

АЛГОРИТМИЧЕСКАЯ ОБРАБОТКА СТРУКТУРНОЙ ИНФОРМАЦИИ

В.А. Скоробогатов

§ 1. Структурные модели и графы

Для представления и изучения таких структурных моделей, как молекулярные структуры, вторичные структуры рибонуклеиновых кислот и электронные схемы используются графы [6,8].

Графом G называют пару множеств $G = (V, X)$, где V — конечное множество элементов, называемых вершинами, и множество X элементов x , называемых ребрами, $X \subseteq V \times V$ — подмножество множества всевозможных пар вершин. Граф имеет геометрическую интерпретацию в виде диаграммы, на которой вершины изображаются точками, а ребра — линиями, соединяющими пары вершин из множества V . Граф $G = (V, X)$ называют двудольным, если его множество вершин представимо в виде $V = V_1 \cup V_2$ и в множестве X не содержится пар (v', v'') таких, что $(v', v'') \in V_1$ либо $(v', v'') \in V_2$. Множества вершин V_1 и V_2 называют долями графа G .

Граф $G = (V, X)$ называется мультиграфом, если его вершины могут соединяться несколькими параллельными ребрами — звеньями, называемыми мультиребрами. В случае помеченных графов их вершинам и ребрам могут быть сопоставлены некоторые метки, множество которых можно представлять в виде алфавитов. Если требуется представление объекта в виде моделей, характеризующих его с различной степенью точности, то можно применить формальное преобразование эквивалентности графов по определенным правилам, множества которых называются грамматиками. В этом случае структурная модель представляется в виде помеченного графа, заданного над алфавитом с грамматикой: $G = (V, X, A(V, X), \varphi(V, X), \Gamma)$, где $A(V, X)$ — алфавиты вершинных и реберных меток, $\varphi(V, X)$ — правила сопоставления меток элементам графа, а Γ — правила преобразования меток.

Рассмотрим примеры представления в виде графов некоторых конкретных объектов.

Л о г и ч е с к и е с х е м ы представимы в виде двудольных неориентированных помеченных графов.[7].Схема состоит из следующих элементов полюсов (входных, выходных), узловых точек, соответствующих соединениям и логическим элементам. Схеме ставится в соответствие граф следующим образом. Вершины из первой доли соответствуют полюсам схемы и имеют пометки: входной, выходной, внутренний; вершины из второй доли – элементам схемы и имеют пометки соответствующих элементов схемы. Ребра графа соединяют вершины, принадлежащие только различным долям и соответствуют соединениям полюсов и элементов. Получаем граф $G = (V, V', X, A_{\Pi}, A_{\Delta}, \Phi_{\Pi}, \Phi_{\Delta})$, где $A_{\Pi} = \{V_{\text{вх}}, V_{\text{вых}}, V_{\text{вн}}\}$, алфавит A_{Δ} состоит из множества обозначений логических элементов.

Х и м и ч е с к и е г р а ф ы. Различают несколько видов химических графов. Молекулярный граф – неориентированный помеченный мультиграф, поставленный в соответствие структурной формуле молекулы так, что атомам химической структуры соответствуют вершины графа, связям – ребра, кратным связям в структуре соответствуют мультиребра в графе. Молекулярный граф может представлять молекулу с различной степенью подробности. Наиболее подробное представление – поатомное; менее подробное это представление с точностью до углеродного каркаса, когда вершины графа сопоставляются только углеродным атомам. Различные специалисты могут представлять одну и ту же молекулу по-разному, хотя и одинаково понятно для других. Отличия могут возникать из-за применения различных меток.

Реакционные графы для изомеров описывают взаимные превращения изомеров. Помеченные вершины в них соответствуют изомерам, а ребра – различным взаимным превращениям.

Существуют и другие виды химических графов.

В математической биологии рассматриваются структуры р и б о н у к л е и н о в ы х к и с л о т (РНК). Это линейные биополимеры, выполняющие важные функции в генетических управляющих системах. РНК способны образовывать двойные спирали из полинуклеотидных последовательностей. Исследование вторичных структур РНК позволяет установить особенности ее пространственного строения, которое, в свою очередь, связано с функционированием этого класса биологических молекул. Вершинам графа вторичных структур РНК соответствуют нуклеотиды: $A(V) = \{A, U, G, C\}$, а ребрам – связи, оп-

ределяемые по правилам комплементарности, вытекающие из свойств, приводящих к образованию спирали, либо по последовательности соединения нуклеотидов в молекуле [9].

Большинство представленных моделей-графов имеет много общего. В частности, небольшие степени вершин по сравнению с их числом (для химграфов $d(v) \leq 6$, для биологических структур ≤ 3). В схемах возможны достаточно большие значения степеней, но по сравнению с числом вершин оно также невелико. Указанные системы образуют класс графов с особыми метрическими свойствами, которые накладывают определенные ограничения на значения метрических характеристик этих графов, что может быть использовано для построения эффективных алгоритмов.

