1986

Some Results on V-ary Asymmetric Tries

Wojciech Szpankowski
Purdue University, spa@cs.purdue.edu

Report Number:
86-582

SOME RESULTS ON V-ARY
ASYMMETRIC TRIES

Wojciech Szpankowski

CSD-TR-582
March 1986

SOME RESULTS ON V-ARY ASYMMETRIC TRIES

*Wojciech Szpankowski**

Department of Computer Sciences

Purdue University

West Lafayette, IN 47907, U.S.A.

* and Technical University of Gdansk, Poland.

Abstract

Tries (radix search tries) appear in many applications. It is assumed that a trie is built over an alphabet $U = \{\sigma_1, \ldots, \sigma_V\}$ (V-ary trie), and keys are strings of elements from U. The occurrence of an element σ_i in a key is represented by a probability p_i (asymmetric trie). Our main interest is to compute the depth of a leaf and the external path length. By solving a system of recurrences we find an exact formula for all factorial moments of these two parameters. Then, using the Mellin transform technique, we derive asymptotic approximations, and we prove that the m-th factorial moment of the depth of insertion for a trie with n keys is equal to $\alpha\, ln^m n + \beta\, ln^{m-1} n + O(ln^{m-2} n)$, where the constants α and β are functions of the probabilities p_i, $i = 1, \ldots, V$. In particular, we show that for symmetric tries the variance of the depth is $O(1)$, while for asymmetric tries it is $\alpha \ln n + O(1)$. These results extend previous analyses by Knuth [12], Flajolet and Sedgewick [6], Jacquet and Regnier [10], and Kirschenhofer and Prodinger [11].

List of symbols

a_n, $\hat{a}_n$

b_i, c_i , e_i

e_{b-1}

d_n^m, l_n^m

D_n, L_n

$f_r(n)$, $F(n)$

$H_n(z)$, $D_n(z)$

h_i, H_{b-1}

$j_\Sigma = j_1 + ... + j_V$

m, n, r

$$\begin{bmatrix} n \\ \mathrm{j} \end{bmatrix} = \frac{n!}{j_n! \cdots j_V!}$$

p_i, V

T_n, $S(n,r)$, $T(n,r)$

x_n , $X(z)$

z_k^r

α, β_i, γ_i, $\phi(z)$, Φ

σ_i, σ_n^2 , ξ_k

$\Gamma(z)$

$$\int_{c-i\infty}^{c+i\infty}$$

1. INTRODUCTION

Digital searching is a well-known technique for storing and retrieving information using lexicographical (digital) structure of words. Let U be an alphabet containing V elements, $U = \{\sigma_1, \sigma_2, \ldots, \sigma_V\}$, and we define a set S which consists of finite number, say n, of (possibly infinite) strings (keys) from U. A *trie* or *radix search trie* is such a V-ary digital search tree that edges are labelled by elements from U and leaves (external nodes) contain the keys [1],[8], [12]. The access path from the root to a leaf is a minimal prefix of the information contained in the leaf. An important variant of tries is obtained using a sequential storage algorithm for subtries with the size less than or equal to a fixed bound b. In other words, each external mode is capable of storing at most b keys. Such a trie will be called *b-trie* [3], [5].

Tries find many applications in computer sciences. A trie is used as an index to access data in secondary memory, e.g. in dynamic hashing algorithms [3], [5]. Tries are also applied in a number of other problems such as: tree-type conflict resolution algorithms for broadcast communications, radix exchange sort, polynomial factorization, simulation, Huffman's algorithm [1], [4], [5], [7], [8], [12], [13] , and so on. In these applications such tree parameters as *external path length* and *depth of a leaf (depth of insertion)* are used to compute important performance measures of the algorithms.

In most analysis binary tries are studied with uniform distribution of keys, that is, for $V = 2$ it is assumed that the probability of occurrence of σ_1, and σ_2 is the same and equal to 0.5. This was overcome by Knuth [12] who analyzed radix exchange sort algorithm with $V > 2$. Nevertheless, uniform distribution of elements from U was assumed [12]. This simplification is dropped in this paper, and we assume that a sequence of elements from U (keys) is an independent sequence of Bernoulli trials with probabilities p_i, $i = 1,2, \ldots, V$. More precisely, if x_k is the k-th element of a key, then $Pr\{x_k = \sigma_i\} = p_i$. Such an approach is known as *Bernoulli model* [3], [5], [10].

Previous analyses of trees have been mainly restricted either to the worst case or the average case analyses. Such a restriction was mainly imposed by difficulties to obtain other quantities of interests. However, it is well known that the mean is a very poor measure, and there is a need for higher moments of quantities of interest (these moments might be further used to study qualitative properties of trees). For example, the variance of the depth of a leaf provides information on *how well a tree is balanced*. Indeed, if a tree is ideally balanced, then all depths are the same and the variance is equal to zero. For other trees, by the Tchebyshev inequality, the smaller the variance , the more balanced the tree is, since the depth (random variable) is "closer" to its average value. Moreover, the third centralized moment of the depth is a measure of the skewness property of the distribution, and so on.

This paper presents a thorough analysis of *$b-$tries* from the depth of a leaf point of view, that is, we study all factorial moments of the depth for an asymmetric V-ary *$b-$trie*. An exact closed form expression for all factorial moments of the depth of insertion are provided through a solution of a system of recurrences. Furthermore, using the Mellin transform technique we derive an asymptotic approximation for these moments. In particular, we prove that the m-th factorial moment of the depth of insertion is $\alpha \ln^m n + \beta \ln^{m-1} n + O(\ln^{m-2} n)$ where α and β are functions of the probabilities p_i, $i = 1,2, \ldots, V$. We shall also show that the variance of the depth is either equal to $\alpha \ln n + O(1)$ for an asymmetric trie or $O(1)$ for a symmetric trie. The previous analyses were restricted to the average case for symmetric V-ary tries [12], [6]. Recently, Kirschenhofer and Prodinger [11] studied the variance of the depth of insertion for symmetric binary tries, while Jacquet and Regnier [10] obtained the limiting distribution for the depth of insertion for binary tries. This paper extends all of these analyses.

