

УДК 37.026; 519.17; 54.022

ЗАДАЧА О КЛАССИФИКАЦИИ ИЗОМЕРОВ КАК ПРИМЕР ВОЗМОЖНОЙ ИНТЕГРАЦИИ МАТЕМАТИКИ И ХИМИИ

Костин С.В., старший преподаватель, МГТУ МИРЭА, Москва, Россия
E-mail: kostinsv77@mail.ru

Аннотация. Отмечена необходимость более глубокой интеграции различных учебных дисциплин. Продемонстрирована плодотворность использования теории графов в химии (в частности, для описания структуры молекул и для нахождения количества изомеров химических соединений).

Ключевые слова: теория графов, структурная формула молекулы, изомеры, А. Кэли.

ISOMER CLASSIFICATION PROBLEM AS AN EXAMPLE OF INTEGRATION OF MATHEMATICS AND CHEMISTRY

Kostin S., senior lecturer, MSTU MIREA, Moscow, Russia
E-mail: kostinsv77@mail.ru

Annotation. Necessity of deeper integration of different education subjects is noted. High efficiency of graph theoretical methods in chemistry (e. g. when applied to the description of molecule structure and to enumeration of chemical isomers) is demonstrated.

Keywords: graph theory, structure of molecule, isomers, A. Cayley.

1. Введение

В настоящее время широко обсуждается вопрос о возможности и желательности более глубокой интеграции различных учебных дисциплин. Необходимость такой интеграции, с одной стороны, продиктована внутренними потребностями науки (хорошо известна, например, продуктивность математических методов в самых разных, иногда очень далеких от математики, областях деятельности). С другой стороны, более тесная интеграция различных дисциплин способствует развитию глубины и неформальности мышления, расширяет кругозор обучающихся, позволяет им приобрести умение взглянуть на проблему «со стороны» и применить для ее решения методы других наук. Все эти качества крайне необходимы современному исследователю.

Одним из тех разделов математики, который находит обширные и разнообразные применения как внутри самой математики, так и во многих других науках (физике, химии, биологии, экономике, кибернетике, лингвистике и др.), является теория графов. В качестве примера укажем на использование теории графов в математике (например,

при изображении диаграммы Хассе отношения порядка), теоретической физике (например, в так называемых фейнмановских диаграммах, служащих для описания элементарных актов взаимодействия частиц), электротехнике (например, при анализе структуры электрических цепей), программировании (например, для изображения граф-схем алгоритмов) и экономике (например, для решения задачи об оптимальном распределении производственных ресурсов).

Важные применения теория графов находит и в различных разделах химии (в классической теории строения молекул, в химической кинетике, в квантовой химии и т. д.). В данной статье рассматривается использование теории графов для описания структуры молекул, для нахождения брутто-формул молекул и для решения задачи о количестве изомеров химических соединений.

2. Необходимые сведения из теории графов

Пусть V и E — множества, причем $V \neq \emptyset$.

Пусть $V^{(2)}$ — множество всех двухэлементных подмножеств множества V .

Пусть $\theta: E \rightarrow V^{(2)}$ — отображение множества E во множество $V^{(2)}$.

Определение 1. Упорядоченная тройка $G = \langle V, E, \theta \rangle$ называется *графом*. \square

Элементы множества V называются *вершинами* графа G . Элементы множества E называются *ребрами* графа G . Отображение θ называется *отображением инцидентности*.

Если $\theta(e) = \{v_1, v_2\}$, то вершины v_1 и v_2 называются *концами* ребра e . Говорят, что ребро e *заканчивается* в вершинах v_1 и v_2 . Говорят также, что ребро e *соединяет* вершины v_1 и v_2 .

Если существует ребро, соединяющее вершины v_1 и v_2 , то вершины v_1 и v_2 называются *смежными* вершинами. Если ребра e_1 и e_2 имеют хотя бы один общий конец (то есть если $\theta(e_1) \cap \theta(e_2) \neq \emptyset$), то ребра e_1 и e_2 называются *смежными* ребрами. Если вершина v является концом ребра e (то есть если $v \in \theta(e)$), то вершина v и ребро e называются *инцидентными*.

