The Properties of Random Trees

Technical Report 86-024

1986

John H. Halton

The University of North Carolina at Chapel Hill
Department of Computer Science
New West Hall 035 A
Chapel Hill, N.C. 27514

THE PROPERTIES
OF
RANDOM TREES

JOHN H. HALTON

The University of North Carolina
Department of Computer Science
New West Hall 035 A
Chapel Hill, NC 27514 USA

23 NOVEMBER 1986

DEDICATION

Patrick D. Halton celebrated his hundredth year on the 11th of August 1978. His patience, understanding, support, encouragement, and love have been a constant inspiration to me; and his unquestioning faith and trust in me have allowed me to persevere and survive through times of doubt and discouragement. He died peacefully on the 29th of May 1979. He had long been fascinated by trees of all kinds, and it is therefore particularly fitting that this paper is, hereby, humbly dedicated to him, with my most sincere and heartfelt love and admiration.

ABSTRACT

Consider an *s-ary tree* (in which every node has no more than s children). Each node holds a single datum, including a *key*. These are the *occupied*, *internal*, or *closed* nodes of the data-structure. Augment the tree, following D. E. Knuth, by adding a set of *unoccupied*, *external*, *open*, or *free* nodes, so that every internal node now has just s children and every external node has no children. We assume that there is an unambiguous *rule*, depending only on the key-values at the internal nodes of the tree, whereby a new datum, with a new key value, will be inserted at one of the external nodes; this node then becomes internal and acquires s new external nodes as children. We further assume that the rule and the *statistical distribution* of data are such that every external node has equal probability of being selected for insertion of a new datum, at every stage. Various statistics of such trees are now obtained explicitly, in a systematic manner which may be extended to higher moments. The principal result is that the *average level* of both internal and external nodes in a given tree is asymptotic *in probability* to $\frac{s}{s-1}\log m$ as $m \to \infty$, where m is the number of internal nodes in the tree. Since the corresponding average level for a k-level fully balanced tree (with $m = \frac{s^k - 1}{s - 1}$) is asymptotic to $k \sim \log_s m = \frac{1}{\log s}\log m$ as $m \to \infty$, we conclude that, unless the distribution of data is far from the rather plausible assumption made here, it is highly improbable that the considerable cost of rebalancing trees when constructing data-bases will ever be justified in practice.

THE PROPERTIES OF RANDOM TREES

by

John H. Halton

1. INTRODUCTION

The underlying problem which we consider is the construction of an efficient data-storage structure of arbitrary size, when the data are identified by a *key*, which may be thought of as one or several real numbers. A sequence of such data is received and successively inserted in an initially-null structure, according to a *rule* depending only on the (possibly multi-dimensional) *order* of the keys, not on their magnitudes or other parts of the data. When we consider the *statistics* of such data, it is reasonable to assume that every possible order of the incoming data is equally probable. We seek to devise a structure such that the work of *insertion*, *deletion*, and *retrieval* of data is a slowly-growing function of the number m of data to be handled.

A favorite structure, balancing speed of insertion, deletion, and retrieval is a *tree*. When the key consists of a single real number, so that all key-values are linearly ordered, we may choose a *binary tree*, in which every node has 0, 1, or 2 *children*, and every node but one (the *root*) has just one *parent* (the root has none). If we call the nodes of such a tree *internal* (or *occupied* or *closed*) *nodes*, we may augment the tree with additional *external* (or *unoccupied* or *open*) *nodes*, in such a way that all internal nodes have just two children and all external nodes have none (see Knuth [68]). Each internal node contains the key of just one datum; the key belonging to the first datum received in sequence being placed at the root of the tree; and, thereafter, we proceed recursively, comparing each new key with keys stored at successive nodes encountered in a traversal of the tree, beginning at the root, moving to the right child if the new key exceeds the key found at the current node, and to the left child if not; when an external node is reached, the new key is placed there. At every stage, every right child has a key greater than that at its parent, every left child has a smaller key than its parent. Figure 1 below shows the augmented binary trees with m internal and n external nodes, for m = 0, 1, 2, 3, and 4. All topologically distinct trees are shown. Internal nodes are shown as filled (black) circles and external nodes as open (white) circles. Note that, in every case, $n = m + 1$. This is generally true (as is proved by Knuth [68]) for binary trees, and indeed is a special case of a general result for s-ary trees.

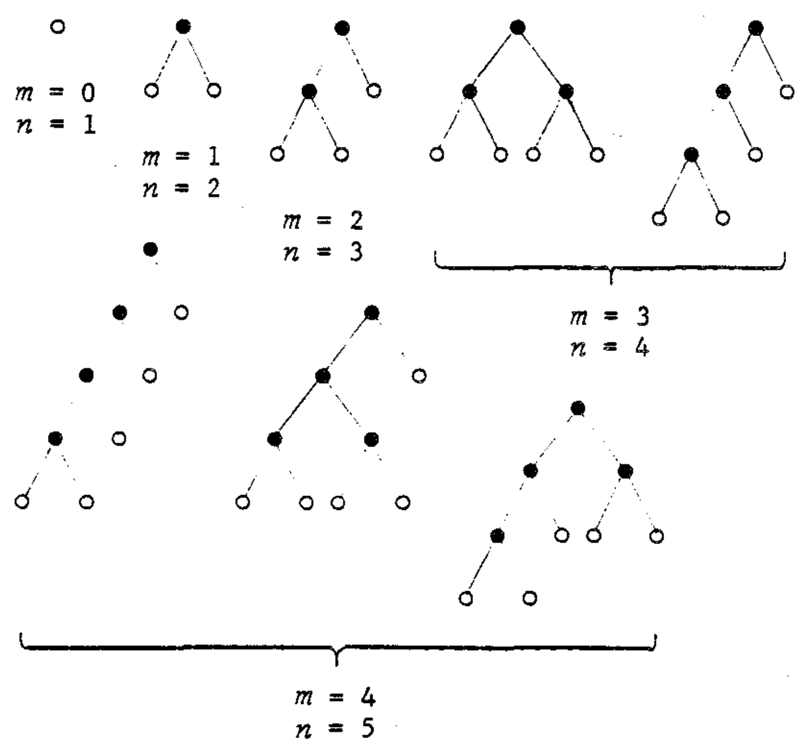


Figure 1.

Every node has a *level*, defined as the number of steps (edges) in a direct path from the root to the node; the root thus has level 0. The *height* of the tree is the maximum level over all internal nodes. Knuth [68] calls the sum of the levels of all external nodes of a tree the *external path length* of the tree (we denote this by $E_m^{(1)}$, when there are m internal nodes in the tree) and the sum of the levels of all internal nodes the *internal path length* (we denote this by $F_m^{(1)}$). Given a tree, with m internal nodes, the work required to insert a new datum at level h is essentially proportional to the number of comparisons required to find an external node at which to place it, and a little reflection shows that this is just h. If, as we shall argue later, all external nodes are equally likely candidates for insertion of a random datum, it follows that the expected (average) amount of work required to insert a datum is proportional to $E_m^{(1)}/n$, where n is the number of external nodes in the tree. Similarly, the work required to build the entire tree is proportional to $F_m^{(1)}$. The work required to search for a datum without success is essentially the same as the work required to insert the datum sought and not found: the average amount of work required by an unsuccessful search is thus proportional to $E_m^{(1)}/n$. The work required to find a given datum is proportional to one more than the level at which it is found; so that the average work required to find a datum is proportional to $1 + F_m^{(1)}/m$.

When the key consists of more than one real number, the ordering becomes multi-dimensional, and a binary tree does not suffice for efficient storage and retrieval. This motivates the concept, familiar from graph theory, of an *s-ary tree*, in which every node has 0, 1, 2, ..., s children.

As before, we may augment the m internal nodes of such a tree with n external nodes, so that every internal node has just s children and every external node has none. Again, each internal node holds the key of a single datum. The insertion rule will not be specified, except as stated earlier. Level, height, internal and external path lengths are all defined as for binary trees, and the reasoning leading to the formulae for average work required for various operations holds without any change. It is clear that the quantities

$$X_m^{(1)} = E_m^{(1)}/n \tag{1}$$

and

$$Y_m^{(1)} = F_m^{(1)}/m \tag{2}$$

are central to these considerations. Figure 2 below is the counterpart of Figure 1, for general s.

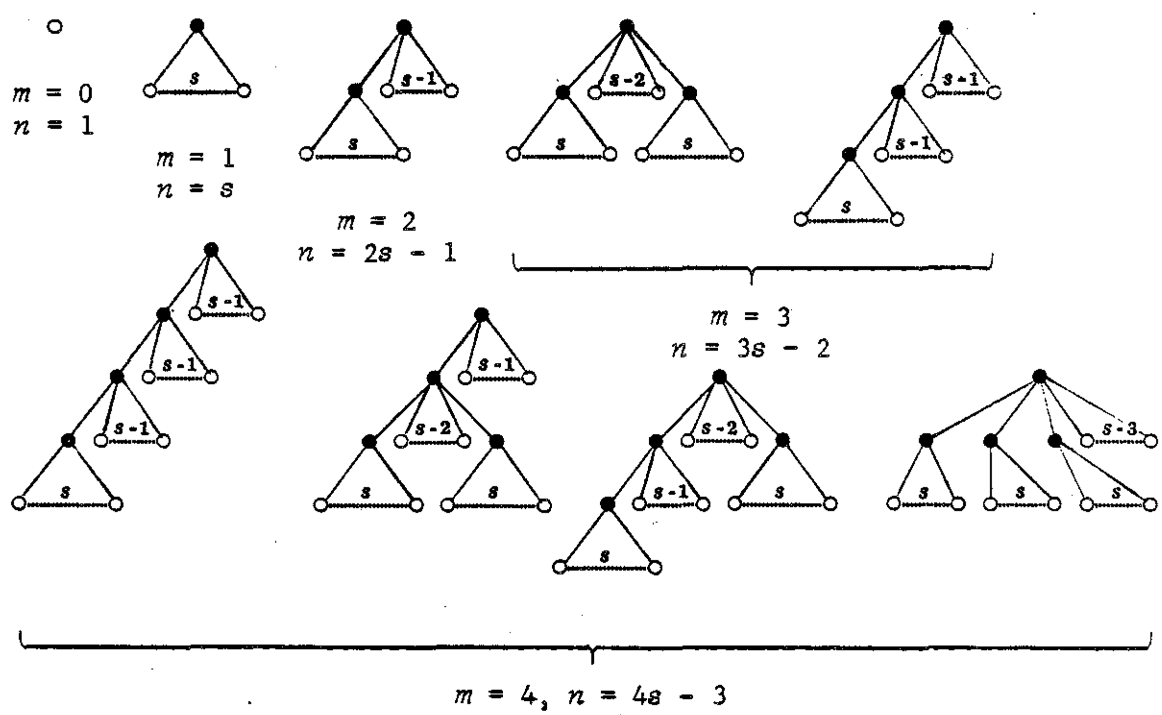


Figure 2.

