The Smallest Component Size in Decomposable Structures

L. Dong, Z. Gao, D. Panario, L.B. Richmond

Carleton University, Ottawa, Canada

April 14, 2008

The Exponential Formula for Decomposable Structures

We consider labelled structures built upon components by the multi-set construction. Let f_k (c_k) be the number of labeled structures (components) of size k, and consider their exponential generating functions

$$F(z) = \sum_{k \geq 0} f_k \frac{z^k}{k!}, \text{ and } C(z) = \sum_{k \geq 1} c_k \frac{z^k}{k!}.$$

The Exponential Formula says

$$F(z) = \exp(C(z)).$$

The rth Smallest Component Size in a Random Structure

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 will derive results about the limiting distribution of $X_n(r)$ when the component generating function $C(z)$ is of *algebraic-logarithmic* type.
To do that, we will study 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-Richmond (2001),Arratia-Barbour-Tavaré (2003), and Dong-Gao-Panario (2007).

Decomposable Structures with a Restricted Pattern

Let J be a set of positive integers, $N = \{0, 1, 2, \cdots\}$, and S be a function from J to N. We say that a decomposable structure has a *restricted pattern* S when the number of components of size j in the structure is specified as $S(j)$ for each $j \in J$.
The following notation will be used throughout the talk.

- $|J|$ denotes the number of elements in J,
- $\hat{j} = \max\{j : j \in J\}$,
- $m = n - \sum_{j \in J} jS(j)$ denotes the degree of freedom of a structure of size n and with a restricted pattern S.

We also use

$$C(z; J) = \sum_{j \in J} c_j \frac{z^j}{j!}.$$

Generating Function for Structures with a Restricted Pattern

The following is a simple extension of the Exponential Formula.
Lemma 1: Let $S : J \mapsto N$ be a given restricted pattern. The exponential generating function of labeled structures with a restricted pattern S is

$$F(z;S) = \exp\left(C(z) - C(z;J)\right) \prod_{j \in J} \frac{c_j^{S(j)} z^{jS(j)}}{(j!)^{S(j)} S(j)!}.$$

The $\triangle(\nu, \theta)$ region

For constants ν and θ with $\nu > 0$, $0 < \theta < \pi/2$, define the Δ region

$$\triangle(\nu, \theta) = \{z : |z| < 1 + \nu, z \neq 1, |arg(z - 1)| > \theta\}$$

.

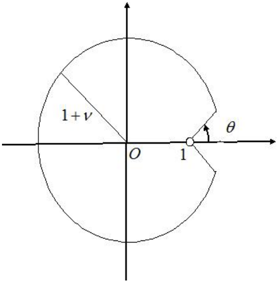

Figure: The Δ region

.

Component Generating Functions of *Alg-Log* Type

In this talk we assume that our component generating function $C(z)$ is of *algebraic-logarithmic type* at singularity $\rho > 0$, that is, $C(\rho z)$ is analytic in $\triangle(\nu, \theta)$ and

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

as $z \to 1$ in $\triangle(\nu, \theta)$.
Here α is called the algebraic exponent, and β the logarithmic exponent.
The special case $\alpha = 0, \ \beta = 1$ (called the *logarithmic* type) has been studied extensively before.

Alg-Log Type with $0 < \alpha < 1$

We first consider the case that the algebraic exponent α satisfies $0 < \alpha < 1$. In this case,

$$
\begin{aligned}
C(\rho z) &= c + d(1-z)^{\alpha}\left(\ln\frac{1}{1-z}\right)^{\beta}(1+o(1)), \\
F(\rho z) &= e^{c}\left(1 + d(1-z)^{\alpha}\left(\ln\frac{1}{1-z}\right)^{\beta}\right).
\end{aligned}
$$

Flajolet-Odlyzko's transfer theorem gives

$$
\begin{aligned}
[z^n]C(\rho z) &\sim \frac{d}{\Gamma(-\alpha)}(\log n)^{\beta} n^{-1-\alpha}, \\
[z^n]F(\rho z) &\sim \frac{d\exp(c)}{\Gamma(-\alpha)}(\log n)^{\beta} n^{-1-\alpha}.
\end{aligned}
$$

