Acta Informatica 21, 229–237 (1984)

Acta Informatica

A Probabilistic Analysis of the Height of Tries and of the Complexity of Triesort*

Luc Devroye

School of Computer Science, McGill University, 805 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6

Summary. We consider binary tries formed by using the binary fractional expansions of $X_1, \ldots, X_n$, a sequence of independent random variables with common density f on $[0,1]$. For H_n, the height of the trie, we show that either $E(H_n) \sim 2\log_2 n$ or $E(H_n) = \infty$ for all $n \geqq 2$ according to whether $\int f^2(x)\,dx$ is finite or infinite. Thus, the average height is asymptotically twice the average depth (which is $\sim \log_2 n$ when $\int f^2(x)\,dx < \infty$). The asymptotic distribution of H_n is derived as well.

If f is square integrable, then the average number of bit comparisons in triesort is $n\log_2 n + 0(n)$, and the average number of nodes in the trie is $0(n)$.

1. Introduction

Tries were introduced by Fredkin in 1960. In its simplest (binary) form, a trie is a binary tree used to store data $X_1, \ldots, X_n$ in the following manner: each X_i, considered as a countable string of 0's and 1's, defines an infinite path in the binary tree ("0" indicates a left turn, "1" a right turn); the trie defined by $X_1, \ldots, X_n$ is the smallest binary tree T for which the paths truncated at the leaves of T are all pairwise different. The X_i's are then associated with the leaves of T.

In a trie, the following quantities are of interest:

1) D_{ni}: the depth of X_i (distance from the root to the leaf corresponding to X_i in the trie formed by $X_1, \ldots, X_n$);

2) A_n: $n^{-1} \sum_{i=1}^{n} D_{ni}$, the *average depth*;

3) H_n: $\max_{1 \leqq i \leqq n} D_{ni}$, the *height*.

The distribution of D_{ni}, A_n and H_n depends upon the distribution of $X_1, \ldots, X_n$. We assume throughout that the X_i's are *independent identically distributed*

* Research of the author was supported in part by FCAC Grant EQ-1678

random variables with *density* f on $[0,1]$. Clearly, the countable string of 0's and 1's that we need for X_i is the binary fractional expansion of X_i. Under this assumption, the following is known:

Theorem 1 (Devroye, 1982). *Either*

$$E(A_n)=\infty \qquad \text{for all } n \geqq 2$$

or

$$\lim_{n\to\infty} E(A_n)/\log_2 n = 1$$

according to whether $\int f^2(x)\,dx=\infty$ *or* $\int f^2(x)\,dx<\infty$.

Theorem 2 (Yao, 1980). *If f is the uniform density on* $[0,1]$, *then there exist constants* c_1, c_2 *such that*

$$0<c_1 \leqq E(H_n)/\log_2 n \leqq c_2 < \infty.$$

Strictly speaking, Yao (1980) showed Theorem 2 only when the number of X_i's is N, a Poisson random variable with mean n, but the "de-Poissonization" step is simple. Regnier (1982) improved Theorem 2, also under the Poisson model and for f uniform, and showed that $E(H_n)\sim 2\log_2 n$. Flajolet and Steyaert (1982) considered our model with f uniform, and obtained a few terms in the asymptotic expansion of $E(H_n)-2\log_2 n$.

Theorem 1 implies that only one of two possible situations can occur: either tries are asymptotically optimal (i.e., $E(A_n)/\log_2 n\to 1$ as $n\to\infty$) or they are disastrous (i.e., $E(A_n)=\infty$ for all $n\geqq 2$), according to whether the density f is in L_2 or not. Implicitly, Theorem 1 characterizes the L_2 densities: f is in L_2 if and only if the expected length of the largest common left substring of X_1 and X_2 is finite. Yao's result about $E(H_n)$ for the uniform density is extendible to all densities in L_2, as we will see below. In fact, we will show that for all densities in L_2, $E(H_n)\sim 2\log_2 n$; in other words, the average height is approximately twice the average depth. The machinery used to obtain this result (a combination of a Poissonization argument and the Lebesgue density theorem) is strong enough to allow us to obtain much finer results such as the asymptotic distribution of H_n. All of these results are now stated.