Если ребра e_1 и e_2 соединяют одни и те же вершины (то есть если $\theta(e_1) = \theta(e_2)$), то ребра e_1 и e_2 называются *кратными* ребрами.

Замечание 1. Приведенное выше определение графа не является самым общим из возможных. Дело в том, что в теории графов иногда рассматривают графы с петлями

(петля — это ребро с совпадающими концами). Графы с петлями необходимы, например, в теории автоматов, поскольку диаграмма Мура автомата может содержать петли. Однако поскольку наша статья посвящена использованию графов для описания строения молекул, а в этом случае петли обычно не нужны, мы ограничились выше определением, которое не предполагает наличие у графа петель. Можно сказать, что приведенное выше определение — это определение мультиграфа (мультиграф — это граф, который может содержать кратные ребра, но при этом не содержит петель). □

Замечание 2. Приведенное выше определение — это определение так называемого обычного, или неориентированного, графа. В теории графов рассматриваются также ориентированные графы (орграфы). Ориентированные графы тоже находят важные применения в химии, например они используются в химической кинетике при описании химических реакций. В данной статье мы ориентированные графы не рассматриваем. □

Пусть $v \in V$ — вершина графа G .

Определение 2. Количество ребер, заканчивающихся в вершине v (то есть количество ребер $e \in E$ таких, что $v \in \theta(e)$), называется *степенью* вершины v в графе G . □

Обозначение: $\deg_G v$. □

Пусть $v_1, v_2 \in V$ — различные вершины графа G .

Определение 3. Количество ребер, соединяющих вершины v_1 и v_2 (то есть количество ребер $e \in E$ таких, что $\theta(e) = \{v_1, v_2\}$), называется *взаимной степенью* вершин v_1 и v_2 в графе G . □

Обозначение: $\deg_G(v_1, v_2)$. □

Определение 4. Упорядоченный набор $P = \langle v_0, e_1, v_1, e_2, v_2, \dots, v_{s-1}, e_s, v_s \rangle$ ($s \in \mathbb{N}$), состоящий из чередующихся вершин и ребер графа G , называется *путем* в графе G , если при всех $i \in [1..s]$ имеет место равенство $\theta(e_i) = \{v_{i-1}, v_i\}$. □

Путь P удобно записывать следующим образом:

$$v_0 \xrightarrow{e_1} v_1 \xrightarrow{e_2} v_2 \longrightarrow \dots \longrightarrow v_{s-1} \xrightarrow{e_s} v_s. \quad 1)$$

Запись (1) означает, что из вершины v_0 по ребру e_1 мы идем в вершину v_1 , из вершины v_1 по ребру e_2 мы идем в вершину v_2 , ..., из вершины v_{s-1} по ребру e_s мы идем в вершину v_s .

Вершина v_0 называется *началом* пути P . Вершина v_s называется *концом* пути P . Говорят, что путь P *соединяет* вершины v_0 и v_s . Число s называется *длиной* пути P .

Если начало пути P совпадает с концом (то есть если $v_0 = v_s$), то путь P называется *замкнутым*. Замкнутый путь, не содержащий повторяющихся ребер, называется *циклом*.

Определение 5. Если для любых двух различных вершин графа G существует путь, соединяющий эти вершины, то граф G называется *связным* графом. \square

Определение 6. Если в графе G не существует ни одного цикла, то граф G называется *ациклическим* графом или *лесом*. \square

Замечание 3. Если граф G является ациклическим графом, то граф G не содержит кратных ребер. Действительно, если бы в графе G было несколько ребер, соединяющих вершины v_1 и v_2 , то в графе G существовал бы цикл длины 2, идущий из вершины v_1 в вершину v_2 по одному ребру и возвращающийся назад в вершину v_1 по другому ребру. Существование такого цикла противоречит ациклическости графа G .

Одним из важнейших понятий теории графов является понятие «дерево». Дадим определение этого понятия.

Определение 7. Граф G называется *деревом*, если:

- 1) граф G является связным графом;
- 2) граф G является ациклическим графом.

