Asymptotics of Smallest Components Sizes in Decomposable Combinatorial Structures of Alg-Log Type

Daniel Panario
School of Mathematics and Statistics
Carleton University
daniel@math.carleton.ca

Based on joint work with
Li Dong, Zhicheng (Jason) Gao and Bruce Richmond

CanaDAM'09, May 2009

Decomposable structures

A decomposable combinatorial structure consists of simpler objects called components which by themselves can not be further decomposed.

For example, permutations decompose into cycles, graphs into connected components, and polynomials over finite fields into irreducible factors.

The size of an object is the number of its elements. For example, a permutation on n objects, a graph on n vertices and a monic n-degree polynomial over a finite field have all size n.

Related work

- Riddel and Uhlenbeck (1953):
  exponential formula in chemistry and physics;
- Foata & Schützenberger (1970):
  exponential formula for labelled objects;
- Bender & Goldman (1971):
  exponential formula for unlabelled objects;
- Joyal (1981) and Bergeron, Labelle & Leroux (1994):
  theory of species;
- Wilf (1990): exponential formula with cards, hands and decks;
- Flajolet (1990's):
  the symbolic method and exp–log/alg–log classes;
- Arratia, Barbour & Tavaré (1990's): probabilistic approach;

Related work (cont.)

- Labelle & Leroux (1996), Cameron (1997): $A(C(z))$;
- Goncharov (1942, 1962) and Shepp & Lloyd (1966): cycle distribution in permutations;
- Stepanov (1969): distributions in random mappings;
- Knuth & Trabb-Pardo (1976): permutations and numbers;
- Flajolet & Soria (1990, 1993): number of components in exp–log and alg–log classes;
- Gourdon (1996): largest size of components in exp–log and alg–log;
- Panario & Richmond (2001): smallest size of components in exp–log.

Factorization pattern

The factorization pattern of a decomposable combinatorial object of size n is an n-tuple $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that the object has α_i components of size i, and $\sum_{i=1}^{n} i\alpha_i = n$.

The restricted pattern of a decomposable object is a mapping $S : J \mapsto \mathbb{N}$, where J is a set of components' sizes, $\mathbb{N}$ is the set of nonnegative integers, and $S(j)$ is the number of components of size j. If we do not have a restricted pattern, J is the empty set.

Notation: $S = \left[\prod_{j\in J} j^{S(j)}\right]$.

We derive asymptotic expressions for the probability that a random decomposable object has a given restricted pattern and a restricted size on its rth smallest component.

Log and alg-log components

We turn the set of all structures of size n into a probability space using the uniform distribution. Let $X_n(r)$ be the size of the rth smallest component in a random structure of size n. We derive results about the limiting distribution of $X_n(r)$ when the component generating function $C(z)$ is of algebraic-logarithmic type. We also study the asymptotic properties of structures with a restricted pattern.

When the component generating function $C(z)$ is of logarithmic type, the smallest component size has been studied by Panario and Richmond (2001), Arratia, Barbour and Tavaré (2003), and Dong, Gao and Panario (2007).

Generating functions

Notation:

$$C(z) = \sum_{k\geq 1} C_k \frac{z^k}{k!}, \qquad C(z) = \sum_{k\geq 1} C_k z^k,$$

$$C(z;J) = \sum_{j\in J} C_j \frac{z^j}{j!}, \qquad C(z;J) = \sum_{j\in J} C_j z^j,$$

$$\mathscr{S} = \begin{cases} \prod_{j\in J} \left(\frac{C_j}{j!}\right)^{S(j)} \frac{1}{S(j)!} & \text{for labeled case,} \\ \prod_{j\in J} \binom{C_j+S(j)-1}{S(j)} & \text{for unlabeled case.} \end{cases}$$

Let $m' = \sum_{j\in J} jS(j)$ denote the size of the restricted pattern, and $m = n - m'$ be the total size of the unrestricted components.

Generating function with a restricted pattern

Lemma. Let $S\colon J \mapsto \mathbb{N}$ be a given restricted pattern and $F(z;S)$ be the generating function of structures with the restricted pattern S. For the labeled case, we have

$$F(z;S) = \mathscr{S} z^{m'} \exp\left(C(z) - C(z;J)\right).$$

For the unlabeled case, we have

$$F(z;S) = \mathscr{S} z^{m'} \exp\left(\sum_{k\geq 1} \frac{C(z^k) - C(z^k;J)}{k}\right).$$

We use $|J|$ to denote the number of elements in the set J, and $\hat{j}$ to denote the maximum element in J.