Theorem 3 [Asymptotic distribution of H_n]. *If* $\int f^2(x)\,dx<\infty$ *and* $\alpha = n^2\int f^2(x)\,dx/2$, *then*

$$\lim_{n\to\infty} |P(H_n \leqq \log_2\alpha + x) - \exp(-\alpha/2^{\mathrm{int}(\log_2\alpha+x)})| = 0, \qquad \text{all } x\in R.$$

(Here int(.) denotes the integer part of (.).)

Theorem 4 [Expected height]. *Let* $\int f^2(x)\,dx<\infty$, *and let* γ *be Euler's constant. Then*

$$-1 \leqq \liminf_{n\to\infty} E(H_n)-(\ln\alpha+\gamma)/\ln 2 \leqq \limsup_{n\to\infty} E(H_n)-(\ln\alpha+\gamma)/\ln 2 \leqq 1.$$

If $\int f^2(x)\,dx=\infty$, *then* $E(H_n)=\infty$ *for all* $n\geqq 2$.

Theorems 3 *and* 4 *qualify how close* H_n is to $2\log_2 n$. In Theorem 3, we show that the distribution of $H_n - 2\log_2 n - \log_2(\frac{1}{2}\int f^2(x)\,dx)$ is close to a suitably discretized version of the extreme-value distribution $\exp(-\exp(-x))$ (Johnson and Kotz, 1970, pp. 272-295). One of the corollaries of Theorem 4 is that

$$E(H_n) \sim 2\log_2 n \tag{1}$$

for all f in L_2. The integral of f^2 influences the values of H_n only in the constant term.

As a by-product of some of the Lemmas proved in Section 2, we will analyze the complexity of triesort for all densities f on $[0, 1]$ in Section 3. We use the terminology "trie search" for searching for an element in a trie, and "triesort" for sorting by first constructing a trie and then traversing the trie in preorder. Other terms have been used in the literature such as digital tree search and radix sort.

2. Proofs

Lemma 1 [A density theorem]. *Let f be a nonnegative integrable function on* $[0, 1]$, *and let A_{ni} be the set of all x in* $\left[\frac{i-1}{n}, \frac{i}{n}\right)$, $1 \leqq i \leqq n$. *Then,*

$$\lim_{n\to\infty} n \sum_{i=1}^{n} \Big(\int_{A_{ni}} f(x)\,dx\Big)^2 = \int_0^1 f^2(x)\,dx.$$

Proof. By Jensen's inequality,

$$n \sum_{i=1}^{n} \Big(\int_{A_{ni}} f(x)\,dx\Big)^2 \leqq \sum_{i=1}^{n} \int_{A_{ni}} f^2(x)\,dx = \int_0^1 f^2(x)\,dx. \tag{2}$$

For the lower bound corresponding to (2), we will need the Lebesgue density theorem (Wheeden and Zygmund, 1977) in the following form: let

$$f_n^*(x) = \inf_{0 < r \leqq 1/n} \min\left(r^{-1}\int_x^{x+r} f(y)\,dy,\; r^{-1}\int_{x-r}^{x} f(y)\,dy\right);$$

then, if $0 \leqq f$, $\int f(x)\,dx < \infty$, $f_n^*(x) \to f(x)$ for almost all x as $n \to \infty$; in particular, $f(x) \geqq f_n^*(x)$ for almost all x. We have for almost all $x \in A_{ni}$: $\big(\int_{A_{ni}} f(x)\,dx\big)^2 \geqq f_n^{*2}(x)/n^2$, and thus, by Fatou's Lemma,

$$\begin{aligned}\liminf_{n\to\infty} n \sum_{i=1}^{n} \Big(\int_{A_{ni}} f(x)\,dx\Big)^2 &\geqq \liminf_{n\to\infty} n \sum_{i=1}^{n} n^{-1} \int_{A_{ni}} f_n^{*2}(x)\,dx \\ &= \liminf_{n\to\infty} \int_0^1 f_n^{*2}(x)\,dx \geqq \int_0^1 \liminf_{n\to\infty} f_n^{*2}(x)\,dx = \int_0^1 f^2(x)\,dx.\end{aligned}$$

