АЛГОРИТМЫ АНАЛИЗА СТРУКТУРНОЙ ИНФОРМАЦИИ
(Вычислительные системы)
1984 год Выпуск 103

УДК 519.1

НАХОЖДЕНИЕ МАКСИМАЛЬНЫХ ОБЩИХ ПОДГРАФОВ В СЕМЕЙСТВЕ ГРАФОВ

Е.Ю.Денищик

Описываемый подход основан на реализации многоместной операции модульного произведения и позволяет решать следующие задачи.

1. Пусть задано множество обыкновенных графов $M = \{G_i\}_n$, требуется найти максимальный по вложению общий для всех $G_i \in M$, возможно, несвязный подграф H такой, что $H \subseteq G_i$, $i=1,\dots,n$, $n \geq 2$.

2. Пусть задано семейство химических соединений $S = \{S_i\}_n$, представленных помеченными или частично-помеченными мультиграфами. Найти общий для соединений из семейства S структурный фрагмент $F = \bigcap\limits^{a} S_i$ с учетом исходной разметки.

3. Пусть задано множество графов или мультиграфов, возможно помеченных, требуется проверить их изоморфность.

4. Пусть задано семейство химических соединений, требуется проверить однородность этого семейства, т.е. определить, не содержатся ли в нем соединения, чья структура существенно отличается от структуры большинства соединений этого семейства.

Решение подобных задач для $n = 2$, основанное на применении двуместной операции модульного произведения к относительным разбиениям графов и последующем нахождении в графе-произведении клик максимального порядка, описано в [1-4].

Обобщением двуместной операции модульного произведения графов является многоместная операция, введенная в [4]. Легко видеть, что порядок графа соответствий t графов порядка p равен p^t, ввиду чего эта операция представляет чисто теоретический интерес. Поэтому задачи типа 1-3 для $n > 2$ решаются попарным сравнением графов. Вводя многоместную операцию модульного произведения на семействе относительных разбиений исходных графов, можно находить общие подграфы или структурные фрагменты о д н о в р е м е н н о для нескольких графов семейства существенно быстрее, чем попарным сравнением.

Пусть $C = \{G^l(V^l, X^l), l = \overline{1,n},\ n \ge 2\}$ - семейство, состоящее из n графов, и пусть $\hat{C} = \{G^l(v_0^l) : l = \overline{1,n},\ n \ge 2\}$ - множество разбиений графов (см. [3]) семейства C по отношению к вершинам из заданного набора $(v_0^1, \ldots, v_0^l, \ldots, v_0^n)$, $v_0^l \in V^l$, $l = \overline{1,n}$, который будет определен ниже. Допустим, что длина разбиений неодинакова, и пусть k - минимальная длина по семейству.

ОПРЕДЕЛЕНИЕ. Модульным произведением семейства относительных разбиений будем называть неограф $\Gamma(W, E) = \mathop{\nabla}\limits_{l=1}^{n} \hat{G}^l(v_0^l)$ со множеством вершин $W = \bigcup\limits_{i=1}^{k} W_i$, $W_i = \{w = (v_i^1, \ldots, v_i^l, \ldots, v_i^n) \mid v_i^l \in V_0^l(v_0^l),\ l = \overline{1,n},\ n \ge 2\}$, и множеством ребер E, которое определяется следующим образом. Две вершины w, $w' \in W$ смежны тогда и только тогда, когда выполняется условие $\forall l: l = \overline{1,n}$, $v^l \ne v'^l$, $(v^l, v'^l) \in X^l$ либо $(v^l, v'^l) \notin X^l$.

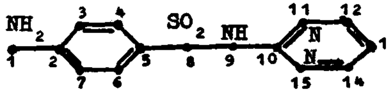
Решение задачи о нахождении общего подграфа в семействе графов с применением многоместной операции модульного произведения состоит в следующем. Будем считать, что известно разбиение множества вершин графа $G^l(V^l, X^l)$ на классы эквивалентности $\hat{V}_\theta^l = \{\hat{V}_{\theta_i}^l\}$. Класс определен как множество вершин, для которых существует автоморфизм φ, переводящий φ_j в φ_k: $\hat{V}_{\theta_i}^l = \{v_j^l \in V^l \mid v_j^l = \varphi v_k^l\}$. Такое разбиение на классы эквивалентности называется орбитальным [2].