Component generating functions of alg-log type

Let

$$\Delta(\nu, \theta) = \{z : |z| \leq \rho(1 + \nu), z \neq \rho, |Arg(z - \rho)| \geq \theta\}$$

for some constants $\rho > 0$, $\nu > 0$, and $0 < \theta < \pi/2$. We say that $C(z)$ is of alg-log type at singularity ρ, if $C(z)$ is analytic in $\Delta(\nu, \theta)$, and for some constants c and d,

$$C(z) = c + d(1 - z/\rho)^{\alpha} \left(\ln \frac{1}{1 - z/\rho}\right)^{\beta} (1 + o(1)),$$

as $z \to \rho$ in $\triangle(\nu, \theta)$.

We call α the algebraic exponent, and β the logarithmic exponent.

The special case $\alpha = 0$, $\beta = 1$ is the logarithmic type.

Alg-log type with $0 < \alpha < 1$

We first consider the case that the algebraic exponent α satisfies $0 < \alpha < 1$. In this case, as $z \to \rho$ in $\Delta(\rho, \phi)$ and for labeled structures, we have

$$F(z, \emptyset) = e^c + de^c(1 - z/\rho)^\alpha \left(\ln \frac{1}{1 - z/\rho}\right)^\beta (1 + o(1)).$$

Flajolet and Odlyzko's singularity analysis implies

$$[z^n]F(z; \emptyset) \sim e^c \frac{d}{\Gamma(-\alpha)} \rho^{-n} n^{-\alpha-1} (\ln n)^\beta,$$

and

$$\frac{C_n}{n!} \sim \frac{d}{\Gamma(-\alpha)} \rho^{-n} n^{-\alpha-1} (\ln n)^\beta.$$

Labeled structures

Theorem. Assume $0 < \alpha < 1$, $\hat{j} = O(m/\ln m)$, and

$$\sum_{j \in J} j^{-\alpha}(\ln j)^{\beta} = o(m^{1-\alpha}(\ln m)^{\beta-1}).$$

Then, as $m \to \infty$, and uniformly in all S

$$[z^n]F(z;S) \sim de^c \frac{\mathscr{S}}{\Gamma(-\alpha)} m^{-\alpha-1}(\ln m)^{\beta}\rho^{-m}\exp\left(-C(\rho;J)\right). \quad \square$$

A random structure is chosen uniformly at random from a given family of structures:

$$P(S,n) = \frac{[z^n]F(z;S)}{[z^n]F(z;\emptyset)}$$

is the probability that a random decomposable combinatorial structure of size n has a restricted pattern S.

Corollary

Corollary. For labeled structures with algebraic exponent $0 < \alpha < 1$ such that $\hat{j} = \max\{j; j \in J\} = O(m/\ln m)$, and

$$\sum_{j \in J} \frac{(\ln j)^\beta}{j^\alpha} = o\left(m^{1-\alpha}(\ln m)^{\beta-1}\right),$$

we have, as $m \to \infty$,

$$P(S,n) \sim \left(\frac{n}{m}\right)^{1+\alpha} \left(\frac{\ln m}{\ln n}\right)^\beta \prod_{j \in J} \frac{1}{S(j)!} \left(C_j \rho^j / j!\right)^{S(j)} e^{-C_j \rho^j / j!}.$$

If, in addition, the restricted pattern $S = \left[\prod_{j \in J} j^{S(j)}\right]$ also satisfies $\sum_{j \in J} j S(j) = o(n)$, that is, $m \sim n$, we have as $n \to \infty$,

$$P(S,n) \sim \prod_{j \in J} P\left(\left[j^{S(j)}\right], n\right) \sim \prod_{j \in J} \frac{1}{S(j)!} \left(C_j \rho^j / j!\right)^{S(j)} e^{-C_j \rho^j / j!}.$$

The rth smallest component size

Let $X_n^{[r]}$ and $X_n^{[r]}(S)$ denote the sizes of the rth smallest component in a random structure of size n without or with a restricted pattern S, respectively.

We define the set $N_k = \{1, 2, \dots, k\}$ where k is a positive integer and $d(k) = \sum_{j \in J \cap N_k} S(j)$. We only need to consider $d(k) < r$ since $P(X_n^{[r]}(S) > k) = 0$ when $d(k) \geq r$.