2. NOTATION AND PRELIMINARY RESULTS

Let us consider a set T_n of all *$b-$tries* with n keys over an alphabet $U = \{\sigma_1, \ldots, \sigma_V\}$. We assume that a key $x = \{x_1, x_2, \ldots, x_k, \ldots\}$ is a sequence of elements from U which form an independent sequence of Bernoulli trials. That is, for any k the probability $Pr\{x_k = \sigma_i\} = p_i$, $\sum_{i=1}^{V} p_i = 1$, and p_i does not depend on k. A trie $t \in T_n$ is called V-ary asymmetric trie since the alphabet U contains V elements which are distributed according to the probabilities p_i, $i = 1,2, \ldots, V$. Two parameters of tries are of particular interest: the depth of a leaf (depth of insertion) and the external path length. *The depth of a leaf* (successful search path [12]) is the number of internal nodes in the trie on the path from the root to the leaf. *The external path length* of a trie is the sum of the depths of all its leaves. We shall study properties of these two parameters in a set of all tries with n keys, i.e., in the set T_n.

There is a simple relationship between the depth of insertion and the external path length. To figure it out, we introduce some notations. For a trie $t \in T_n$, let $H_n^k(t)$ denote the number of keys at depth k in $t \in T_n$. Then $H_n^t(z) = \sum_{k=0}^{\infty} H_n^k(t) z^k$, is the generating function for $H_n^k(t)$, $t \in T_n$. Averaging $H_n^t(z)$ over all tries t in T_n we define a new generating function $H_n(z) = E\{H_n^t(z) \mid t \in T_n\}$. Similarly, let $L_n(t)$ and $D_n(t)$ denote the external path length and the depth of insertion for a given trie t in T_n, respectively. By L_n and D_n we define random variables representing the external path length and the depth of a leaf for all tries t in T_n, that is, $L_n = E\{L_n(t) \mid t \in T_n\}$ and $D_n = E\{D_n(t) \mid t \in \mid T_n\}$. The probability generating function of D_n is denoted by $D_n(z)$, that is, in T_n

$$D_n(z) = E\ z^{D_n} = \sum_{k=0}^{\infty} Pr\{D_n = k\} z^k, \tag{2.1}$$

where $Pr\{D_n = k\}$ is the probability that a randomly chosen key in a randomly selected trie $t \in T_n$ is at the depth k.

In this paper we deal with the factorial moments of D_n and L_n. The m-th factorial moments, $d_n^{\underline{m}}$ and $l_n^{\underline{m}}$, of D_n and L_n respectively are defined as :

$$d_n^{\underline{m}} = E\{D_n(D_n - 1) \cdots (D_n - m + 1)\} \tag{2.2a}$$

$$l_n^{\underline{m}} = E\{L_n(L_n - 1) \cdots (L_n - m + 1)\} \tag{2.2b}$$

The following lemma establishes a relationship between $H_n(z)$, $D_n(z)$ and the factorial moments :

LEMMA 1. For any natural m and n the following holds

$$H_n(1) = n, \quad \frac{d^m H_n(z)}{dz^m}|_{z=1} = H^{(m)}(1) = l_n^{\underline{m}}, \tag{2.3}$$

$$D_n(z) = H_n(z)/n, \tag{2.4}$$

$$d_n^{\underline{m}} = l_n^{\underline{m}}/n. \tag{2.5}$$

Proof: Eqs. (2.3) follow from the definition of $H_n(z)$ and (2.2b). Relationship (2.4) is a consequence of the fact that

$$Pr\{D_n(t) = k\} = \frac{H_n^k(t)}{n},$$

that is, for a given trie $t \in T_n$ the probability that a randomly chosen key is at depth k is equal to the fraction of keys at depth k (i.e., $H_n^k(t)$) and all keys (i.e., n). Finally, (2.5) follows immediately from (2.4).

□

Remark. For simplicity of notation we denote $l_n^{\underline{0}} \stackrel{def}{=} n$, hence by (2.3) $H^{(m)}(1) = l_n^{\underline{m}}$ for all $m \geq 0$.

By Lemma 1 the analysis of b*-tries* is reduced to study $H_n(z)$. There is no explicit formula for $H_n(z)$, but a rather nice recurrence. Let $\mathbf{j} = (j_1, j_2, \ldots, j_V)$ be a vector such that $j_1 + j_2 + \cdots + j_V = n$. Then $\binom{n}{\mathbf{j}} = \binom{n}{j_1, \ldots, j_V} = \frac{n!}{j_1! j_2!, \ldots, j_V!}$ and by $\sum_{j_\Sigma = n} f(j_1, \ldots, j_V)$ we denote a sum of $f(j_1, \ldots, j_V)$ over all $\mathbf{j}$ such that $j_1 + j_2 + \cdots + j_V = n$. We prove

LEMMA 2. For any natural n and b, $H_n(z)$ satisfies the following recurrence

$$H_n(z) = n \quad \text{for} \quad n \le b \tag{2.6a}$$

$$H_n(z) = z \sum_{j_\Sigma = n} \begin{bmatrix} n \\ \mathbf{j} \end{bmatrix} p_1^{j_1} p_2^{j_2} \cdots p_V^{j_V} \, [H_{j_1}(z) + \cdots + H_{j_V}(z)] \quad \text{for} \quad n > b. \tag{2.6b}$$

Proof. Let $t \in T_n$. Then obviously $H_n^t(z) = n$ for $n \le b$. For $n > b$ consider V subtries of the root, each with $j_1, j_2, \ldots, j_V$ keys, $j_1 + j_2 + \cdots + j_V = n$. Then

$$H_n^t(z) = z \ \{H_{j_1}^{(t)}(z) + \cdots + H_{j_V}^t(z)\}.$$

But $H_n(z) = E\{H_n^t(z) \mid t \in T_n\}$ and the probability of $j_1, j_2, \ldots, j_V$ keys in each of the subtrees is equal to $\begin{bmatrix} n \\ \mathbf{j} \end{bmatrix} p_1^{j_1}, \cdots p_V^{j_V}$. Hence, (2.6) follows.

□

3. EXACT ANALYSIS FOR THE FACTORIAL MOMENTS

By Lemma 1 to compute factorial moments of the depth of insertion we need factorial moments of the external path length l_n^m, which are the m-th derivatives of $H_n(z)$. Using Lemma 2 we prove that l_n^m $m = 1,2, \ldots,$ satisfy a system of recurrence equations.

