Math. Proc. Camb. Phil. Soc. (2006), **140**, 537 © 2006 *Cambridge Philosophical Society*
doi:10.1017/S030500410600925X *Printed in the United Kingdom*

Logical limit laws for logarithmic structures

By DUDLEY STARK

*School of Mathematical Sciences, Queen Mary, University of London, Mile End, London, E*1 4*NS.*
e-mail: d.s.stark@maths.qmul.ac.uk

(*Received* 04 *February* 2005; *Revised* 3 *June* 2005)

Abstract

We prove monadic second order limit laws for classes of logarithmic adequate relational structures by using probabilistic methods.

1. *Introduction*

Given a class of finite relational structures $\mathcal{K}$, suppose that ν_n is the uniform measure over all structures of size n in $\mathcal{K}$. We are interested in the existence of $\lim_{n\to\infty} \nu_n(\Phi)$ for sentences Φ in monadic second order logic and we say that $\mathcal{K}$ has a monadic second order limit law if, for every such sentence Φ, the limit exists. In this paper we will use probabilistic methods to give sufficient conditions for certain relational structures to have monadic second order limit laws, extending results in [**4**].

We define c_n to be the number of structures in $\mathcal{K}$ of size $n \geqslant 1$ and define $c_0 = 1$. Compton [**6**, **7**] gave sufficient conditions for the class of relational structures $\mathcal{K}$ to have monadic second order limit laws involving only the asymptotic behavior of the c_n. Following Bell and Burris [**4**], we define the class RT_ρ to be those sequences of numbers c_n such that $c_n > 0$ for n large enough and such that $\lim_{n\to\infty} c_{n-1}/c_n = \rho$. A result of Compton we will use [**4**, theorem 7·2] is

Theorem 1 (Compton). *If* $c_n \in \mathrm{RT}_\rho$ *where* $0 < \rho < 1$, *and if there exist* K *and* M *such that*

$$\frac{c_{n-k}}{c_n} \leqslant M\rho^k \quad \text{for } K \leqslant k \leqslant n,$$

then $\mathcal{K}$ *has a monadic second order limit law.*

The next result appeared as [**4**, corollary 8·1].

Corollary 1 (Bell, Burris). *Suppose* $c_n \in \mathrm{RT}_\rho$ *and* $c_n/c_{n+1} \leqslant \rho$ *for* n *large enough. Then* $\mathcal{K}$ *has a monadic second order limit law.*

A class of finite relational structures $\mathcal{K}$ is said to be *adequate* if it is closed under disjoint union of components, and, up to isomorphism, it has only finitely many structures of each size. For an adequate class of relational structures we define a_j to be the number of component structures of size j. The relationship between c_n and a_j

https://doi.org/10.1017/S030500410600925X Published online by Cambridge University Press

for an adequate relational structure is given by

$$\sum_{n=0}^{\infty} c_n s^n = \prod_{j=1}^{\infty} (1 - s^j)^{-a_j}. \tag{1·1}$$

The conditions of Corollary 1 are shown in [**4**] to be satisfied for sequences $a_j \in \mathrm{RT}_\rho$ when $0 < \rho < 1$, as long as $\liminf_{j\to\infty} j a_j \rho^j > 1$. Given sequences a_j, b_j, we write $a_j \sim b_j$ to mean $\lim_{j\to\infty} a_j/b_j = 1$. [**4**, corollary 9·4] states that the conditions of Corollary 1 are satisfied when

$$a_j \sim C y^j j^{-\mu} \tag{1·2}$$

for constants $\mu < 1 < y$ and $C > 0$, or $\mu = 1 < y$ and $C > 1$. We will call such a class of relational structures with $\mu = 1$ *logarithmic* and call those with $\mu < 1$ *expansive*, in accordance with the terminology in [**5**] for related combinatorial structures called assemblies and multisets.

A variant of condition (1·2) for expansive structures is shown in [**8**] to imply monadic second order limit laws and related asymptotic enumeration results. In [**8**] the requirement on the a_j is relaxed to

$$a_j \asymp y^j j^{-\mu}, \tag{1·3}$$

where we write $a_j \asymp b_j$ for sequences a_j, b_j if the ratios a_j/b_j are bounded above and below by positive constants. (A weaker condition than (1·3) is also shown to suffice.)

