

ON A RECURRENT METHOD FOR COUNTING GRAPHS WITH MARKED VERTICES

MATHEMATICS

1969

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196901.44547>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

UDC 519.1

MATHEMATICS

V. A. LISKOVETS

ON A RECURRENT METHOD FOR COUNTING GRAPHS WITH MARKED VERTICES

(Presented by Academician V. M. Glushkov on 24 VI 1968)

In the present note a certain principle is formulated for counting graphs with marked vertices, which makes it possible to find the number of strongly connected directed graphs and graphs of certain other types. Asymptotic and other properties of the quantities obtained are studied.

§ 1. Let $\{y_i\}$ be a fixed set of elements, $i = 1, 2, \dots$; $Y = Y_n = \{y_1, \dots, y_n\}$, $n = 1, 2, \dots$. If P is some class of graphs with distinguished (labeled) vertices $(^{1,3})$, then by $P(M)$ we denote the class of all graphs from P whose set of vertices is the set M . If $A \in P(M)$, $N \subseteq M$, then by $A(N)$ we denote the subgraph of the graph A induced on N . We adhere to the terminology of $(^2)$. Everywhere we restrict ourselves to graphs without loops and multiple edges or arcs, having the elements y_i as vertices.

The following **principle of graph counting** is valid. Let P and Q be two classes of graphs with marked vertices, and let in $Y = Y_n$ some family of subsets $\Phi_n = \Phi$, $n = 1, 2, \dots$, $Y \in \Phi$, be distinguished in such a way that the following conditions are satisfied:

1. There exists a single-valued mapping $\varphi : P(Y) \rightarrow \Phi$ such that if $A \in P(Y)$, $\varphi(A) = Z$, then $A(Z) = B \in Q(Z)$.
2. Conversely, for all $Z \in \Phi_n$ and all $B \in Q(Z)$, the graph B corresponds in the indicated way ($B = A(\varphi(A))$) to exactly $a(n, t)$ graphs from $P(Y)$, where t is the number of elements in Z ; $a(n, t)$ depends only on n and t , but not on the form of B and Z ; $a(n, n) = 1$.

Then, if $p(n)$ and $q(n)$ denote the numbers of all graphs with arbitrary n vertices from the classes P and Q , respectively, and $\beta(n, t)$ is the number of sets in Φ_n containing exactly t elements, the relation

$$q(n) + \sum_{t=1}^{n-1} \beta(n, t) a(n, t) q(t) = p(n), \quad n = 1, 2, \dots \quad (1)$$

holds.

The proof of this assertion is obvious.

When $p(n)$, $a(n, t)$, and $\beta(n, t)$ are known, equality (1) is a linear recurrence relation for $q(n)$, containing all $q(i)$, $i = 1, \dots, n$. The initial conditions are set automatically at $n = 1$.

Conditions 1 and 2 mean that graphs of the class Q are embedded in a definite way as unambiguous components in graphs from P , for which the enumeration problem is solved. Therefore in what follows the indicated principle will be called the **principle of embedding**. Usually the class P is chosen naturally for Q , and Q is distinguished in it by a condition of the type: $A(\varphi(A))$ is a maximal component of A with a certain property for some preselected vertex; then Φ is the set of all subsets of Y containing this vertex, and $\beta(n, t) = C_{n-1}^{t-1}$ (a binomial coefficient).

The principle of embedding generalizes the method applied by Gilbert ⁽³⁾ for counting connected graphs with marked vertices. In it, it is no longer required that $A(\varphi(A))$ be separated independently in A . Usually $a(n, t) = p(n-t)a'(n, t)$, $t < n$, with an obvious combinatorial meaning. For connected graphs $a'(n, t) = 1$. This made it possible to obtain in ⁽³⁾ a convenient explicit expression for their number by means of generating functions (formal

logarithm). The same can be done for $q(n)$ when

$$a(n, t) = a_1(n-t)a_2(t)a_3(n)$$

and ordinary $\beta(n, t)$. Otherwise it is not possible to obtain an explicit solution of equation (1) in general form, unless one counts as an obvious expression $q(n)$ in the form of a certain determinant of order n . From the computational point of view, however, relation (1) is quite satisfactory. The extension of the inclusion principle to the case of enumeration by the number of vertices and arcs is obvious (cf. (3)). Other generalizations of this principle are also possible.