Theorem. Let $S\colon J \mapsto \mathbb{N}$ be a restricted pattern with $\hat{j} = O(m/\ln m)$ and $\sum_{j \in J} j^{-\alpha}(\ln j)^{\beta} = o\left(m^{1-\alpha}(\ln m)^{\beta-1}\right)$.
Assume $r - d(k) = O(\ln m)$ and $k = o(m(\ln m)^{\frac{1}{\alpha-1}})$.
As $m \to \infty$, we have

$$P(X_n^{[r]}(S) > k) \sim P(S, n) \exp(-C(\rho; N_k \backslash J)) \sum_{j=0}^{r-1-d(k)} \frac{C^j(\rho; N_k \setminus J)}{j!}. \quad \square$$

More results

Corollary. Let $r = O(\ln n)$ and $k = o\left(n(\ln n)^{\frac{1}{\alpha-1}}\right)$. As $n \to \infty$, if there is no restricted pattern, we have

$$P(X_n^{[r]} > k) \sim \exp(-C(\rho; N_k)) \sum_{j=0}^{r-1} \frac{C^j(\rho; N_k)}{j!}.$$

Theorem. Let $S\colon J \mapsto \mathbb{N}$ be a restricted pattern satisfying $\hat{j} = o\left(m(\ln m)^{\frac{1}{\alpha-1}}\right)$ and $\sum_{j\in J} \frac{(\ln j)^\beta}{j^\alpha} = o\left(m^{1-\alpha}(\ln m)^{\beta-1}\right)$. When $r = \sum_{j\in J} S(j) + 1$, we have in the labeled case

$$E(X_n^{[r]}(S)) \sim P(S, n) m \exp(-c + C(\rho; J)),$$

and $E(X_n^{[1]}) \sim ne^{-c}$.

We have equivalent results for unlabelled structures.

The case $\alpha = -p < 0$

In the following we consider the case

$$C(z) = d(1 - z/\rho)^{-p} + c + o(1)$$

as $z \to \rho$ in the open disk $|z| < \rho$, where c, d, and p are constants with $p, d > 0$, and $C(z)$ is analytic in the open disc $|z| < \rho$.

We define $h(z) = d(1 - z)^{-p} + b \ln \frac{1}{1-z}$. The asymptotics of $[z^n] \exp(h(z))$ was studied by Wright (1949) and Hayman (1956).

Condition: $\hat{j} = \max\{j : j \in J\}$, $\hat{j} = o\left(m^{\frac{1}{p+1}}\right)$, $\sum_{j \in J} j^{p-1} = o\left(m^{\frac{p}{6(p+1)}}\right)$, and $\sum_{j \in J} j^{p} = o\left(m^{\frac{1}{p+1}}\right)$.

Finally, let $R = R(m)$ be defined by

$$Rh'(R) = m, \ 0 < R < 1.$$

The case $\alpha = -p < 0$

We can extend the result of Wright and Hayman to generating functions including a restricted pattern S.

Theorem. Let $\rho < 1$; suppose the pattern S satisfies the previous conditions, and let R be defined as before. Then as $m \to \infty$ and for unlabeled structures, we have

$$[z^n]F(z;S) \sim \frac{1}{\sqrt{2\pi p(p+1)d}} \left(\prod_{j \in J} \binom{C_j + S(j) - 1}{S(j)} e^{-C_j \rho^j} \right) (R\rho)^{-m} \left(\frac{m}{pd}\right)^{-\frac{p+2}{2(p+1)}} e^{C(R\rho)+r_0},$$

where $r_0 = \sum_{k \geq 2} C(\rho^k)/k$. □

Corollary. Suppose that in addition to the previous conditions the restricted pattern $S = \left[\prod_{j\in J} j^{S(j)}\right]$ also satisfies

$$\sum_{j\in J} jS(j) = o\left(n^{1/(p+1)}\right).$$

Then, as $n \to \infty$,

$$\begin{aligned} P(S,n) &\sim \prod_{j\in J} P\left(\left[j^{S(j)}\right], n\right) \\ &\sim \prod_{j\in J} \binom{C_j + S(j) - 1}{S(j)} \rho^{jS(j)} \left(1-\rho^j\right)^{C_j}. \qquad \square \end{aligned}$$

We also have results for rth smallest sizes as well as for labeled structures.