3. Структурная формула молекулы

Структурную формулу молекулы можно рассматривать как помеченный граф, каждой вершине которого приписан символ определенного химического элемента.

Приведем соответствующие определения.

Пусть $G = \langle V, E, \theta \rangle$ — граф.

Пусть $\text{Mend} = \{\text{H}, \text{He}, \text{Li}, \dots\}$ — множество символов всех химических элементов из таблицы Д. И. Менделеева.

Пусть $\varepsilon: V \rightarrow \text{Mend}$ — отображение множества вершин V графа G во множество Mend .

Определение 8. Упорядоченная пара $M = \langle G, \varepsilon \rangle$ называется *структурной формулой молекулы*. \square

Граф G называется *носителем* структурной формулы молекулы $M = \langle G, \varepsilon \rangle$.

Химический символ $\varepsilon(v)$ называется *меткой* вершины $v \in V$.

В дальнейшем для краткости вместо словосочетания «структурная формула молекулы» мы будем иногда говорить просто «молекула».

Пусть $M = \langle G, \varepsilon \rangle$ — молекула с носителем $G = \langle V, E, \theta \rangle$ и пусть $M' = \langle G', \varepsilon' \rangle$ — молекула с носителем $G' = \langle V', E', \theta' \rangle$.

Определение 9. Молекулы M и M' называются *изоморфными*, если существует взаимно однозначное отображение $\varphi: V \rightarrow V'$ множества вершин V графа G на множество вершин V' графа G' такое, что:

1) отображение φ сохраняет взаимные степени любых двух вершин, то есть

$$(\forall v_1, v_2 \in V, v_1 \neq v_2): \deg_G(v_1, v_2) = \deg_{G'}(\varphi(v_1), \varphi(v_2)); \quad 2)$$

2) отображение φ сохраняет метки всех вершин, то есть

$$(\forall v \in V): \varepsilon(v) = \varepsilon'(\varphi(v)). \quad 3) \quad \square$$

Обозначение: $M \cong M'$. \square

4. Алканы и монохлоралканы

Пусть $M = \langle G, \varepsilon \rangle$ — молекула с носителем $G = \langle V, E, \theta \rangle$.

Определение 10. Молекула M называется *алканом*, если:

- 1) граф G является деревом;
- 2) степень $\deg_G v$ каждой вершины $v \in V$ равна либо 1, либо 4;
- 3) если $\deg_G v = 1$, то $\varepsilon(v) = \text{H}$, а если $\deg_G v = 4$, то $\varepsilon(v) = \text{C}$. \square

Здесь H и C — химические символы водорода и углерода.

Замечание 4. Наряду с термином «алканы» в химии используются также равносильные термины «предельные ациклические углеводороды», «насыщенные ациклические углеводороды», «парафины». \square

Определение 11. Молекула M называется *монохлоралканом*, если:

- 1) граф G является деревом;
- 2) степень $\deg_G v$ каждой вершины $v \in V$ равна либо 1, либо 4;

3) если $\deg_G v = 1$, то $\varepsilon(v) \in \{H, Cl\}$, а если $\deg_G v = 4$, то $\varepsilon(v) = C$;

4) существует ровно одна вершина $v \in V$ такая, что $\varepsilon(v) = Cl$. \square

Здесь H, C и Cl — химические символы водорода, углерода и хлора.

Замечание 5. Монохлоралкан получается из алкана в результате замены одного атома водорода на атом хлора. Наряду с монохлоралканами, содержащими один атом хлора, существуют дихлоралканы, содержащие два атома хлора, трихлоралканы, содержащие три атома хлора, и т. д. \square

5. Вывод брутто-формулы алкана

Пусть алкан содержит n атомов углерода. Сколько в этом случае он содержит атомов водорода, то есть какова его брутто-формула? Эта задача может быть достаточно легко решена с помощью теории графов.