При построении модульного произведения выбирается по представителю из каждого класса $\hat{V}_{\theta_i}^l$ и рассматриваются разбиения графов только относительно этих выбранных вершин, что позволяет учитывать симметрии исходных графов. Таким образом, в случае транзитивных графов для решения задачи достаточно построить одно модульное произведение, а для тождественных графов, у которых $|V_\theta| = p^l$, где p^l - порядок l-го графа, его требуется строить $\prod\limits_{l=1}^{n} p^l$ раз.

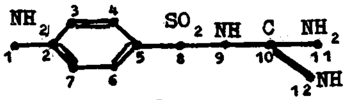
В случае помеченных графов метрически-эквивалентные классы пересекаются с меточно-эквивалентными; получающееся при этом разбиение не мельче орбитального.

Для сокращения числа разбиений, или количества модульных произведений, которые необходимо рассматривать для нахождения общей части химических структур, предлагается не строить разбиений от-

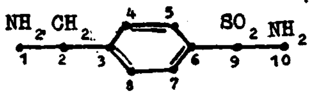
ПРИМЕР I.



Сульфазин



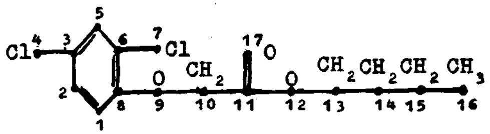
Сульгин



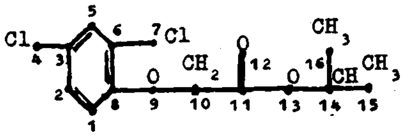
Марфанил

Подстановка, определяющая общий подграф:
2 3 4 5 6 7 8
2 3 4 5 6 7 8
3 4 5 6 7 8 9
Время счета
I мин 6 сек.
Память I40 К.

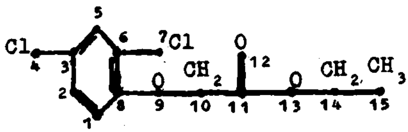
ПРИМЕР 2.



