CROATICA CHEMICA ACTA
CCACAA **78** (4) 563–567 (2005)

ISSN-0011-1643
CCA-3049
Note

Counting Disconnected Structures: Chemical Trees, Fullerenes, *I*-graphs, and others*

Marko Petkovšek[a] and Tomaž Pisanski[a,b,]**

[a] *IMFM, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia*

[b] *University of Primorska Slovenia, Koper, Slovenia*

RECEIVED FEBRUARY 3, 2005; REVISED MARCH 4, 2005; ACCEPTED MARCH 8, 2005

Keywords
chemical trees
enumeration
fullerenes
I-graphs

When counting isomers with a given number of atoms one usually assumes that we want to count (connected) molecules. In this note we present a method that can be used for counting disconnected structures if counts of connected structures are given. The method can be used also in the reverse direction. If the numbers of all structures are known, the number of connected structures can be readily determined.

INTRODUCTION

We assume that the reader is familiar with graph-theoretic language; for a reference, see Ref. 1. In order to present the problem addressed in this paper, we ask the following simple question. What is the number of regular graphs of valence 1 on n vertices? Let b_n denote the number of such graphs. Clearly, the only connected regular graph of valence 1 is K_2, the complete graph on two vertices that consists of a single edge. If we let a_n denote the number of connected regular graphs of valence 1 on n vertices, then we have the following table:

n	0	1	2	3	4	5	6	7	8	9	10 …
a_n	0	0	1	0	0	0	0	0	0	0	0 …

It is not hard to see that the only regular 1-valent graphs are disjoint unions of some number of K_2-s. Hence

n	0	1	2	3	4	5	6	7	8	9	10 …
b_n	1	0	1	0	1	0	1	0	1	0	1 …

Instead of writing the sequence a_n we can provide its generating function:

$$F(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + ...$$

Similarly, we can define:

$$G(x) = b_0 + b_1x + b_2x^2 + b_3x^3 + ...$$

In our particular example we have

$$F(x) = x^2$$

and

$$G(x) = 1 + x^2 + x^4 + x^6 + ... = 1/(1–x^2)$$

There is a similar question for regular 2-valent graphs. The only connected ones are the cycles. Hence:

n	0	1	2	3	4	5	6	7	8	9	10 …
a_n	0	0	0	1	1	1	1	1	1	1	1 …

---

* Dedicated to Dr. Edward C. Kirby on the occasion of his 70th birthday.

** Author to whom correspondence should be addressed. (E-mail: Tomaz.Pisanski@fmf.uni-lj.si)

and

$$F(x) = x^3 + x^4 + x^5 + ... = x^3/(1-x)$$

One has to think a bit before we can compile the table of all regular 2-valent graphs.

n	0	1	2	3	4	5	6	7	8	9	10...
b_n	1	0	0	1	1	1	2	2	3	4	5 ...

It is not at all obvious what should the corresponding generating function $G(x)$ look like. In this paper we will show how to apply the theory of combinatorial species for answering this question. The formula relating $F(x)$ to $G(x)$ is general and gives a method of calculating b_n if a_n is known and *vice versa*.

Long time ago we extended our computer system Vega[2] to handle similar problems. We will present examples that will show how our method can be used to obtain new enumeration results.

THE METHOD

If $F(x)$ is the generating function of the counting sequence of connected structures, then the corresponding generating function $G(x)$ of the counting sequence of all structures is given by

$$G(x) = \exp\sum_{k\geq 1}\frac{F(x^k)}{k} \quad (1)$$

(*cf.* Ref. 3). Applying a variant of Möbius inversion to this formula it is also possible to express $F(x)$ in terms of $G(x)$:

$$F(x) = \sum_{k\geq 1}\frac{\mu(k)}{k}\log G(x^k) \quad (2)$$

where μ denotes the well-known Möbius function (*cf.* Ref. 3).

Now we give some examples.

Example 1: Let a_n be the number of connected regular 1-valent graphs on n vertices and b_n the number of all regular 1-valent graphs on n vertices. Then $F(x) = x^2$ and

$$G(x) = \exp\sum_{k\geq 1}\frac{x^{2x}}{k} = \exp(-\log(1-x^2)) = \frac{1}{1-x^2}$$

in agreement with our observation above.