При решении некоторых классов задач приходится иметь дело с операциями над парами или даже семействами структур. В этих случаях результатом операции может быть граф значительно более сложный, чем исходные структуры, и важно уметь находить его особенности, благодаря которым можно единообразно упрощать вычисления за счет сокращения числа анализируемых вариантов. Это позволяет сделать знания симметрий и применение алгоритмических схем рекурсивного разбора в качестве основы для построения алгоритмов.

§ 2. Исследования метрических свойств графов

Многие свойства графов определяются расстояниями между определенными подмножествами вершин, т.е. длинами кратчайших связывающих цепей.

Учет метрических свойств графов позволяет, например, сократить перебор при распознавании изоморфизма [10].

Метрические свойства графов выражаются при помощи относительных разбиений графов [1, 10].

Под относительным разбиением $\hat{G}(u)$ графа $G = (V, X)$ по отношению к $u \in V$ (или по отношению к любому $V \subseteq V$) понимается упорядоченное множество классов $V_j(u)$ такое, что $\hat{V}(u) = \{V_j(u) \mid j = 0, k(u), \forall v \in V, \leftrightarrow d(v, u) = j\}$, класс $V_j(u)$ называется j -слоем для u .

Пусть $v \in V_j(u)$, тогда $V_j(u)$ — собственный слой v , а слои $V_{j-1}(u)$ и $V_{j+1}(u)$ являются смежными слева и справа, если в каждом из них содержится хотя бы по одной вершине, смежной с v . Очевидно, что для любой v и любого j всегда есть левый смежный слой для v .

Можно говорить о смежных слоях для слоя $V_j(u)$, если хотя бы для одной v существует смежный слой слева и хотя бы для одной v' - смежный слой справа. Для разбиений связанных графов очевидно, что левый смежный слой всегда существует. Относительные разбиения порождают разбиения степени вершины v из некоторого слоя на левую, собственную и правую частичные степени, которые соответственно равны числу смежных с v слева, в собственном слое и справа. Левая степень всегда отлична от нуля. Собственная степень может быть равна нулю, если данная вершина не принадлежит нечетному циклу, ребро которого принадлежит данному слою.

Если в некотором $\hat{G}(v)$ у всех вершин собственные степени равны нулю, то граф двудольный [1].

Вершина v в слое V_j , $j \neq k$, с нулевой правой степенью называется тупиковой, множество всех таких вершин - тупиковым множеством разбиения, а объединение тупиков всех разбиений - тупиковым множеством графа. Тупик графа не содержит точек сочленения [1] и позволяет определить особый класс гладких графов - графов без тупиков.

Подграф $\langle W_{i,k}(V_0) \rangle$, $V_0 \subset V$, порожденный множеством $W_{i,k}(V_0) = \bigcup_{j=0}^{k-1} V_{i+j}(V_0)$, назовем k -слойным поясом, или k -поясом. При $k=1$ получаем 1-пояс или 1-слой. При $k=2$ получаем 2-пояс, порожденный парой смежных слоев и т.д. При $i=e(v)$ подграф $G_e(v) = \langle W_e(v), i(v) \rangle$ называется оболочкой графа для v . Если $k=\max_{v \in V} \{e(v)\} = d(G)$, то G_d - диаметральная оболочка графа. Аналогично определяются радиальная оболочка.

Интересные для приложений сведения о графе можно получить, изучая его "центральные" слои, т.е. слой по отношению к его центру. При этом понятие центра может быть обобщено путем использования вместо эксцентриситетов некоторых других метрических характеристик.

Множеству относительных разбиений можно сопоставить матрицу слоев $\lambda^a(G) = \|a_{i,j}\|$, $i = \overline{1, \binom{p}{n}}$, $j = \overline{1, d(G)}$, таким образом, что $a_{i,j}$ равно числу вершин в j -м слое относительного разбиения по отношению к i -му множеству V_i^0 , $n = |V_1^0|$. Если $V_j(V_i^0) = \emptyset$, то $a_{i,j} = 0$. Упорядочив строки $\lambda^a(G)$ по уменьшению длины (числу ненулевых элементов), а затем упорядочив строки одинаковой длины лексикографич-

чески, получим каноническую матрицу $\lambda^n(G)$. При $n = 1$ имеем матрицу слоев одновершинных разбиений.

По аналогии с вершинными матрицами $\lambda_G^n(V)$ можно рассматривать реберные части матриц слоев $\lambda_G^n(X)$, в которых элементам соответствуют значения разрезов между слоями (число ребер, соединяющих смежные слои): $\lambda_G^n(X) = \|b_{i,j}\|$, $b_{i,j} = |X_{j-1,j}|$, $X_{j-1,j}$ - множество ребер, соединяющих слои V_{j-1} и V_j . Полная одновершинная матрица слоев может быть представлена в виде $\lambda_G^1 = [\lambda_G^1(V), \lambda_G^1(X)]$. Можно также рассматривать матрицы следующего вида. Вершинные матрицы слоев $\lambda_G^2(V, x)$, полученные по отношению к ребрам; $\lambda_G^2(V, \bar{x})$ по отношению к парам несмежных вершин. Очевидно, что $\lambda_G^2(V, x) \cup \lambda_G^2(V, \bar{x}) = \lambda_G^2(V)$. Матрицы слоев можно представлять в виде, например, для графа $G \cong \bar{P}_5$: $\lambda^1(\bar{P}_5) = \|1, 2, 3 \ (2, 2); 4, 5 \ (3, 1)\|$ или, если номера вершин (строк) не важны, то $\|3 \ (2, 2); 2 \ (3, 1)\|$, эта запись означает, что матрица слоев состоит из трех строк (2, 2) и двух строк (3, 1).