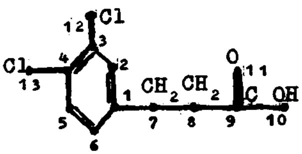
Дикопур



Изопропиловый эфир

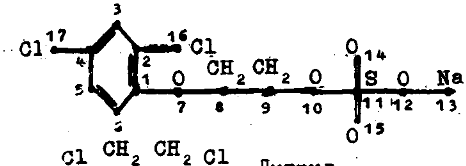


Этиловый эфир

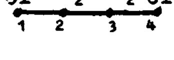


4-Д

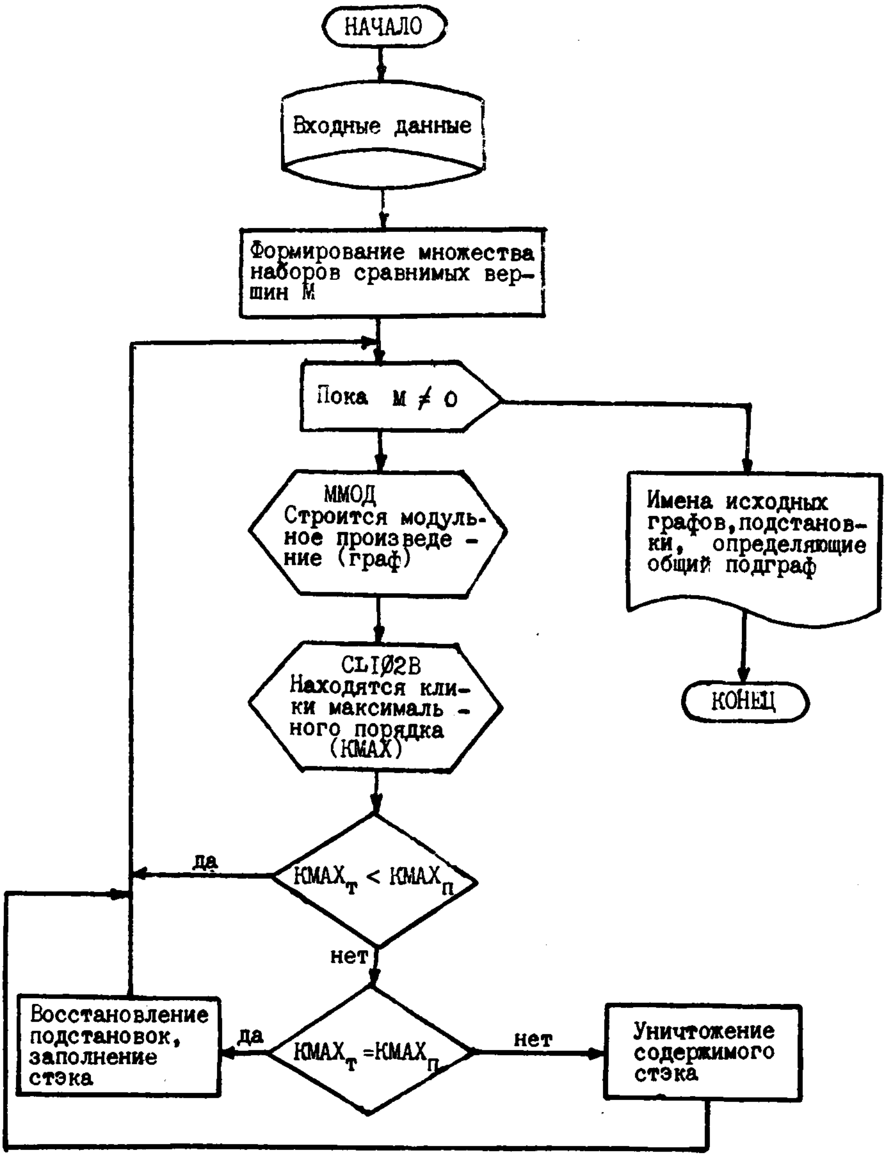
Семейство не содержит общего структурного фрагмента, большего, чем атом.
Время счета 44 сек
Память 2I2К.



Сессон



Дитрил



Блок-схема управляющей программы

носительно висячих вершин, имеющих метку CH_3, поскольку общая одновершинная компонента с меткой CH_3 неинформативна.

Нахождение не строго общих частей, но структурно похожих, достигается заданием эквивалентности некоторых вершинных меток. Например, если требуется установить наличие в исходной структуре шестичленного цикла, неважно, замещенного или нет, можно задать эквивалентность углерода и гетероатома.

Реализация включает в себя несколько программ: транслирующая программа ОГРА-3.0 (автор Кочетова А.А.), программа нахождения симметрий (автор Хворостов П.В.), программа нахождения клик максимального порядка CLIO2B (автор Бессонов Ю.Е.), программы построения многоместной операции модульного произведения MMOD и программу анализа и управления SPRING, блок-схема которой приведена на рисунке (автор Денищик Е.Ю.).

Все программы написаны на языке ПЛ/I для ОС ЕС; их суммарный объем ~1350 операторов этого языка. Объектные модули MMOD и SPRING занимают ~30 К.

Приведены два примера работы программного комплекса.

Следует отметить, что чем сильнее отличаются друг от друга исходные структуры семейства, тем быстрее работают программы.

Литература

1. 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.

2. СКОРОБОГАТОВ В.А., ХВОРОСТОВ П.В. Орбиты, клики, канонизация. -В кн.: Методы и программы решения оптимизационных задач на графах и сетях. Тезисы докладов I Всесоюзного совещания. Новосибирск, 1980.

3. СКОРОБОГАТОВ В.А. Относительные разбиения и слои графов. -В кн.: Вопросы обработки информации при проектировании систем (Вычислительные системы, вып. 69). Новосибирск, 1977, с.3-10.

4. СКОРОБОГАТОВ В.А., БЕССОНОВ Ю.Е. Об одном семействе схем рекурсивного разбора графов. -В кн.: Машинные методы обнаружения закономерностей, анализа структур и проектирования (Вычислительные системы, вып. 92). Новосибирск, 1982, с. 3-49.

Поступила в ред.-изд.отд.
29 мая 1984 года