THEOREM 1. The m-th factorial moment of L_n is given by the following recurrence

$$l_n^m = 0 \quad \text{for} \quad n \le b \tag{3.1a}$$

$$l_n^m = m! \sum_{i=1}^{m} (-1)^{m-i} \frac{l_n^{i-1}}{(i-1)!} + \sum_{j_\Sigma = n} \begin{bmatrix} n \\ \mathbf{j} \end{bmatrix} p_1^{j_1} \cdots p_V^{j_V} (l_{j_1}^m + \cdots + l_{j_V}^m) \quad n > b, \tag{3.1b}$$

where $l_n^0 = n$.

Proof: The proof is by induction. Let $m = 1$. Then differentiating (2.6b) at $z = 1$, by (2.3) we find, after some algebra,

$$l_n^1 = n + \sum_{j_\Sigma = n} \binom{n}{\mathbf{j}} p_1^{j_1}, \ldots, p_V^{j_V} [l_{j_1}^1 + \cdots + l_{j_V}^1] \; n > b,$$

which is of the form (3.1b) for $m = 1$ with $l_n^0 = n$. Let (3.1b) holds for all $k < m$. Differentiating (2.6b) m times we find after some algebra,

$$\begin{aligned} H_n^{(m)}(z) = m \sum_{j_\Sigma = n} \binom{n}{\mathbf{j}} p_1^{j_1} \cdots p_V^{j_V} [H_{j_1}^{(m-1)}(z) + \cdots + H_{j_V}^{(m-1)}(z)] + \\ z \sum_{j_\Sigma = n} \binom{n}{\mathbf{j}} p_1^{j_1} \cdots p_V^{j_V} [H_{j_1}^{(m)}(z) + \cdots + H_{j_V}^{(m)}(z)]. \end{aligned} \tag{3.2}$$

Assume now in (3.2) $z = 1$. Then the second term of (3.2) is equal to the second term of (3.1b). On the other hand, using the induction assumption for $k = m-1$ we find that the first term of (3.2) for $z = 1$ is equal to

$$m \sum_{j_\Sigma = n} \binom{n}{\mathbf{j}} p_1^{j_1} \cdots p_V^{j_V} [l_{j_1}^{m-1} + \cdots + l_{j_V}^{m-1}] = m\, [l_n^{m-1} + (m-1)! \sum_{i=1}^{m-1} (-1)^{m-i} \frac{l_n^{i-1}}{(i-1)!}],$$

and after some manipulation we prove that the above is equivalent to the first term in (3.1b).

□

By Theorem 1 computation of l_n^m is equivalent to a solution of the following recurrence

$$\begin{aligned} & l_n^m = 0 \quad n \le b \\ & l_n^m = a_n^{(m)} + \sum_{j_\Sigma = n} \binom{n}{\mathbf{j}} p_1^{j_1} \cdots p_V^{j_V} (l_{j_1}^m + \cdots + l_{j_V}^m), \end{aligned} \tag{3.3}$$

where

$$a_n^{(m)} = m! \sum_{i=1}^{m} (-1)^{m-i} \frac{l_n^{i-1}}{(i-1)!}. \tag{3.4}$$

Note that $a_n^{(m)}$ does *not* depend upon l_n^m, but it is a function of l_n^i for $i = 0, 1, \ldots, m-1$. Hence, to solve (3.3) we must first find l_n^1 ($a_n^{(1)} = n$), then using it we compute l_n^2, and so on. In that sense, (3.4) is a system of recurrences. To solve it we need a solution of the recurrence of type (3.3) with a general (additive) term $a_n^{(m)}$. This is done in the next subsection.

Solution of a recurrence equation

Let us consider a generalization of the recurrence (3.3), namely:

$$\text{given: } x_0, x_1, \ldots, x_b,$$
$$\text{solve: } x_n = a_n + \sum_{j_\Sigma = n} \binom{n}{\mathbf{j}} p_1^{j_1} \cdots p_V^{j_V}(x_{j_1} + \cdots + x_{j_V}), \quad n > b, \tag{3.5}$$

where a_n is a given sequence , and $\sum_{i=1}^{V} p_i = 1$. This type of recurrence was studied in [13], [14] for $b = 1$. Here, we generalize the solution for $b > 1$. Let $X(z)$ be the exponential generating function for x_n, i.e. $X(z) = \sum_{n=0}^{\infty} x_n \frac{z^n}{n!}$. Then, multiplying (3.5) by $z^n/n!$ and summing from $n = 0$ to infinity we obtain:

$$X(z) - \sum_{i=1}^{V} X(p_i z)e^{(1-p_i)z} = A(z) - \sum_{i=0}^{b} z^i \beta_i, \tag{3.6}$$

where

$$\beta_i = -\{x_i - a_i - \sum_{j_\Sigma = i} \binom{i}{\mathbf{j}} p_1^{j_V} \cdots p_V^{j_V} \sum_{k=1}^{V} L_{j_k}\}/i!, \quad i = 1,2, \ldots, b. \tag{3.7}$$

and $A(z)$ is the exponential generating function for a_n. Let now $\phi(z) = X(z)e^{-z}$. Then (3.6) becomes

$$\phi(z) - \sum_{i=1}^{V} \phi(p_i z) = A(z)e^{-z} - e^{-z} \sum_{i=1}^{b} z^i \beta_i. \tag{3.8}$$

To solve (3.8) we need coefficients of the Taylor expansion of $A(z)e^{-z}$. Let for a sequence a_n its *binomial inverse relationship* $\hat{a}_n$ be defined as below [15], [13],

$$\hat{a}_n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k a_k. \tag{3.9a}$$

Note that (3.9a) implies [15], [12],

$$\hat{x}_n = \frac{\hat{a}_n - \sum_{r=0}^{b} (-1)^r \begin{bmatrix} k \\ r \end{bmatrix} a_r}{1 - \sum_{i=1}^{V} p_i^k}. \tag{3.14}$$

Proof: Eq. (3.13) follows immediately from (3.12) if one notices that under the above assumption and (3.7) $\beta_r = a_r / r!$. Eq. (3.14) is a consequence of (3.9b) and (3.13).

□

Remark. Recurrence (3.5) is of its own interest, and it finds many applications in analysis of algorithms, e.g., in radix exchange sort [8], [12], in the performance evaluation of conflict resolution algorithms [4], [13], [14], and so on (see also [6] and [8]).