More Notation

- $[z^n]G(z)$ denotes the coefficient of z^n in the generating function $G(z)$.
- $X_n(r)$ denotes the size of the rth smallest component of a random decomposable combinatorial structure of size n.
- $N_k = \{1, 2, \dots, k\}$.
- $$K(S) = \prod_{j \in J} \frac{\left(c_j \rho^j / j!\right)^{S(j)}}{S(j)!} \exp\left(-c_j \rho^j / j!\right).$$

Theorem 1: Let S be a restricted pattern such that $|J| = o\left(m(\log m)^{\frac{-1}{1-\alpha}}\right)$ and $\hat{j} = O(m/\log m)$. Then, as $m \to \infty$, the probability that a random structure of size n has the pattern S is given by

$$\frac{[z^n]F(z;S)}{[z^n]F(z)} \sim K(S)\left(\frac{\log m}{\log n}\right)^{\beta}\left(\frac{n}{m}\right)^{\alpha+1},$$

where the asymptotics is uniform over all J.

The Probability of Having a Restricted Pattern

When the restricted pattern S is small such that $m \sim n$, we have

$$\frac{f_n(S)}{f_n} \sim \prod_{j \in J} \frac{\left(c_j \rho^j / j!\right)^{S(j)}}{S(j)!} \exp\left(-c_j \rho^j / j!\right).$$

This result can be restated as follows.

Corollary 1: Let $Z_n(j)$ be the number of components of size j in a random structure of size n. Suppose $|J| = o\left(n(\log n)^{\frac{-1}{1-\alpha}}\right)$ and $\hat{j} = O(n/\log n)$, then $(Z_n(j) : j \in J)$ are asymptotically independent Poisson random variables with mean $c_j \rho^j / j!$ for each $j \in J$.

Similar result for the logarithmic type was obtained by Arratia-Stark-Tavaré (1995) and Dong-Gao-Panario (2007).

Sketch of the proof of Theorem 1

From Lemma 1, we have

$$\begin{aligned} [z^n]F(z;S) &= \rho^{-n}[z^n]F(\rho z;S) \\ &= \rho^{-n}K(S)[z^m]\exp\left(C(\rho z) - C(\rho z;J) + C(\rho;J)\right). \end{aligned}$$

Using Cauchy's integral formula and the conditions on J, one can prove that

$$[z^m]\exp\left(C(\rho z) - C(\rho z;J) + C(\rho;J)\right) \sim [z^m]\exp\left(C(\rho z)\right).$$

Thus

$$[z^n]F(z;S) \sim K(S)\frac{d\exp(c)}{\Gamma(-\alpha)}(\log m)^{\beta}m^{-1-\alpha}\rho^{-n}.$$

Sketch of the proof of Theorem 1

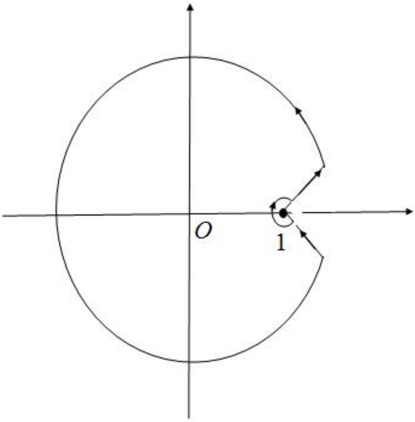

Figure: The contour

The rth Smallest Component Size

To keep track of the rth smallest component size, we consider a pattern $S\colon N_k \mapsto N$, where $S(j)$ is specified below. We note that each structure with its rth smallest component size greater than k corresponds to a structure with a pattern S such that $\sum_{i\in N_k} S(i) \le r-1$. Hence we have

$$P(X_n(r) > k) = \sum \left\{ \frac{f_n(S)}{f_n} : \sum_{i\in N_k} S(i) \le r-1 \right\}.$$

When $r = O(\log n)$ and $k = o\left(n(\log n)^{\frac{-1}{1-\alpha}}\right)$, S satisfies the conditions of Theorem 1, and $m \sim n$.