Известны следующие свойства матриц слоев $G \cong N \rightarrow \lambda^1(G) = \lambda^1(N)$; $\lambda^1(G) = \lambda^1(N) \not\rightarrow G \cong N$, даже для тождественных графов (с тривиальной группой). Это верно и для деревьев. Если $\{\lambda^1(G)\}$ - семейство всех 1-матриц графа, то $\forall i \ \lambda^1(N) = \lambda^1(G) \not\rightarrow G \cong N$, $\lambda_G^1(X) = \lambda_N^1(X) \not\rightarrow G \cong N$ и $\lambda^1(G) = \lambda^1(N) \not\rightarrow \lambda^j(G) = \lambda^j(N)$, $i, j \in \{1, 2\}$.

Следует заметить, что свойства рассматриваемых матриц пока изучены слабо. В настоящее время не известны примеры графов G и N , для которых не выполнялось бы $\lambda_G^2(V, x) = \lambda_N^2(V, x) \rightarrow G \cong N$ и $\lambda_G^1(V) = \lambda_N^1(V) \& \lambda_G^2(V, x) = \lambda_N^2(V, x) \rightarrow G \cong N$.

Матрицы слоев, помимо того, что являются сильными двумерными характеристиками графов, еще обеспечивают простые и единообразные вычисления целого ряда других характеристик [11].

Степень точности характеристики графов этими характеристиками [12] основывается на понятии 1-спектра графа G : $1(G)$, который определяется как часть матрицы слоев, состоящая из всех попарно различных строк. 1-спектр позволяет получить критерий изометричности графов [13]. Говорят, что граф N изометричен из G : $G \rightsquigarrow N$ тогда и только тогда, когда для любой вершины $v \in V(G)$ существует однозначное отображение $\phi_v: V(G) \xrightarrow{\text{на}} V(N)$ такое, что для любой вершины $u \in V(G)$ выполняется $d(u, v) = d(\phi_v u, \phi_v v)$, $\phi_v u, \phi_v v \in V(N)$. Если $G \rightsquigarrow N$ и $N \rightsquigarrow G$, то $G \leftrightarrow N$, т.е. графы G и N изометричны. Необходимые и достаточные условия изометричности графов, которые формули-

руются в следующем виде: $1(G) = 1(H) \iff G \iff H$, показаны в [12]. Изометричные графы могут иметь различные вершинные матрицы слоев первого порядка. Графы с одинаковыми матрицами изометричны, у них одинаково и число всех возможных соответствий изометричности, на которое влияет только число одинаковых строк в матрицах.

Матрицы слоев применяются для вычисления ряда метрических характеристик графов, основанных на системах расстояний в графе. Для использования этих характеристик может быть применена методика и программа МЕТСНАР [11,16], позволяющие находить матрицы слоев, 1-спектры и ряд других метрических характеристик графов. Некоторые из этих характеристик находят применение в социологии [14] и химии [15].

§ 3. Исследование и роль симметрий графов

Граф считается симметричным, если его группа автоморфизмов имеет нетривиальные орбиты. Говорят, что множество вершин графа образует его орбиту, если оно замкнуто относительно автоморфизмов, принадлежащих некоторой одной орбите группы автоморфизмов графа. Тогда вершины графа, принадлежащие одной и той же орбите, называются симметричными. Поскольку для любой пары вершин из орбиты существует автоморфизм, который переводит эти вершины друг в друга, то для многих вычислений такие вершины эквивалентны. Отсюда понятна важная роль симметрий графа: они позволяют сокращать время решения задач путем устранения эквивалентных вариантов вычислений при реализации переборных схем. В некоторых случаях можно рассматривать множество вершин — представителей орбит графа, а не все множество вершин.

Симметрии графов играют важную роль при автоматической генерации структур [18,37] для создания каталогов графов, позволяющих упростить процедуры проверки гипотез, или при установлении строения вещества по его спектральным характеристикам [38,39] и во многих других областях исследований алгоритмов над графами.

Для определения орбит группы автоморфизмов обыкновенных графов и помеченных мультиграфов разработаны алгоритмы [2,18], основанные на построении и анализе разбиений $\hat{V}^*$ и $\hat{V}_*$ множества вершин графа $G = (V, X)$ на классы эквивалентности, для которых выполнены соотношения $\hat{V}_* \leq \hat{V}_\Theta \leq \hat{V}^*$, где $\hat{V}_\Theta$ — разбиение множества вершин на классы, совпадающие с орбитами группы автоморфиз-

мов. Эти алгоритмы применяются и для исследования теоретико-групповых свойств графов [19-21], и при решении практических задач в радиоэлектронике - поиск структурно-подобных неисправностей в схемах при генерации тестов [22]; и в химии - предсказание биологической активности соединений [17,33,34].