We infer from this that

$$n = (s - 1)m + 1. \tag{3}$$

Indeed, on the one hand, since the augmented s-ary tree with m internal and n external nodes is a tree, it is well known (see, e.g., Knuth [68], §2.3.4.1, or Aho, Hopcroft, and Ullman [83], §7.1) that it has $m + n - 1$ edges; on the other hand, every edge points from an internal node to one of its children (external nodes have no children; internal nodes have exactly s children each), so there are just sm of them: (3) follows.

Various applications of s-ary trees have been suggested (e.g., see Muntz and Uzgalis [70], Finkel and Bentley [74], and Bentley [75, 79]). In all cases, the postulate that all (multi-dimensional) orderings of the data are equally probable is quite plausible. A full discussion of this matter is postponed.

We shall further generalize the quantities $E_m^{(1)}$, $F_m^{(1)}$; $X_m^{(1)}$, and $Y_m^{(1)}$ defined above to the sum of the p-th powers of the levels of all external nodes of a tree, which we shall call the *p-th external sum* of the tree and denote by $E_m^{(p)}$, the sum of the p-th powers of the levels of all internal nodes, which we shall call the *p-th internal sum* of the tree and denote by $F_m^{(p)}$, and the corresponding averages,

$$X_m^{(p)} = E_m^{(p)}/n \tag{4}$$

and

$$Y_m^{(p)} = F_m^{(p)}/m. \tag{5}$$

Averaging over all the nodes of a tree give one kind of expected behavior; but it is more interesting to ask how trees in general behave; so that we need to average again over all trees generated by random data.

General techniques will be developed below, which may be used to obtain the mathematical expectations, and higher moments, of the four statistics appearing in (4) and (5). These will be computed on the assumption that *insertion of a new datum is equally probable at every external node.*

In particular, we will explicitly obtain the following results:

$$E[X_m^{(1)}] = T_m^{(1)}, \tag{6}$$

$$E[Y_m^{(1)}] = \frac{m + \theta}{m} T_m^{(1)} - 1 - \theta, \tag{7}$$

$$E[X_m^{(2)}] = [T_m^{(1)}]^2 + T_m^{(1)} - T_m^{(2)}, \tag{8}$$

$$E[Y_m^{(2)}] = \frac{m + \theta}{m} \left\{[T_m^{(1)}]^2 - (1 + 2\theta)T_m^{(1)} - T_m^{(2)}\right\} + (1 + \theta)(1 + 2\theta), \tag{9}$$

$$\mathrm{var}[X_m^{(1)}] = (2 + \theta)\frac{m}{m + \theta} - T_m^{(2)} - \frac{1}{m + \theta} T_m^{(1)}, \tag{10}$$

and

$$\mathrm{var}[Y_m^{(1)}] = \left(\frac{m + \theta}{m}\right)^2 \mathrm{var}[X_m^{(1)}]; \tag{11}$$

where

$$\theta = \frac{1}{s - 1} \tag{12}$$

and

$$T_m^{(q)} = (1 + \theta)^q \left\{\frac{1}{(1 + \theta)^q} + \frac{1}{(2 + \theta)^q} + \frac{1}{(3 + \theta)^q} + \dots + \frac{1}{(m + \theta)^q}\right\}. \tag{13}$$

Some special cases of these results do occur in the literature, mainly for binary trees. When $s = 2$, $\theta = 1$, and $T_m^{(q)} = 2^q[2^{-q} + 3^{-q} + 4^{-q} + \dots + (m + 1)^{-q}]$; Booth and Colin [60], Windley [60], and Hibbard [62] have all independently obtained the equivalents of (6) and (7), and

Lynch [65] and Knuth [73] have corresponding equivalents of (10) and (11). Wilson [76] has results similar to (6) and (7), and also has the variance, for $s = 3$.

We proceed to derive asymptotic results for $m \to \infty$.

$$T_m^{(1)} \sim (1 + \theta) \log m + u_1(\theta) \tag{14}$$

and $$T_m^{(2)} \sim u_2(\theta); \tag{15}$$

where $$(1 + \theta)(\gamma - 1) \leqslant u_1(\theta) \leqslant (1 + \theta)\gamma, \; \gamma = 0.5772156649\ldots$$

and $$(1 + \theta)^2(\frac{\pi^2}{6} - 1) \leqslant u_2(\theta) \leqslant (1 + \theta)^2 \frac{\pi^2}{6}. \tag{16}$$

Whence $$\mathrm{E}[X_m^{(1)}] = (1 + \theta) \log m + O(1), \tag{17}$$

$$\mathrm{E}[Y_m^{(1)}] = (1 + \theta) \log m + O(1), \tag{18}$$

$$\mathrm{E}[X_m^{(1)}] - \mathrm{E}[Y_m^{(1)}] = 1 + \theta + O(\frac{\log m}{m}), \tag{19}$$

$$\mathrm{E}[X_m^{(2)}] = (1 + \theta)^2(\log m)^2 + O(\log m), \tag{20}$$

$$\mathrm{E}[Y_m^{(2)}] = (1 + \theta)^2(\log m)^2 + O(\log m), \tag{21}$$

$$\mathrm{var}[X_m^{(1)}] = 2 + \theta - u_2(\theta) + O(\frac{\log m}{m}), \tag{22}$$

$$\mathrm{var}[Y_m^{(1)}] = 2 + \theta - u_2(\theta) + O(\frac{\log m}{m}), \tag{23}$$

$$\mathrm{E}[X_m^{(2)}] - (\mathrm{E}[X_m^{(1)}])^2 = (1 + \theta) \log m + O(1), \tag{24}$$

and $$\mathrm{E}[Y_m^{(2)}] - (\mathrm{E}[Y_m^{(1)}])^2 = (1 + \theta) \log m + O(1). \tag{25}$$

The last two expressions may be viewed as the in-tree variance of the node-levels, in an average tree.

Previous authors do not seem to have examined the asymptotics of the results they have obtained. As a result, they have failed to make the following observations, which would appear to be crucial to important strategic decisions in setting up a data-base structure and its algorithms.

Chebyshev's inequality (see, e.g., Feller [68] or Tucker [67]) states that, if a random variable Q has finite expectation $E[Q]$ and variance $\mathrm{var}[Q]$, then, for any $\varepsilon > 0$,

$$\mathrm{Prob}[|Q - E[Q]| \geqslant \varepsilon\, E[Q]] \leqslant \mathrm{var}[Q]/\varepsilon^2(E[Q])^2. \qquad (26)$$

Taking $Q = X_m^{(1)}$, we derive that, by (17) and (22),

$$\mathrm{Prob}[|X_m^{(1)} - E[X_m^{(1)}]| \geqslant \varepsilon\, E[X_m^{(1)}]] \leqslant \mathrm{var}[X_m^{(1)}]/\varepsilon^2(E[X_m^{(1)}])^2$$
$$\sim \kappa(\theta)/\varepsilon^2(\log m)^2 \to 0 \quad \text{as} \quad m \to \infty, \qquad (27)$$

where
$$\kappa(\theta) = [2 + \theta - u_2(\theta)]/(1 + \theta)^2 = O(1). \qquad (28)$$

Taking $Q = Y_m^{(1)}$, we derive, in exactly the same way, by (18) and (23), that

$$\mathrm{Prob}[|Y_m^{(1)} - E[Y_m^{(1)}]| \geqslant \varepsilon\, E[Y_m^{(1)}]] \to 0 \quad \text{as} \quad m \to \infty. \qquad (29)$$

Similarly, for the in-tree distribution of levels (in an average tree), we see that, if x and y denote the levels of random external and internal nodes, respectively, then, by (17) and (24),

$$\mathrm{Prob}[|x - E[X_m^{(1)}]| \geqslant \varepsilon\, E[X_m^{(1)}]] \leqslant \mathrm{var}[x]/\varepsilon^2(E[X_m^{(1)}])^2$$
$$\sim 1/(1 + \theta)\varepsilon^2 \log m \to 0 \quad \text{as} \quad m \to \infty, \qquad (30)$$

and, similarly, by (18) and (25),

$$\mathrm{Prob}[|y - E[Y_m^{(1)}]| \geqslant \varepsilon\, E[Y_m^{(1)}]] \to 0 \quad \text{as} \quad m \to \infty. \qquad (31)$$

These results mean that the random variables $X_m^{(1)}/E[X_m^{(1)}]$, $Y_m^{(1)}/E[Y_m^{(1)}]$, $x/E[X_m^{(1)}]$, and $y/E[Y_m^{(1)}]$ tend to 1 *in probability* as $m \to \infty$. The *Central Limit Theorem* (see ibid., or Halton [85]) does not directly apply, but we may expect that at least the distributions of the level-averages $X_m^{(1)}$ and

$Y_m^{(1)}$ approximate the normal for large m. The critical points of the normal distribution are 3.090232 for probability 10^{-3} and 4.753424 for probability 10^{-6}, for example. Roughly doubling these for safety, we may infer that

$$\text{Prob}[X_m^{(1)} > E[X_m^{(1)}] + 6.18(\text{var}[X_m^{(1)}])^{\frac{1}{2}}] < 10^{-3} \tag{32}$$

and $$\text{Prob}[X_m^{(1)} > E[X_m^{(1)}] + 9.51(\text{var}[X_m^{(1)}])^{\frac{1}{2}}] < 10^{-6}, \tag{33}$$

with similar results for the $Y_m^{(1)}$.

For comparison, consider an ideally balanced tree, with s^j internal nodes at level j, for $j = 0, 1, 2, \ldots, h$. Then

$$m = 1 + s + s^2 + \ldots + s^h = \frac{s^{h+1} - 1}{s - 1}, \quad n = s^{h+1}; \tag{34}$$

so that $$h = \log_s n - 1 = \frac{\log[m(s - 1) + 1]}{\log s} - 1. \tag{35}$$

Since all external nodes of such a tree are at the same level,

$$X_m^{(1)} = h + 1. \tag{36}$$

Using (6), (10), (144), and (145), we may now calculate some values of m, ideal-tree $X_m^{(1)}$, and the bound (33):

s	2	4	10	100
m	127	85	111	101
	7 < 14.56	4 < 11.54	3 < 11.11	2 < 10.56
m	8191	5461	11111	10101
	13 < 23.33	7 < 17.54	5 < 16.56	3 < 15.55
m	1048575	1398101	1111111	1010101
	20 < 33.05	11 < 24.95	7 < 21.68	4 < 20.21
m	67108863	89478485	111111111	101010101
	26 < 41.36	14 < 30.49	9 < 26.80	5 < 24.86

(37)

One final statistic is available to us for comparison, in the case of $s = 2$. Adel'son-Vel'skii and Landis [62] (see also Knuth [73], §6.2.3) have devised the concept of a *balanced tree* as one which, at every node, has the heights of the left and right sub-trees differing by no more than one; and they have a very elegant algorithm for *rebalancing* such a tree with every insertion, at a cost of insertion times about five times as long as for simple insertion (see empirical discussion in Knuth). Knuth points out that the Fibonacci tree is the least ideal kind of balanced tree; here, the tree of index k has $n = m + 1 = F_k$ $(k = 2, 3, 4, \ldots)$, where F_k is the *Fibonacci number* of index k, satisfying, for all integers k,

$$\left.\begin{array}{c} F_0 = 0,\ F_1 = F_2 = 1,\ F_3 = 2,\ F_4 = 3,\ F_5 = 5,\ F_6 = 8,\ \ldots; \\ F_k = F_{k-1} + F_{k-2}. \end{array}\right\} \quad (38)$$