The rth Smallest Component Size

Corollary 2: Suppose $r = O(\log n)$ and $k = o\left(n(\log n)^{\frac{-1}{1-\alpha}}\right)$. Then, as $n \to \infty$,

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

Corollary 3: The expected size of the smallest component is asymptotic to ne^{-c}.

The Case $\alpha = -p < 0$

In the following we consider the case

$$C(z) = d(1 - z/\rho)^{-p} + b \ln \frac{1}{1 - z/\rho} + c + o(1) \quad \text{as } z \to \rho.$$

For convenience we define $h(z) = d(1 - z)^{-p} + b \ln \frac{1}{1-z}$.
The asymptotics of $[z^n] \exp(h(z))$ has been studied by Wright (1949) and Hayman (1956). In particular, it is known that, for $0 < p < 2$,

$$[z^n] \exp(h(z)) \sim \frac{1}{\sqrt{2\pi p(p+1)d}} \exp\left((1+p)d \left(\frac{n}{pd}\right)^{\frac{p}{p+1}} + \frac{pd}{2} \left(\frac{n}{pd}\right)^{\frac{p-1}{p+1}} + \frac{b-1}{p+1} \ln \frac{n}{pd} \right)$$

The Case $\alpha = -p < 0$

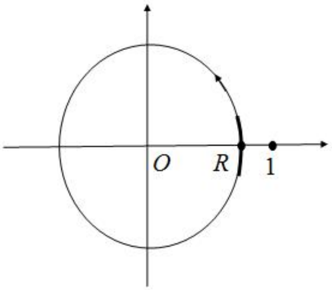

Figure: The contour through the saddle point R, where

$$R(1-R)^{-p-1} = \frac{m}{pd}$$

.

The Case $\alpha = -p < 0$

We can extend the result of Wright and Hayman to generating functions including a restricted pattern S, using the saddle point method.

Theorem 2: For each $0 < p < 2$, there is a positive constant $0 < \eta < 1/(p+1)$ such that if a pattern S satisfies $\hat{j} = O\left(m^{\eta}\right)$, then

$$[z^n]F(z;S) \sim \frac{K(S)}{\sqrt{2\pi p(p+1)d}} \exp\left((1+p)d\left(\frac{m}{pd}\right)^{\frac{p}{p+1}} + \frac{pd}{2}\left(\frac{m}{pd}\right)^{\frac{p-1}{p+1}} + \frac{b-1}{p+1}\ln\frac{m}{pd} + c\right)$$

The Case $\alpha = -p < 0$

Corollary 4: Let $Z_n(j)$ be the number of components of size j in a random structure of size n. Suppose $\hat{j} = O(n^\eta)$, then $(Z_n(j) : j \in J)$ are asymptotically independent Poisson random variables with mean $c_j \rho^j / j!$ for each $j \in J$.

Corollary 5: Suppose $r = O(\log n)$ and $k = O\left(n^\eta\right)$. Then, as $n \longrightarrow \infty$, we have

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

Corollary 6: The expected size of the smallest component is asymptotic to the constant

$$\sum_{k \geq 0} \exp(-C(\rho; N_k)).$$

Examples

- 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 $\alpha = 1/2$:

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

  The expected size of the smallest subtree is asymptotic to n/e.
- A *fragmented permutation* is a set of permutations. The component generating function for fragmented permutations is $C(z) = \dfrac{z}{1-z} = \dfrac{1}{1-z} - 1$. We have $C(\rho, N_k) = k$, and the expected size of the smallest component is asymptotic to $\sum_{k \geq 0} e^{-k} = \dfrac{e}{e-1}$.

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 $\alpha < 0$, 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.