Lemma 2 [Poissonization inequalities]. *Let n be an integer, and let n_1, n_2 be real numbers such that* $0 < n_1 < n < n_2 < \infty$. *Let* $A_i = [(i-1)/2^l, i/2^l)$ *where l is an* *integer and* $1 \leqq i \leqq 2^l$. *Let* $p_i = \int_{A_i} f(x)\,dx$. *If* $N(\lambda)$ *is a Poisson random variable with parameter* λ, *then*

$$P(H_n \leqq l) \leqq P(N(n_1) \geqq n) + \prod_{i=1}^{2^l} (1 + n_1 p_i)\, e^{-n_1 p_i}$$

and

$$P(H_n \leqq l) \geqq \prod_{i=1}^{2^l} (1 + n_2 p_i)\, e^{-n_2 p_i} - P(N(n_2) \leqq n).$$

Proof. For integer k, we let $B_l(k)$ be the event that none of the sets A_i, $1 \leqq i \leqq 2^l$ have more than one of the points $X_1, \ldots, X_k$. It is clear that the events $B_l(n)$ and $[H_n \leqq l]$ are equivalent. We thus have the following implications between events:

$$[H_n \leqq l] = B_l(n);$$

$$[H_n \leqq l] \subseteq [N(n_1) \geqq n] \cup B_l(N(n_1));$$

$$B_l(N(n_2)) \subseteq [H_n \leqq l] \cup [N(n_2) \leqq n].$$

Also, by a property of the Poisson distribution, for all $\lambda > 0$,

$$P(B_l(N(\lambda))) = \prod_{i=1}^{2^l} (1 + \lambda p_i)\, e^{-\lambda p_i}.$$

This concludes the proof of Lemma 2.

Lemma 3 [Exponential inequalities]. *For all* $x \geqq 0$,

$$e^{-x^2/2} \leqq (1+x)\, e^{-x} \leqq e^{-x^2/2(1+x)}.$$

Proof. Note that

$$(1+x)(\log(1+x) - x) = \left(-2 + \frac{1}{1+\xi}\right) \frac{x^2}{2} \leqq -\frac{x^2}{2}, \qquad 0 \leqq \xi \leqq x,$$

and that

$$\log(1+x) - x = -\frac{1}{(1+\xi)^2} \frac{x^2}{2} \geqq -\frac{x^2}{2}, \qquad 0 \leqq \xi \leqq x.$$

Lemma 4 [Tail of the Poisson distribution]. *Let* $\varepsilon \in (0, \frac{1}{2})$, $n_1 = n(1-\varepsilon)$, $n_2 = n(1+\varepsilon)$. *Then, if* $N(\lambda)$ *is a Poisson random variable with parameter* λ,

$$P(N(n_2) \leqq n) \leqq e^{-n\varepsilon^2/4}$$

and

$$P(N(n_1) \geqq n) \leqq e^{-n\varepsilon^2/2}.$$

Proof. If X is gamma (n) distributed, then, by inequalities for the tail of the gamma distribution (Devroye, 1981)

$$P(N(n_2) \leqq n) = P(X \geqq n_2) = P(X - n \geqq n\,\varepsilon) \leqq e^{-n\varepsilon^2(1-\varepsilon)/2}$$

and

$$P(N(n_1) \geqq n) = P(X \leqq n_1) = P(X - n \leqq -n\,\varepsilon) \leqq e^{-n\varepsilon^2/2}.$$

Proof of Theorem 3. Let ε, n_1 and n_2 be as in Lemma 4. Combining Lemmas 2, 3 and 4 gives for integer l,

$$P(H_n \leqq l) \geqq \prod_{i=1}^{2^l} \exp(-(n_2 p_i)^2/2) - \exp(-n\varepsilon^2/4)$$

$$\geqq \exp\left(-\tfrac{1}{2} n_2^2 2^{-l} \int_0^1 f^2(x)\,dx\right) - \exp(-n\varepsilon^2/4) \quad \text{(by (2))}$$

$$= \exp(-\alpha(1+\varepsilon)^2/2^l) - \exp(-n\varepsilon^2/4).$$

Now,

$$P(H_n \leqq \log_2 \alpha + x) = P(H_n \leqq \operatorname{int}(\log_2 \alpha + x))$$

$$\geqq \exp(-\alpha(1+\varepsilon)^2/2^{\operatorname{int}(\log_2 \alpha + x)}) - o(1)$$

$$\geqq \exp(-\alpha/2^{\operatorname{int}(\log_2 \alpha + x)}) \cdot \exp(-2(2\varepsilon+\varepsilon^2)/2^x) - o(1), \tag{3}$$

and the right hand side of (3) is arbitrarily close to $\exp(-\alpha/2^{\operatorname{int}(\log_2 \alpha + x)})$ by the choice of ε.