Examples

Example 1. A rooted labeled tree consists of a set of components (subtrees). The component generating function $C(z)$ is of alg-log type at the singularity $1/e$ with algebraic exponent $\alpha = 1/2$, $\beta = 0$, $d = -\sqrt{2}$ and $c = 1$:

$$C(z) = 1 - \sqrt{2}(1 - ez)^{1/2} + O(1 - ez).$$

Consider the restricted pattern $S = [2^s t^3]$, where $s = \lfloor \sqrt{n} \rfloor$, and $t = \lfloor \ln n \rfloor$. We have

$$P(S, n) \sim \frac{1}{3!} \frac{\left(C_t e^{-t}/t!\right)^3}{\exp\left(C_t e^{-t}/t!\right)} \frac{1}{s!} \frac{\left(C_2 e^{-2}/2!\right)^s}{\exp\left(C_2 e^{-2}/2!\right)}.$$

Noting

$$C_2/2! = 1 \text{ and } C_t e^{-t}/t! \sim \frac{1}{\sqrt{2\pi}} t^{-3/2} \text{ as } t \to \infty,$$

we obtain

$$P(S,n) \sim \frac{1}{6}(2\pi)^{-3/2}(\ln n)^{-9/2}\frac{1}{s!}e^{-2s}\exp\left(-e^{-2}\right).$$

If in addition to the restricted pattern S we want information related to the 5th $(r = 5)$ smallest component size, we get

$$P(X_n^{[5]}(S) > k) \sim P(S,n)\exp(-C(\rho; N_k \backslash J)) \sum_{j=0}^{4-d(k)} \frac{(C(\rho; N_k \setminus J))^j}{j!},$$

for $k = o\left(n(\ln n)^{-2}\right)$. Since here $d(k) = 0$ when $k < 2$ and $d(k) \geq s$ when $k \geq 2$, we have $P(X_n^{[5]}(S) > k) = 0$ when $k \geq 2$, and

$$P\left(X_n^{[5]}(S) > 1\right) \sim P(S,n)\exp(-1/e)\sum_{j=0}^{4}\frac{e^{-j}}{j!}.$$

The expected size of the smallest subtree is asymptotic to n/e.

Example 2. Fragmented permutation are the combinatorial class $\mathscr{F} = SET(SEQ_{\geq 1}(\mathscr{Z}))$. *They correspond to unordered collections of permutations. The generating function of fragmented permutations is* $F(z) = \exp\left(\frac{z}{1-z}\right)$. *The coefficients form the sequence* $A000262$ *in Sloane's Encyclopedia of Integer Sequences that counts sets of lists.*

The component generating function is $C(z) = \dfrac{z}{1-z} = \dfrac{1}{1-z} - 1$ *and so it is of alg-log type with* $\alpha = -1$, $\rho = 1$, $d = 1$ *and* $c = -1$. *We have*

$$[z^n]F(z) = \frac{F_n}{n!} \sim \frac{e^{-1/2}e^{2\sqrt{n}}}{2\sqrt{\pi}n^{3/4}},$$

and the saddle point is

$$R = 1 - \frac{1}{\sqrt{n}} + \frac{1}{2n} + O(n^{-3/2}).$$

Now consider fragmented permutations with the restricted pattern $S = [2^s t^3]$, where $s = \lfloor n^{1/4} \rfloor$ and $t = \lfloor \ln n \rfloor$. The probability that a random fragmented permutation has this restricted pattern S is

$$P(S,n) \sim \frac{1}{3!} (C_t/t!)^3 \exp(-C_t/t!) \frac{1}{s!} (C_2/2!)^s \exp(-C_2/2!) .$$

We note that $C_j = j!$, and we obtain

$$P(S,n) \sim \frac{1}{6} \frac{1}{s!} e^{-2}.$$

We have $C(\rho, N_k) = k$, and the expected size of the smallest component is asymptotic to $\sum_{k \geq 0} e^{-k} = \frac{e}{e-1}$.

Summary

Let Y_n denote the number of components in a random structure of size n, and X_n be the smallest component size.

- When $C(z)$ is of logarithmic type, Y_n is asymptotically normal with expected value and variance both proportional to $\ln n$. $E(X_n)$ is also proportional to $\ln n$.
- When $C(z)$ is of alg-log type with algebraic exponent $0 < \alpha < 1$, we have $P(Y_n = k) \sim \dfrac{e^{-c}c^{k-1}}{(k-1)!}$. That is, $Y_n - 1$ is asymptotically Poisson with mean c. $E(X_n)$ is asymptotic to ne^{-c}.
- When $C(z)$ is of alg-log type with algebraic exponent $\alpha = -p < 0$, Y_n is also asymptotically normal with mean and variance both proportional to $n^{p/(p+1)}$. $E(X_n)$ is asymptotic to a constant.