Example 2: Let a_n be the number of connected regular 2-valent graphs on n vertices and b_n the number of all regular 2-valent graphs on n vertices. Then $F(x) = x^3 / (1-x)$ and

$$G(x) = \exp\sum_{k\geq 1}\frac{x^{3k}}{k(1-x^k)} =$$
$$1 + x^3 + x^4 + x^5 + 2x^6 + 2x^7 + 3x^8 + 4x^9 + 5x^{10} + \cdots$$

Example 3: Let a_n be the number of connected graphs on n vertices and b_n the number of all graphs on n vertices. Then

$$F(x) = x + x^2 + 2x^3 + 6x^4 + 21x^5 + 112x^6 + 853x^7 + 11117x^8 + 261080x^9 + 11716571x^{10} + \cdots$$

(*cf.* Ref. 4) and

$$G(x) = \exp\sum_{k\geq 1}F(x^k)/k$$
$$= 1 + x + 2x^2 + 4x^3 + 11x^4 + 34x^5 + 156x^6 + 1044x^7 + 12346x^8 + 274668x^9 + 12005168x^{10} + \cdots$$

(*cf.* Ref. 4).

Example 4: The number of fullerene isomers is well known (*cf.* Ref. 4). There is one on 20 atoms, none on 22 atoms, one on 24 atoms, *etc*. The values of $a_{20}, a_{22}, a_{24}, ..., a_{60}$ are

1,0,1,1,2,3,6,6,15,17,40,45,89,116,199,271,
437,580,924,1205,1812,...

Thus there are 1812 non-isomorphic fullerenes on 60 atoms, one being the renowned buckminster-fullerene. Using our method we can easily compute $b_{60} = 1892$.

It is now a simple matter to apply the method to a number of sequences that were produced for connected structures. For instance, in Ref. 1 the initial numbers of chemical trees

1,1,1,1,2,3,5,9,18,35,75,159,355,802,1858,
4347,10359,24894,...

are presented (*cf.* Ref. 4). The corresponding numbers of chemical forests on n carbon atoms are easily computed using the same procedure.

Generically, b_n is a polynomial in $a_1, a_2, ..., a_n$ and *vice versa*. Expanding the right-hand side of (1) into power series we obtain the explicit formula

$$b_n = \sum_{j=0}^{n}\frac{1}{j!}\sum_{i_1k_1+\cdots+i_jk_j=n}\frac{a_{i_1}\cdots a_{i_j}}{k_1\cdots k_j}$$

where the inner sum is taken over all j-tuples $(i_1,...,i_j)$ and $(k_1,...,k_j)$ of positive integers whose dot product equals n. If we first differentiate (1) w.r.t. x, then expand both sides into power series we obtain the recursive formula

$$b_n = \frac{1}{n}\sum_{k=1}^{n}b_{n-k}\sum_{j|k}ja_j \quad \text{for} \quad n\geq 1 \quad (3)$$

Here we list the first five b's as polynomials in the a's:

$b_0 = 1$
$b_1 = a_1$
$b_2 = (a_1 + a_1^2)/2 + a_2$
$b_3 = (2a_1 + 3a_1^2 + a_1^3)/6 + a_1a_2 + a_3$
$b_4 = (6a_1 + 11a_1^2 + 6a_1^3 + a_1^4)/24$
$\quad + (a_2 + a_1a_2 + a_1^2a_2 + a_2^2)/2 + a_1a_3 + a_4$

In the initial terms of these polynomials (corresponding to $a_2 = a_3 = ... = 0$) we see Stirling cycle numbers $s_{n,k}$ (a.k.a. unsigned Stirling numbers of the first kind). This can be explained combinatorially as follows: When $a_2 = a_3 = ... = 0$, each connected component consists of a single point. There are a_1 types (or: colors) of points to choose from, so in order to construct all unlabelled structures on n points we have to take all combinations with repetitions of n elements out of a_1 possible types. This can be done in

$$\binom{a_1+n-1}{n} = \frac{a_1(a_1+1)\cdots(a_1+n-1)}{n!} = \sum_{k=0}^{n} s_{n,k}a_1^k/n!$$

ways, giving the initial term of b_n.

In general, we have $b_n = a_n + B_n(a_1,a_2,...,a_{n-1})$. All the coefficients of B_n are nonnegative, and the products $a_p^q a_r^s\cdots$ have the property that $pq + rs + \cdots \le n$.