Next,

$$P(H_n \leqq l) \leqq \exp(-n\varepsilon^2/2) + \exp\left(-\sum_{i=1}^{2^l} (n_1 p_i)^2/2(1+n_1 p_i)\right)$$

$$\leqq \exp(-n\varepsilon^2/2) + \exp\left(-\sum_{i=1}^{2^l} (n_1 p_i)^2/2(1+\varepsilon)\right) \cdot \exp\left(\frac{1}{2} \sum_{i=1}^{2^l} (n_1 p_i)^2 I_{[n_1 p_i > \varepsilon]}\right)$$

where I is the indicator function of an event. By Lemma 1,

$$2^l \sum_{i=1}^{2^l} p_i^2 = \int_0^1 f^2(x)\,dx + o(1) \qquad \text{as } l \to \infty.$$

Thus,

$$P(H_n \leqq l) \leqq \exp(-n\varepsilon^2/2)$$

$$+ \exp\left(-\alpha \frac{(1-\varepsilon)^2}{1+\varepsilon} 2^{-l}(1+o(1))\right) \exp\left(\frac{1}{2} \sum_{i=1}^{2^l} (n_1 p_i)^2 I_{[n_1 p_i > \varepsilon]}\right) \tag{4}$$

We can argue as for (3), and thus make $P(H_n \leqq \log_2 \alpha + x)$ arbitrarily close to $\exp(-\alpha/2^{\operatorname{int}(\log_2 \alpha + x)})$ for all n large enough by choice of ε, if we can show that for all $\varepsilon > 0$,

$$\sum_{i=1}^{2^l} (n_1 p_i)^2 I_{[n_1 p_i > \varepsilon]} = o(1). \tag{5}$$

If f^* is the maximal function corresponding to f (Wheeden and Zygmund, 1977, pp. 105), then $n_1 p_i \leqq 2n f^*(x)/2^l$ for all $x \in A_i$. Thus,

$$\sum_{i=1}^{2^l} (n_1 p_i)^2 I_{[n_1 p_i > \varepsilon]}$$

$$\leqq \sum_{i=1}^{2^l} 2^l \int_{A_i} (2n f^*/2^l)^2 I_{[2n f^*/2^l > \varepsilon]}$$

$$= \left(\frac{4n^2}{2^l}\right) \int f^{*2} I_{[f^* > \varepsilon 2^{l-1}/n]}$$

$$= o(1) \tag{6}$$

when $l=\mathrm{int}(\log_2 \alpha + x)$. Here we used the fact that $n^2/2^l$ remains bounded, that $2^l/n \to \infty$ and that $\int f^{*2} < \infty$ when $\int f^2 < \infty$. This concludes the proof of Theorem 3.

Proof of Theorem 4. In this proof, we will repeatedly use the fact that a random variable with distribution function $\exp(-\exp(-x))$ has mean γ (Johnson and Kotz, 1970), i.e.

$$\int_0^\infty (1-e^{-e^{-x}})\,dx - \int_{-\infty}^0 e^{-e^{-x}}\,dx = \gamma.$$

From the first chain of inequalities in the Proof of Theorem 3, we see that

$$\begin{aligned}
E(H_n) &= \sum_{l=0}^{\infty} P(H_n > l) = \sum_{l=0}^{n-1} P(H_n > l) \\
&\leqq n\exp(-n\varepsilon^2/4) + \sum_{l=0}^{\infty} (1-\exp(-\alpha(1+\varepsilon)^2/2^l)) \\
&\leqq o(1) + \int_{-1}^{\infty} (1-\exp(-\alpha(1+\varepsilon)^2 \exp(-t\ln 2)))\,dt \\
&= o(1) + \int_{-\ln 2 - \ln\alpha(1+\varepsilon)^2}^{\infty} (1-e^{-e^{-u}})\frac{du}{\ln 2} \\
&= o(1) + 1 + (\gamma + \ln(\alpha(1+\varepsilon)^2)/\ln 2.
\end{aligned}$$

This shows the limit supremum half of Theorem 4, because ε can be chosen arbitrarily small.