§ 2. Despite its simplicity, the inclusion principle can be applied to finding many previously unknown quantities. This mainly concerns the enumeration of directed graphs of various kinds possessing certain (stronger) connectedness properties.

We shall call a directed graph **initially connected** with respect to a chosen vertex (y_1) if all its other vertices are reachable from it (by directed paths). Let $c(n)$, $a(n)$, $s(n)$ denote, respectively, the number of (weakly) connected, initially connected (with respect to the chosen vertex), and strongly connected ([2], p. 14) directed graphs, where n is the number of labeled vertices, and let $c_0(n)$ denote the number of connected undirected graphs. Let $a(n, N)$, $s(n, N)$ be the number of the corresponding graphs with n vertices and N arcs. Application of the inclusion principles immediately gives, as in [3],

$$c(n) + \sum_{t=1}^{n-1} C_{n-1}^{t-1} 2^{(n-t)(n-t-1)} c(t) = 2^{n(n-1)},$$

$$c_0(n) + \sum_{t=0}^{n-1} C_{n-1}^{t-1} 2^{(n-t)(n-t-1)/2} c_0(t) = 2^{n(n-1)/2}.$$

Proposition 1.

$$a(n) + \sum_{t=1}^{n-1} C_{n-1}^{t-1} 2^{(n-1)(n-t)} a(t) = 2^{n(n-1)}, \quad (2)$$

$$a(n, N) + \sum_{t=1}^{n-1} \sum_{T=0}^{t(t-1)} C_{n-1}^{t-1} C_{(n-1)(n-t)}^{N-T} a(t, T) = C_{n(n-1)}^N. \quad (3)$$

Corollary.

$$a(n) = 2^{n(n-1)/2} c_0(n); \quad a(n, N) = \sum_{T=0}^N C_{n(n-1)/2}^{N-T} c_0(n, T).$$

This assertion is not difficult to prove directly as well.

Theorem 1*.

$$s(n) + \sum_{t=1}^{n-1} C_{n-1}^{t-1} \lambda_t(n-t) s(t) = a(n), \quad (4)$$

where

$$\lambda_t(m) + \sum_{k=0}^{m-1} C_m^k 2^{(m-1)(m-k)} \lambda_t(k) = 2^{m(m+t-1)}. \quad (5)$$

The proof splits into two stages, each time applying the inclusion principle and solving an auxiliary combinatorial problem; $\lambda_1(n) = 2^{-n} a(n+1)$.

Corollary. 2 enters into the canonical factorization of the number $s(n)$ with the same exponent as in $(n-1)!$

Finding successively from (2) and (5) $a(n)$ and $\lambda_t(k)$ and substituting their values into (4), we compute the values $s(n)$:

$$s(3) = 18, \quad s(4) = 1606, \quad s(5) = 2^3 \cdot 70635, \quad s(6) = 2^3 \cdot 91846847, \dots$$

* This solves the “labeled” variant of problem 1 from [6].

Theorem 2.

$$s(n, N) + \sum_{t=1}^{n-1} \sum_{T=0}^{t(t-1)} C_{n-1}^{t-1} \lambda_t(n-t, N-T) s(t, T) = a(n, N),$$

where

$$\lambda_t(m, M) + \sum_{k=0}^{m-1} \sum_{K=0}^{k(k-1)} C_m^k C_{(m-1)(m-k)}^{M-K} \lambda_t(k, K) = C_{m(m+t-1)}^M.$$

From the computational point of view, the scheme for finding $s(n, N)$ is rather complicated and, for large n , requires on the order of $1/90n^6$ multiplications for a single step.