From (2) we obtain similarly

$$a_n = \sum_{k|n}\frac{\mu(k)}{k}\sum_{(i_1+\cdots+i_j)k=n}\frac{(-1)^{j+1}}{j}b_{i_1}\cdots b_{i_j}$$

where the inner sum is taken over all j-tuples $(i_1,...,i_j)$ of positive integers whose sum equals n/k. From (3) we also obtain the recursive formula

$$a_n = b_n - \frac{1}{n}\left(\sum_{\substack{j|n\\ j\neq n}} ja_j + \sum_{k=1}^{n-1} b_{n-k}\sum_{j|k} ja_j\right) \quad \text{for } n \ge 1$$

Here we list the first six a's as polynomials in the b's:

$a_0 = 0$
$a_1 = b_1$
$a_2 = b_2 - (b_1 + b_1^2)/2$
$a_3 = b_3 - b_1b_2 - (b_1 - b_1^3)/3$
$a_4 = b_4 - b_1b_3 + b_1^2b_2 - (b_2 + b_2^2)/2 + (b_1^2 - b_1^4)/4$
$a_5 = b_5 - b_1b_4 + b_1^2b_3 + b_1b_2^2 - b_1^3b_2 - b_2b_3$
$\quad + (b_1^5 - b_1)/5$

Again we have $a_n = b_n - A_n(b_1,b_2,...,b_{n-1})$, and the products $a_p^q a_r^s\cdots$ in A_n have the property that $pq + rs + \cdots \le n$.

A RELATED PROBLEM

Finally, let us turn to another problem. Let us start with a sequence a_n counting connected objects of certain type on n elements. Assume that a disconnected object on n elements can be formed by selecting k identical connected objects with m elements where $m \cdot k = n$. Let b_n count the total number of elements. Then $b_n = \sum_{k|n} a_k$ and, by Möbius inversion, $a_n = \sum_{k|n} \mu(n/k)b_k$.

Example 5: Let a_n be the number of connected vertex-transitive graphs on n vertices. Then b_n is the total number of vertex-transitive graphs on n vertices.

Example 6: *The generalized Petersen graph* $G(n,k)$ is a graph with vertex set

$$V(G(n,k)) = \{u_0,u_1,...,u_{n-1},v_0,v_1,...,v_{n-1}\}$$

and edge set

$$E(G(n,k)) = \{u_iu_{i+1},u_iv_i,v_iv_{i+k} : i = 0,...,n-1\}$$

Here subscripts are to be read modulo n. Note that $G(n,k)$ is isomorphic to $G(n,n-k)$, and $G(n,n/2)$ is not simple. Therefore, for $n \ge 3$, we consider only graphs $G(n,k)$ where $k < n/2$.

Generalized Petersen graphs constitute a standard family of graphs which represents a generalization of the renowned Petersen graph $G(5,2)$. This important and well--known family of graphs introduced in 1969 by Mark Watkins[5] possesses a number of interesting properties. For example, $G(n,r)$ is vertex transitive if and only if $n = 10$, $r = 2$ or $r^2 \equiv \pm 1 \pmod n$. It is a Cayley graph if and only if $r^2 \equiv 1 \pmod n$. It is arc-transitive only in the following seven cases: (n,r) = (4,1), (5,2), (8,3), (10,2), (10,3), (12,5), (24,5). The family contains some very important graphs, such as the n-prism $G(n,1)$, the Dürer graph $G(6,2)$, the Möbius-Kantor graph $G(8,3)$, the dodecahedron $G(10,2)$, the Desargues graph $G(10,3)$, *etc*.

The generalized Petersen graphs form a special case of the so-called I-graphs, see Ref. 6. The I-graph $I(n,j,k)$ is a graph with vertex set

$$V(I(n,j,k)) = \{u_0,u_1,...,u_{n-1},v_0,v_1,...,v_{n-1}\}$$

and edge set

$$E(I(n,j,k)) = \{u_iu_{i+j},u_iv_i,v_iv_{i+k} : i = 0, ..., n-1\}$$

Since $I(n,j,k) = I(n,k,j)$ we usually assume that $j \le k$. Clearly $G(n,k) = I(n,1,k)$. Following the usual representation of these graphs where we draw vertices u_i on one circle and vertices v_i on another concentric circle (with smaller radius), we call u_i and v_i *the vertices on the outer rim* and *the vertices on the inner rim*, respectively. The edges between the two rims are called *spokes*. The class a

b

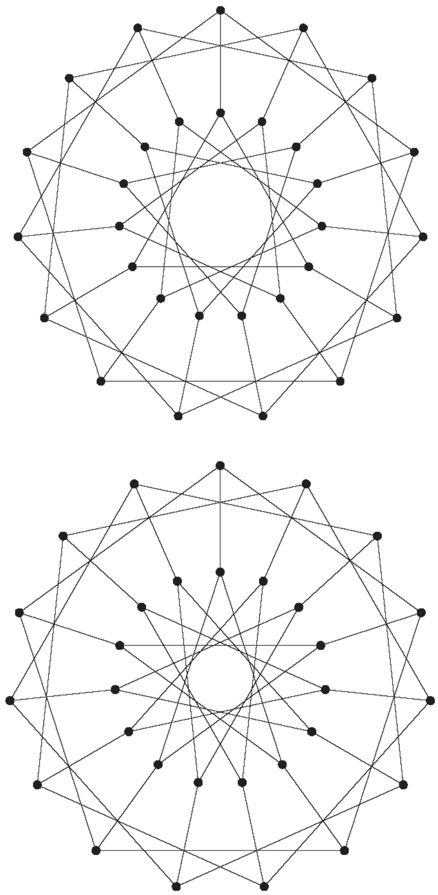

Figure 1. *I*-graphs *I*(15,3,5) (a) and *I*(15,3,6) (b). The former is connected and the latter is not.