For the other half of Theorem 4, we use Fatou's lemma and a tail estimate. We have

$$\begin{aligned}
&\liminf_{n\to\infty} E(H_n) - \log_2 \alpha \\
&= \liminf_{n\to\infty} \left[\int_0^\infty P(H_n - \log_2 \alpha > t)\,dt - \int_{-\infty}^0 P(H_n - \log_2 \alpha < t)\,dt\right] \\
&\geqq \int_0^\infty \liminf_{n\to\infty} P(H_n - \log_2\alpha > t)\,dt - \limsup_{n\to\infty} \int_{-\infty}^0 P(H_n - \log_2 \alpha < t)\,dt.
\end{aligned}$$

Now,

$$\begin{aligned}
P(H_n - \log_2 \alpha > t) &\geqq 1 - \exp(-\alpha/2^{\mathrm{int}(\log_2\alpha + t)}) + o(1) \\
&\geqq 1 - \exp(-2^{-t}) + o(1) \sim \int_0^\infty (1-e^{-e^{-u}})\,du/\ln 2.
\end{aligned}$$

Also, as we will show,

$$\begin{aligned}
&\int_{-\infty}^0 P(H_n - \log_2 \alpha < t)\,dt \leqq o(1) + \int_{-\infty}^0 \exp(-\tfrac{1}{2}2^{-t})\,dt \cdot (1+o(1)) \\
&\sim \left(\int_{-\infty}^0 e^{-e^{-u}}\,du + \int_0^{\ln 2} e^{-e^{-u}}\,du\right)\Big/ \ln 2 \\
&\leqq \left(\int_{-\infty}^0 e^{-e^{-u}}\,du + \ln 2\right)\Big/ \ln 2. \qquad (7)
\end{aligned}$$

A combination of these bounds shows that

$$\liminf_{n\to\infty}(E(H_n)-\log_2\alpha)\geqq\frac{\gamma}{\ln 2}-1.$$

We need show the first inequality in (7). Since $H_n<n$, we see that the integration interval can be taken as $[-n,0]$ without loss of generality. Let $\beta=\frac{1}{3}\log_2\alpha$. Let ε be an arbitrary positive number smaller than 1. Then, clearly, by (4),

$$\begin{aligned}&P(H_n-\log_2\alpha<t)\leqq P(H_n\leqq\overline{\log_2\alpha+t})\\&\leqq\exp(-n\varepsilon^2/2)+\exp\left(-\alpha\frac{(1-\varepsilon)^2}{1+\varepsilon}2^{-\overline{\log_2\alpha+t}}(1-\varepsilon)\right)\end{aligned}$$

for all $t\geqq-\frac{1}{2}\log_2\alpha+A$, where A is a positive number depending upon ε and f only (this requires a careful verification of the $o(1)$ terms in (4) and (6)). Thus,

$$\begin{aligned}&\int_{-\infty}^{0}P(H_n-\log_2\alpha<t)\,dt\\&\leqq n\exp(-n\varepsilon^2/2)+\int_{-\infty}^{0}\exp\left(-\alpha\frac{(1-\varepsilon)^3}{1+\varepsilon}2^{-\overline{\log_2\alpha+t}}\right)dt\\&+nP(H_n\leqq\overline{\log_2\alpha-\tfrac{1}{2}\log_2\alpha+A}).\end{aligned}$$

The first term is $o(1)$, the second term is not greater than

$$\int_{-\infty}^{0}\exp\left(-\frac{(1-\varepsilon)^3}{1+\varepsilon}\cdot 2^{-(t+1)}\right)dt,$$

which, by our choice of ε, is arbitrarily close to $\int_{-\infty}^{0}\exp(-2^{-(t+1)})\,dt$. For the third term, we use, once again, bound (4) with $l=\overline{\frac{1}{2}\log_2\alpha+A}$:

$$\begin{aligned}&n\left[\exp(-n\varepsilon^2/2)+\exp\left(-\sum_{i=1}^{2^l}\frac{(n_1p_i)^2}{2(1+n_1p_i)}\right)\right]\\&\leqq n\left[\exp(-n\varepsilon^2/2)+\exp\left(-2^l\frac{(n_1/2^l)^2}{2(1+n_1/2^l)}\right)\right]\end{aligned}$$

(by Jensen's inequality and the convexity of $v^2/(1+v)$)

$$\begin{aligned}&\leqq n\left[\exp(-n\varepsilon^2/2)+\exp\left(-\frac{n_1^2}{2(n_1+2^{A+1}n\sqrt{\int f^2})}\right)\right]\\&\quad(\text{since } 2^p<2^{A+1}\sqrt{\alpha}<2^{A+1}n\sqrt{\textstyle\int f^2})\\&=o(1).\end{aligned}$$

3. Triesort

Tries can be used to sort n elements as follows:

(1) Construct the trie sequentially by inserting $X_1,\dots,X_n$, one element at a time. This takes C_n comparisons.