Let us next consider the number $s_0(n)$ of strongly connected tournaments ⁽⁴⁾ with n (labeled) vertices, i.e., strongly connected complete antisymmetric graphs ^(2, pp. 14, 124). Then by our method, or by direct reasoning (see ⁽⁴⁾), one easily obtains the relation

$$s_0(n) + \sum_{t=1}^{n-1} C_n^t 2^{(n-t)(n-t-1)/2} s_0(t) = 2^{n(n-1)/2}. \quad (6)$$

In an analogous way some other classes of initially connected or strongly connected graphs are also enumerated. The same method is applicable also to the enumeration of connected initial automata, which are of independent interest. Here we use this only to derive several identities.

Let $g_{r,d}(n)$ denote the number of abstract automata without outputs with a set of states Y_n and r input signals, each of whose states is reachable from at least one of the given d states y_1, \dots, y_d , $1 \leq d \leq n$. Application of the inclusion principle with $\beta(n, t) = C_{n-d}^{t-d}$ gives in this case

$$g_{r,d}(n) + \sum_{t=d}^{n-1} C_{n-d}^{t-d} n^{r(n-t)} g_{r,d}(t) = n^{nr}. \quad (7)$$

We note that $g_{r,1}(n)/(n-1)!$ is the number of pairwise nonisomorphic (with respect to states) connected initial automata ⁽⁵⁾. For $r = 2$ this solves the 19th problem posed by Harary ⁽⁶⁾.

Direct computation for $r = 1$ gives

$$g_{1,1}(n) = n!, \quad g_{1,2}(n) = \frac{1}{2}n!(n+2), \quad g_{1,3}(n) = \frac{1}{24}n!(3n^2 + 19n + 24).$$

Therefore

$$\sum_{t=1}^n C_{n-1}^{t-1} n^{n-t} t! = n^n. \quad (8)$$

This is the identity found by Riordan ⁽⁷⁾. Further, from (7), with the help of (8), one obtains a new identity ($d = 2$),

$$\sum_{t=1}^n C_{n-1}^{t-1} n^{n-t} t(t+1)! = 2n^{n+1}, \quad (9)$$

and with the help of (8) and (9) (for $r = 1$, $d = 3$),

$$\sum_{t=1}^n C_{n-1}^{t-1} n^{n-t} t(3t+1)(t+2)! = 24n^{n+2}.$$

Apparently, in general the identity holds:

$$\sum_{t=1}^n C_{n-1}^{t-1} n^{n-t} f_k(t)(t+k)! = (2k)!n^{n+k}, \quad k = 0, 1, \dots,$$

where $f_k(t)$ is a certain polynomial of degree k with integer coefficients. In particular (if this is true for $k = 3, 4$), $f_3(t) = 15t^2(t+1)$, $f_4(t) = 7t(15t^3 + 30t^2 + 5t - 2)$. In the general case the form of $f_k(t)$ is unknown.

§ 3. Owing to the presence of a free term in (1), the inclusion principle in many cases makes it possible simply to obtain asymptotic estimates of the quantities under investigation. Namely, often $q(n) \sim p(n)$ as $n \rightarrow \infty$, which usually follows easily from (1), since $q(n) \leq p(n)$, and therefore

$$q(n) \geq p(n) - \sum_{t=1}^{n-1} \beta(n, t) \alpha(n, t) p(t)$$

and so on. In this way one can then obtain further estimates. If $q(n) \approx p(n)$, finding the asymptotic behavior of $q(n)$ is much more difficult (as is the case with $g_{r,d}(n)$ in (7)).

Proposition 2 (cf. (1)). $c_0(n) = 2^{n(n-1)/2}(1 - n2^{-n+1} - C_n^3 2^{-3n+7} - O(n^4 2^{-4n}))$; $c(n) = 2^{n(n-1)}(1 - n2^{-2n} - O(n^2 2^{-4n}))$ as $n \rightarrow \infty$.