Solution of the system of recurrence (3.1)

We now use (3.13) and (3.14) to solve (3.1) or equivalently (3.3). Assume first $m = 1$. Then, by (3.4) $a_n^{(1)} = n$, and $\hat{a}_n^{(1)} = \delta_{n1}$ [15]. Hence, by (3.13) and (3.14)

$$l_n^1 = \sum_{k=2}^{n} (-1)^k \begin{bmatrix} n \\ k \end{bmatrix} \frac{\sum_{r=1}^{b} (-1)^{r+1} \begin{bmatrix} k \\ r \end{bmatrix} r}{1 - \sum_{i=1}^{V} p_i^k}, \tag{3.15a}$$

and

$$\hat{l}_n^1 = \sum_{r=1}^{b} (-1)^{r+1} \begin{bmatrix} k \\ r \end{bmatrix} \frac{r}{1 - \sum_{i=1}^{V} p_i^k}. \tag{3.15b}$$

Let now $m = 2$ Then by (3.4) $a_n^{(2)} = 2[\, l_n^1 - n\,]$, and $\hat{a}_n^{(2)} = 2\hat{l}_n^1 - 2\delta_{n,1}$ [15]. Using (3.13) and (3.15b) we immediately find, after some algebra, that

$$l_n^2 = 2 \sum_{r=1}^{b} (-1)^{r+1}\, r \sum_{k=2}^{n} (-1)^n \begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ r \end{bmatrix} \frac{\sum_{i=1}^{V} p_i^k}{(1 - \sum_{i=1}^{V} p_i^k)^2}. \tag{3.16a}$$

and

$$a_n = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} (-1)^k \hat{a}_k, \tag{3.9b}$$

(i.e. $\hat{\hat{a}}_n = a_n$), and the exponential generating function $\hat{A}(z)$ of $\hat{a}_n$ is given by

$$\hat{A}(-z) = A(z)e^{-z}. \tag{3.10}$$

Using this and (3.8) with ϕ_k denoting the coefficient of $\phi(z)$ at z^k in its Taylor expansion, we finally obtain

$$\phi_k = (-1)^k \frac{\hat{a}_k - \sum_{r=0}^{\min\{b,k\}} (-1)^r \begin{bmatrix} k \\ r \end{bmatrix} r!\beta_r}{k!(1 - \sum_{i=1}^{V} p_i^k)}, \quad k \neq 1. \tag{3.11}$$

Finally, noting that $x_n = n! \sum_{k=0}^{n} \phi_k/(n-k)!$ we prove that

LEMMA 3. For any n, the recurrence (3.5) possesses the following solution

$$x_n = x_0 + n(x_1 - x_0) + \sum_{k=2}^{n} (-1)^k \begin{bmatrix} n \\ k \end{bmatrix} \frac{\hat{a}_k - \sum_{r=0}^{\min\{b,k\}} (-1)^r \begin{bmatrix} k \\ r \end{bmatrix} r!\beta_r}{1 - \sum_{i=1}^{V} p_i^k}. \tag{3.12}$$

□

Solution (3.12) simplifies if $x_0 = x_1 = \cdots = x_b = 0$. In this case, in fact, we are also able to compute $\hat{x}_n$ which is necessary for the solution of our problem (3.1). We prove that

COROLLARY 1. If $x_0 = x_1 = \cdots = x_b = 0$, then

$$x_n = \sum_{k=2}^{n} (-1)^k \begin{bmatrix} n \\ k \end{bmatrix} \frac{\hat{a}_k - \sum_{r=0}^{b} (-1)^r \begin{bmatrix} k \\ r \end{bmatrix} a_r}{1 - \sum_{i=1}^{V} p_i^k}, \tag{3.13}$$

and

$$l_n^2 = 2 \sum_{r=1}^{b} (-1)^{r+1} \binom{k}{r} r \frac{\sum_{i=1}^{V} p_i^k}{(1-\sum_{i=1}^{V} p_i^k)^2} \tag{3.16b}$$

To compute l_n^3 we need l_n^1 as well as l_n^2 (Eqs. (3.15b), (3.16b)), and so on. Generalizing, we obtain:

PROPOSITION 1. For all n the m-th factorial moments, l_n^m, of the external path length in a b-trie is given by

$$l_n^m = m! \sum_{r=1}^{b} (-1)^{r+1}\, r \sum_{k=2}^{n} (-1)^k \binom{n}{k} \binom{k}{r} \frac{[\sum_{i=1}^{V} p_i^k]^{m-1}}{[1 - \sum_{i=1}^{V} p_k^k]^m}, \quad n > b, \tag{3.17}$$

and $l_n^m = 0$ for $n \le b$.

Proof: The proof uses induction and the same type of reasoning as above for (3.15) and (3.16). Note that $\hat{a}_n^{(m)} = m! \sum_{i=1}^{m} (-1)^{m-1}\, l_n^{i-1}/(i-1)!$ and l_n^{i-1}, $i = 1, \ldots, m$ are available by induction assumption and (3.14). The details are left to the reader.

□

Remarks

(i) For symmetric case formula (3.17) is simple. Assuming $p_1 = p_2 = \cdots = p_V = 1/V$ in (3.17) we obtain

$$l_n^m = m! \sum_{r=1}^{b} (-1)^{r+1}\, r \sum_{k=2}^{n} (-1)^n \binom{n}{k} \binom{k}{r} \frac{V^{(1-k)(m-1)}}{(1 - V^{1-k})^m}, \tag{3.18}$$

(ii) Using (3.12) we might obtain a closed form for the generating function $H_n(z)$. Indeed, it is well known that $H_n(z)$ is completely described around $z = 1$ by all its (finite) factorial moments, and

$$H_n(z) = \sum_{m=0}^{\infty} \frac{l_n^m}{m!} (z-1)^m .$$

Let $P(k) = \sum_{i=1}^{V} p_i^k$. Then, using the above and (3.17) we easily find,

$$H_n(z) = \sum_{r=1}^{b} (-1)^{r+1} r \sum_{k=2}^{n} (-1)^n \binom{n}{k} \binom{k}{r} \frac{1}{P(k)} \frac{1-P(k)}{1-zP(k)}, \tag{3.19}$$

and (3.19) is valid for z closed to 1.