(2) Traverse the trie in preorder and note all the leaves (X_i's) as they are visited. This takes time proportional to the number of nodes N_n in the trie.

Since C_n and N_n are appropriate measures of the complexity of this algorithm, we will not bother to analyze other quantities. It is clear that $E(C_n)=E(N_n)=\infty$ for all $n\geqq 2$ if $\int f^2(x)\,dx=\infty$, so we will assume throughout that f is in L_2.

Theorem 5. *If* $\int f^2(x)\,dx<\infty$, *then*

$$E(C_n)=n\log_2 n+0(n)$$

and

$$E(N_n)\leqq n(1+\sqrt{18\int f^2(x)\,dx}).$$

Proof. Since

$$C_n=D_{11}+\ldots+D_{nn},$$

we leave

$$E(C_n)=\sum_{i=1}^{n}E(D_{ii})\geqq\sum_{i=1}^{n}(\mathrm{int}(\log_2(2i-1)-1)=n\log_2 n+0(n)$$

and

$$E(C_n)\leqq\sum_{i=1}^{n}\left(\log_2 i+1+\left(\gamma+\frac{1}{2n-2}\right)\Big/\ln 2+192\int f^2(x)\,dx\right)$$
$$\leqq n\log_2 n+0(n)$$

where we used inequality (2) of Devroye (1982).

Because an internal node indicates that an interval of the type $[(i-1)/2^l, i/2^l)$ has at least two elements, we have the following equality for $E(N_n)$:

$$E(N_n)=n+\sum_{l=0}^{\infty}\sum_{i=1}^{2^l}(1-(1-p_{li})^n-n\,p_{li}(1-p_{li})^{n-1})$$

where p_{li} is the integral of f over $[(i-1)/2^l, i/2^l)$. Now,

$$(1-p_{li})^n+n\,p_{li}(1-p_{li})^{n-1}\geqq\max(0,1-n\,p_{li}+n\,p_{li}(1-n\,p_{li}))$$
$$\geqq\max(0,1-(n\,p_{li})^2).$$

Choose an integer $L\geqq 0$, and note that by (2) and the last inequality,

$$E(N_n)\leqq n+\sum_{l=0}^{L}2^l+\sum_{l=L+1}^{\infty}n^2\sum_{i=1}^{2^l}p_{li}^2$$
$$\leqq n+2^{L+1}+n^2\int f^2(x)\,dx\sum_{l=L+1}^{\infty}2^{-l}$$
$$=n+2^{L+1}+n^2\int f^2(x)\,dx/2^L.$$

If we take $L=\mathrm{int}(\log_2\sqrt{n^2\int f^2(x)\,dx/2})$, then trivial bounding techniques give

$$E(N_n)\leqq n+n\sqrt{\int f^2(x)\,dx}\,(\sqrt{2}+2\sqrt{2}),$$

which was to be shown.

References

1. Devroye, L.: Laws of the iterated logarithm for order statistics of uniform spacings. Ann. Probability **9**, 860–867 (1981)
2. Devroye, L.: A note on the average depth of tries. Computing **28**, 367–371 (1982)
3. Flajolet, Ph., Steyaert, J.M.: A branching process arising in dynamic hashing, trie searching and polynomial factorization. Proceedings of the 9th ICALP Colloquium, Lecture Notes in Computer Science, Vol. 140, pp. 239–251. Berlin-Heidelberg-New York: Springer 1982
4. Fredkin, E.H.: Trie memory. Comm. ACM **3**, 490–500 (1960)
5. Johnson, N.L., Kotz, S.: Distributions in statistics: Continuous univariate distributions –1. New York: John Wiley 1970
6. Regnier, M.: On the average height of trees in digital search and dynamic hashing. Information Processing Lett. **13**, 64–66 (1982)
7. Yao, A.C.: A note on the analysis of extendible hashing. Information Processing Lett. **11**, 84–86 (1980)
8. Wheeden, R.L., Zygmund, A.: Measure and Integral. New York: Marcel Dekker 1977

Received March 24, 1982/December 13, 1983