A little thought shows that the external path length of a Fibonacci tree, $\mathcal{E}_k = E^{(1)}_{F_k - 1}$ satisfies the recurrence relation (with $\mathcal{E}_2 = 0$, $\mathcal{E}_3 = 2$, $\mathcal{E}_4 = 5$)

$$\mathcal{E}_k = \mathcal{E}_{k-1} + \mathcal{E}_{k-2} + F_k, \quad (39)$$

It is easily verified that this has the solution

$$\mathcal{E}_k = \tfrac{3}{5}(k - 1)F_k + \tfrac{1}{5}(k - 5)F_{k-1}. \quad (40)$$

It is well known (and easily checked) that (*Binet's* formula)

$$F_k = \frac{1}{\sqrt{5}}\,(\alpha^k - \beta^k), \qquad \alpha = \frac{1 + \sqrt{5}}{2},\ \beta = \frac{1 - \sqrt{5}}{2}; \quad (41)$$

whence we see that

$$\begin{aligned} \mathcal{E}_k &= \frac{1}{5\sqrt{5}}\left\{3(k - 1)(\alpha^k - \beta^k) + (k - 5)(\alpha^{k-1} - \beta^{k-1})\right\} \\ &= \frac{1}{5\sqrt{5}}(3 + \frac{1}{\alpha})\ k\ \alpha^k + O(\alpha^k) \sim 0.3236\ k\ \alpha^k \text{ as } k \to \infty, \end{aligned} \quad (42)$$

since $\alpha = 1.618$, $\beta = -\,0.618$. This implies that, for Fibonacci trees,

$$X^{(1)}_m \sim 0.7236\ k \sim 1.5037 \log m, \quad (43)$$

since $m = F_k - 1 \sim \alpha^k/\sqrt{5}$. By contrast, (36) gives 1.4427 (i.e., 1/log 2) as the factor of $\log m$.

To summarize, we may conclude that:

(1) on the assumption that *at every stage, a new datum has equal probability of being inserted at any of the external nodes of a tree,* it is possible (with adequate automated formula-manipulation assistance) to calculate all the moments of the distribution of average work for insertion, search, and full-tree construction, by the techniques developed here;

(2) asymptotic results indicate that, for large enough trees (in the sense of numerous enough nodes), the probability that the in-tree average work for search or insertion exceeds the mathematical expectation, $\frac{s}{s-1}\log m$, by any appreciable percentage, is negligible;

(3) it follows that any *rebalancing scheme* is of doubtful utility, in view of the additional work entailed, when the tree becomes large enough, even when outlying cases are to be avoided.

The thrust of the argument presented here is that absolute worst-case situations become of such extremely small probability that extra work to avoid them is not economically justifiable, for trees having, say, a hundred or more nodes. Some authors have, nevertheless, studied the *heights* of random binary trees (as measures of the worst-case statistics), under various assumptions of randomness (see Stepanov [69], Kemp [79], Renyi and Szekeres [67], Yao [80], Robson [79, 82, 83], Flajolet and Odlyzko [82], de Bruijn, Knuth, and Rice [72], Mahmoud and Pittel [84], Pittel [84], Devroye [84, 86], for example).

2. THE STRUCTURE OF AN s-ARY TREE

We consider an s-ary tree, augmented with external nodes, so that there are m internal nodes (each with s children) and $n = (s - 1)m + 1$ external nodes (by (3)), each without children. At each level k ($k = 0, 1, 2, \ldots$), let the number of internal and external nodes be μ_{mk} and ν_{mk}, respectively. We observe that, when $m = 0$, $\mu_{00} = 0$ and $\nu_{00} = 1$, while, if $m > 0$,

$$\mu_{m0} = 1 \quad \text{and} \quad \nu_{m0} = 0; \tag{44}$$

since the root is the only node at level 0 and is the first to be occupied. Also, since a tree of m nodes cannot reach level m,

$$\text{if} \quad k \geqslant m, \quad \mu_{mk} = \nu_{m(k+1)} = 0. \tag{45}$$

Of course,

$$\sum_{k=0}^{\infty} \mu_{mk} = m \tag{46}$$

and

$$\sum_{k=0}^{\infty} \nu_{mk} = (s - 1)m + 1. \tag{47}$$

Since every internal node has just s (internal and external) children, we see that, for $k \geqslant 1$,

$$s\mu_{m(k-1)} = \mu_{mk} + \nu_{mk}. \tag{48}$$

Following Knuth [68], we define the external and internal path lengths and generalize them to the p-th external and internal sums of the tree, for $p \geqslant 0$,

$$E_m^{(p)} = \sum_{k=0}^{\infty} \nu_{mk} k^p \quad \text{and} \quad F_m^{(p)} = \sum_{k=0}^{\infty} \mu_{mk} k^p, \tag{49}$$

so that

$$E_m^{(0)} = (s - 1)m + 1, \quad F_m^{(0)} = m, \tag{50}$$

by (46) and (47), and $E_0^{(0)} = 1$, $F_0^{(0)} = E_0^{(p)} = F_0^{(p)} = 0$ $(p > 0)$. (51)

By (46), (48), and (49), we see that, when $m > 0$ and $p > 0$,

$$E_m^{(p)} = \sum_{k=0}^{\infty} \nu_{mk} k^p = \sum_{k=1}^{\infty} \nu_{mk} k^p = \sum_{k=1}^{\infty} [s\mu_{m(k-1)} - \mu_{mk}]$$

$$= s\sum_{j=0}^{\infty} \mu_{mj}(j+1)^p - F_m^{(p)} = (s-1)F_m^{(p)} + s\sum_{q=0}^{p-1} \binom{p}{q} F_m^{(q)}. \quad (52)$$

The corresponding averages are defined in (4) and (5) and are the focus of our investigation.

We may note that, when $s = 2$ and $p = 1$, (52) with (50) reduces to

$$E_m^{(1)} = F_m^{(1)} + 2m, \quad (53)$$

which is obtained directly by Knuth [68], §2.3.4.5, by induction; and he proceeds in [73], §6.2.1, to derive that

$$Y_m^{(1)} = (1 + \frac{1}{m})\, X_m^{(1)} - 2, \quad (54)$$

in our notation (Knuth writes C'_m for our $X_m^{(1)}$ and C_m for our $Y_m^{(1)} + 1$). He attributes (54) to Hibbard [62]. He also gives (3) and (52) for general s but only $p = 1$, as exercises ([68], §2.3.4.5).

3. RANDOM STORAGE OF DATA IN TREES

For single-keyed data, let the input sequence of keys be $[\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_m]$ and let ρ denote the (unique) permutation of $[1, 2, 3, \ldots, m]$ for which

$$\alpha_{\rho(1)} < \alpha_{\rho(2)} < \alpha_{\rho(3)} < \ldots < \alpha_{\rho(m)}. \quad (55)$$

Assumption 1: *The random input of single-keyed data is so structured that every ordering permutation* ρ *satisfying* (55) *is equally likely.*

Lemma 1: *Given the ordering permutation* ρ *of the first* m *data, the only possible permutations* ρ' *of* $[1, 2, 3, \ldots, m, m+1]$ *compatible with* (55) *are those which place* $\rho'(m+1)$ *in one of the* $m+1$ *intervals formed by* $\rho(1), \rho(2), \ldots, \rho(m)$.

Proof. ⟦We require that, as in (55),

$$\alpha_{\rho'(1)} < \alpha_{\rho'(2)} < \alpha_{\rho'(3)} < \ldots < \alpha_{\rho'(m)} < \alpha_{\rho'(m+1)}; \qquad (56)$$

so that m of the $\rho'(j)$ must be the $\rho(i)$, in the order of (55). If $\rho'(k) = m + 1$ ($k = 1, 2, 3, \ldots,$ or $m + 1$); then $\rho'(1) = \rho(1)$, $\rho'(2) = \rho(2)$, ..., $\rho'(k - 1) = \rho(k - 1)$, $\rho'(k + 1) = \rho(k)$, $\rho'(k + 2) = \rho(k + 1)$, ..., $\rho'(m + 1) = \rho(m)$, proving the lemma.⟧

COROLLARY 1.1: *Given the ordering* (55) *of the first m data keys, the* $(m + 1)$*-st key* $\alpha_{\rho(m+1)}$ *has equal probability of falling into any of the* $m + 1$ *intervals formed by the earlier keys (in order)* $\alpha_{\rho(1)}, \alpha_{\rho(2)}, \ldots, \alpha_{\rho(m)}$.

⟦Of all possible permutations ρ' specifying the ordering (56) of all $m + 1$ keys, only the $m + 1$ permutations defined in Lemma 1 are possible, if the ordering (55) of the first m keys is given. By Assumption 1 and the definition of conditional probability, these $m + 1$ permutations are themselves equally probable.⟧

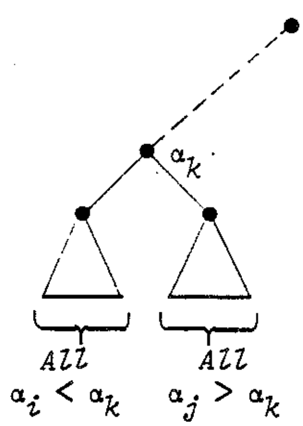


Figure 3.

LEMMA 2: *All keys in the left sub-tree of any node are less than the key at that node, and that is less than all keys in the right sub-tree.*

⟦(See Figure 3). The insertion rule ensures that any key finding its way into the right sub-tree must pass through a comparison at the node holding α_k (say) and exceed it; similarly, any key entered in the left sub-tree must be less than α_k in value.⟧

COROLLARY 2.1: *If an m-node binary tree is formed by entering m data with keys* $\alpha_1, \alpha_2, \ldots, \alpha_m$ *(entered successively, in the order stated), and the tree is augmented with* $m + 1$ *external (open) nodes; then each open node corresponds to one of the* $m + 1$ *intervals described in Lemma 1 (that is, an* $(m + 1)$*-st datum will be entered at that open node if it falls in the corresponding interval).*

⟦If $\alpha_i < \alpha_j$, then, by Lemma 2, either (i) α_i is entered at an (internal) node in the left sub-tree of the node holding α_j, or (ii) α_j is entered at a node in the right sub-tree of the node holding α_i, or (iii) there is a node holding a key α_k such that $\alpha_i < \alpha_k < \alpha_k$, and α_i is in the left and α_j is in the right sub-trees of the node holding α_k. If, further, we know that there is no key (entered in the tree) which lies between α_i and α_j, then case (iii) is excluded entirely; and, in case (i), α_i is at the rightmost (internal, i.e., occupied) node of the left sub-tree of the node holding α_j, so that it is at the last of a string of right-children of the root of that sub-tree; while, in case (ii), α_j is similarly at the leftmost node of the right sub-tree of the node holding α_i. In either case, (i) the right child of the node holding α_i or (ii) the left child of the node holding α_j, is an open node (since its parent is the *last* occupied node, going to the right (or left, respectively) and will be filled by a new key if and only if that key lies between α_i and α_j. Since there are just $m + 1$ open nodes and just $m + 1$ intervals between the ordered keys, this suffices to prove the result.⟧ Case (i) is illustrated in Figure 4; case (ii) is entirely analogous.