4. ASYMPTOTIC APPROXIMATIONS

To find asymptotic approximation for l_n^m we transform (3.17) into a more suitable form. Applying Newton's formula to $\left[\sum_{i=1}^{V} p_i^k \right]^{m-1}$ in (3.17) we obtain

$$l_n^m = m! \sum_{j_\Sigma = m-1} \binom{m-1}{\mathrm{j}} \sum_{r=1}^{b} (-1)^{r+1} r \sum_{k=2}^{n} (-1)^k \binom{n}{k} \binom{k}{r} \frac{\alpha^k}{(1-\sum_{i=1}^{V} p_i^k)^m}, \tag{4.1a}$$

where

$$\alpha = \prod_{s=1}^{V} p_s^{j_s}, \tag{4.1b}$$

and $\mathrm{j} = (j_1, \ldots, j_V)$ such that $j_1 + j_2 + \cdots + j_V = m-1$. The most difficult to compute is the inner sum in (4.1). Let us denote the sum by ($S(n,r)$, that is,

$$S(n,r) \overset{def}{=} \sum_{k=2}^{n} (-1)^k \binom{n}{k} \binom{k}{r} \frac{\alpha^k}{(1-\sum_{i=1}^{V} p_k^k)^m}, \tag{4.2}$$

where $r = 1,2, \ldots, b$. To evaluate (4.2) we may use either Rice's method [6], [7] or Mellin transform technique [6], [7], [9], [13], [16]. We apply here the latter method. It turns out that $S(n+r,r)$ is easier to evaluate than $S(n,r)$. In Appendix A we prove that $S(n,r) = T(n,r) + O(1)$ where

$$T(n+r),r) = (-1)^r \frac{n+r}{r!} \alpha[1+O(n^{-1})] \int_{(½ - [2-r]^+)} \frac{\Gamma(z)(n\alpha)^{r-1-z}}{(1 - \sum_{i=1}^{V} p_i^{r-z})^m} dz, \qquad (4.3)$$

and $\Gamma(z)$ is the gamma function [16], $a^+ = \min\{0,a\}$ and the integral notation $\int_{(c)} f(\cdot)$ stands for $\frac{1}{2\pi_i} \int_{c-i\infty}^{c+i\infty} f(\cdot)$. Formula (4.3) holds also for $r=0$. The line of integration is either $(-3/2 - i\infty, -3/2 + i\infty)$ for $r=0$, or $(-1/2 - i\infty, -1/2 + i\infty)$ for $r=1$ or $(1/2 - i\infty, 1/2 + i\infty)$ for $r>1$. Using (4.11) and(4.3) we also find that

$$l_n^m = -n \sum_{r=1}^{b} \frac{1}{(r-1)!} \int_{(½ - [2-r]^+)} \frac{\Gamma(z)(n-r)^{-z} \left[\sum_{i=1}^{V} p_i^{r-z} \right]^{m-1}}{\left[1 - \sum_{i=1}^{V} p_i^{r-z} \right]^m} dz + O(1), \qquad (4.4)$$

however, (4.3) is more suitable for the evaluation of the integral.

The evaluation of the counter integral in (4.3) is routine: one goes from $(c, -iN_1)$ to (c, iN_1) to (N_2, iN_1) to $(N_2, -iN_1)$ to $(c, -iN_1)$ in a negative sense, where $c = ½ - [2-r]^+$. For $N_1 \to \infty$ the horizontal parts of the integral vanish since $\Gamma(t + iN_1) = O(1 + iN_1|^{t-½}e^{-t - \pi N_1/2})$ [16], while the vertical component over $(N_2, -iN_1)$ decays due to the factor n^{r-1-z} [12], [16]. Hence the required integral is minus the sum of residuals of the function under the integral to the right of the vertical line fixed at point $c = ½ - [2-r]]^+$. Note also that the function under the integral is analytical except the roots of the denominator and possible poles of the gamma function (only for $r=1$ in our case). Let us denote by z_k^r $r = 1,2, \ldots, b$, $k = 0,\pm1, \ldots,$ the roots of the following equation

$$1 - \sum_{i=1}^{V} p_i^{r-z} = 0. \qquad (4.5)$$

Naturally, z_k^r are roots of the denominator of multiplicity m. The equation (4.5) is difficult to study, however, the following properties might be established (see also [12], [4]).

(i) There exists only one real solution of (4.5), namely $z_0^r = r-1$. To prove this claim note that for x real the function $f(x) = \sum_{i=1}^{V} p_i^{r-x} - 1$ is an increasing function of x in $(-\infty, +\infty)$ and $f(-\infty) = -1, f(+\infty) = \infty$. Since $f(x)$ is continuous, hence the claim is proved.

(ii) In the complex domain there are infinite numbers of roots of type $z_k^r = x_n^r + iy_k^r$, $x_k^r \geq r-1$. These roots are uniformly discrete in the sense that there exists such δ that two distinct roots with parameter k and k' are separated by δ, that is, $|z_{k'}^r - z_k^r| > \delta$ [4].

Let $g_r(z)$ denote the function under the integral, that is,

$$g_r(z) = \frac{\Gamma(z)(n\alpha)^{r-1-z}}{(1 - \sum_{i=1}^{V} p_i^{r-z})^m}. \tag{4.6}$$

Then, for $k \neq 0$ the roots z_k^r are poles of order m for $g_r(z)$, and for $r=1$, $z_0^1 = 0$ is the pole of order $m+1$, since zero is an additional singularity of the gamma function. Hence, by the Cauchy's theorem [9], [16] the integral is the sum of residues of $g_r(z)$, that is

$$\int_{\frac{1}{2} - [2-r]^+} g_r(z)dz = -res\ g_r(r-1) - \sum_{z_k^r \neq z_0^r} res\ g_r(z_k^r). \tag{4.7}$$

It turns out that the second sum in (4.7) is a fluctuating function with a very small amplitude [5], [6], [7], [12], [14], so the leading factor is represented by the first term in (4.7).

The most difficult to compute is the residue at $z=0$ for $r=1$. We use Taylor expansion of the functions involved in $g_1(z)$ to obtain $res\ g_1(0)$. Let γ_k, e_k and b_k, $k = -1, 0, 1, \ldots,$ be coefficients of the Taylor expansion of these functions around $z=0$, that is,

$$\Gamma(z) = \gamma_{-1}\, z^{-1} + \gamma_0 + \gamma_1\, z + \cdots + \gamma_{m-1}\, z^{m-1} + O(z^m), \tag{4.8a}$$

$$(n\alpha)^{-z} = e_0 + e_1 z + \cdots + e_m\, z^m + O(z^{m+1}), \tag{4.8b}$$

$$(1 - \sum_{i=1}^{V} p_i^{1-z})^m = z^m\, [\, b_0 + b_1 z + \cdots + b_m z^m + O(z^{m+1})\,]\,. \tag{4.8c}$$

Then, the following algorithm is used to compute $res\;\, g_1(0)$ [9]:

Step 1. For $n = -1, 0\,, \ldots, m-1$ compute

$$a_n = \sum_{k=-1}^{n} \gamma_k\, e_{n-k}\,. \tag{4.9a}$$

Step 2. Let $c_{-1} = a_{-1}/b_0$, then recursively for $n = 0, 1\,, \ldots, m-1$

$$c_n = \frac{1}{b_0} \{ a_n - \sum_{k=-1}^{n-1} c_k\, b_{n-k} \}. \tag{4.9b}$$

Step 3.

$$res\;\, g_1(0) = c_{m-1}. \tag{4.9c}$$

□

The coefficients in (4.8a,b) are easy to compute. It is well known that [2]

$$\gamma_{-1} = 1\,; \quad \gamma_0 = -\gamma\,; \quad \gamma_1 = \frac{1}{2}\,[\frac{\pi^2}{6} + \gamma]; \quad \cdots \tag{4.10}$$

where $\gamma = 0.5772$ is the Euler constant, and

$$e_k = \frac{(-1)^k}{k!}\, \ln^k n\alpha. \tag{4.11}$$

Using the Leibnitz formula, we prove also that

$$b_k = \frac{1}{(m+k)!} (-1)^m \sum_{j_\Sigma = m+k} \begin{bmatrix} m+k \\ \mathbf{j} \end{bmatrix} h_{j_1}\, h_{j_2} \cdots h_{j_m}, \tag{4,12a}$$

where $h_0 = 0$ and

$$h_n = (-1)^n \sum_{i=1}^{V} p_i \, ln^n \, p_i, \quad n > 0. \tag{4.12b}$$

In particular, we find that

$$\begin{array}{llll} m = 1 & b_0 = -h_1 & b_1 = -h_2/2 & \\ m = 2 & b_0 = h_1^2 & b_1 = h_1 h_2 & b_2 = \frac{1}{4} h_2^2 + \frac{1}{3} h_1 h_3 \\ m = 3 & b_0 = -h_1^3 & b_1 = -\frac{3}{2} h_1^2 h_2 & b_2 = -\frac{3}{4} h_2^2 \, h_1 - \frac{1}{2} h_1^2 h_3 \end{array} \tag{4.13}$$

The algorithm (4.9) is also used to compute the other residues subject to the following changes:

(i) replace z by $w = z - z_k^r$,

(ii) set $\gamma_{-1} = 0$,

(iii) start step 2 with $c_0 = a_0/b_0$.

Naturally, the coefficients in (4.8) depend now upon z_k^r.

Using the above approach, we may compute the exact asymptotic approximation for l_n^m with accuracy $O(1)$, and d_n^m with accuracy $O(n^{-1})$ (see Lemma 1). Note that by (4.1a) and (4.7)

$$l_n^m = m! \, n \sum_{j_\Sigma = m-1} \binom{m-1}{j} \alpha \sum_{r=1}^{b} [res \;\; g_r(r-1)/(r-1)! + \sum_{z_k^r \neq z_0^r} res \;\; g_r(z_k^r)/(r-1)!] + O(1). \tag{4.14}$$

Let us denote

$$f(n) = m! \sum_{j_\Sigma = m-1} \binom{m-1}{j} \alpha \sum_{z_k^r \neq z_0^r} res \;\; g_r(z_k^r)/(r-1)!. \tag{4.15}$$

Then, we prove

THEOREM 2. The m-th factorial moment of the depth of insertion d_n^m is a polynomial of degree m in $\ln n$. More precisely,

$$d_n^m = \sum_{k=0}^{m} \xi_k \, ln^k n + f(n) + O(n^{-1}), \tag{4.16}$$

where ξ_k are constants which do not depend on n.

Proof: By (2.5) d_n^m is equal to l_n^m/n, and l_n^m is obtained through Eq. (4.14). By (4.9c) the residues $res\ g_r(r-1)$ are equal to c_{m-1} . But c_{m-1} depends on $\ln n$ only through the coefficients e_k, $k = 0, 1, \ldots, m-1$ for $r>1$, and in addition for $r=1$ e_m appears in c_{m-1}. Hence, for $r=1$ $a_{m-1} = \frac{(-1)^m}{m!} ln^m n\alpha + O(ln^{m-1} n\alpha)$ and for other $r>1$ $a_{m-1} = O(ln^{m-1} n\alpha)$. By (4.9b) c_{m-1} is a linear combination of a_k, $k=-1, 0, \ldots, m$, therefore c_{m-1} is a polynomial in $\ln n$, and by the above the coefficient ξ_m in (4.16) is not equal to zero, hence from this and (4.14), (4.15) the theorem follows.

□

The computation of the coefficients ξ_k in (4.16) are rather troublesome, however, we might prove that

PROPOSITION 2. For any m, and n large enough

$$d_n^m = \frac{1}{h_1^m} ln^m n + \frac{m}{h_1^m} ln^{m-1} n [\gamma + \frac{m}{2}\frac{h_2}{h_1} - (m-1)h_1 - H_{b-1} - h_1^m F(n)] + O(ln^{m-2} n) \tag{4.17}$$

where H_{b-1} is the $(b-1)$-st harmonic number [12], and

$$F(n) = \sum_{r=1}^{b} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{\Gamma(z_k^r) n^{r-1-z_k^r}}{(r-1)^1 \, h_1^m(z_k^r)}, \tag{4.18}$$

with $h_1(z_k^r)$ being a coefficient defined as (see (4.12))

$$h_1(z_k^r) = -\sum_{i=1}^{V} p_i^{r-zk} ln\, p_i . \tag{4.19}$$

Proof: By (4.16) we must compute ξ_m and ξ_{m-1}. Note that according to (4.12)

$b_0 = (-1)^m h_1^m$ and $b_1 = (-1)^m \frac{m}{2} h_1^{m-1} h_2$. Then, (4.9) implies for $r=1$

$$res\ \ g_1(0) = c_{m-1} = \frac{a_{m-1}}{b_0} - \frac{a_{m-2}b_1}{b_0} + O(ln^{m-2}n),$$

and after some algebra one finds

$$res\ \ g_1(0) = \frac{1}{m!}\,\frac{1}{h_1^m}\, ln^m n\alpha + \frac{1}{(m-1)!h_1^m}\, ln^{m-1} n\alpha[\gamma + \frac{m}{2}\,\frac{h_2}{h_1}] + O(ln^{m-2} n\alpha).$$

On the other hand, for $r>1$ we find

$$res\ \ g_r(z_0^r) = \frac{a_{m-1}}{b_0} + O(ln^{m-2} n\alpha) = \frac{-(r-2)!}{(m-1)!h_1^m}\ \ln^{m-1} n\alpha + O(ln^{m-2} n\alpha),$$

and for $k \neq 0$

$$res\ \ g_r(z_k^r) = \frac{1}{(m-1)!}\ \frac{\Gamma(z_k^r)(n\alpha)^{r-1-z_k^r}}{h_1^m(z_k^r)}\ \ln^{m-1} n\alpha + O(ln^{m-2} n\alpha).$$

Finally, using the above and (4.14) we obtain (4.17).