Действительно, пусть алкан содержит m атомов водорода. Сумма степеней всех вершин графа равна удвоенному числу ребер (эту теорему теории графов часто называют «лемма о рукопожатиях»), поэтому

$$\sum_{v \in V} \deg_G v = 4n + m = 2|E|. \quad 4)$$

С другой стороны, граф G является деревом. Из теории графов известно, что в любом дереве количество ребер на единицу меньше, чем количество вершин, то есть

$$|E| = |V| - 1 = n + m - 1. \quad 5)$$

Из формул (4) и (5) находим $m = 2n + 2$. Таким образом, общая брутто-формула алкана имеет вид $C_n H_{2n+2}$. Монохлоралкан получается из алкана в результате замены одного атома водорода на атом хлора. Поэтому общая брутто-формула монохлоралкана имеет вид $C_n H_{2n+1} Cl$.

6. Изомеры алканов и монохлоралканов

Пусть M и M' — алканы.

Определение 12. Алканы M и M' называются *структурными изомерами*, если:

1) алканы M и M' имеют одинаковую брутто-формулу $C_n H_{2n+2}$;

2) молекулы M и M' не являются изоморфными. \square

Аналогично определяется понятие структурной изомерии для монохлоралканов.

В дальнейшем для краткости вместо словосочетания «структурные изомеры» мы будем иногда говорить просто «изомеры».

Введем следующие обозначения:

f_n — количество изомеров алканов с брутто-формулой C_nH_{2n+2} ;

g_n — количество изомеров монохлоралканов с брутто-формулой $C_nH_{2n+1}Cl$.

Замечание 6. Числа f_n и g_n определены и имеют ясный химический смысл при $n \in N$. Однако по чисто математическим соображениям удобно считать, что эти числа определены также при $n = 0$, а именно, что $f_0 = 1$, $g_0 = 1$. □

Замечание 7. Если заменить в молекуле монохлоралкана атом хлора Cl на атом кислорода O и присоединить к этому атому кислорода еще один дополнительный атом водорода H (иначе говоря, если заменить атом хлора Cl на гидроксильную группу –OH), то мы получим одноатомный алкановый (то есть предельный ациклический) спирт с брутто-формулой $C_nH_{2n+1}OH$. Поэтому число g_n дает не только количество изомеров монохлоралканов, но и количество изомеров одноатомных алкановых спиртов. □

Замечание 8. В химии рассматривают различные типы структурной изомерии: изомерия углеродного скелета, изометрия положения заместителя, изомерия положения кратной связи, межклассовая изомерия.

В случае алканов структурная изомерия может быть связана только с различным строением углеродного скелета.

В случае монохлоралканов структурная изомерия может быть связана, во-первых, с различным строением углеродного скелета и, во-вторых, с изомерией положения заместителя, то есть с возможностью различного расположения в молекуле атома хлора.

Аналогично, в случае одноатомных алкановых спиртов структурная изомерия может быть связана, во-первых, с различным строением углеродного скелета и, во-вторых, с изомерией положения заместителя, то есть с возможностью различного расположения в молекуле гидроксильной группы.

Замечание 9. Наряду со структурной изомерией, в химии рассматривают также так называемую пространственную изомерию или стереоизомерию (геометрическую и оптическую). В нашей статье мы пространственную изомерию не рассматриваем. □

7. Задача о перечислении изомеров алканов и монохлоралканов

При небольших значениях n числа f_n и g_n можно найти с помощью простого перебора.

Например, при $n = 5$ существует $f_5 = 3$ изомеров алкана с брутто-формулой C_5H_{12} . Это пентан, изопентан (по номенклатуре ИЮПАК его называют 2-метилбутан) и неопентан (по номенклатуре ИЮПАК его называют 2,2-диметилпропан). На рис. 1 изображены углеродные скелеты этих молекул.