The goal of this paper is to show that a similar weakening of (1·2) for logarithmic structures implies logical limit laws. Both [**8**] and this paper make use of the identity (2·6) below, though the analysis is necessarily different. In particular, the distribution of the random variable T_n appearing in (2·6) does not converge to a normal distribution under centering and rescaling for logarithmic structures as it does for expansive structures.

Our requirement for a class of relational structures to be logarithmic is that (1·3) holds with $\mu = 1$, so that

$$0 < \inf_j j a_j y^{-j} \leqslant \sup_j j a_j y^{-j} < \infty. \tag{1·4}$$

It is appropriate for us to require $y > 1$, since $y \leqslant 1$ implies that $a_j = 0$ for large enough j. For a relational structure satisfying condition (1·4), we define quantities

$$C_1 = \inf_j j a_j y^{-j} \quad \text{and} \quad C_2 = \sup_j j a_j y^{-j}. \tag{1·5}$$

THEOREM 2. *For any class of logarithmic adequate relational structures with $C_1 > 1$ and $y > 1$, the sequence c_n satisfies the conditions of Corollary 1 with $\rho = y^{-1}$. It follows that adequate logarithmic multisets with $C_1 > 1$ have second order monadic limit laws.*

Requiring $C_1 > 1$ agrees with the condition $C > 1$ used in [**4**].

In the course of verifying the condition $c_{n+1}/c_n \leqslant y^{-1}$ for large enough n needed to apply Corollary 1, we prove Theorem 3 which has a weaker hypothesis on the a_j than Theorem 2 and could be of independent interest.

THEOREM 3. *Suppose that there exists $C_1 \geqslant 1$ and $y > 1$ such that*

$$a_j \geqslant C_1 y^j / j, \quad j \geqslant 1.$$

Then $c_n/c_{n+1} \leqslant y^{-1}$ for all $n \geqslant 1$.

https://doi.org/10.1017/S030500410600925X Published online by Cambridge University Press

Combinatorial objects called multisets and assemblies related to the structures defined by (1·4) have been the study of numerous papers, including [**3**] on random logarithmic assemblies. We will use methods related to [**3**, section 5] to prove Theorem 2 and Theorem 3. Logarithmic multisets which have parameters satisfying conditions similar to (1·2) with $\mu = 1$ together with an analytic condition on a generating function of the form (1·1) were investigated in [**9**]. The component counting processes of logarithmic combinatorial objects are studied in great detail in the recent monograph [**1**].

2. *Preliminaries*

The identity (1·1) may be rewritten as

$$\sum_{n=0}^{\infty} c_n s^n = \exp\left(\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} \frac{a_k}{l} s^{kl}\right) = \exp\left(\sum_{j=1}^{\infty} a_j^{\star} s^j\right), \tag{2·1}$$

where

$$a_j^{\star} = \sum_{kl=j} \frac{a_k}{l}. \tag{2·2}$$

Lemma 1 is a result for a_j satisfying (1·4) similar to [**4**, lemma 5·2], of which required the a_j to satisfy (1·2).

Lemma 1. *Suppose* $y > 1$. *If the sequence* a_j *satisfies* $\inf_j j a_j y^{-j} \geqslant C_1 > 0$, *then,* $\inf_j j a_j^{\star} y^{-j} \geqslant C_1$. *If the sequence* a_j *satisfies* $\sup_j j a_j y^{-j} < \infty$, *then,* $\sup_j j a_j^{\star} y^{-j} < \infty$.

Proof. We have

$$a_j^{\star} = a_j + \sum_{\substack{kj=j \\ k<j}} \frac{a_k}{l} \geqslant a_j$$

giving $\inf_j j a_j^{\star} y^{-j} \geqslant C_1$. Moreover,

$$\begin{aligned} j a_j^{\star} &= \sum_{d|j} d\, a_d \\ &\leqslant j a_j + \sum_{d \leqslant j/2} d\, a_d \\ &\leqslant C_2 y^j + \sum_{d \leqslant j/2} C_2 y^d \\ &\leqslant C_2 y^j + C_2 y^{j/2}/(1 - y^{-1}) \\ &\leqslant C_2 \frac{2 - y^{-1}}{1 - y^{-1}} y^j, \end{aligned}$$

resulting in

$$\sup_j j a_j^{\star} y^{-j} < \infty.$$

https://doi.org/10.1017/S030500410600925X Published online by Cambridge University Press