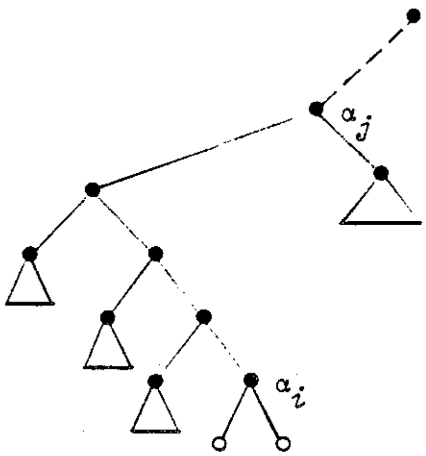


Figure 4.

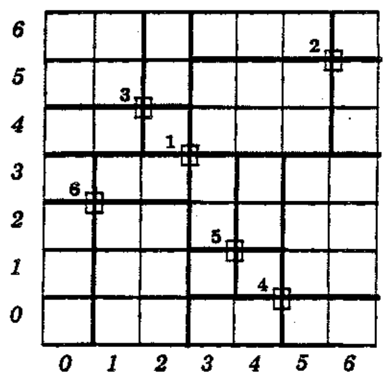
Theorem 1: *If m data are formed into a binary tree according to the insertion rule defined earlier, and the tree is augmented with m + 1 open (external) nodes; then a further random datum is equally likely to be inserted at any of the open nodes.*

⟦By Corollary 1.1, all intervals are equally likely candidates for the placement of the new key into the order of the previous data; by Corollary 2.1, each interval corresponds to a single open node. The theorem follows.⟧

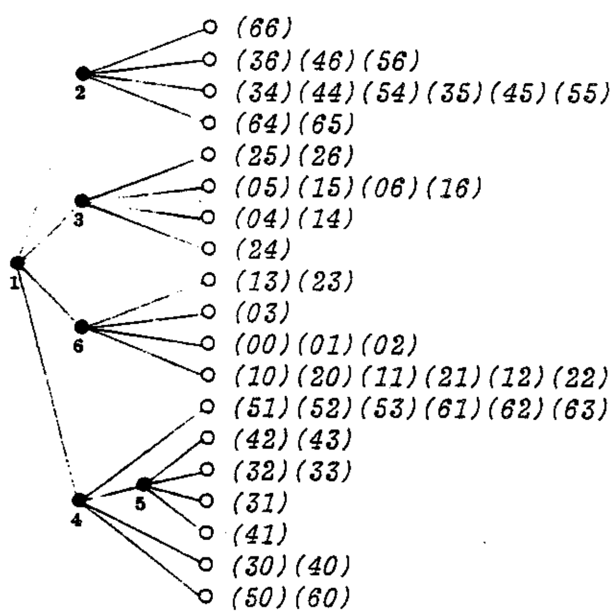
Even for binary trees, it is possible to devise alternative probability structures to that used in Assumption 1 above, or its consequence, Theorem 1. For example, we may define equivalence-classes of binary trees, and say that all equivalence-classes are equally likely (see, e.g., the work of Renyi and Szekeres [67], de Bruijn, Knuth, and Rice [72], Meir and Moon [78], Kemp [79], Odlyzko [79], Flajolet, Raoult, and Vuillemin [79], Flajolet and Steyaert [80], or Flajolet and Odlyzko [82]). When we attempt to generalize to s-ary trees, the alternatives multiply. We shall make the following assumption, by analogy with Theorem 1.

Assumption 2: *The random input of the s-ary tree is so structured that, if m data are formed into a tree according to a suitable insertion rule, and the tree is augmented with* $n = (s - 1)m + 1$ *external (open) nodes; then a further random datum is equally likely to be inserted at any of the n open nodes.*

We note in passing that this is not the same probability as is naturally generated by quad-trees and similar structures (see, e.g., Finkel and Bentley [74] or Bentley [75, 79]), in which each datum has d keys $(\alpha_1, \alpha_2, \ldots, \alpha_d)$ and a 2^d-ary tree is generated by simultaneous ordering of each key. In Figure 5, a simple example shows the difference, when $m = 6$ and $d = 2$.



Six double-key data entered in order '1', '2', '3', '4', '5', '6'. Italic numbers denote intervals generated by these data, in each coordinate.



Corresponding tree, indicating which 'squares' go with which open nodes.

Figure 5.

It is natural to extend the structure of Corollary 1.1 and assume that *all intervals generated by each key are equally probable, within any key, and independent, between keys.* This means that the "squares" in the first part of the figure would be equally probable. But this particular situation yields the tree shown in the second part of the figure, where we see that the sets of "squares" corresponding to the various open nodes vary in number from one to six (there are clearly $7 \times 7 = 49$ "squares" and only 19 open nodes). [The assumed insertion rule is similar to that for binary trees: at each node, search moves to one of the four children, "NE" or "++", if the keys to be inserted are both greater than their counterparts at the node being examined, "NW" or "-+", if the first key is less and the second greater, "SW" or "--" if both are less, and "SE" or "+-", if the first key is greater and the second is less.] The regions of the ordering-diagram (top of Figure 5) corresponding to the open nodes of the graph generated by the same data (bottom of Figure 5) are outlined in thicker borders, and it is clear that their boundaries are generated in a very natural way, each successive datum falling into such a region and quadrisecting it. Since the "squares" are not really square, but formally define order only; it is plausible to argue that the equal status of each of these regions (corresponding one-to-one to the open nodes) is more analogous to the equal status of the intervals into which single-key data dissect the line, than is the conferring of equal status to each "square", which is a knee-jerk application of the Cartesian product, taking no notice of the order in which the data are entered.

4. STATISTICAL RELATIONSHIPS

By Assumption 2, in an m-node s-ary tree, each of the open nodes has probability

$$1/n = 1/[(s - 1)m + 1], \tag{57}$$

by (3), of being the next node filled. We may define the mathematical expectations of the parameters μ_{mk} and ν_{mk} defined in §2 to be

$$M_{mk} = E[\mu_{mk}] \quad \text{and} \quad N_{mk} = E[\nu_{mk}]. \tag{58}$$

By (44) - (48), if $k \geqslant 1$ and $m \geqslant 1$, we have that

$$M_{m0} = 1 \quad \text{and} \quad N_{m0} = 0; \tag{59}$$

$$\text{if} \quad k \geqslant m, \quad M_{mk} = N_{m(k+1)} = 0; \tag{60}$$

$$\sum_{k=0}^{\infty} M_{mk} = m, \tag{61}$$

$$\sum_{k=0}^{\infty} N_{mk} = (s - 1)m + 1; \qquad (62)$$

and
$$N_{mk} = sM_{m(k-1)} - M_{mk}. \qquad (63)$$

Consider now an $(m - 1)$-node tree to which an m-th node is added at level h. Then, clearly, since $m \geq 1$,

$$\mu_{mk} = \mu_{(m-1)k} + \delta_{hk}, \qquad (64)$$

where δ_{hk} is the Kronecker symbol,

$$\delta_{hk} = \begin{cases} 1 & \text{if } h = k \\ 0 & \text{if } h \neq k \end{cases}; \qquad (65)$$

and so, by (48) and (58), if $k \geq 1$ and $m \geq 1$,

$$M_{mk} - M_{(m-1)k} = E[\delta_{hk}] = \sum_{h=0}^{\infty} \delta_{hk} \frac{E[\nu_{(m-1)h}]}{(s - 1)(m - 1) + 1}$$

$$= \frac{sM_{(m-1)(k-1)} - M_{(m-1)k}}{(s - 1)(m - 1) + 1}.$$

Collecting terms, we thus see that

$$M_{mk} = \alpha_m M_{(m-1)k} + \beta_m M_{(m-1)(k-1)}, \qquad (66)$$

where
$$\alpha_m = \frac{m - 1}{m - 1 + \theta} \quad \text{and} \quad \beta_m = \frac{1 + \theta}{m - 1 + \theta}, \qquad (67)$$

by (12).

Also, inserting the m-th node at level h reduces $\nu_{(m-1)h}$ by one and increases $\nu_{(m-1)(h+1)}$ by s; so that, similarly, if $k \geq 1$ and $m \geq 1$,

$$\nu_{mk} = \nu_{(m-1)k} - \delta_{hk} + s\delta_{(h+1)k}; \qquad (68)$$

whence, just as in getting (66) from (64), we obtain that

$$N_{mk} = \alpha_m N_{(m-1)k} + \beta_m N_{(m-1)(k-1)}, \qquad (69)$$

with the same coefficients α_m and β_m as before. The difference in the values of M_{mk} and N_{mk} originates in the differing initial conditions,

$$M_{00} = 0,\ M_{m0} = 1\ (m \geq 1) \quad \text{and} \quad N_{00} = 1,\ N_{m0} = 0\ (m \geq 1). \qquad (70)$$

Now let us define the functions

$$\boldsymbol{E}_m(t) = \sum_{k=0}^{\infty} \mathrm{N}_{mk} \mathrm{e}^{ikt} \tag{71}$$

and

$$\boldsymbol{F}_m(t) = \sum_{k=0}^{\infty} \mathrm{M}_{mk} \mathrm{e}^{ikt}; \tag{72}$$

so that, by (49) and (58),

$$\boldsymbol{E}_m(t) = \sum_{k=0}^{\infty} \mathrm{N}_{mk} \sum_{p=0}^{\infty} \frac{(ikt)^p}{p!} = \sum_{p=0}^{\infty} \frac{(it)^p}{p!} \mathrm{E}[E_m^{(p)}], \tag{73}$$

and, similarly,

$$\boldsymbol{F}_m(t) = \sum_{p=0}^{\infty} \frac{(it)^p}{p!} \mathrm{E}[F_m^{(p)}]. \tag{74}$$

We now note that (see Figure 2 and (60))

$$\begin{aligned} &\mathrm{N}_{10} = 0,\ \mathrm{N}_{11} = s,\ \mathrm{N}_{1k} = 0\ (k \geqslant 2), \\ \text{and}\quad &\mathrm{M}_{10} = 1,\ \mathrm{M}_{1k} = 0\ (k \geqslant 1); \end{aligned} \tag{75}$$

whence

$$\boldsymbol{E}_1(t) = s\mathrm{e}^{it} \quad \text{and} \quad \boldsymbol{F}_1(t) = 1. \tag{76}$$