Proposition 3. $s_0(n) = 2^{n(n-1)/2} - 2n2^{(n-1)(n-2)/2} + n(n-1)2^{(n-2)(n-3)/2} - \gamma(n)$, where

$$0 \leq \gamma(n) < 16(C_n^3 + 64)2^{(n-3)(n-4)/2}.$$

This proposition refines the result (4) and, in particular, makes it possible to estimate the number $\bar{s}_0(n) = 2^{n(n-1)/2} - s_0(n)$ of not strongly connected tournaments.

Theorem 3. $s(n) \geq 2^{n(n-1)} - 2(n+4)2^{(n-1)^2}$; $s(n) = 2^{n(n-1)}(1 - n2^{-n+2} + n(2n-1)2^{-2n+2} + O(n^3 2^{-3n}))$ as $n \rightarrow \infty$.

The fact that $s(n) \sim 2^{n(n-1)}$ also follows from (8).

Theorem 3 also gives an asymptotic estimate for the number $\bar{s}(n)$ of not strongly connected graphs (see also (9)). Let $d(n)$ denote the number of semicomplete graphs with n labeled vertices, i.e., directed graphs having at least one vertex with zero out- or indegree. It is not difficult to show that $d(n) \sim 2n2^{(n-1)^2}$.

Corollary 1. Asymptotically almost all not strongly connected graphs are semicomplete and connected (weakly).

It follows from Theorem 3 that already for $n \geq 13$ more than 99% of directed graphs are strongly connected (and connected, respectively, for $n \geq 6$).

Let $G_0(n)$ denote the number of all pairwise nonisomorphic undirected graphs with n vertices, and $G(n)$ the corresponding number of directed graphs. Denote by $S(n)$ the number of pairwise nonisomorphic strongly connected graphs. As was shown by Oberschelp ⁽¹⁰⁾,

$$G_0(n) = \frac{2^{n(n-1)/2}}{n!} (1 + n(n-1)2^{-n+1} + O(n^4 2^{-2n})),$$

which strengthens the well-known result of Pólya ^(1,6). Similarly one can obtain that

$$G(n) = \frac{2^{n(n-1)}}{n!} (1 + O(n^2 2^{-2n})).$$

Hence it follows

Corollary 2.

$$S(n) = \frac{2^{n(n-1)}}{n!} (1 - n2^{-n+2}(1 + o(1))).$$

Let us note that the estimate $s(n, N) \sim C_n^N(n-1)$ as $n \rightarrow \infty$ is valid when $N = [n(\ln n + \alpha)]$, $\alpha \rightarrow \infty$, which in fact follows from ⁽¹¹⁾, where graphs with loops and the case $\alpha = \text{const}$ are considered.

The results presented can evidently be interpreted in the language of (random) matrices of 0's and 1's (cf. ^(8,9)), random graphs of a specified type (see ^(4,11)), or random Markov chains.

Institute of Mathematics
Academy of Sciences of the Belorussian SSR

Received
24 VI 1968

REFERENCES

1. G. W. Ford, G. E. Uhlenbeck, Proc. Nat. Acad. Sci. U.S.A., 42, No. 3, 122 (1956); 43, No. 1, 163 (1957).
2. K. Berge, *Theory of Graphs and Its Applications*, IL, 1962.
3. E. N. Gilbert, Canad. J. Math., 8, No. 3, 405 (1956).
4. J. W. Moon, L. Moser, Canad. Math. Bull., 5, No. 1, 61 (1962).
5. V. M. Glushkov, UMN, 16, issue 5 (101), 3 (1961).
6. F. Harary, Magyar tud. akad. Mat. kutató int. közl., A5, No. 1-2, 63 (1960).
7. J. Riordan, Ann. Math. Statist., 33, No. 1, 178 (1962).
8. J. W. Moon, L. Moser, Studia sci. math. Hung., 1, No. 1-2, 153 (1966).
9. B. R. Heap, Numer. Math., 8, 114 (1966).
10. W. Oberschelp, Math. Ann., 174, No. 1, 53 (1967).
11. I. Palásti, Studia sci. math. Hung., 1, No. 1-2, 205 (1966).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.