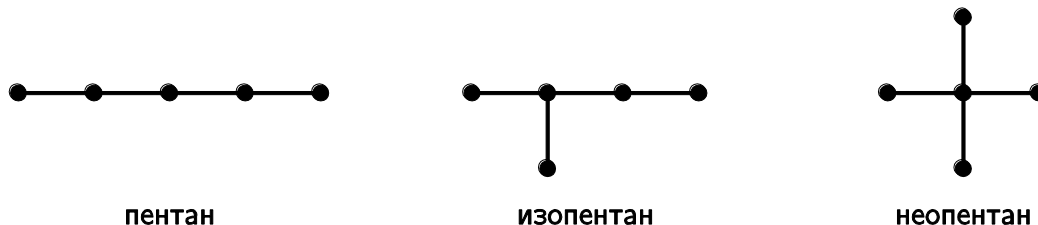


Рис. 1

Приведем еще один пример. При $n = 4$ существует $g_4 = 4$ изомеров одноатомного алканового спирта с брутто-формулой C_4H_9OH . Это бутанол (по номенклатуре ИЮПАК его называют 1-бутанол), изобутанол (по номенклатуре ИЮПАК его называют 2-метил-1-пропанол), вторбутанол (по номенклатуре ИЮПАК его называют 2-бутанол), и третбутанол (по номенклатуре ИЮПАК его называют 2-метил-2-пропанол). На рис. 2 изображены углеродные скелеты этих молекул. Атом углерода, к которому присоединена гидроксильная группа $-OH$, обведен в кружок.

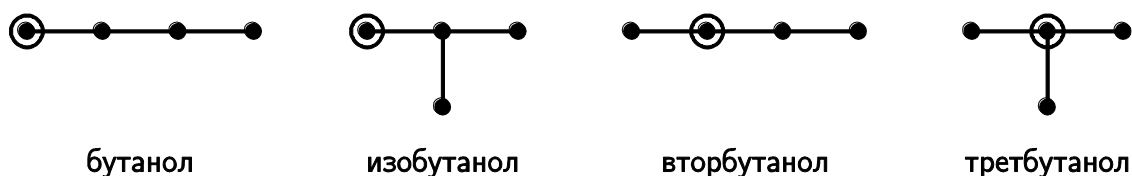


Рис. 2

К сожалению, с ростом n прямой перебор всех возможных изомеров становится очень трудоемким или даже невозможным. В связи с этим крайне полезными были бы, скажем, рекуррентные формулы, позволяющие находить числа f_{n+1} и g_{n+1} , если известны числа f_i и g_i при всех $i \in [0..n]$.

Впервые задача о перечислении изомеров алканов была поставлена в 1875 году английским математиком Артуром Кэли. В своей работе [1] он нашел числа f_n при $n \in [1..13]$. Однако, как выяснилось позднее, последние два числа f_{12} и f_{13} были найдены Кэли неверно (он получил для них значения 357 и 799, тогда как правильные значения 355 и 802).

Любопытно, что в 1965 году вышла книга по теории графов американских математиков Р. Басакера и Т. Саати, в которой были воспроизведены результаты Кэли без необходимых исправлений. В 1973 году (практически через 100 лет после исходной работы Кэли) был издан русский перевод этой книги [2], в котором также были повторены неточные значения чисел f_{12} и f_{13} (хотя в предисловии к русскому переводу редактор перевода отметил, что при переводе были «исправлены имевшиеся в оригинале неточности и опечатки»).

Хотя Кэли не удалось полностью решить задачу о перечислении изомеров алканов, он решил некоторые другие важные задачи теории графов, связанные с перечислением деревьев. В частности, он доказал фундаментальную теорему о том, что количество различных помеченных деревьев порядка n равно n^{n-2} . Сейчас эта теорема по праву носит его имя.

Что касается задачи о перечислении изомеров алканов, то она оказалась весьма сложной. Полное решение этой задачи смог получить лишь в 1937 году венгерский математик Дьёрдь Пойа (он хорошо известен также благодаря своим книгам, посвященным математическому творчеству). Для решения этой задачи Пойа применил разработанный им метод, который сейчас называют «теория перечисления Пойа».

К сожалению, подробное изложение найденного Пойа решения задачи перечисления изомеров алканов заняло бы очень много места. (Перевод фундаментальной статьи Пойа на русский язык [3] имеет объем более ста страниц.) Поэтому мы ограничимся лишь тем, что выпишем окончательные рекуррентные формулы. Эти рекуррентные формулы отсутствуют в явном виде в статье Пойа, однако они могут быть получены исходя из приведенных в статье функциональных уравнений для производящих функций последовательностей f_n и g_n ($n = 0, 1, 2, \dots$).

