C)(D)(E)	0,4	
	(AB)(C)(DE)	0,45	
$1,5 \leq C < 3$	(AB)(CDE)	0,42	$l_t = \frac{l(G_t)}{q_t}$
$3 \leq C$	(ABC)(DE)	0,33	

Рис.1 Таксономия



Таксоны:  
(ABCDE), (FGH)

Эталоны:  
 X -  $\min \Phi_1(v)$   
 O -  $\max \Phi_2(v)$   
 Diamond -  $\min \Phi_2(v)$   
 Square -  $\min \Phi_3(v)$

Рис.2. Эталоны

купность наиболее независимых параметров - представителей таксонов.

Простые примеры, иллюстрирующие работу алгоритма таксономии ВЕГ и определения эталонов представлены на рис.1,2.

Алгоритмы таксономии ВЕГ и поиска эталонов и ядер таксонов послужили основой разработки программ (РС IBM, Турбо-Паскаль), включенных в систему поиска общих фрагментов химических соединений, определяющих их свойства, создаваемую совместно сотрудниками ВХТИ АН Болгарии и ИМ СО АН СССР.

#### Л и т е р а т у р а

1. ДУДА Р., ХАРТ П. Распознавание образов и анализ сцен. - М.: Мир, 1976. - 512 с.

2. БРАВЕРМАН Э.М., МУЧНИК И.Б. Структурные методы обработки эмпирических данных. - М.: Наука, 1983. - 464 с.

3. СТЬЮПЕР Э., БРЮГГЕ У., ДЖУРС П. Машинный анализ связи химической структуры и биологической активности. - М.: Мир, 1982. - 233 с.

4. ЗАГОРУЙКО Н.Г., ЁЛКИНА В.Н., ЛБОВ Г.С. Алгоритмы обнаружения эмпирических закономерностей. - Новосибирск, Наука, 1985. - 112 с.

Поступила в ред.-изд.отд.

12 августа 1991 года