

The Enumeration of Labeled Graphs
by Number of Cutpoints

WPI-CS-TR-97-01

Stanley M. Selkow*

*Computer Science Department
Worcester Polytechnic Institute
Worcester, MA 01609 USA
email: sms@wpi.edu*

* Much of this work was done at the Département d'Informatique,
Université de Marne-la-Vallée, Noisy-le-Grand, France

ABSTRACT: A generating function is developed to express the number of labeled graphs with a fixed number of points and cutpoints in terms of the generating function of the number of blocks. An asymptotic bound is derived for the number of connected graphs with any number of cutpoints.

ACKNOWLEDGEMENT: The referees' comments improved the exposition of this paper.

1. INTRODUCTION AND DEFINITIONS

A *cutpoint* of a graph is a point whose removal increases the number of components, and a *noncutpoint* is a point which is not a cutpoint. A *trivial graph* is a graph with exactly one point, and a *block* is a maximal nontrivial connected graph without a cutpoint. We only count labeled graphs.

According to Harary and Palmer [3], in 1950 Uhlenbeck posed the problem of counting blocks, and Riddell [4] and Ford and Uhlenbeck [1] derived an expression relating the number of blocks to the number of connected graphs. The derivation of this expression is an exercise in Goulden and Jackson [2]. An *exponential generating function (EGF)* of the series $\langle A_1, A_2, \dots \rangle$ is

the formal power series $A(z) = \sum_{n \geq 1} A_n \frac{z^n}{n!}$. For any non-negative integer n , the

operator $[z^n]$ applied to polynomial $P(z)$ yields the coefficient of z^n in $P(z)$,

and D_z denotes the derivative with respect to z . For integers $p, q \geq 0$, we let

denote the falling factorial function defined recursively as $p^0 = 1$ and

$p^q = p * (p - 1)^{q-1}$ for $q > 0$.

p^q

The EGF for graphs is $G(z) = \sum_{n \geq 1} G_n \frac{z^n}{n!}$, where $G_n = 2^{\binom{n}{2}}$ is the number of

graphs of n points, $C(z) = \sum_{n \geq 1} C_n \frac{z^n}{n!}$ is the EGF of connected graphs, and

$B(z) = \sum_{n \geq 1} B_n \frac{z^n}{n!}$ is the EGF of blocks (with $B_1 = 0$). Because all blocks are

connected, and all connected graphs are graphs, $B_n \leq C_n \leq G_n$ for all $n \geq 1$. Our goal is to find a semiexponential generating function for

$$S(x, z) = \sum_{m \geq 0} \sum_{n \geq 2} S_{m, n} x^m \frac{z^n}{n!}$$

where $S_{m, n}$ is the number of connected graphs with $n \geq 2$ points and $m \geq 0$ cutpoints. Some small values of $S_{m, n}$ are shown in TABLE I.

It is easy to see that $B(z)$, $C(z)$ and $G(z)$ can be expressed in terms of $S(x, z)$ as:

$$B(z) = [x^0] S(x, z) = S(0, z)$$

$$C(z) = z + \sum_{m \geq 0} [x^m] S(x, z) = z + S(1, z)$$

$$G(z) = e^{z + S(1, z)}.$$

The number of graphs with n points and m cutpoints is $[x^m \frac{z^n}{n!}] e^{z + S(x, z)}$. Given

an enumeration of blocks by their numbers of points and edges, our results can also be extended in a straightforward way to enumerate connected graphs by their numbers of points, cutpoints and edges.

In SECTION 2 we characterize the generating function $S(x,z)$ and develop a recurrence for $S_{m,n}$. In SECTION 3 we examine the asymptotic growth of $S_{m,n}$.

2. THE GENERATING FUNCTION $S(x,z)$

Since a graph with one cutpoint may be considered to be rooted at that cutpoint, the number of connected graphs with one cutpoint can be expressed as a function of $B(z)$.

THEOREM 1: $[x]S(x,z) = z(e^{B'(z)} - B'(z) - 1)$.