Возможно, это покажется несколько странным, но более простая рекуррентная формула получается не для чисел f_n , а для чисел g_n (то есть для количества изомеров монохлоралканов с брутто-формулой $C_nH_{2n+1}Cl$ или для количества изомеров одноатомных алкановых спиртов с брутто-формулой $C_nH_{2n+1}OH$):

$$g_{n+1} = \frac{1}{6} \left[\sum_{i+j+k=n} g_i g_j g_k + 3 \sum_{i+2j=n} g_i g_j + 2 \sum_{3i=n} g_i \right]. \quad 6)$$

В формуле (6) индексы суммирования i, j, k принимают целые неотрицательные значения. С помощью формулы (6) можно найти числа g_n при всех $n \in \mathbb{N}$, если учесть, что, по определению, $g_0 = 1$.

Более сложная рекуррентная формула получается для чисел f_n (то есть для количества изомеров алканов с брутто-формулой C_nH_{2n+2}):

$$f_{n+1} = g_{n+1} + \frac{1}{24} \left[\sum_{i+j+k+l} g_i g_j g_k g_l + 6 \sum_{i+2j+2k=n} g_i g_j g_k + 8 \sum_{i+3j=n} g_i g_j + 3 \sum_{2i+2j=n} g_i g_j + 6 \sum_{4i=n} g_i + 12 \sum_{2i=n+1} g_i - 12 \sum_{i+j=n+1} g_i g_j \right]. \quad (7)$$

В формуле (7) индексы суммирования i, j, k, l принимают целые неотрицательные значения. С помощью формулы (7) можно найти числа f_n при всех $n \in \mathbb{N}$. Обратим внимание на то, что для того, чтобы найти число f_{n+1} , надо предварительно найти все числа $g_i, i \in [0..n+1]$. Внешне формулы (6) и (7) выглядят весьма громоздко. Однако автор данной статьи потратил немного времени и запрограммировал эти формулы на компьютере. Результаты расчетов при $n \in [0..20]$ приведены в таблице 1. Расчет на компьютере всех чисел f_n и $g_n, n \in [0..100]$, занимает всего несколько секунд. А ведь еще 140 лет назад великому Кэли приходилось считать вручную и он допустил арифметическую ошибку уже в числе $f_{12} \dots$

Таблица 1

n	f_n	g_n	n	f_n	g_n
1	1	1	2	1	1
3	1	2	4	2	4
5	3	8	6	5	17
7	9	39	8	18	89
9	35	211	10	75	507
11	159	1238	12	355	3057
13	802	7639	14	1858	19241
15	4347	48865	16	10359	124906
17	24894	321198	18	60523	830219
19	148284	2156010	20	366319	5622109

8. Заключение

Мы рассмотрели только один пример интеграции различных дисциплин, а именно, использование математических методов (теории графов) в химии. Однако даже этот один пример, по нашему мнению, очень убедительно показывает, насколько полезной и плодотворной может быть указанная интеграция.

При этом очень важно понимать, что процесс интеграции различных дисциплин, как правило, носит взаимовыгодный характер. Исследования Кэли и Пойа по теории графов изначально были порождены потребностями химии, однако в результате они оказались полезны не только для химии, но и значительно стимулировали развитие самой математики.

Автор будет благодарен читателям за любые комментарии или замечания по затронутым в данной статье вопросам.

Список литературы

1. Cayley A. Ueber die analytischen Figuren, welche in der Mathematik Bäume genannt werden und ihre Anwendung auf die Theorie chemischer Verbindungen // Ber. deutsch. chem. Ges. 1875. V. 8. P. 1056–1059.
2. Басакер Р., Саати Т. Конечные графы и сети. М.: Наука, 1974. 368 с.
3. Пойа Д. Комбинаторные вычисления для групп, графов и химических соединений // Перечислительные задачи комбинаторного анализа: Сб. переводов. М.: Мир, 1979. С. 36–138.