□

Two moments play usually an important role in tries analysis, namely: the average and the variance, σ_n^2, of the depth of insertion. Using the above approach we obtain immediately

PROPOSITION 3. (i) The average depth of insertion is given by

$$d_n = \frac{1}{h_1}\ ln\ n + \frac{1}{h_1}[\gamma + \frac{h_2}{2h_1} - H_{b-1}] - F(n) + O(n^{-1}). \qquad (4.20)$$

(ii) The variance, σ_n^2, of the depth of insertion is

$$\sigma_n^2 = \frac{h_2 - h_1^2}{h_1^3} \ln n + C + F_1(n) + O(n^{-1}), \qquad (4.21)$$

where

$$C = \frac{1}{h_1}[\gamma(2e_{b-1} - 3) + 3H_{b-1}] + \frac{1}{h_1^2}[4\gamma_1 + 2H_{b-1}\gamma - h_1^2/4 - 2h_2 - \gamma^2 - H_{b-1}^2]$$
$$+ \frac{h_2}{h_1^3}(\gamma + H_{b-1}) + 2h_2^2/h_1^4, \qquad (4.22a)$$

$$F_1(n) = \frac{2}{h_1} F(n)[\gamma + h_2/2h_1 - H_{b-1} - h_1/2] - F^2(n) \qquad (4.22b)$$

and $e_{b-1} = \sum_{k=0}^{b-1} 1/k!$, while γ_1 is given by (4.10).

Proof: Eq. (4.20) follows directly from the previous analysis with m=1. To compute the variance note that $\sigma_n^2 = d_n^2 + d_n - (d_n)^2$. The details are left to the reader.

□

Remarks

(i) *Balance of a tree.* The variance of the depth of insertion might be considered as a measure of *how well a tree is balanced.* In the high-balanced trees the depth of a leaf is the same (or almost the same) for all leaves. Then, the variance of the depth is equal to zero. For other trees the depth is a random variable, however, the smaller the variance is, the more balanced the tree is. Indeed, by Tchebyshev inequality, we know that $Pr\{|D_n - d_n| > \delta\} < \sigma_n^2/\delta^2$. For example, let $\delta = 3\sigma$, then $Pr\{|D_n - d_n| > 3\sigma_n\} < 1/9$, and it says that with probability 0.11 the depth lies in the interval $(d_n - 3\sigma_n, d_n + 3\sigma_n)$, hence the smaller σ is, the smaller the interval is. This also means that for small σ the average of the depth of insertion is a good measure of the depth, while for larger σ, it is very poor performance issue. Let us apply this to tries. By (4.21) we see that for symmetric tries $h_2 - h_1^2 = 0$, hence $\sigma_n^2 = O(1)$ and does not depend on n. We may claim that symmetric tries are of an order of magnitude better balanced than asymmetric tries. Let $V = 2$, then for $p = 0.5$

(symmetric trie) $\sigma_n^2 \approx 3.43$, while for $p = 0.1$ $\sigma_n^2 = 12.64 \ln n + O(1)$ and for $p = 0.3$ $\sigma_n^2 = 0.66 \ln n + O(1)$. The Tchebyshev inequality implies that with probability 0.11 the depth of insertion for a symmetric trie with $V = 2$ lies in $(d_n - 5.5, d_n + 5.5)$, while with the same probability the depth is in the interval $(d_n - 10\sqrt{\ln n}, d_n + 10\sqrt{\ln n})$ for $p = 0.1$ and in $(d_n - 2.4\sqrt{\ln n}, d_n + 2.4\sqrt{\ln n})$ for $p = 0.3$ and large n.

(ii) *Optimization problems.* Let us consider l_n^m as a function of $\mathbf{p} = (p_1, p_2, \ldots, p_V)$. Then a question arises what is an optimal choice of p? It is intuitively clear that the average depth of insertion is minimized for the symmetric case. However, using (3.17) it is easy to notice that l_n^m and d_n^m are minimized for *all* n and m if the trie is a symmetrical one, that is, $p_1 = p_2 = \cdots = p_V = 1/V$. Naturally, the bigger the V is, the smaller the average depth of insertion is, however, the data structure becomes more complicated. Moreover, formula (4.20) shows that the bigger the b is, the smaller the average depth of insertion is, however, the impact of b is of the secondary importance since the leading factor in (4.20) does not depend on b.

(iii) The most difficult to evaluate in (4.20) and (4.21) is the function $F(n)$. However, it turns out that this fluctuating function has a very small amplitude. This is easy to see for a symmetric trie. Then, equation (4.5) possesses the following roots

$$z_k^r = r - 1 + 2\pi ik/ln\ V \qquad k = 0, \pm 1, \pm 2,$$

and by (4.18) $F(n)$ is $\sum_{r=1}^{b} f_r(n)/(r-1)!$, where

$$f_r(n) = \frac{2}{\ln^m V} \sum_{k=1}^{\infty} \mathrm{Re}\{\Gamma(r-1 + 2\pi ik/ln\ V)\exp[-2\pi ik\ \log_V n]\}, \qquad (4.23)$$

where Re z is a real part of z. This function (for m=1) was studied by many authors, e.g. see Knuth [12] and [7], [8] , [13], [14]. In particular, it is easy to notice that

$f_r(n)$ is a periodic function of $\log_V n$, that is, $f_r(np) = f_r(n)$. Moreover, the function is bounded, and for $m=1$ Knuth computed that for $V=2$ $f_0(n) < 1.72 \cdot 10^{-7}$, and for $V=5$ $f_0(n) < 8.5 \cdot 10^{-4}$. Therefore, in many computations $F(n)$ may be safely ignored.