§ 4. Структурное подобие и модульные произведения графов

Одной из структурных характеристик, при помощи которой может быть оценено структурное подобие графов, является количественный и фактический состав максимальных или близких к ним общих подграфов для графов данного семейства.

Была предложена методика распознавания изоморфизма и установления изоморфного вложения графов при помощи операции модульного произведения, или построения графа соответствий, основным свойством которого является то, что каждая его клика (максимальный полный подграф) соответствует общему подграфу в исходных графах [23].

Этот подход легко обобщается на случай нахождения общих подграфов для всего семейства, однако его практическая реализация на ЭВМ затруднена из-за значительной трудоемкости алгоритмов поиска клик [24] в графе соответствий, поскольку порядок последнего равен произведению порядков исходных графов. В связи с этим был предложен метод, позволяющий снизить трудоемкость нахождения общих подграфов путем использования метрических свойств графов и их симметрий [20]. При этом анализ всего графа соответствий заменяется анализом некоторого множества его подграфов, определяемых модульными произведениями относительных разбиений.

Если нет необходимости находить все изоморфные общие подграфы, то для построения относительных разбиений выбираются только вершины-представители орбит графа и в графе соответствий находятся не все клики, а только кликовая база [21,25], т.е. множество клик, не переводящихся друг в друга автоморфизмами.

Для установления корреляции между структурой вещества и его свойствами разработана методика выбора информативных структурных характеристик [26] из множества признаков структурного подобия, которые определяются наибольшими общими подграфами всех или части молекул данного класса. Формирование признаков структурного подобия можно производить на классах соединений, которые уже обладают некоторыми общими свойствами, например, близкими спектральными

характеристиками [35] или имеют похожие брутто-формулы связей [36] или какие-либо другие подобные структурные дескрипторы. Получив множество таких признаков в классах веществ методами распознавания образов, выбирается подмножество наиболее информативных характеристик, по сочетанию которых можно определить наличие интересующего свойства. При выборе информативных признаков учитывается, что в информативное подмножество должен войти тот признак, который часто встречается у представителей одного класса и редко у представителей другого класса. Отсюда следует, что не нужно находить общие фрагменты в молекулах из разных классов, это достаточно делать только для молекул из одного и того же класса. Для формирования систем информативных структурных характеристик могут применяться алгоритмы из [27].

Методы структурного подобия могут применяться и для определения строения неизвестных соединений [40].

§ 5. Исследование и применение методов рекурсивного разбора графов

В общем виде операция разборки – последовательного превращения исходного графа в конечную совокупность "простых" графов – была определена в [28]. Изменяя правила "упрощения", можно строить различные алгоритмы выделения "особых" подмножеств графа либо вычисления его характеристик. Частные случаи этих операций используются для построения алгоритмов нахождения клик [3,29,30] и независимых множеств [31] графа.

Для решения класса задач выделения в графе множеств вершин, обладающих заданными метрическими свойствами, применяются схемы рекурсивного разбора графов – множество алгоритмов, объединенных общей структурой [3,4,29]. В основу разбора положена операция разбиения графа на совокупность поясов относительно некоторой выделенной вершины v . Каждый k -пояс есть подграф, порожденный k соседними слоями в разбиении графа относительно вершины v [1]. Каждый алгоритм из семейства задает процесс последовательного разбиения графа на k -пояса (исходный граф разбивается на k -пояса; каждый полученный k -пояс вновь разбивается на k -пояса и т.д.) и определяется следующими параметрами: k – величина пояса; l – длина (т.е. число слоев без единиц) в общем случае неполного относительного разбиения; O – правило остановки, определяющее допус-

тимостью продолжения разбора каждого k -пояса; В – правило выбора, определяющее вершину, относительно которой строится разбиение; Γ – правило преобразования k -пояса, которое некоторым образом упрощает граф^{*}; Н – правило накопления, которое анализирует и накапливает результаты разбора.

Семейство схем рекурсивного разбора обозначается как РАЗБОР (k, l, O, B, Γ, H).

Процесс разбора протекает следующим образом: исходный граф анализируется правилом О. Если дальнейший разбор возможен, то правило В выделяет вершину, относительно которой строится разбиение на k -пояса; выбирается очередной k -пояс и преобразуется правилом Γ . Преобразованный граф анализируется правилом О. Если дальнейший разбор возможен, то вступает в действие правило В и т.д., если разбор невозможен, то правило Н анализирует и накапливает результаты (множества, графы, числа и т.п.). Если все k -пояса оказались рассмотренными, то процесс заканчивается. В противном случае выбирается очередной нерассмотренный k -пояс, подвергается действию правила Γ , анализируется правилом О и т.д.

Приведенное представление схем разбора дает возможность систематизировать известные алгоритмы анализа графов. Принцип замещения параметров разбора позволяет строить новые алгоритмы, а также создает удобства при машинных экспериментах, так как набор различных вариантов правил О, В, Γ и Н может быть реализован в виде библиотеки процедур.