Now, by (69) with (60) and (70), if $m \geqslant 2$,

$$\begin{aligned} \boldsymbol{E}_m(t) &= \sum_{k=1}^{m} \mathrm{N}_{mk} \mathrm{e}^{ikt} = \alpha_m \sum_{k=1}^{m} \mathrm{N}_{(m-1)k} \mathrm{e}^{ikt} + \beta_m \sum_{k=1}^{m} \mathrm{N}_{(m-1)(k-1)} \mathrm{e}^{ikt} \\ &= \alpha_m \sum_{k=1}^{m-1} \mathrm{N}_{(m-1)k} \mathrm{e}^{ikt} + \beta_m \mathrm{e}^{it} \sum_{j=1}^{m-1} \mathrm{N}_{(m-1)j} \mathrm{e}^{ijt} \quad [\![j = m-1]\!] \\ &= (\alpha_m + \beta_m \mathrm{e}^{it}) \boldsymbol{E}_{m-1}(t); \end{aligned} \tag{77}$$

whence

$$\boldsymbol{E}_m(t) = \prod_{h=1}^{m} (\alpha_h + \beta_h \mathrm{e}^{it}), \tag{78}$$

since, by (67), $\alpha_1 = 0$ and $\beta_1 = \frac{1+\theta}{\theta} = s$ (thus including the first equation of (76) as a special case of (78)). Similarly, we see that

$$\begin{aligned} \boldsymbol{F}_m(t) &= 1 + \sum_{k=1}^{m} \mathrm{M}_{mk}\mathrm{e}^{\mathrm{i}kt} \\ &= 1 - \alpha_m + \alpha_m \sum_{k=0}^{m-1} \mathrm{M}_{(m-1)k}\mathrm{e}^{\mathrm{i}kt} + \beta_m \mathrm{e}^{\mathrm{i}t} \sum_{j=0}^{m-1} \mathrm{M}_{(m-1)j}\mathrm{e}^{\mathrm{i}jt} \\ &= (1 - \alpha_m) + (\alpha_m + \beta_m \mathrm{e}^{\mathrm{i}t})\boldsymbol{F}_{m-1}(t); \end{aligned} \tag{79}$$

whence

$$\boldsymbol{F}_m(t) = \sum_{j=1}^{m} (1 - \alpha_j) \prod_{h=j+1}^{m} (\alpha_h + \beta_h \mathrm{e}^{\mathrm{i}t}), \tag{80}$$

as is easily verified.

Applying Maclaurin's theorem,

$$\phi(t) = \sum_{p=0}^{\infty} \frac{t^p}{p!}[(\frac{\partial}{\partial t})^p \phi(t)]_{t=0}, \tag{81}$$

to (73) and (74), we see that

$$\mathrm{E}[E_m^{(p)}] = (-\mathrm{i})^p [(\frac{\partial}{\partial t})^p \boldsymbol{E}_m(t)]_{t=0} \tag{82}$$

and

$$\mathrm{E}[F_m^{(p)}] = (-\mathrm{i})^p [(\frac{\partial}{\partial t})^p \boldsymbol{F}_m(t)]_{t=0}. \tag{83}$$

By (61), (62), (71), and (72), we have that

$$\boldsymbol{E}_m(0) = \frac{m+\theta}{\theta} \quad \text{and} \quad \boldsymbol{F}_m(0) = m. \tag{84}$$

(This is also obtained, by a little algebra, from (78) and (80).) Now, we see that

$$\begin{aligned} \mathrm{E}[E_m^{(1)}] &= -\mathrm{i}[\frac{\partial}{\partial t} \prod_{h=1}^{m} (\alpha_h + \beta_h \mathrm{e}^{\mathrm{i}t})]_{t=0} = \sum_{h=1}^{m} \frac{\beta_h}{\alpha_h + \beta_h} \boldsymbol{E}_m(0) \\ &= \frac{m+\theta}{\theta} \sum_{h=1}^{m} \frac{1+\theta}{h+\theta} = \frac{m+\theta}{\theta} T_m^{(1)}, \end{aligned} \tag{85}$$

by (13). Similarly,

$$E[E_m^{(2)}] = \frac{m+\theta}{\theta}\left\{\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\beta_i}{\alpha_i+\beta_i}\frac{\beta_j}{\alpha_j+\beta_j} + \sum_{i=1}^{m}\left[\frac{\beta_i}{\alpha_i+\beta_i} - \left(\frac{\beta_i}{\alpha_i+\beta_i}\right)^2\right]\right\}$$

$$= \frac{m+\theta}{\theta}\{[T_m^{(1)}]^2 + T_m^{(1)} - T_m^{(2)}\}, \qquad (86)$$

$$E[E_m^{(3)}] = \frac{m+\theta}{\theta}\left\{\sum_{h=1}^{m}\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{1+\theta}{h+\theta}\frac{1+\theta}{i+\theta}\frac{1+\theta}{j+\theta} + 3\sum_{h=1}^{m}\sum_{i=1}^{m}\left[\frac{1+\theta}{h+\theta}\right.\right.$$

$$\left. - \left(\frac{1+\theta}{h+\theta}\right)^2\right]\frac{1+\theta}{i+\theta} + \sum_{h=1}^{m}\left[\frac{1+\theta}{h+\theta} - \left(\frac{1+\theta}{h+\theta}\right)^3\right] + 3\sum_{h=1}^{m}\left[\left(\frac{1+\theta}{h+\theta}\right)^3\right.$$

$$\left.\left. - \left(\frac{1+\theta}{h+\theta}\right)^2\right]\right\} = \frac{m+\theta}{\theta}\{[T_m^{(1)}]^3 + 3[T_m^{(1)}]^2 - 3T_m^{(2)}T_m^{(1)}$$

$$+ T_m^{(1)} - T_m^{(3)} + 3T_m^{(3)} - 3T_m^{(2)}\}$$

$$= \frac{m+\theta}{\theta}\{[T_m^{(1)}]^3 + 3[T_m^{(1)}]^2 - 3T_m^{(2)}T_m^{(1)} + 2T_m^{(3)} - 3T_m^{(2)} + T_m^{(1)}\}, \qquad (87)$$

and so on.

The direct calculation of the corresponding $E[F_m^{(p)}]$ is rather laborious; but, fortunately, we have the relation (52), leading directly to the corresponding relation for the expectations,

$$E[E_m^{(p)}] = (s-1)E[F_m^{(p)}] + s\sum_{q=0}^{p-1}\binom{p}{q}E[F_m^{(q)}]. \qquad (88)$$

From this, we obtain that, since $E[F_m^{(0)}] = m$, by (50), $E[E_m^{(1)}] = (s-1)E[F_m^{(1)}] + sm$, whence, by (85),

$$E[F_m^{(1)}] = (m+\theta)T_m^{(1)} - m(1+\theta); \qquad (89)$$

and $\mathrm{E}[E_m^{(2)}] = (s - 1)\mathrm{E}[F_m^{(2)}] + s\{m + 2\mathrm{E}[F_m^{(1)}]\}$, whence, by (86),

$$\mathrm{E}[F_m^{(2)}] = (m + \theta)\{[T_m^{(1)}]^2 - (1 + 2\theta)T_m^{(1)} - T_m^{(2)}\} + m(1 + \theta)(1 + 2\theta); \tag{90}$$

and $\mathrm{E}[E_m^{(3)}] = (s - 1)\mathrm{E}[F_m^{(3)}] + s\{m + 3\mathrm{E}[F_m^{(1)}] + 3\mathrm{E}[F_m^{(2)}]\}$, whence, by (87),

$$\mathrm{E}[F_m^{(3)}] = (m + \theta)\{[T_m^{(1)}]^3 - 3T_m^{(2)}T_m^{(1)} + 2T_m^{(3)} - 3\theta[T_m^{(1)}]^2 + 3\theta T_m^{(2)} + (1 + 6\theta + 6\theta^2)T_m^{(1)}\} - m(1 + \theta)(1 + 6\theta + 6\theta^2); \tag{91}$$

and so on.

By (3), (4), and (5), since $n = \frac{m + \theta}{\theta}$, we see that we immediately obtain (6) - (9), as announced, as well as

$$\mathrm{E}[X_m^{(3)}] = [T_m^{(1)}]^3 + 3[T_m^{(1)}]^2 - [3T_m^{(2)} - 1]T_m^{(1)} - 3T_m^{(2)} + 2T_m^{(3)} \tag{92}$$

and

$$\mathrm{E}[Y_m^{(3)}] = \frac{m + \theta}{m}\{[T_m^{(1)}]^3 - 3\theta[T_m^{(1)}]^2 - [3T_m^{(2)} - 1 - 6\theta - 6\theta^2]T_m^{(1)} + 3\theta T_m^{(2)} + 2T_m^{(3)}\} - (1 + \theta)(1 + 6\theta + 6\theta^2); \tag{93}$$

with a clear path to higher internal and external sums and averages, by increasingly, but not intolerably, laborious calculations.

To continue the analysis, we should now consider the higher moments of the quantities $E_m^{(p)}$, $F_m^{(p)}$, $X_m^{(p)}$, and $Y_m^{(p)}$. Since things rapidly get highly complicated, we shall only explicitly calculate the *variances* of $X_m^{(1)}$ and $Y_m^{(1)}$. The method used, however, is clearly extensible to other cases.

We have, by (4) and (49), with (85), that

$$\begin{aligned}\mathrm{var}[X_m^{(1)}] &= \mathrm{E}[\{X_m^{(1)}\}^2] - \{\mathrm{E}[X_m^{(1)}]\}^2 \\ &= \left(\frac{\theta}{m + \theta}\right)^2 \mathrm{E}\left[\left\{\sum_{k=0}^{\infty} k\nu_{mk}\right\}^2\right] - [T_m^{(1)}]^2 \\ &= \left(\frac{\theta}{m + \theta}\right)^2 \sum_{i=1}^{m}\sum_{j=1}^{m} ij\mathrm{E}[\nu_{mi}\nu_{mj}] - [T_m^{(1)}]^2. \end{aligned} \tag{94}$$

Referring to the relation (68), we see that

$$E[\nu_{mi}\nu_{mj}] = E[\{\nu_{(m-1)i} - \delta_{hi} + s\delta_{(h+1)i}\}\{\nu_{(m-1)j} - \delta_{hj} + s\delta_{(h+1)j}\}], \quad (95)$$

if the m-th node of the tree is inserted at level h. As before, we average first over the m-th node and then over all trees, and write S_{mij} for the quantity (95). Then

$$\begin{aligned} S_{mij} &= S_{(m-1)ij} - E[\nu_{(m-1)i}\delta_{hj}] + sE[\nu_{(m-1)i}\delta_{(h+1)j}] - E[\delta_{hi}\nu_{(m-1)j}] \\ &\quad + E[\delta_{hi}\delta_{hj}] - sE[\delta_{hi}\delta_{(h+1)j}] + sE[\delta_{(h+1)i}\nu_{(m-1)j}] \\ &\quad - sE[\delta_{(h+1)i}\delta_{hj}] + s^2E[\delta_{(h+1)i}\delta_{(h+1)j}] \\ &= \frac{m-1-\theta}{m-1+\theta} S_{(m-1)ij} + \frac{1+\theta}{m-1+\theta}[S_{(m-1)i(j-1)} + S_{(m-1)(i-1)j}] \\ &\quad + \frac{\theta}{m-1+\theta}\Big\{N_{(m-1)i}[\delta_{ij} - \frac{1+\theta}{\theta}\delta_{i(j-1)}] + N_{(m-1)(i-1)}[\frac{1+\theta}{\theta}\delta_{(i-1)j} \\ &\quad - (\frac{1+\theta}{\theta})^2 \delta_{ij}]\Big\}. \end{aligned} \quad (96)$$