Proof. To enumerate graphs with one cutpoint, we identify the unlabeled roots of $k \geq 2$ blocks. There are n ways to choose the root of a block of n points. Thus, the EGF of rooted blocks is $zD_z B(z)$, and $D_z B(z)$ counts rooted blocks with the root unlabeled. There are $\frac{(B'(z))^k}{k!}$ k -multisets of rooted blocks with unlabeled roots. Since the new cutpoint can belong to any number $k \geq 2$ of blocks, the number of multisets being counted is

$$\sum_{k \geq 2} \frac{(B'(z))^k}{k!} = \sum_{k \geq 0} \frac{(B'(z))^k}{k!} - B'(z) - 1 = e^{B'(z)} - B'(z) - 1$$

and the theorem follows from adding a factor of z to label the root.

Although extending THEOREM 1 to graphs with two cutpoints is straightforward, it seems to be very difficult to handle much larger values of m . This is because the cutpoint of THEOREM 1 corresponds to a root, but with several cutpoints the root can be any cutpoint which belongs to at least one block containing no other cutpoint. However, $S(x,z)$ may be characterized as in the following relation.

THEOREM 2: $(zD_z - xD_x)S(x,z) = zB'(z + \Phi S(x,z))$, where $\Phi = x((1-x)D_x + zD_z)$.

Proof. Both sides of the equation enumerate connected nontrivial graphs rooted at a noncutpoint. Every nontrivial graph has at least two noncutpoints, and a graph counted by $S_{m,n}$ has $n-m$ of them. Thus there are

$$\sum_{m \geq 0} \sum_{n \geq 2} (n-m) S_{m,n} x^m \frac{z^n}{n!} = (zD_z - xD_x)S(x,z)$$

ways to choose a nontrivial connected graph and root it at a noncutpoint.

A noncutpoint of a nontrivial connected graph G belongs to exactly one block. Assume that the block containing the root has $k \geq 2$ points, $\{v_1, \dots, v_k\}$, where v_k is the root of G . There are kB_k rooted blocks with k points. For each $v_i \in \{v_1, \dots, v_k\}$, remove $\{v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k\}$ from G , and let H_i be the (possibly trivial) component rooted at v_i . If H_i is trivial, it is counted by z . H_k is trivial. If H_i is not trivial, then v_i is a cutpoint of G . If v_i is a cutpoint of H_i ,

then $x D_x S(x, z)$ counts the number of choices for H_i times the number of ways to root it at a cutpoint. If v_j is not a cutpoint of H_i , then $(z D_z - x D_x) S(x, z)$ counts the number of choices for H_i times the number of ways to root it at a noncutpoint. Finally, we add a factor of x since v_j becomes a cutpoint of G .

Connected graphs rooted at a noncutpoint which belongs to a block of $k \geq 2$ points are counted by

$$k B_k \frac{z(z + x D_x S(x, z) + x(z D_z - x D_x) S(x, z))^{k-1}}{k!}.$$

Since k can be any integer greater than or equal to 2, combining these possibilities yields

$$\begin{aligned} & \sum_{k \geq 2} k B_k \frac{z(z + x D_x S(x, z) + x(z D_z - x D_x) S(x, z))^{k-1}}{k!} \\ &= z \sum_{k \geq 2} B_k \frac{(z + x D_x S(x, z) + x(z D_z - x D_x) S(x, z))^{k-1}}{(k-1)!} \\ &= z B'(z + x D_x S(x, z) + x(z D_z - x D_x) S(x, z)) \end{aligned}$$

which yields the theorem.

For $m=0$, the relation of THEOREM 2 reduces to the identity $z B'(z) = z B'(z)$. For $m \geq 1$, the recurrence permits the computation of $S_{m,n}$ for all $n \geq 2$. This task is made easier for $n \gg m$ by noticing that at most m points of the block containing the root can be cutpoints. The recurrence of the following COROLLARY was used to compute the values of $S_{m,n}$ displayed in TABLE I and used in TABLE II.

COROLLARY 1: For $m \geq 1$ and $n \geq 2$,

$$S_{m,n} = \frac{n!}{n-m} \sum_{k=2}^{n-m} B_k \sum_{q=1}^{\min(k-1, m)} \left[x^m z^{n+q-k} \right] \frac{(\Phi S(x, z))^q}{q!(k-1-q)!}.$$

Proof. As in THEOREM 2, we enumerate connected nontrivial graphs, with at least one cutpoint, rooted at a noncutpoint. These are counted by the

terms of $(z D_z - x D_x) S(x, z)$ with positive exponents of x , or $\sum_{m \geq 1} \sum_{n \geq 2} (n-m) S_{m,n} x^m \frac{z^n}{n!}$.