Результат применения правила выбора в общем случае зависит от нумерации вершин графа, поэтому для двух изоморфных графов при заданных параметрах схем процесс разбора может происходить по-разному. Процесс разбора графа с занумерованными вершинами называется реализацией схемы.

Для описания структуры реализации схемы рекурсивного разбора используется дерево разбора: исходному графу соответствует корневая вершина дерева; k -поясам исходного графа соответствуют вершины дерева, смежные с корневой, и т.д.; вершина дерева является висячей, если соответствующий ей k -пояс удовлетворяет условию остановки.

^{*}) Преобразование может заключаться, например, в удалении некоторых вершин или ребер; правило Γ может отсутствовать [5], необходимость его введения обнаружилась при построении алгоритма выделения кликовой базы [21].

Характеристикой сложности разбора и функцией сложности графа называются соответственно величины $\eta(G)$ и $\xi(G)$ – наибольший порядок дерева разбора и наибольшее число операций при выполнении разбора графа G , взятые по всем нумерациям его вершин. Указанные величины связаны соотношением: $\xi(G) = O(e(G) \cdot \eta(G))$, где $e(G)$ – суммарная максимальная трудоемкость (число операций) при выполнении относительного разбиения и правил O , B , F и H . Таким образом, $\xi(G)$ и $\eta(G)$ можно считать характеристиками графа G , выражающими его сложность, по отношению к операциям разбора.

Интересно получить и исследовать зависимости сложности разбора и функций сложности графа от параметров графа, которые позволили бы оценивать сложность решения практических задач, поскольку графы, возникающие в приложениях, обычно характеризуются вполне определенными значениями параметров.

Такого рода исследования были проведены в [4,5] для схем рекурсивного разбора, предназначенных для построения на их основе алгоритмов поиска клик.

Поскольку клика имеет диаметр, равный 1, и поэтому всегда находится в некотором 2-поясе, то для выделения клик необходимо положить $k = 2$ и достаточно*) положить $l = 2$. Для того чтобы разбор не продолжался бесконечно, необходимо принять правило остановки в виде:

O_1 : если 2-поис – полный граф, то его разбор закончить.

Правило выбора можно определить в двух вариантах:

B_1 : выбрать вершину в 2-поясе с наименьшим номером и несмежную с какой-нибудь другой вершиной в данном 2-поясе;

B_2 : выбрать вершину в 2-поясе с наименьшим номером среди всех вершин с минимальной степенью.

Пусть F_1 есть тождественное преобразование, а H_1 заносит множества вершин 2-пояса в накопитель информации.

Результатом разбора графа в соответствии с данной схемой будет множество его полных подграфов, состоящих из клик, и, возможно, некоторых частей клик. Функция сложности графа и характеристика сложности разбора для этой схемы связаны соотношением: $\xi(G) \leq$

*) В [3] для выделения клик рассматриваются полные относительные разбиения. данная схема разбора с трудом поддается анализу, хотя удалось доказать, что для графов с большим диаметром ее трудоемкость меньше, чем трудоемкость схем с $l = 2$.

$\leq O(p^2(G) \cdot \eta(G))$, где $p(G)$ — число вершин графа G . Данное соотношение показывает, что для оценки функции сложности графа достаточно исследовать величину $\eta(G)$.

Доказана монотонность функции характеристики сложности разбора — ее значение на подграфах из G не превосходит значения на графе G . Методика получения верхних оценок характеристики сложности разбора, зависящих от параметров графа [4], основана на анализе рекуррентного соотношения:

$$\Psi(a_1, \dots, a_n) = \Psi(a'_1, \dots, a'_n) + \Psi(a''_1, \dots, a''_n), \quad (1)$$

где $a_1, \dots, a_n$ — верхние границы n различных числовых функций на G , обладающих определенными свойствами монотонности; a'_i и a''_i ($i = \overline{1, n}$) — верхние границы значений указанных функций соответственно на первом и втором 2-поясах G .

Функция Ψ , удовлетворяющая соотношению (1), ограничивает сверху функцию $\eta(G)$.

В соответствии с данной методикой были получены верхние оценки характеристики сложности разбора, зависящие от порядка графа, от его плотности и от отношения порядка графа к его плотности [5].

На основе схем рекурсивного разбора были разработаны алгоритмы: перечисление всех клик, нахождение максимальной клики, нахождение пересечений^{*)} двух графов, выделение неавтоморфных клик.

Алгоритм перечисления всех клик определен как схема

РАЗБОР $(2, 2, O_2, B_2, P_1, N_2)$,

где правило O_2 устроено так, чтобы, во-первых, в результате разбора не получались подграфы клик и, во-вторых, чтобы полученные в результате 2-простые графы не подвергались дальнейшему разбору (2-простыми называются графы, максимальная степень вершин дополнения которых не превосходит 2). Правило N_2 перечисляет все клики в 2-простых графах и накапливает те из них, которые являются кликами в G .

Алгоритм нахождения максимальной клики определен как схема РАЗБОР $(2, 2, O_2, B_2, P_1, N_3)$, в которой N_3 ищет за полиномиальное время максимальную клику в 2-простом графе и заносит ее в память, если там не содержится большей клики.