We note that, by (47), $\Sigma_{j=1}^{m}S_{mij} = \Sigma_{j=1}^{m}E[\nu_{mi}\nu_{mj}] = E[\Sigma_{j=1}^{m}\nu_{mi}\nu_{mj}] = \frac{m+\theta}{\theta}E[\nu_{mi}] = \frac{m+\theta}{\theta}N_{mi}$ and $\Sigma_{i=1}^{m}N_{mi} = \frac{m+\theta}{\theta}$. Thus, (94) and (96) yield that

$$\begin{aligned} \mathrm{var}[X_m^{(1)}] &= (\frac{\theta}{m+\theta})^2 \sum_{i=1}^{m}\sum_{j=1}^{m}\Big\{\frac{m-1-\theta}{m-1+\theta} ijS_{(m-1)ij} + \frac{2(1+\theta)}{m-1+\theta} i(j+1)S_{(m-1)ij} \\ &\quad + \frac{\theta}{m-1+\theta}[i^2N_{(m-1)i} - 2\frac{1+\theta}{\theta} i(i+1)N_{(m-1)i} + (\frac{1+\theta}{\theta})^2(i \\ &\quad + 1)^2N_{(m-1)i}]\Big\} - [T_m^{(1)}]^2 \\ &= \left(1 - \frac{1}{(m+\theta)^2}\right)\left(\mathrm{var}[X_{m-1}^{(1)}] + [T_{m-1}^{(1)}]^2\right) + \frac{1}{(m+\theta)^2}E[X_{m-1}^{(2)}] \\ &\quad + 2(\frac{1+\theta}{m+\theta})E[X_{m-1}^{(1)}] + (\frac{1+\theta}{m+\theta})^2 - [T_m^{(1)}]^2. \end{aligned} \quad (97)$$

This reduces to:

$$\operatorname{var}[X_m^{(1)}] = \left(1 - \frac{1}{(m+\theta)^2}\right)\operatorname{var}[X_{m-1}^{(1)}] + \frac{1}{(m+\theta)^2}[T_{m-1}^{(1)} - T_{m-1}^{(2)}]; \qquad (98)$$

and we observe that

$$\prod_{h=k+1}^{m}\left(1 - \frac{1}{(h+\theta)^2}\right) = \prod_{h=k+1}^{m} \frac{(h-1+\theta)}{(h+\theta)}\frac{(h+1+\theta)}{(h+\theta)}$$

$$= \frac{(k+0+\theta)}{(k+1+\theta)}\frac{(k+2+\theta)}{(k+1+\theta)}\frac{(k+1+\theta)}{(k+2+\theta)}\frac{(k+3+\theta)}{(k+2+\theta)}\frac{(k+2+\theta)}{(k+3+\theta)}\frac{(k+4+\theta)}{(k+3+\theta)}$$

$$\cdots \frac{(m-2+\theta)}{(m-1+\theta)}\frac{(m-0+\theta)}{(m-1+\theta)}\frac{(m-1+\theta)}{(m-0+\theta)}\frac{(m+1+\theta)}{(m-0+\theta)}$$

$$= \frac{(k+\theta)(m+1+\theta)}{(k+1+\theta)(m+\theta)}, \qquad (99)$$

with the fractions cross-cancelling in pairs, except for the first and last. Since $X_1^{(1)} = 1$, so that $\operatorname{var}[X_1^{(1)}] = 0$, we see that (98) can be solved in the form

$$\operatorname{var}[X_m^{(1)}] = \sum_{k=2}^{m}\left\{\prod_{h=k+1}^{m}\left(1 - \frac{1}{(h+\theta)^2}\right)\right\}\frac{1}{(k+\theta)^2}[T_{k-1}^{(1)} - T_{k-1}^{(2)}]$$

$$= \sum_{k=2}^{m}\frac{m+1+\theta}{m+\theta}\left(\frac{1}{k+\theta} - \frac{1}{k+1+\theta}\right)[T_{k-1}^{(1)} - T_{k-1}^{(2)}]. \qquad (100)$$

Now note that, for any sequence $f_0, f_1, f_2, \ldots$

$$\sum_{k=2}^{m}\left(\frac{1}{k+\theta} - \frac{1}{k+1+\theta}\right)f_{k-1} = \sum_{k=2}^{m}\frac{1}{k+\theta}f_{k-1} - \sum_{j=3}^{m+1}\frac{1}{j+\theta}f_{j-2}$$

$$= \sum_{k=2}^{m}\frac{1}{k+\theta}(f_{k-1} - f_{k-2}) + \frac{1}{2+\theta}f_0 - \frac{1}{m+1+\theta}f_{m-1}. \qquad (101)$$

Further, we evaluate the telescoping series:

$$\sum_{k=2}^{m}\frac{1}{(k+\theta)(k-1+\theta)} = \sum_{k=2}^{m}\left(\frac{1}{k-1+\theta} - \frac{1}{k+\theta}\right) = \frac{1}{1+\theta} - \frac{1}{m+\theta} \qquad (102)$$

and $$\sum_{k=2}^{m} \frac{1}{(k+\theta)(k-1+\theta)^2} = \sum_{k=2}^{m}\left(\frac{1}{(k-1+\theta)^2} - \frac{1}{(k+\theta)(k-1+\theta)}\right)$$

$$= \frac{1}{(1+\theta)^2} T_{m-1}^{(2)} - \frac{1}{1+\theta} + \frac{1}{m+\theta} . \qquad (103)$$

Successively taking $f_j = T_j^{(1)}$ and $T_j^{(2)}$, and noting that $T_0^{(q)} = 0$ and that $T_{k-1}^{(q)} - T_{k-2}^{(q)} = (\frac{1+\theta}{k-1+\theta})^q$, we see that (100), with the help of (101) - (103), becomes

$$\operatorname{var}[X_m^{(1)}] = \frac{m+1+\theta}{m+\theta}\left\{\sum_{k=2}^{m} \frac{1+\theta}{(k+\theta)(k-1+\theta)} - \frac{1}{m+1+\theta} T_{m-1}^{(1)}\right.$$

$$\left. - \sum_{k=2}^{m} \frac{(1+\theta)^2}{(k+\theta)(k-1+\theta)^2} + \frac{1}{m+1+\theta} T_{m-1}^{(2)}\right\}$$

$$= \frac{m+1+\theta}{m+\theta}\left\{1 - \frac{1+\theta}{m+\theta} - \frac{1}{m+1+\theta} T_{m-1}^{(1)} - T_{m-1}^{(2)} + (1+\theta)\right.$$

$$\left. - \frac{(1+\theta)^2}{m+\theta} + \frac{1}{m+1+\theta} T_{m-1}^{(2)}\right\}$$

$$= (2+\theta)(1 + \frac{1}{m+\theta})(1 - \frac{1+\theta}{m+\theta}) - \frac{1}{m+\theta} T_{m-1}^{(1)} - T_{m-1}^{(2)}$$

$$= (2+\theta)\frac{m}{m+\theta} - \frac{1}{m+\theta} T_m^{(1)} - T_m^{(2)},$$

confirming equation (10). Fortunately, we can get $\operatorname{var}[Y_m^{(1)}]$ from this; by way of (52) with $p = 1$, with (3), (4), (5), and (50); namely,

$$X_m^{(1)} = E_m^{(1)} \frac{\theta}{m+\theta} = \frac{1}{m+\theta} F_m^{(1)} + \frac{1+\theta}{m+\theta} F_m^{(0)} = \frac{m}{m+\theta}[Y_m^{(1)} + 1 + \theta], \qquad (104)$$

or $$Y_m^{(1)} = \frac{m+\theta}{m} X_m^{(1)} - 1 - \theta; \qquad (105)$$

whence, $\operatorname{var}[Y_m^{(1)}] = (\frac{m+\theta}{m})^2 \operatorname{var}[X_m^{(1)}]$, confirming equation (11).

5. ASYMPTOTIC RELATIONS

We observe that

$$A(1,\ k,\ m) = \sum_{h=k}^{m-1}\int_0^1 \frac{dz}{h + z + \theta} = \sum_{h=k}^{m-1}[\log(h + z + \theta)]_0^1$$

$$= \sum_{h=k}^{m-1}\{\log(h + 1 + \theta) - \log(h + \theta)\} = \log\frac{m + \theta}{k + \theta}, \quad (106)$$

$$A(q,\ k,\ m) = \sum_{h=k}^{m-1}\int_0^1 \frac{dz}{(h + z + \theta)^q} = \sum_{h=k}^{m-1}\left[\frac{1}{q - 1}\ \frac{-1}{(h + z + \theta)^{q-1}}\right]_0^1$$

$$= \frac{1}{q - 1}\sum_{h=k}^{m-1}\left\{\frac{1}{(h + \theta)^{q-1}} - \frac{1}{(h + 1 + \theta)^{q-1}}\right\}$$

$$= \frac{1}{q - 1}\left\{\frac{1}{(k + \theta)^{q-1}} - \frac{1}{(m + \theta)^{q-1}}\right\} \quad \text{for} \quad q \geqslant 2; \quad (107)$$

and, further,

$$\sum_{h=k+1}^{m}\int_0^1 \frac{dz}{h - z + \theta} = \sum_{h=k+1}^{m}[-\log(h - z + \theta)]_0^1$$

$$= \sum_{h=k+1}^{m}\{\log(h + \theta) - \log(h - 1 + \theta)\} = A(1,\ k,\ m), \quad (108)$$

$$\sum_{h=k+1}^{m}\int_0^1 \frac{dz}{(h - z + \theta)^q} = \sum_{h=k+1}^{m}\left[\frac{1}{q - 1}\ \frac{1}{(h - z + \theta)^{q-1}}\right]_0^1$$

$$= \frac{1}{q - 1}\sum_{h=k+1}^{m}\left\{\frac{1}{(h - 1 + \theta)^{q-1}} - \frac{1}{(h + \theta)^{q-1}}\right\} = A(q,\ k,\ m) \quad \text{for} \quad q \geqslant 2. \quad (109)$$

Now note that the function $1/(x + \theta)^q$, with $q \geqslant 1$, is *decreasing* and *concave upward*. Therefore, as we see in Figure 6, the horizontal segment LB lies below the arc AB and the segment BM lies above the arc BC; furthermore, the chord BC lies below BM and above the arc BC, while the arc lies above the