The identity (2·1) arises in the enumeration of decomposable labelled combinatorial objects called assemblies described in [**2**]. Let m_j be the number of possible components on a set of size j of an assembly. Let N_n be the total number of representatives on n labelled objects. Then the relationship between the m_j and the N_n with $N_0 = 1$ given by [**2**, (121)] is

$$\sum_{n=0}^{\infty} \frac{N_n}{n!} = \exp\left(\sum_{j=1}^{\infty} \frac{m_j}{j!} z^k\right).$$

The equations (2·1) and (2·2) therefore describe an assembly with $m_j = j! a_j^{\star}$ and $N_n = n! c_n$ (allowing m_j to be nonintegral).

We use the connection between logarithmic structures and assemblies in the following way. Let $x > 0$ be any fixed number, let Z_j be independent Poisson (λ_j) distributed random variables, where

$$\lambda_j = a_j^{\star} x^j, \tag{2·3}$$

and define T_n to be

$$T_n = Z_1 + 2Z_2 + 3Z_3 + \cdots + nZ_n. \tag{2·4}$$

The identity (125) of [**2**] implies that the c_n satisfy

$$c_n = x^{-n} \exp\left(\sum_{j=1}^{n} \lambda_j\right) \mathbb{P}(T_n = n), \tag{2·5}$$

for any fixed $x > 0$. Although (2·5) is valid for any $x > 0$, in this paper we will always take $x = y^{-1} < 1$. With this choice of x the *count ratio*, defined to be c_n / c_{n+1}, equals

$$\frac{c_n}{c_{n+1}} = y^{-1} e^{-\lambda_{n+1}} \frac{\mathbb{P}(T_n = n)}{\mathbb{P}(T_{n+1} = n+1)}. \tag{2·6}$$

We now begin to estimate the ratio of probabilities in (2·6). Assume that $j\lambda_j \geqslant C_1$ for some constant $C_1 > 0$ for all $j \geqslant 1$. Let Z_j' be independent Poisson (C_1/j) distributed random variables and let Z_j'' be independent random variables which are also independent of the Z_j' and are Poisson $(\lambda_j - C_1/j)$ distributed when $\lambda_j > C_1/j$ and are identically 0 when $\lambda_j = C_1/j$. Define

$$T_n' = \sum_{j=1}^{n} jZ_j' \tag{2·7}$$

and

$$T_n'' = \sum_{j=1}^{n} jZ_j'', \tag{2·8}$$

so that $T_n \stackrel{\mathcal{D}}{=} T_n' + T_n''$.

Lemma 2 is central to our analysis of (2·6).

LEMMA 2. *With T_n, T_n' and T_n'' defined by* (2·4), (2·7) *and* (2·8), *respectively, for all* $n \geqslant 0$,

https://doi.org/10.1017/S030500410600925X Published online by Cambridge University Press

$$\frac{\mathbb{P}(T_{n+1}=n+1)}{\mathbb{P}(T_n=n)} = \mathbb{P}(Z_{n+1}=0) + \frac{\mathbb{P}(T_n=0)\mathbb{P}(Z_{n+1}=1)}{\mathbb{P}(T_n=n)} + \frac{\mathbb{P}(Z_{n+1}=0)\mathbb{P}(T_n''=n+1)\mathbb{P}(T_n'=0)}{\mathbb{P}(T_n=n)} + \frac{\mathbb{P}(Z_{n+1}=0)}{\mathbb{P}(T_n=n)}\sum_{k=0}^{n}\mathbb{P}(T_n''=k)[\mathbb{P}(T_n'=n+1-k)-\mathbb{P}(T_n'=n-k)].$$