*) Пересечением двух графов называется максимальный по включению подграф, изоморфно входящий как в первый граф, так и во второй (см. также § 3).

Алгоритм нахождения пересечений двух графов реализован в виде схемы РАЗБОР $(2, 2, O_2, B_2, F_1, H_2)$, применяемой к графу соответствий [5]. В частных случаях алгоритм нахождения пересечений может быть использован для распознавания изоморфизма двух графов и изоморфного вхождения одного графа в другой.

Алгоритм выделения максимального множества неавтоморфных клик (кликерной базы [21]) определен как схема

РАЗБОР $(2, 2, O_2, B_1, F_2, H_2)$,

в которой правило O_2 прекращает разбор 2-пооя, если он полный, либо входит в подграф, разбор которого к данному моменту закончен; правило F_2 работает следующим образом: если 2-пояс G_2 разбивается на 2-пооя G'_2 и G''_2 относительно вершины v , то в G''_2 удаляются все вершины, определенным образом эквивалентные вершине v [21]; правило H_2 накапливает клики графа G . Данный алгоритм может также использоваться для нахождения неавтоморфных пересечений двух графов.

Были получены следующие теоретические оценки вычислительной сложности предложенных алгоритмов.

1. Число операций при выполнении алгоритма перечисления всех клик ограничено сверху величиной $O(p^2(G) \cdot 3^{p(G)/3})$.

2. Число операций при выполнении алгоритма перечисления всех клик и нахождения максимальной клики ограничено сверху величиной

$O(p^2(G) \cdot x_0^{p(G)})$, а также величиной $O\left(p^2(G) \cdot \left(\frac{p(G)}{m(G)}\right)^{m(G)}\right)$ при $p(G)/m(G) > 1$ ($x_0 > 0$ - корень уравнения $x^n - x^{n-1} = 1$, $n = \left\lceil \frac{p(G)}{m(G)} \right\rceil$).

3. Объем оперативной памяти при выполнении алгоритмов имеет порядок $O(p^2(G))$.

Экспериментальное исследование предложенного алгоритма перечисления всех клик показало его более высокое быстродействие по сравнению с хорошо известным алгоритмом Брона-Кербоша [29] и его модификацией [5]. Эксперименты с алгоритмом выделения кликовой базы подтверждают его корректность и показывают, что на каждую клику из базы расходуется $O(p^2(G))$ операций.

Построенные на основе рекурсивного разбора алгоритмы реализованы программно и нашли применение при решении ряда прикладных задач.

Процедура выделения максимальных неавтоморфных клик в составе программного комплекса COMMAX [26,32] используется для проведения исследований по созданию математических методов прогнозирования биологической активности химических соединений. Процедура нахождения пересечений двух графов используется для изучения зависимостей между структурой органических соединений и их свойствами. Процедура перечисления клик графа применяется при расчете структур биологических молекул [9]. В частности, с помощью этой программы был установлен важный результат: для определенных классов биологических молекул рассчитанная структура совпадает со структурой, полученной экспериментально.

На основании результатов работ по алгоритмическим исследованиям графов и их приложений создан пакет программы, который предназначен для анализа структурной информации, представленной в виде графов различных типов, и может применяться для проведения расчетов в области теории графов и ее применений в электронике, химии, биологии, социологии и других областях. Входная информация для всех прикладных программ пакета представляется в виде описаний на входном языке OGRA-30 [41]. Программная система пакета включает библиотеку инструкций, библиотеку процедур и библиотеку прикладных программ, разработанных не только в Институте математики, но также и в других организациях, в том числе и за рубежом. Прикладные программы пакета позволяют находить наибольшие общие подграфы для пар графов семейства и всех графов семейства одновременно: находить орбиты графов; группу автоморфизмов; канонические матрицы смежности и каноническую нумерацию вершин; устанавливать изоморфизм между графами двух семейств; осуществлять метрический и топологический анализ, а также содержат ряд сервисных программ.

Приведенный краткий обзор результатов, полученных в Институте математики СО АН СССР в области алгоритмических исследований методов решения прикладных задач на графах, позволяет сделать заключение о том, что создана алгоритмическая основа, которая может служить для получения новых теоретических результатов, справочных данных [19], методов, алгоритмов и программ, а также для компоновки пакетов программ, предназначенных для конкретных применений.

В заключение автор приносит благодарность Бессонову Ю.Е. и Хворостову П.В. за помощь в подготовке материалов для статьи.