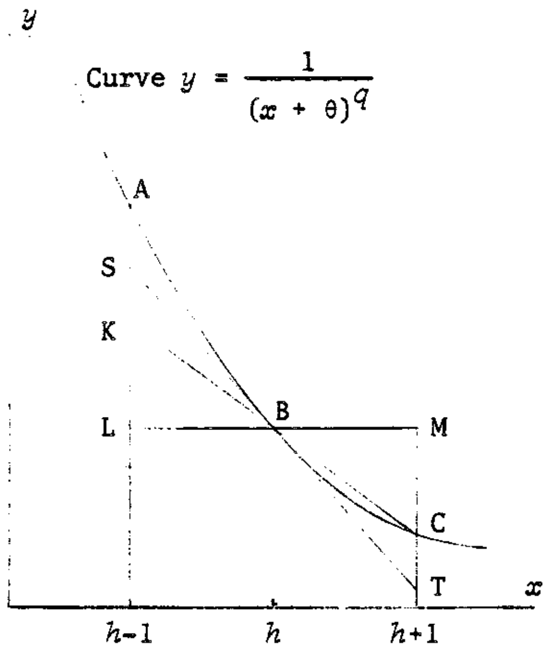


Figure 6.

tangent SBT (with the portion SB lying above the extension KB of the chord BC). Thus, first,

$$\frac{1}{(h+\theta)^q} < \int_{h-1}^{h} \frac{dx}{(x+\theta)^q} = \int_0^1 \frac{dz}{(h-z+\theta)^q} \tag{110}$$

and

$$\frac{1}{(h+\theta)^q} > \int_{h}^{h+1} \frac{dx}{(x+\theta)^q} = \int_0^1 \frac{dz}{(h+z+\theta)^q}; \tag{111}$$

secondly,

$$\frac{3/2}{(h+\theta)^q} - \frac{1/2}{(h+1+\theta)^q} < \int_0^1 \frac{dz}{(h-z+\theta)^q} \tag{112}$$

and

$$\frac{1/2}{(h+\theta)^q} + \frac{1/2}{(h+1+\theta)^q} > \int_0^1 \frac{dz}{(h+z+\theta)^q}; \tag{113}$$

and thirdly, since the derivative of the function $1/(x+\theta)^q$ is $-q/(x+\theta)^{q+1}$,

$$\frac{1}{(h+\theta)^q} + \frac{q/2}{(h+\theta)^{q+1}} < \int_0^1 \frac{dz}{(h-z+\theta)^q} \tag{114}$$

and

$$\frac{1}{(h+\theta)^q} - \frac{q/2}{(h+\theta)^{q+1}} < \int_0^1 \frac{dz}{(h+z+\theta)^q}. \tag{115}$$

Each of these inequalities may now be summed from $h = k + 1$ to $h = m$, yielding by (13) that, respectively,

$$T_m^{(q)} - T_k^{(q)} = (1+\theta)^q \sum_{h=k+1}^{m} \frac{1}{(h+\theta)^q} < (1+\theta)^q A(q, k, m), \tag{116}$$

$$T_m^{(q)} - T_k^{(q)} > (1+\theta)^q A(q, k+1, m+1), \tag{117}$$

$$T_m^{(q)} - T_k^{(q)} + \frac{(1+\theta)^q}{2}\left(\frac{1}{(k+1+\theta)^q} - \frac{1}{(m+1+\theta)^q}\right)$$

$$< (1+\theta)^q A(q, k, m), \quad (118)$$

$$T_m^{(q)} - T_k^{(q)} - \frac{(1+\theta)^q}{2}\left(\frac{1}{(k+1+\theta)^q} - \frac{1}{(m+1+\theta)^q}\right)$$

$$> (1+\theta)^q A(q, k+1, m+1), \quad (119)$$

with (118) clearly better than (116), and (119) better than (117),

$$T_m^{(q)} - T_k^{(q)} + \frac{q}{2(1+\theta)}[T_m^{(q+1)} - T_k^{(q+1)}] < (1+\theta)^q A(q, k, m), \quad (120)$$

$$T_m^{(q)} - T_k^{(q)} - \frac{q}{2(1+\theta)}[T_m^{(q+1)} - T_k^{(q+1)}]$$

$$< (1+\theta)^q A(q, k+1, m+1). \quad (121)$$

By (106) and (107), (118) and (119) simplify to

$$T_m^{(q)} - T_k^{(q)} < (1+\theta)^q[A(q, k, m) - \frac{q}{2} A(q+1, k+1, m+1)] \quad (122)$$

and

$$T_m^{(q)} - T_k^{(q)} > (1+\theta)^q[A(q, k+1, m+1) + \frac{q}{2} A(q+1, k+1, m+1)], \quad (123)$$

making (123) our best lower bound for $T_m^{(q)} - T_k^{(q)}$. Using this, we see that (120) yields that

$$T_m^{(q)} - T_k^{(q)} < (1+\theta)^q[A(q, k, m) - \frac{q}{2} A(q+1, k+1, m+1)$$

$$- \frac{q(q+1)}{4} A(q+2, k+1, m+1)], \quad (124)$$

which is clearly better than (122); and, using (124), we see that (121) becomes

$$T_m^{(q)} - T_k^{(q)} < (1+\theta)^q[A(q, k+1, m+1) + \frac{q}{2} A(q+1, k, m)$$

$$- \frac{q(q+1)}{4} A(q+2, k+1, m+1)$$

$$- \frac{q(q+1)(q+2)}{8} A(q+3, k+1, m+1)]. \quad (125)$$

If we write

$$A(q, k, m) = B(q, k) - B(q, m), \tag{126}$$

so that, by (106) and (107),

$$B(1, r) = -\log(r + \theta) \tag{127}$$

and

$$B(q, r) = \frac{1}{q - 1}\frac{1}{(r + \theta)^{q-1}} \quad \text{for } q \geqslant 2; \tag{128}$$

then, for any θ, we may interpret (123) - (125) as stating that the sequence

$$U_r^{(q)} = T_r^{(q)} + (1 + \theta)^q[B(q, r + 1) + \frac{q}{2} B(q + 1, r + 1)] \tag{129}$$

increases monotonically as $r \to \infty$, while the sequences

$$V_r^{(q)} = T_r^{(q)} + (1 + \theta)^q[B(q, r) - \frac{q}{2} B(q + 1, r + 1) - \frac{q(q + 1)}{4} B(q + 2, r + 1)] \tag{130}$$

and

$$W_r^{(q)} = T_r^{(q)} + (1 + \theta)^q[B(q, r + 1) + \frac{q}{2} B(q + 1, r) - \frac{q(q + 1)}{4} B(q + 2, r + 1) - \frac{q(q + 1)(q + 2)}{8} B(q + 3, r + 1)] \tag{131}$$

decrease monotonically as $r \to \infty$. Now, we note that $(\xi - 1)^{-q} = \xi^{-q}(1 - \frac{1}{\xi})^{-q}$ and therefore

$$\begin{aligned}(\xi - 1)^{-q} - \xi^{-q} - \frac{q}{2}\,\xi^{-q-1} - \frac{q(q + 1)}{4}\,\xi^{-q-2} &= \frac{q}{2}\,\xi^{-q-1} + \frac{q(q + 1)}{4}\,\xi^{-q-2} + \frac{q(q + 1)(q + 2)}{6}\,\xi^{-q-3} + \ldots \\ &> 0 \quad \text{for} \quad \xi > 0;\end{aligned} \tag{132}$$

whence, by (128), with $\xi = r + 1 + \theta$,

$$\begin{aligned}W_r^{(q)} - U_r^{(q)} &= \frac{1}{2}(1 + \theta)^q[(\xi - 1)^{-q} - \xi^{-q} - \frac{q}{2}\,\xi^{-q-1} - \frac{q(q + 1)}{4}\,\xi^{-q-2}] \\ &> 0,\end{aligned} \tag{133}$$

so that both $U_r^{(q)}$ and $W_r^{(q)}$ *converge* to respective limits:

$$U_r^{(q)} \nearrow u_q(\theta), \quad W_r^{(q)} \searrow w_q(\theta) \geqslant u_q(\theta). \tag{134}$$

Further, when $q = 1$, we note that

$$-\log \frac{\xi - 1}{\xi} - \xi^{-1} - \frac{1}{4}\xi^{-2} = \frac{1}{4\xi^2} + \frac{1}{3\xi^3} + \frac{1}{4\xi^4} + \ldots > 0, \tag{135}$$

and, when $q \geqslant 2$,

$$\frac{1}{q-1}[(\xi - 1)^{-q+1} - \xi^{-q+1}] - \xi^{-q} - \frac{q}{4}\xi^{-q-1}$$

$$= \frac{q}{4}\xi^{-q-1} + \frac{q(q+1)}{6}\xi^{-q-2} + \frac{q(q+1)(q+2)}{24}\xi^{-q-3} + \ldots > 0; \tag{136}$$

so that, similarly,

$$V_r^{(q)} - U_r^{(q)} > 0, \tag{137}$$

whence

$$V_r^{(q)} \searrow v_q(\theta) \geqslant u_q(\theta). \tag{138}$$

Indeed, we further observe that

$$W_r^{(q)} - U_r^{(q)} < \frac{1}{2}(1 + \theta)^q\ (r + \theta)^{-q} \to 0 \quad \text{as} \quad r \to \infty, \tag{139}$$

$$V_r^{(1)} - U_r^{(1)} < (1 + \theta)\log \frac{r + 1 + \theta}{r + \theta} \to 0 \quad \text{as} \quad r \to \infty, \tag{140}$$

$$V_r^{(q)} - U_r^{(q)} < (1 + \theta)^q \frac{1}{q-1}(r + \theta)^{-q+1} \to 0 \quad \text{as} \quad r \to \infty, \quad \text{for} \quad q \geqslant 2; \tag{141}$$

so that it follows that

$$u_q(\theta) = v_q(\theta) = w_q(\theta). \tag{142}$$

Some numerical calculations yield the following values for the limit:

s	2	4	10	100	∞
θ	1.000000	0.333333	0.111111	0.010101	0
$q = 1$	-0.845569	0.176045	0.453270	0.566386	0.577216
$q = 2$	2.579736	1.947728	1.744310	1.653890	1.644934
$q = 3$	1.616455	1.329923	1.242978	1.205694	1.202057

(143)

Finally, we can now deduce from all this that, by (129),

$$T_m^{(q)} < u_q(\theta) - (1+\theta)^q[B(q, m+1) + \frac{q}{2} B(q+1, m+1)]; \qquad (144)$$

by (130),

$$T_m^{(q)} > u_q(\theta) - (1+\theta)^q[B(q, m) - \frac{q}{2} B(q+1, m+1) - \frac{q(q+1)}{4} B(q+2, m+1)]; \qquad (145)$$

and, by (131),

$$T_m^{(q)} > u_q(\theta) - (1+\theta)^q[B(q, m+1) + \frac{q}{2} B(q+1, m+1) - \frac{q(q+1)}{4} B(q+2, m+1) - \frac{q(q+1)(q+2)}{8} B(q+3, m+1)]. \qquad (146)$$