Proof. We may express the ratio of probabilities on the left-hand side of the equality above as

$$\frac{\mathbb{P}(T_{n+1}=n+1)}{\mathbb{P}(T_n=n)} = \frac{\mathbb{P}(T_n=n+1)\mathbb{P}(Z_{n+1}=0)+\mathbb{P}(T_n=0)\mathbb{P}(Z_{n+1}=1)}{\mathbb{P}(T_n=n)} = \mathbb{P}(Z_{n+1}=0) + \frac{\mathbb{P}(T_n=0)\mathbb{P}(Z_{n+1}=1)}{\mathbb{P}(T_n=n)} + \frac{\mathbb{P}(Z_{n+1}=0)\,(\mathbb{P}(T_n=n+1)-\mathbb{P}(T_n=n))}{\mathbb{P}(T_n=n)}.$$

Moreover, we have

$$\begin{aligned}&\mathbb{P}(T_n=n+1)-\mathbb{P}(T_n=n)\\ &\quad=\sum_{k=0}^{n+1}\mathbb{P}(T_n''=k)\mathbb{P}(T_n'=n+1-k)-\sum_{k=0}^{n}\mathbb{P}(T_n''=k)\mathbb{P}(T_n'=n-k)\\ &\quad=\sum_{k=0}^{n}\mathbb{P}(T_n''=k)[\mathbb{P}(T_n'=n+1-k)-\mathbb{P}(T_n'=n-k)]+\mathbb{P}(T_n''=n+1)\mathbb{P}(T_n'=0).\end{aligned}$$

3. *Bounding the count ratio*

In this section we prove Theorem 3.

The assumption $ja_jy^{-j}\geqslant C_1\geqslant 1$ implies $j\lambda_j\geqslant C_1$ by Lemma 1, (2·3) and our choice $x=y^{-1}$. The point probabilities $\mathbb{P}(T_n=k)$ satisfy the recursion

$$\mathbb{P}(T_n'=k)=\frac{1}{k}\sum_{j=1}^{k\wedge n}C_1\mathbb{P}(T_n'=k-j),\quad k\geqslant 1$$

with

$$\mathbb{P}(T_n'=0)=\exp\left(-\sum_{j=1}^{n}\frac{C_1}{j}\right), \tag{3·1}$$

where $k\wedge n$ is the minimum of k and n (see [**2**, (153)]). It follows by induction that, for $k\in[0,n]$,

$$\mathbb{P}(T_n'\leqslant k)=\exp\left(-\sum_{l=1}^{n}\frac{C_1}{l}\right)\prod_{j=1}^{k}\left(1+\frac{C_1}{j}\right). \tag{3·2}$$

Moreover, we get, for $k\in[1,n-1]$,

$$\mathbb{P}(T_n'=k+1)-\mathbb{P}(T_n'=k)=\frac{C_1(C_1-1)}{k(k+1)}\mathbb{P}(T_n'\leqslant k-1)\geqslant 0. \tag{3·3}$$

https://doi.org/10.1017/S030500410600925X Published online by Cambridge University Press

The difference for $k=0$ is

$$\mathbb{P}(T'_n = 1) - \mathbb{P}(T'_n = 0) = (C_1 - 1)\mathbb{P}(T'_n = 0) \geqslant 0. \tag{3·4}$$

When $k=n$, we have

$$\begin{aligned}
&\mathbb{P}(T'_n = n+1) - \mathbb{P}(T'_n = n) \\
&\quad = \frac{C_1}{n+1}\sum_{i=1}^{n}\mathbb{P}(T'_n = n+1-i) - \frac{C_1}{n}\sum_{i=1}^{n}\mathbb{P}(T'_n = n-i) \\
&\quad = \frac{C_1(C_1-1)}{n(n+1)}\mathbb{P}(T'_n \leqslant n-1) - \frac{C_1}{n+1}\mathbb{P}(T'_n = 0) \\
&\quad \geqslant -\frac{C_1}{n+1}\mathbb{P}(T'_n = 0).
\end{aligned} \tag{3·5}$$