Л и т е р а т у р а

1. СКОРОВОГАТОВ В.А. Относительные разбиения и слои графов. - В кн.: Вопросы обработки информации при проектировании систем (Вычислительные системы, вып.69). Новосибирск, 1977, с.3-10.
2. СКОРОВОГАТОВ В.А., ХВОРОСТОВ П.В. Анализ симметрий графов. - В кн.: Методы и программы решения оптимизационных задач на графах и сетях. Ч.II. Тезисы докл. II Всесоюз. совещания. Новосибирск, 1982, с.138-140.
3. БЕССОНОВ Ю.Е., СКОРОВОГАТОВ В.А. Применение относительных разбиений для поиска клик. - В кн.: Автоматизация проектирования в микроэлектронике. Теория. Методы. Алгоритмы (Вычислительные системы, вып.77). Новосибирск, 1978, с.26-33.
4. БЕССОНОВ Ю.Е., СКОРОВОГАТОВ В.А. О рекурсивном разборе графов. - В кн.: Алгоритмические основы обработки структурной информации (Вычислительные системы, вып.85). Новосибирск, 1981, с.3-20.
5. БЕССОНОВ Ю.Е., СКОРОВОГАТОВ В.А. Об одном семействе схем рекурсивного разбора графов. - В кн.: Машинные методы обнаружения закономерностей, анализ структур и проектирования (Вычислительные системы, вып.92). Новосибирск, 1982, с.3-49.
6. ХАРАРИ Ф. Теория графов. - М.: мир, 1973.
7. БЕССОНОВ Ю.Е., СКОРОВОГАТОВ В.А. Использование структурных особенностей схем при оптимизации размещений. - В кн.: Алгоритмические основы обработки структурной информации (Вычислительные системы, вып.85). Новосибирск, 1981, с.21-34.
8. БЕРК Л. Теория графов и ее применение. - М.: ИЛ, 1962.
9. ОМЫЛЫНЧУК Л.В., БЕССОНОВ Ю.Е., КОЛЧАНОВ Н.А. Метод расчета низкоэнергетических вторичных структур РНК, использующий алгоритм поиска клик. - В кн.: Машинные методы обнаружения закономерностей (Вычислительные системы, вып.88). Новосибирск, 1981, с.137-146.
10. СКОРОВОГАТОВ В.А. О распознавании изоморфизма неориентированных графов. - В кн.: Вычислительные системы, вып.33. Новосибирск, 1969, с.34-37.
11. СКОРОВОГАТОВ В.А., ХВОРОСТОВ П.В. Анализ метрических свойств графов. - В кн.: Методы обнаружения закономерностей с помощью ЭВМ (Вычислительные системы, вып.91). Новосибирск, 1981, с.3-20.
12. СКОРОВОГАТОВ В.А. Матрицы слов и изометричность графов. - В кн.: Автоматизация проектирования в микроэлектронике. Теория. Методы. Алгоритмы (Вычислительные системы, вып.77). Новосибирск, 1976, с.20-24.
13. CHARTRAND G., STEWART M. Isometric graphs. - Lect.Waz.Lincei.Rend.A. sci.fis.mat. l'natur., 1975(1976), v.59, N 6, p.154-157.
14. ХЕВИК Т., ГЛЕДИЧ Н.П. Структурные параметры графов. Теоретические исследования. - В кн.: Математика в социологии. М., 1977, с.151-169.
15. BONCHEV D., TRINAJSTIC N. Information theory, distance matrix and molecular branching. - The Journal of Chemical Physics, 1977, v.67, N 10.15.

16. ДЕНИЩИК Е.Ю., СКОРОВОБОГАТОВ В.А., ХВОРОСТОВ П.В. О вычислении метрических характеристик графов. - В кн.: Методы и программы решения оптимизационных задач на графах и сетях. Тезисы докл. Всесоюз. совещ., 3-5 сентября 1980 г. Новосибирск, ЦД СО АН СССР. Новосибирск, 1980, с.23-24.

17. СКОРОВОБОГАТОВ В.А. О нахождении общих частей в семействах графов химических структур. - В кн.: Использование вычислительных машин в спектроскопии молекул и химических исследованиях. Тезисы докл. У Всесоюз. конф. Новосибирск, 1980, с.8-9.

18. СКОРОВОБОГАТОВ В.А., ХВОРОСТОВ П.В. Об анализе комбинаторных симметрий химических структур. - Там же, с.10-11.

19. ХВОРОСТОВ П.В. Симметрии кубических графов. - в кн.: Машинные методы обнаружения закономерностей, анализ структур и проектирования (вычислительные системы, вып.92). Новосибирск, 1982, с.80-141.

20. СКОРОВОБОГАТОВ В.А. Нахождение общих частей в семействах графов. - В кн.: Прикладные задачи на графах и сетях. Материалы Всесоюз. совещ. Новосибирск, 1981, с.117-132.

21. БЕССОНОВ Ю.Б., СКОРОВОБОГАТОВ В.А., ХВОРОСТОВ П.В. Алгоритмы нахождения кликовой связи графа. - В кн.: Методы и программы решения оптимизационных задач на графах и сетях. 1.1. Тезисы докл. П Всесоюз. совещ. Новосибирск, 1982, с.16-18.

22. ХВОРОСТОВ П.В. Генерация подстановок при построении тестов. - в кн.: Вычислительная техника и дискретная математика. Тезисы докл. региональной науч.-техн. конф. Новосибирск, 1983, с.27.

23. МИЗАНГ В.Г. Сведение проблемы изоморфизма и изоморфного вложения к задаче нахождения неплотности. - в кн.: Труды III Всесоюз. конф. по проблемам теоретической кибернетики, Новосибирск, 1974, с.124.