Let the block containing the root have k points, q of which are cutpoints, $k > q \geq 1$. Each subgraph rooted at one of the q cutpoints of the block containing the root is counted by $\Phi S(x, z)$. There are $k-q$ choices for the root of the graph, and B_k ways to arrange the lines of the block containing the root. Combining these terms yields

$$\sum_{q=1}^{k-1} \frac{(\Phi S(x, z))^q}{q!} (k-q) B_k \frac{z^{k-q}}{(k-q)!}$$

Allowing k to vary from 2 to $n-m$,

$$\sum_{m \geq 1} \sum_{n \geq 2} \frac{n-m}{n!} S_{m,n} x^m z^n = \sum_{k=2}^{n-m} B_k \sum_{q=1}^{\min(k-1, m)} \frac{(\Phi S(x, z))^q}{q!(k-1-q)!} z^{k-q}$$

and the COROLLARY follows from comparing coefficients.

3. ASYMPTOTIC VALUES OF $S_{m,n}$

Since $S_{m,n}=0$ unless $0 \leq m \leq n-2$, then for any n , $\sum_{m \geq 0} S_{m,n}$ is distributed among the values $\{S_{0,n}, \dots, S_{n-2,n}\}$. Because the only graph with n points and $n-2$ cutpoints is a path of length n , and the number of labeled paths of length n is $n!$ divided by the order of the symmetry group of the path, then $S_{n-2,n} = \frac{n!}{2}$ for $n \geq 2$.

The derivation of other values of $S_{m,n}$ appears to be more difficult, so we consider asymptotics. Since it follows that $\sum_{m \geq 0} S_{m,n} = C_n \sim G_n = 2^{\binom{n}{2}}$, then $\sum_{m \geq 0} S_{m,n} \sim 2^{\binom{n}{2}}$. Almost all graphs are blocks [3], $S_{0,n} \sim G_n$, so $S_{0,n} \sim \sum_{m \geq 0} S_{m,n} \sim 2^{\binom{n}{2}}$, and as $n \rightarrow \infty$, almost all of $\sum_{m \geq 0} S_{m,n}$ is concentrated in the first term. In fact, the following theorem shows that the series $\{\sum_{k \geq i \geq 0} S_{i,n}\}_{k=0}^m$ yields increasingly tight approximations to $C_n = \sum_{k \geq i \geq 0} S_{i,n} + o(S_{k,n})$.

THEOREM 3: For any $m \geq 1$, $S_{m,n} = o(S_{m-1,n})$.

Proof. Any connected graph with n points and m cutpoints must have a cutpoint, w , belonging to blocks with a total of $p \geq \frac{n}{m}$ points. The contribution of such a graph to $S_{m,n}$ consists of four factors:

- $\binom{n}{p}$, the number of ways to choose labels for the p points,
- $S_{1,p}$, the number of ways to distribute edges and labels among the p points,
- the number of ways to distribute edges and labels among the $n-p$ other points,
- the number of ways to choose and identify cutpoints connecting the sets of p points and $n-p$ points.

By making w a noncutpoint, the only factor above that changes is that $S_{1,p}$ is replaced by $S_{0,p}$. Since $S_{0,p} \sim G_p$, the theorem follows from the observation that $S_{1,p} = o(S_{0,p})$.

It will be seen (THEOREM 4) that since B_n grows so rapidly with n , then for any fixed $m \geq 0$, as $n \rightarrow \infty$ almost all connected graphs with n points and m cutpoints consist of $m+1$ blocks, with m of the blocks being single edges and $n-m$ of the points belonging to a single large block. Each of the m simple blocks has a distinct vertex of the big block as an endpoint. There are

$\binom{n}{m} = \frac{n^m}{m!}$ ways to choose the labels for the m points not belonging to the big

block, $(n-m)^m$ choices of points in the big block for attaching the m little

blocks, and $B_{n-m} \sim 2^{\binom{n-m}{2}}$ ways to arrange the edges and labels on the vertices of the big block. Combining terms,

$$S_{m,n} \sim \frac{n^m}{m!} (n-m)^m 2^{\binom{n-m}{2}} = \frac{n^{2m}}{m!} 2^{\binom{n-m}{2}} \sim \frac{n^{2m}}{m!} 2^{\binom{n-m}{2}}.$$

LEMMA 1: For any $q > m \geq 0$ and $n \geq 0$, $[x^m z^n](\Phi S(x, z))^q = 0$.