APPENDIX A. *Derivation of Eq. (4.3).*

We show here how to obtain (4.3) for $r=1$. As mentioned before, we compute $S(n+1,r)$ instead of $S(n,r)$, where

$$S(n,r) = \sum_{k=2}^{n} (-1)^k \begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ r \end{bmatrix} \frac{\alpha^k}{(1 - \sum_{i=1}^{V} p_i^k)^m} \tag{A1}$$

To find a series expansion of $(1 - \sum_{i-1}^{V} p_i^k)^{-m}$ we use the following arguments. Let $a < 1$, then differentiating m-times the geometric series formula $\sum_{k=0}^{\infty} a^k = 1/(1-a)$ we obtain

$$\sum_{k=m}^{\infty} k(k-1) \cdots (k-m+1)a^{k-m} = \frac{m}{(1-a)^{m+1}}. \tag{A2}$$

Let $k^{\underline{l}} = k(k-1) \cdots (k-l+1)$. Then using (A2) in (A1) we find

$$S(n+1,1) = \frac{1}{(m-1)!} \sum_{k=2}^{n+1} (-1) \begin{bmatrix} n+1 \\ k \end{bmatrix} k \sum_{l=m-1}^{\infty} l^{\underline{m-1}} \sum_{j_\Sigma = l+m-1} \begin{bmatrix} l+m-1 \\ \mathrm{j} \end{bmatrix} \prod_{s=1}^{V} (\alpha p_s^{j_s})^k$$

Denoting $\Phi = \alpha \prod_{s=1}^{V} p_s^{j_s}$ and using the identity [15]

$$\sum_{k=1}^{n+1} (-1)^k \begin{bmatrix} n+1 \\ k \end{bmatrix} k \, \Phi^k = -(n+1)\Phi[(1-\Phi)^n - 1]$$

we obtain

$$S(n+1,1) = -\frac{n+1}{(m-1)!}\sum_{l=m-1}^{\infty} l^{\underline{m-1}} \sum_{j_\Sigma = l+m-1} \binom{l+m-1}{\mathrm{j}} \Phi\,[(1-\Phi)^n - 1]$$

Define now a new variable $x = n\Phi$, and note that $(1-\Phi)^n = (1-\frac{x}{n})^n = e^{-x} + x^2 O(n^{-1})$.

We approximate $S(n,r)$ by $T(n,r)$, where

$$T(n+1,1) = -\frac{n+1}{(n-1)!}\sum_{l=m-1}^{\infty} l^{\underline{m-1}} \sum_{j_\Sigma = l+m-1} \binom{l+m-1}{\mathrm{j}} \Phi\,(e^{-x} - 1). \tag{A3}$$

Using Mellin transform [2], [7], [9], [16] we represent $e^{-x} - 1$ as

$$e^{-x} - 1 = \frac{1}{2\pi i}\int_{-\frac{1}{2}+i\infty}^{-\frac{1}{2}+i\infty} \Gamma(z)x^{-z}\,dz = \int_{(-\frac{1}{2})} \Gamma(x)x^{-z}\,dz \tag{A4}$$

where $\int_{(c)}$ stands for $\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}$. By (A3), (A4) and definition of Φ, after some algebra, we finally obtain,

$$T(n+1,1) = -(n+1)\alpha \int_{(-\frac{1}{2})} \frac{\Gamma(z)(n\alpha)^{-z}}{(1-\sum_{i=1}^{V} p_i^{1-z})^m}\,dz. \tag{A5}$$

Using the above approach we prove also that for $r > 1$ (see also [13], [14])

$$T(n+r,r) = (-1)^r \binom{n+r}{r} \int_{(\frac{1}{2})} \frac{\Gamma(z)n^{-z}\alpha^{r-z}}{(1-\sum_{i=1}^{r} p_i^{r-z})^m}\,dz \tag{A6}$$

and for $r = 0$

$$T(n,0) = \int_{(-3/2)} \frac{\Gamma(z)(n\alpha)^{-z}\,dz}{(1-\sum_{i=1}^{V} p_i^{-z})^m}. \tag{A7}$$

To justify our approximation we must find a relationship between $T(n,r)$ and $S(n,r)$. But, using the idea from [14], we immediately prove that

$$T(n,r) = S(n,r) + O(1) \tag{A8}$$

Taking into account (A5)-(A8) we finally obtain (4.3).

REFERENCES

[1] Aho, A., Hopcroft, J. and Ullman, J., Data structures and algorithms, Addison-Wesley, 1983.

[2] Frdelyi, A., Higher transcendental functions, McGraw-Hill Book, 1953.

[3] Fagin R., Nievergelt, J., Pippenger, N. and Strong H., Extendible hashing: A fast access method for dynamic files, *ACM TODS* 4, 1979, 315-344.

[4] Fayolle, Ph., Flajolet, Ph, Hofri, M. and Jacquet, Ph., Analysis of a stack algorithm for random multiple-access communication, *IEEE Trans. on Information Theory*, **IT-31, 2,** 1985, 244-254.

[5] Flajolet, Ph., On the performance evaluation of extendible hashing and trie searching, *Acta Informatica* 20, 1983, 345-369.

[6] Flajolet, Ph. and Sedgewick R., Digital search trees revisited, INRIA Technical Report, 311, 1984, to appear in SIAM *J. Compt.*

[7] Flajolet, Ph., Mathematical methods in the analysis of algorithms and data structures, INRIA Technical Report, 400, 1985.

[8] Gonnet G., Handbook of algorithms and data structures, Addison-Wesley, 1984.

[9] Henrici, P., Applied and computational complex analysis, John Wiley & Sons, New York 1977.

[10] Jacquet, Ph. and Regnier, M., Limiting distributions for trie parameters, INRIA Technical Report, 1985.

[11] Kirschenhofer, P. and Prodinger, H., Some further results on digital trees, *ICALP 86*,

to appear.

[12] Knuth, D., The art of computer programming, Addison-Wesley, 1973.

[13] Szpankowski, W., Analysis of a recurrence equation arising in stack-type algorithms for collision-detecting channels, Proc. Intern. seminar on Computer Networking and Performance Evaluation, Tokyo 1985, 9-3-1-9-3-12.

[14] Szpankowski, W., On a recurrence equation arising in the analysis of conflict resolution algorithms, Purdue University, CDS-TR-556, 1985.

[15] Riordan, J., Combinatorial identities, John Wiley & Sons, 1968.

[16] Whittaker, E. and Watson, G., A course of modern analysis, Cambridge Press, 1935.