24. БЕССОНОВ Ю.Б., СКОРОВОБОГАТОВ В.А. Применение относительно - ных разбиений для поиска клик. - В кн.: Автоматизация проектирования в микроэлектронике. Теория, методы. Алгоритмы (вычислительные системы, вып.77). Новосибирск, 1978, с.26-33.

25. СКОРОВОБОГАТОВ В.А., ХВОРОСТОВ П.В. Оориты, клики, канонизация. - В кн.: Методы и программы решения оптимизационных задач на графах и сетях. Тезисы докл. Всесоюз. совещ. Новосибирск, 1980, с.85-87.

26. ЗАГОРУЙКО Н.Г., СКОРОВОБОГАТОВ В.А. Выбор признаков структурного различия классов химических веществ. - В кн.: Использование вычислительных машин в спектроскопии молекул и химических исследованиях. Тезисы докл. У Всесоюз. конф. Новосибирск, 1983, с.92-93.

27. ЗАГОРУЙКО Н.Г., ЛБОВ Г.С., МАШАРОВ Ю.П. Пакет прикладных программ для обработки таблиц экспериментальных данных ОТЭС-1. - В кн.: Вопросы обработки информации при проектировании систем (вычислительные системы, вып.69). Новосибирск, 1977, с.93-101.

28. ЗЫКОВ А.А. Функции от графов, определяемые линейными уравнениями. (Сообщ. 1, 2, 3). - Изв. Сиб. отд-ния АН СССР, 1959, т. 5, с.3-19; 1960, № 9, с.17-33; 1960, № 12, с.13-27.

29. BRON C., KERBOSEN J. Finding all cliques of an undirected graph. - Comm. ACM, 1973, v.16, N 9, p.575-577.

30. DAS S.R., SHENG C., CHEN Z. An algorithm for finding all maximal complete subgraphs and an estimate of the order of computational complexity. - Comput. Elect. Engng., 1978, v.5, N 4, p.365-368.

31. TARJAN R., TROJANOVSKY A. Finding a maximum independent set. - SIAM J. Computing, 1977, v.6, N 3, p.537-546.

32. ДЕНИЩИК Е.Ю., СКОРОБОГАТОВ В.А., ХВОРОСТОВ П.В. Нахождение общих подструктур в семействах химических соединений. - В кн.: Использование вычислительных машин в спектроскопии молекул и химических исследованиях. Тезисы докл. УИ Всесоюз. конф. Новосибирск, 1983, с.197-198.

33. ДЕНИЩИК Е.Ю., СКОРОБОГАТОВ В.А., ХВОРОСТОВ П.В. Исследование модульных произведений и их применение для нахождения общих подграфов. - В кн.: Методы и программы решения оптимизационных задач на графах и сетях. Ч.П. Тезисы докл. П Всесоюз. совещ. Новосибирск, 1982, с.43-45.

34. СКОРОБОГАТОВ В.А., ХВОРОСТОВ П.В. Анализ структурной химической информации при предсказании биологической активности. - В кн.: Методы и средства обработки сложноструктурированной семантически насыщенной графической информации. Тезисы докл. I Всесоюз. конф. Горький, 1983, с.28.

35. ПИОТТУХ-ПЕЛЕЦКИЙ В.Н. Проблема структурного подобия в задачах идентификации органических соединений и других задачах органической химии. - В кн.: Использование вычислительных машин в спектроскопии молекул и химических исследованиях. Тезисы докл. УИ Всесоюз. конф. Новосибирск, 1983, с.90-91.

36. Об использовании брутто-формул связей для поиска информации об органических соединениях / Мищенко Г.Л., Гладкова Г.А., Голубева Н.Е. и др. - Информационный поиск. НТИ, сер.2, 1979, № 12.

37. БАРАЕВ А.М., ФАРАДЖЕВ И.А. Построение и исследование на ЭМ однородных и однородных двудольных графов. - В кн.: Алгоритмические исследования в комбинаторике. М., 1978, с.25-40.

38. Машинные методы установления строения органических соединений по их масс-спектрам / Дерендяев Б.Г. и др. - В кн.: Использование вычислительных машин в спектроскопии молекул и химических исследованиях. Тезисы докл. IV Всесоюз. конф., Новосибирск, 1983, с.61.

39. МОЛОДЦОВ С.Г., ПИОТТУХ-ПЕЛЕЦКИЙ В.Н. Построение всех неизоморфных структурных формул химических соединений, содержащих заданный набор непересекающихся фрагментов. - Там же, с.174-175.

40. Возможности использования структурной информации банка данных по спектрам ЯМР¹³C при определении строения неизвестных соединений / Грищенко И.В., Богданова Г.Ф., Торопов О.В. и др. - Там же, с.16-17.

41. КОЧЕТОВА А.А., СКОРОБОГАТОВ В.А., ХВОРОСТОВ П.В. Язык описания структурной информации ОГРА-3.0. - В кн.: Машинные методы обнаружения закономерностей, анализа структур и проектирования (Вычислительные системы, вып.92). Новосибирск, 1982, с.70-79.

Поступила в ред.-изд.отд.

13 декабря 1983 года