Proof Since $\Phi = x((1-x)D_x + zD_z)$, then $(\Phi S(x, z))^q$ must be divisible by x^q .

LEMMA 2: For any $m \geq 1$ and $q > n - m$, $[x^m z^n](\Phi S(x, z))^q = 0$.

Proof (by induction on q):

$$[x^m z^n] \Phi S(x, z) = \frac{m S_{m,n} + (n-m+1) S_{m-1,n}}{n!},$$

so $[x^m z^n] \Phi S(x, z) = 0$ if $n \leq m$, which establishes the LEMMA for $q=1$.

$$\begin{aligned} [x^m z^n](\Phi S(x, z))^q &= [x^m z^n] \Phi S(x, z) (\Phi S(x, z))^{q-1} \\ &= \sum_{j,k} [x^j z^k] \Phi S(x, z) [x^{m-j} z^{n-k}] (\Phi S(x, z))^{q-1}. \end{aligned}$$

By the basis, $[x^j z^k] \Phi S(x, z) = 0$ unless $k > j$, but in that case $q-1 > (n-k) - (m-j)$,

so that $[x^{m-j} z^{n-k}] (\Phi S(x, z))^{q-1} = 0$ by the induction hypothesis.

LEMMA 3: For any $m \geq q \geq 1$, $[x^m z^{m+q}] (\Phi S(x, z))^q = \binom{m-1}{q-1}$.

Proof.

$$[x^m z^{m+q}] (\Phi S(x, z))^q = \sum_{\substack{r_1 + \dots + r_q = m \\ s_1 + \dots + s_q = m+q}} \prod_{q \geq k \geq 1} [x^{r_k} z^{s_k}] \Phi S(x, z).$$

By LEMMA 1 and LEMMA 2, every term in the above summation is 0 unless $r_i \geq 1, s_i = r_i + 1, q \geq i \geq 1$, so

$$\begin{aligned}
[x^m z^{m+q}] (\Phi S(x, z))^q &= \sum_{r_1 + \dots + r_q = m} \prod_{q \geq k \geq 1} [x^{r_k} z^{r_k + 1}] \Phi S(x, z) = \sum_{\substack{r_1 + \dots + r_q = m \\ r_1 \geq 1, \dots, r_q \geq 1}} 1 \\
&= \binom{m-1}{q-1},
\end{aligned}$$

which is the number of compositions of m into q parts.

Consequences of LEMMA 3 which will be used in THEOREM 4 are that for any $m \geq 1$, $[x^m z^{2m}] (\Phi S(x, z))^m = 1$ and $[x^m z^{m+1}] \Phi S(x, z) = 1$.

THEOREM 4: For any fixed m ,

$$S_{m,n} \sim \frac{n^{2m}}{m!} 2^{\binom{n-m}{2}}.$$

Proof. The proof is by induction on m , the number of cutpoints. The basis $S_{0,n} \sim 2^{\binom{n}{2}}$ is proved in Harary and Palmer [3]. We now fix an $m \geq 1$ and assume the statement of the theorem for all smaller values. Being interested in asymptotic results, we assume $n \gg m$. From COROLLARY 1,

$$S_{m,n} = \frac{n!}{n-m} \sum_{k=2}^{n-m} B_k \sum_{q=1}^{\min(k-1, m)} [x^m z^{n+q-k}] \frac{(\Phi S(x, z))^q}{q!(k-1-q)!}.$$

We will show that the sum is dominated by the term ($k=n-m$, $q=m$), and now consider three cases:

$k=n-m$ By LEMMA 3,

$$\begin{aligned}
&\frac{n!}{n-m} B_{n-m} \sum_{q=1}^m [x^m z^{m+q}] \frac{(\Phi S(x, z))^q}{m!(n-m-q-1)!} \\
&= \frac{n!}{n-m} B_{n-m} \sum_{q=1}^m \binom{m-1}{q-1} \frac{1}{m!(n-m-q-1)!} \\
&= \frac{n!}{n-m} B_{n-m} \left(\frac{1}{m!(n-2m-1)!} + \sum_{q=1}^{m-1} \binom{m-1}{q-1} \frac{1}{m!(n-m-q-1)!} \right) \\
&\sim \frac{n!}{n-m} \frac{B_{n-m}}{m!(n-2m-1)!} \left(1 + o\left(\frac{1}{n}\right) \right) \\
&\sim \frac{n^{2m}}{m!} 2^{\binom{n-m}{2}}.
\end{aligned}$$

$k=2$ By THEOREM 3 and the induction hypothesis,

$$\frac{n!}{n-m} B_2 [x^m z^{n-1}] \Phi S(x, z) = \frac{n!}{n-m} \left(\frac{m S_{m,n-1} + (n-m) S_{m-1,n-1}}{(n-1)!} \right) \sim n S_{m-1,n-1}$$

$$\sim \frac{n^{2m-1}}{(m-1)!} 2^{\binom{n-m}{2}} = o\left(\frac{n^{2m}}{m!} 2^{\binom{n-m}{2}}\right).$$

$n-m > k > 2$

Before treating this case, a bound must be determined for

$$[x^m z^{n+q-k}] (\Phi S(x, z))^q.$$

LEMMA 4: For any fixed $m \geq q \geq 1$,

$$[x^m z^n] (\Phi S(x, z))^q = \frac{2^{\binom{n-m-q+2}{2}}}{(n-4m)!} \mathcal{O}(1).$$

Proof (by induction on q): For $q=1$,

$$[x^m z^n] \Phi S(x, z) = \frac{m S_{m,n} + (n-m+1) S_{m-1,n}}{n!}$$

By THEOREM 3 and the induction hypothesis of THEOREM 4, this is asymptotic to

$$\frac{S_{m-1,n}}{(n-1)!} \sim \frac{n^{2m-2}}{(n-1)!(m-1)!} 2^{\binom{n-m+1}{2}} = \frac{2^{\binom{n-m+1}{2}}}{(n-2m+1)!} \mathcal{O}(1)$$

Fixing $q > 1$, we assume the LEMMA for all smaller values.

$$[x^m z^n] (\Phi S(x, z))^q = \sum_{j=q-1}^{m-1} \sum_{k=j+q-1}^{n-m+j-1} [x^j z^k] (\Phi S(x, z))^{q-1} [x^{m-j} z^{n-k}] \Phi S(x, z).$$

The term $2^{\binom{n-m-q+2}{2}}$ in the induction hypothesis grows so rapidly that the inner summation of the last equation is dominated by the end values, $k=j+q-1$ and $k=n-m+j-1$.

$$\begin{aligned} & [x^m z^n] (\Phi S(x, z))^q \\ & \sim \sum_{j=q-1}^{m-1} \left(\binom{j-1}{q-2} [x^{m-j} z^{n-j-q+1}] \Phi S(x, z) + [x^j z^{n-m+j-1}] (\Phi S(x, z))^{q-1} \right) \\ & = \sum_{j=q-1}^{m-1} \left(\frac{2^{\binom{n-m+q+2}{2}}}{(n-2m+j-q+2)!} + \frac{2^{\binom{n-m+q+2}{2}}}{(n-m-j-2q+4)!} \right) \mathcal{O}(1) = m \frac{2^{\binom{n-m-q+2}{2}}}{(n-4m+5)!} \mathcal{O}(1) \end{aligned}$$

where the last step is justified by LEMMA 1. This finishes the proof of LEMMA 4.

Using LEMMA 4 to finish the third case of THEOREM 4 (where the constants in the \mathcal{O} -notation depend upon m),