of graphs $I(n,j,k)$ contains the class $G(n,k)$. The I-graphs $I(15,3,5)$ and $I(15,3,6)$ are depicted in Figure 1. The reader is referred to Refs. 7–10 for the motivation of the study of generalized Petersen graphs and *I*-graphs in connection to combinatorial and geometric configurations.

If a_n counts connected *I*-graphs then b_n counts all *I*-graphs.

Example 7: There are only 7 connected arc-transitive *I*-graphs. They are the generalized Petersen graphs G(4,1), G(5,2), G(8,3), G(10,2), G(10,3), G(12,5), G(24,5). Hence

n	0	1	2	3	4	5	6	7	8	9	10…
a_n	0	0	0	0	1	1	0	0	1	0	2 …

There are infinitely many arc-transitive *I*-graphs if we drop the condition of connectivity.

n	0	1	2	3	4	5	6	7	8	9	10…
b_n	1	0	0	0	1	1	0	0	2	0	3 …

If we think of a and b as infinite column vectors $a = (a_1,a_2,...)^T$ *etc*., then there is an infinite triangular 0-1 matrix $\boldsymbol{M}$ such that $b = \boldsymbol{M}a$.

Clearly $(\boldsymbol{M})_{nk} = 1$ if and only if $k \mid n$, and we have

$$b_1 = a_1$$
$$b_2 = a_1 + a_2$$
$$b_3 = a_1 + a_3$$
$$b_4 = a_1 + a_2 + a_4, \quad etc.$$

Let $\boldsymbol{N}$ be the inverse of $\boldsymbol{M}$. Matrices $\boldsymbol{M}$ and $\boldsymbol{N}$ do not depend on a (or b). By Möbius inversion, $(\boldsymbol{N})_{\mathrm{nk}} = \mu(n/k)$ if $k \mid n$, otherwise 0. Hence

$$a_1 = b_1$$
$$a_2 = -b_1 + b_2$$
$$a_3 = -b_1 + b_3$$
$$a_4 = -b_2 + b_4, \quad etc.$$

CONCLUSION

The method explained in this paper has numerous applications not only in mathematics and computer science but also in chemistry. A possible application would be to find connected particles from a mass spectrogram that presents also some particle combinations.

REFERENCES

1. N. Trinajstić, *Chemical Graph Theory*, 2nd edition, CRC Press, 1992.
2. *Vega 0.2 Quick Reference Manual and Vega Graph Gallery*, T. Pisanski (Ed.), Ljubljana, 1995, http://vega.ijp.si/.
3. F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial Species and Tree-like Structures*, Cambridge University Press, 1998.
4. N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at http://www.research.att.com/njas/sequences/ 2005.
5. M. Watkins, *J. Combin. Theory* **6** (1969) 152–164.
6. I. Z. Bouwer, W. W. Chernoff, B. Monson, and Z. Star, *The Foster Census,* Charles Babbage Research Centre, 1988.
7. A. Betten, G. Brinkmann, and T. Pisanski, *Discrete Appl. Math.* **99** (2000) 331–338.
8. M. Boben, T. Pisanski, and A. Žitnik, *J. Combin. Designs*, to appear.
9. M. Boben and T. Pisanski, *Eur. J. Combin.* **24** (2003) 431–457.
10. M. Boben, T. Pisanski, A. Žitnik, and B. Grünbaum, submitted.

SAŽETAK

Prebrojavanje nepovezanih struktura: kemijska stabla, fulereni, *I*-grafovi i druge strukture

Marko Petkovšek i Tomaž Pisanski

Kada se prebrojavaju izomeri određenoga broja atoma, obično se pretpostavlja da se želi doznati broj povezanih struktura. U ovoj su noti autori prikazali metodu koja se rabi za prebrojavanje nepovezanih struktura, ako se pozna broj povezanih izomera. Ta se metoda može rabiti i u obrnutome smislu. Ako je poznat broj svih struktura, broj povezanih struktura je također poznat.