We note that bounds in (144) - (146) equal $u_q(\theta) + T_m^{(q)} - \{U_m^{(q)}, V_m^{(q)}, W_m^{(q)}\}$, respectively, so that (133), (137), and (139) - (141) imply that

$$T_m^{(q)} \sim u_q(\theta) - (1+\theta)^q[B(q, m+1) + \frac{q}{2} B(q+1, m+1)]; \qquad (147)$$

which yields

$$T_m^{(1)} \sim u_1(\theta) + (1+\theta)\ [\log(m+1+\theta) - \frac{1}{2}\frac{1}{m+1+\theta}]$$

$$\sim (1+\theta) \log m + u_1(\theta) + O(\frac{1}{m}); \qquad (148)$$

$$T_m^{(q)} \sim u_q(\theta) - (1+\theta)^q[\frac{1}{q-1}\frac{1}{(m+1+\theta)^{q-1}} + \frac{1}{2}\frac{1}{(m+1+\theta)^q}]$$

$$\sim u_q(\theta) + O(m^{-q+1}) \quad \text{for} \quad q \geqslant 2. \qquad (149)$$

We recall (see, e.g., Copson [44] or Whittaker and Watson [27]) that the Riemann zeta-function is

$$\zeta(q) = 1 + \frac{1}{2^q} + \frac{1}{3^q} + \ldots + \frac{1}{h^q} + \ldots, \tag{150}$$

and Hurwitz's generalization is

$$\zeta(q, 1 + \theta) = \frac{1}{(1+\theta)^q} + \frac{1}{(2+\theta)^q} + \frac{1}{(3+\theta)^q} + \ldots \tag{151}$$

Since $s \geqslant 2$, by (12), $$0 < \theta \leqslant 1; \tag{152}$$

so that $$\zeta(q) - 1 = \zeta(q, 2) \leqslant \zeta(q, 1+\theta) < \zeta(q, 1) = \zeta(q); \tag{153}$$

and we see that $$T_m^{(q)} \nearrow (1+\theta)^q\, \zeta(q, 1+\theta), \text{ as } m \to \infty. \tag{154}$$

In particular, it is known that

$$\zeta(2) = \frac{\pi^2}{6}; \tag{155}$$

and that $$\sum_{h=1}^{m} \frac{1}{h} - \log m = \gamma_m \to \gamma = 0.5772156649\ldots, \tag{156}$$

as $m \to \infty$; γ is Euler's (or Mascheroni's) constant (see, e.g., Abramowitz and Stegun [72] or Mitrinović [66]). It follows from (148) with (153), (154), and (156) that

$$\begin{aligned}(1+\theta)(\gamma - 1) &= (1+\theta) \lim_{m\to\infty}\left(\sum_{h=2}^{m} \frac{1}{h} - \log m\right) \\ &\leqslant (1+\theta) \lim_{m\to\infty}\left(\sum_{h=1}^{m} \frac{1}{h+\theta} - \log m - \frac{1}{m+\theta}\right) \\ &= \lim_{m\to\infty}\{T_m^{(1)} - (1+\theta)\log m\} = u_1(\theta) \\ &\leqslant (1+\theta) \lim_{m\to\infty}\left(\sum_{h=1}^{m} \frac{1}{h} - \log m\right) = (1+\theta)\gamma;\end{aligned} \tag{157}$$

while, from (149) with (153), (154), and (155),

$$(1 + \theta)^2(\frac{\pi^2}{6} - 1) = (1 + \theta)^2[\zeta(2) - 1] \leqslant \lim_{m\to\infty} T_m^{(2)}$$

$$= (1 + \theta)^2\ u_2(\theta) < (1 + \theta)^2\ \zeta(2) = (1 + \theta)^2\ \frac{\pi^2}{6}. \qquad (158)$$

We see that the relation (148) yields (14), the relation (149) leads to (15), and the bounds given by the relations (157) and (158) yield (16). Note, too, that, when $s = 2$ and $s \to \infty$, we get $1 + \theta = 2$ and 1; and

$$u_1(1) = 2(\gamma - 1) \quad \text{and} \quad u_1(\infty) = \gamma, \qquad (159)$$

and

$$u_2(1) = 4(\frac{\pi^2}{6} - 1) \quad \text{and} \quad u_2(\infty) = \frac{\pi^2}{6}, \qquad (160)$$

as is seen in (143).

We now see immediately that (17) follows from (6) and (148) (or (14)); and, since $\frac{\log m}{m} \to 0$, (18) follows from (7) and (148). We also see that

$$\mathrm{E}[X_m^{(1)}] - \mathrm{E}[Y_m^{(1)}] = 1 + \theta - \frac{\theta}{m} T_m^{(1)} \sim 1 + \theta + O(\frac{\log m}{m}), \qquad (161)$$

as in (19). From (8) and (9), with (148) and (149), we see that

$$\mathrm{E}[X_m^{(2)}] \sim [(1 + \theta)\ \log m + O(1)]^2 + [(1 + \theta)\ \log m + O(1)] - O(1) \qquad (162)$$

and

$$\mathrm{E}[Y_m^{(2)}] \sim (1 + \frac{\theta}{m})\Big\{[(1 + \theta)\ \log m + O(1)]^2 - (1 + 2\theta)[(1 + \theta)\ \log m + O(1)] - O(1)\Big\} + (1 + \theta)(1 + 2\theta)$$

$$\sim (1 + \theta)^2(\log m)^2 + O(\log m) + O[(\log m)^2/m], \qquad (163)$$

which yield (20) and (21), since $(\log m)^2/m \to 0$. From (10) we now get that

$$\mathrm{var}[X_m^{(1)}] \sim (2 + \theta)(1 - \frac{\theta}{m + \theta}) - u_2(\theta) + O(\frac{1}{m}) - \frac{1 + \theta}{m + \theta} \log m + O(1)$$

$$\sim 2 + \theta + u_2(\theta) - \frac{1 + \theta}{m + \theta} \log m + O(1), \qquad (164)$$

which yields (22); and then (11) gives us (23) at once. Finally, we observe that

$$E[X_m^{(2)}] - (E[X_m^{(1)}])^2 = T_m^{(1)} - T_m^{(2)}, \tag{165}$$

leading immediately to (24); and that

$$E[Y_m^{(2)}] - (E[Y_m^{(1)}])^2 = (1 + \frac{\theta}{m})\left(- \frac{\theta}{m}[T_m^{(1)}]^2 + T_m^{(1)} - T_m^{(2)}\right) + \theta(1 + \theta), \tag{166}$$

which readily simplifies to (25).

REFERENCES

ABRAMOWITZ, M., STEGUN, I. A. (Editors) *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables.* National Bureau of Standards, Washington, DC (1972).

ADEL'SON-VEL'SKII, G. M., LANDIS, E. M. *Doklady Akademiia Nauk SSSR*, 146 (1962) 263-266.

AHO, A.V., HOPCROFT, J. E., ULLMAN, J. D. *Data Structures and Algorithms.* Addison-Wesley Publishing Co., Reading, MA (1983).

BENTLEY, J. L. *Communications of the Association for Computing Machinery*, 18 (1975) 509-517.

BENTLEY, J. L. *Institute of Electrical and Electronic Engineers Transactions on Software Engineering*, SE-5 (1979) 333-340.

BOOTH, A. D., COLIN, A. J. D. *Information and Control*, 3 (1960) 327-334.

deBRUIJN, N., KNUTH, D. E., RICE, O. *Graph Theory and Computing* (READ, R. C., Editor). Academic Press, Orlando, FL (1972) 15-22.

COPSON, E. T. *An Introduction to the Theory of Functions of a Complex Variable.* Oxford University Press, Oxford, England (1944).

DEVROYE, L. *Acta Informatica*, 21 (1984) 229-237.

DEVROYE, L. *Journal of the Association for Computing Machinery*, 33 (1986) 489-498.

FELLER, W. *An Introduction to Probability Theory and Its Applications.* Volume 1, Third Edition. John Wiley and Sons, New York, NY (1968).

FINKEL, R. A., BENTLEY, J. L. *Acta Informatica*, 4 (1974) 1-9.

FLAJOLET, P., ODLYZKO, A. *Journal of Computer Systems Science*, 25 (1982) 171-213.

FLAJOLET, P., RAOULT, J. C., VUILLEMIN, J. *Theoretical Computer Science*, 9 (1979) 99-125.

FLAJOLET, P., STEYAERT, J. M. *Proceedings, Seventh International Colloquium on Automata, Languages, and Programming*, Amsterdam, Netherlands (1980).

HALTON, J. H. *Statistics of Trees.* University of Wisconsin, Computer Sciences Department, Technical Report 334 (1978).

HALTON, J. H. *A Mini-Course on Probability and Statistics.* Thistle Press, Chapel Hill, NC (1985).

HIBBARD, T. N. *Journal of the Association for Computing Machinery*, 9 (1962) 16-17.

KEMP, R. *Proceedings, Sixth Internal Colloquium on Automata, Languages, and Programming*, Udine, Italy (1979).

KNUTH, D. E. *The Art of Programming. Volume 1: Fundamental Algorithms.* Addison-Wesley Publishing Co., Reading, MA (1968).

KNUTH, D. E. *The Art of Programming. Volume 3: Sorting and Searching.* Addison-Wesley Publishing Co., Reading, MA (1973).

LYNCH, W. C. *Computer Journal*, 8 (1965)

MAHMOUD, H., PITTEL, B. *Society for Industrial and Applied Mathematics Journal on Algebraic and Discrete Methods*, 5 (1984) 69-81.

MEIR, A., MOON, J. W . *CAnadian Journal of Mathematics*, 30 (1978) 997-1015.

MITRINOVIĆ, D. S. *Calculus of Residues*. P. Noordhoff, Groningen, Netherlands (1966).

MUNTZ, R., UZGALIS, R. *Proceedings, Princeton Conference on Information Sciences and Systems*, 4 (1970) 345-349.

ODLYZKO, A. [Periodic oscillations of coefficients of power series that satisfy functional equations] (1979) Referred to in Flajolet and Odlyzko [82] as "to appear".

PITTEL, B. *Journal of Mathematical Analysis and Its Applications*, 103 (1984) 461-480.

RENYI, A., SZEKERES, G. *Australian Journal of Mathematics*, 7 (1967) 497-507.

ROBSON, J. M. *Australian Computing Journal*, 11 (1979) 151-153.

ROBSON, J. M. *Australian Computer Science Communications* (1982) p. 88.

ROBSON, J. M. *On the Height of Binary Search Trees*. Australian National University, Computer Science Department Technical Report (1983)

STEPANOV, V. E. *Theoretical Probability and Its Applications*, 14 (1969) 65-78.

TUCKER, H. G. *A Graduate Course in Probability*. Academic Press, NY (1967).

WHITTAKER, E. T., WATSON, G. N. *A Course of Modern Analysis*. Cambridge University Press, Cambridge, England (1927).

WILSON, L. B. *BIT*, 16 (1976) 332-337.

WINDLEY, P. F. *Computer Journal*, 3 (1960) 84-88.

YAO, A. C. *Information Processing Letters*, 11 (1980) 84-86.