When the preceding inequalities are used in Lemma 2, one obtains

$$\begin{aligned}
\frac{\mathbb{P}(T_{n+1} = n+1)}{\mathbb{P}(T_n = n)} \geqslant \mathbb{P}(Z_{n+1} = 0) &+ \frac{\mathbb{P}(T_n = 0)\mathbb{P}(Z_{n+1} = 1)}{\mathbb{P}(T_n = n)} \\
&+ \frac{\mathbb{P}(Z_{n+1} = 0)\mathbb{P}(T''_n = n+1)\mathbb{P}(T'_n = 0)}{\mathbb{P}(T_n = n)} \\
&- \frac{C_1\mathbb{P}(Z_{n+1} = 0)\mathbb{P}(T''_n = 0)\mathbb{P}(T'_n = 0)}{(n+1)\mathbb{P}(T_n = n)}.
\end{aligned}$$

We note that

$$\begin{aligned}
\frac{C_1\mathbb{P}(Z_{n+1} = 0)\mathbb{P}(T''_n = 0)\mathbb{P}(T'_n = 0)}{(n+1)\mathbb{P}(T_n = n)} &= \frac{C_1\mathbb{P}(Z_{n+1} = 0)\mathbb{P}(T_n = 0)}{(n+1)\mathbb{P}(T_n = n)} \\
&\leqslant \frac{\lambda_{n+1}\mathbb{P}(Z_{n+1} = 0)\mathbb{P}(T_n = 0)}{\mathbb{P}(T_n = n)} \\
&= \frac{\mathbb{P}(Z_{n+1} = 1)\mathbb{P}(T_n = 0)}{\mathbb{P}(T_n = n)},
\end{aligned}$$

which allows us to conclude that

$$\frac{\mathbb{P}(T_{n+1} = n+1)}{\mathbb{P}(T_n = n)} \geqslant \mathbb{P}(Z_{n+1} = 0) = e^{-\lambda_{n+1}}$$

for all $n \geqslant 1$, which, when used in (2·6), implies Theorem 3.

4. *Convergence of the count ratio*

In this section we prove Theorem 2.

By Corollary 1, to prove Theorem 2 it suffices to show that

$$\lim_{n\to\infty}\frac{c_n}{c_{n+1}} = y^{-1} \tag{4·1}$$

and

$$\frac{c_n}{c_{n+1}} \leqslant y^{-1} \text{ for } n \geqslant 1. \tag{4·2}$$

We have shown (4·2) in Theorem 3; it remains to show (4·1).

As a result of Lemma 1, (2·3), our choice $x = y^{-1}$ and our definition of logarithmic structure, we have

$$\lambda_j \asymp \frac{1}{j}. \tag{4·3}$$

https://doi.org/10.1017/S030500410600925X Published online by Cambridge University Press

Lemma 3 results from (4·3) and [**3**, lemma 3].

Lemma 3. *For any class of adequate logarithmic structures, the T_n satisfy*

$$\mathbb{P}(T_n = n) \asymp \frac{1}{n}.$$

Now we find the asymptotics of the ratio of probabilities in (2·6). The first term on the right-hand side of the identity in Lemma 2 is of order

$$\mathbb{P}(Z_{n+1} = 0) = e^{-\lambda_{n+1}} = 1 + o(1) \tag{4·4}$$

because of (4·3). The second term on the right-hand side of the identity in Lemma 2 is of order

$$\begin{aligned}\frac{\mathbb{P}(T_n = 0)\mathbb{P}(Z_{n+1} = 1)}{\mathbb{P}(T_n = n)} &= \lambda_{n+1} \exp\left(-\sum_{j=1}^{n+1} \lambda_j\right) \frac{1}{\mathbb{P}(T_n = n)} \\ &\leqslant \exp\left(-\sum_{j=1}^{n+1} \frac{C_1}{j}\right) \frac{\lambda_{n+1}}{\mathbb{P}(T_n = n)} \\ &= O\left(n^{-C_1}\right) \\ &= o(1),\end{aligned} \tag{4·5}$$

where we have used (4·3) and Lemma 3. Similarly, using $C_1 > 1$, the third term on the right hand side of the identity in Lemma 2 is of order

$$\frac{\mathbb{P}(Z_{n+1} = 0)\mathbb{P}(T''_n = n+1)\mathbb{P}(T'_n = 0)}{\mathbb{P}(T_n = n)} \leqslant \frac{\mathbb{P}(T'_n = 0)}{\mathbb{P}(T_n = n)} = O(n^{1-C_1}) = o(1). \tag{4·6}$$