$$\begin{aligned} & \frac{n!}{n-m} \sum_{k=3}^{n-m-1} B_k \sum_{q=1}^{\min(k-1, m)} [x^m z^{n+q-k}] \frac{(\Phi S(x, z))^q}{q!(k-q-1)!} \\ & = \frac{n!}{n-m} \sum_{k=3}^{n-m-1} 2^{\binom{k}{2}} \sum_{q=1}^{\min(k-1, m)} \frac{2^{\binom{n-m-k+2}{2}}}{(n-4m-k+q)!(k-q-1)!} \mathcal{O}(1) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=3}^{n-m-1} \sum_{q=1}^{\min(k-1, m)} n^{4m+k-q-1} \frac{2^{\binom{n-m-k+2}{2} + \binom{k}{2}}}{(k-q-1)!} \alpha(1) \\
&= \sum_{k=3}^{n-m-1} \sum_{q=1}^{\min(k-1, m)} \frac{n^{2m}}{m!} 2^{\binom{n-m}{2}} \frac{1}{n^{q-2m-k+1} (k-q-1)! 2^{\frac{(n-m-k+2)k-2n}{2}}} \alpha(1) \\
&= \frac{n^{2m}}{m!} 2^{\binom{n-m}{2}} \frac{1}{n^{q-2m-k} (k-q-1)! 2^{\frac{n}{2}}} \alpha(1) = o\left(\frac{n^{2m}}{m!} 2^{\binom{n-m}{2}}\right)
\end{aligned}$$

which establishes THEOREM 4.

TABLE II gives an indication of the rate of convergence of $S_{m,n}$ to its asymptotic value $\frac{n^{2m}}{m!} 2^{\binom{n-m}{2}}$ for some small values of m and n .

THEOREM 4 suggests an efficient way to generate a random graph with n points, m of which are cutpoints, for $n \gg m$, such that almost all such graphs are chosen with the same probability.

repeat

generate a random graph G of $n-m$ points

until G is a block

Randomly select m points of G

for each of the m points

attach an edge between the point and a new point

Randomly relabel each of the n points

Since $B_n \sim G_n$, the **repeat-until** loop will almost certainly be executed one time. Because the test for being a block (connected with no cutpoints) can be performed in time $O(n^2)$, the expected execution time of the algorithm is

$O(n^2)$, which is optimal, since almost all graphs being counted have $\Omega(n^2)$ edges.

n	$S_{0,n}$	$S_{1,n}$	$S_{2,n}$	$S_{3,n}$
2	1			
3	1	3		
4	10	16	12	
5	238	250	180	60
6	11368	8496	4560	1920
7	1014888	540568	211680	75600
8	166537616	61672192	17186624	4663680
9	50680432112	12608406288	2416430016	469336898
10	29107809374336	4697459302400	597615868800	79132032000
11	32093527159296128	3256012245850496	266262716016000	23121510192000
12	68846607723033232640	4276437400678311936	218583901063537152	12082931084928000

$S_{m,n}$

TABLE I

n	$\frac{C_n}{G_n}$	$\frac{S_{0,n}}{C_n}$	$\frac{S_{0,n}}{2^{\binom{n}{2}}}$	$\frac{S_{1,n}}{n^2 2^{\binom{n-1}{2}}}$	$\frac{S_{2,n}}{\frac{n^4}{2!} 2^{\binom{n-2}{2}}}$	$\frac{S_{3,n}}{\frac{n^6}{3!} 2^{\binom{n-3}{2}}}$	$\frac{S_{4,n}}{\frac{n^8}{4!} 2^{\binom{n-4}{2}}}$
5	.71094	.32692	.23242	.19531	.37500	.00018	.00000
10	.98045	.84379	.82729	.68357	.44526	.22640	.08811
15	.99908	.98731	.98641	.91283	.72892	.49616	.28418
20	.99996	.99928	.99924	.94876	.81045	.61909	.42002
25	.99999+	.99996	.99996	.95994	.84777	.68630	.50724
30	.99999+	.99999+	.99999+	.96666	.87214	.73260	.57142
35	.99999+	.99999+	.99999+	.97143	.88975	.76700	.62111
40	.99999+	.99999+	.99999+	.97500	.90309	.79359	.66067

TABLE II

REFERENCES

- [1] G.W. Ford and G.E. Uhlenbeck, Combinatorial Problems in the Theory of Graphs I, *Proc. Nat. Acad. Sci. U.S.A.*, 42 (1956), 122-128.
- [2] I.P. Goulden and D.M. Jackson, "Combinatorial Enumeration", John Wiley, New York, 1983.
- [3] F. Harary and E. Palmer, "Graphical Enumeration", Academic Press, New York, 1973
- [4] R.J. Riddell, "Contributions to the Theory of Condensation", dissertation, University of Michigan, Ann Arbor, 1951