When $k \in [1, n]$, the identity (3·2) is bounded uniformly by

$$\mathbb{P}(T'_n \leqslant k) = O\left(\left(\frac{k}{n}\right)^{C_1}\right)$$

and (3·1) implies that

$$\mathbb{P}(T'_n = 0) = O(n^{-C_1}).$$

It therefore follows from (3·3), (3·4) and (3·5) that for $k \in [0, n]$,

$$\mathbb{P}(T'_n = k+1) - \mathbb{P}(T'_n = k) = O(n^{-C_1}(k \vee 1)^{C_1 - 2}) = \begin{cases} O(n^{-C_1}) & \text{if } C_1 \leqslant 2 \\ O(n^{-2}) & \text{if } C_1 > 2 \end{cases}$$

uniformly for all $k \in [1, n]$, where $k \vee 1$ is the maximum of k and 1. We have demonstrated that

$$\max_{0 \leqslant k \leqslant n} |\mathbb{P}(T'_n = k+1) - \mathbb{P}(T'_n = k)| = o(n^{-1}). \tag{4·7}$$

Finally, observe that

$$\begin{aligned}&\left|\sum_{k=0}^{n} \mathbb{P}(T''_n = k)[\mathbb{P}(T'_n = n+1-k) - \mathbb{P}(T'_n = n-k)]\right| \\ &\qquad \leqslant \max_{0 \leqslant k \leqslant n} |\mathbb{P}(T'_n = n+1-k) - \mathbb{P}(T'_n = n-k)| \\ &\qquad = o(n^{-1})\end{aligned} \tag{4·8}$$

https://doi.org/10.1017/S030500410600925X Published online by Cambridge University Press

because of (4·7). It now follows from Lemma 3 and (4·8) that

$$\frac{\mathbb{P}(Z_{n+1}=0)}{\mathbb{P}(T_n=n)}\sum_{k=0}^{n}\mathbb{P}(T_n''=k)\left[\mathbb{P}(T_n'=n+1-k)-\mathbb{P}(T_n'=n-k)\right]=o(1). \tag{4·9}$$

Lemma 2, (4·4), (4·5), (4·6) and (4·9) prove that

$$\lim_{n\to\infty}\frac{\mathbb{P}(T_n=n)}{\mathbb{P}(T_{n+1}=n+1)}=1$$

which limit, when used in (2·6) along with (4·3), proves (4·1) and Theorem 2.

Acknowledgements. We thank a referee for an argument shortening the proofs of Theorem 2 and Theorem 3.

REFERENCES

[1] R. A. ARRATIA, A.BARBOUR and S. TAVARÉ. *Logarithmic Combinatorial Structures: A Probabilistic Approach* (EMS monograph in Mathematics, European Mathematical Society Publishing House, 2003).

[2] R. A. ARRATIA and S. TAVARÉ. Independent process approximations for random combinatorial structures. *Adv. Math.* **104** (1994), 90–154.

[3] R. A. ARRATIA, D. STARK and S. TAVARÉ. Total variation asymptotics for Poisson process approximations of logarithmic combinatorial assemblies. *Ann. Probab.* **23** (1995), 1347–1388.

[4] J. P. BELL and S. N. BURRIS. Asymptotics for logical limit laws: when the growth of the components is in an RT class. *Trans. Amer. Math. Soc.* **355** (2003), 3777–3794.

[5] A. D. BARBOUR and B. L. GRANOVSKY. Random combinatorial structures: the convergent case. *J. Combin. Theory Ser. A* **109** (2005), 203–220.

[6] K. J. COMPTON. A logical approach to asymptotic combinatorics I. First order properties. *Adv. Math.* **65** (1987), 65–96.

[7] K. J. COMPTON. A logical approach to asymptotic combinatorics II. Monadic second-order properties. *J. Combin. Theory Ser. A* **50** (1989), 110–131.

[8] B. GRANOVSKY and D. STARK. Asymptotic enumeration and logical limit laws for expansive multisets. To appear in *J. London Math. Soc.*

[9] D. STARK. Total variation asymptotics for independent process approximations of logarithmic multisets and selections. *Random Structures Algorithms* **11** (1997), 51–80.

https://doi.org/10.1017/S030500410600925X Published online by Cambridge University Press