The Annals of Probability
1985, Vol. 13, No. 2, 414–427

ASYMPTOTICAL GROWTH OF A CLASS OF RANDOM TREES

BY B. PITTEL

The Ohio State University

We study three rules for the development of a sequence of finite subtrees $\{t_n\}$ of an infinite m-ary tree t. Independent realizations $\{\omega(n)\}$ of a stationary ergodic process $\{\omega\}$ on m letters are used to trace out paths in t. In the first rule, t_n is formed by adding a node to t_{n-1} at the first location where the path defined by $\omega(n)$ leaves t_{n-1}. The second and third rules are similar, but more complicated. For each rule, the height L_n of the added node is shown to grow, in probability, as ln n divided by h the entropy per symbol of the generic process. A typical retrieval time has the same behavior. On the other hand, lim $\inf_n L_n/\ln n = \sigma_1$, lim $\sup_n L_n/\ln n = \sigma_2$ a.s., where the constants σ_1, σ_2, are, in general, different, depend on the rule in use, and $\sigma_1 < 1/h < \sigma_2$. It is proven along the way that the height of t_n grows as $\sigma_2 \ln n$ with probability one.

1. Introduction; results; comments. Consider a probability space $(\Omega, \mathscr{F}, p)$, where $\Omega = \{\omega: \omega = \{\omega_\nu\}_{\nu=1}^\infty, \omega_n \in S, n \geq 1\}$, $S = \{1, \cdots, m\}$, the σ-field $\mathscr{F}$ is generated by finite-dimensional cylinders, and the probability measure P is such that ω is a stationary ergodic process. One may interpret ω as the nonterminating m-adic expansion of a random number x from (0, 1], or as an infinitely long random text written in an alphabet of m letters. For the first interpretation, it is natural to assume that ω is a sequence of independent trials, with m equally likely outcomes in each trial. If ω is a random text, it is important to take into account that the different letters may appear in ω with different frequencies, and that there may be mutual dependence between the successive letters in ω. A stationary ergodic process ω seems to be a proper probabilistic model in this situation.

Introduce the probability space $(\Omega^\infty, \mathscr{F}^\infty, P^\infty)$, where $\Omega^\infty = \Omega \times \Omega \times \cdots$, $\mathscr{F}^\infty = \mathscr{F} \times \mathscr{F} \times \cdots$, $P^\infty = P \times P \times \cdots$, so Ω^∞ consists of $\omega^\infty = (\omega(1), \omega(2), \cdots)$, where $\omega(n) \in \Omega$. Thus, $\omega(1), \omega(2), \cdots$ are *independent* copies of the generic process $\omega = \{\omega_\nu\}_{\nu=1}^\infty$, and they can be thought of as an infinite sequence of texts.

Let t be the complete *infinite* m-ary tree. We study some rules [6], (Section 6.3), each of which allows us to determine a sequence $\{t_n\}_{n=1}^\infty$ of *finite* sub-trees of t, such that $t_n = t_n(\omega(1), \cdots, \omega(n))$, $t_1 \subset t_2 \subset \cdots$, and t_n has n of its nodes labelled $\omega(1), \cdots, \omega(n)$. ($t_1$ is the root of t labelled $\omega(1)$.)

Rule A. Given the tree t_n, $(n \geq 1)$, introduce its external nodes; they are direct descendants of nodes of t_n, which are not themselves nodes of t_n. The sequence $\omega(n+1) = (\omega_\nu(n+1))_{\nu=1}^\infty$ determines uniquely an infinite path in t: it begins at

Received April 1983; revised August 1984.
AMS 1980 *subject classifications.* Primary 60C05, 60F15; secondary 28D20, 68C25.
Key words and phrases. Random trees, lengths of the paths, ergodic process, asymptotic growth, strong, weak convergence.

the root, and its ν-th link is the i-th left among m links of t going out of the ν-th node of the path if $\omega_\nu(n+1) = i$, $i \in S$, $\nu \geq 1$. Cut off the path when it reaches an external node of t_n, label this node $\omega(n+1)$, and form the next tree t_{n+1} by connecting this new node to t_n.

Thus, in Rule A, t_n has exactly n nodes, all of which are labelled. In contrast, the labelled nodes in Rules B–C are endnodes—i.e., nodes of t_n without direct descendants in t_n, or "leaves" of the tree—and t_n has in general more than n nodes.

Rule B. If the path generated by $\omega(n+1)$ leaves t_n at a node which is not an endnode then, as above, the *next* node of this path is labelled $\omega(n+1)$. Otherwise, the paths of $\omega(n+1)$ and of a certain $\omega(\mu)$, $1 \leq \mu \leq n$, coincide until they leave the tree t_n. The rule prescribes to follow these two paths further until they disengage, and then $\omega(n+1)$ and $\omega(\mu)$ are assigned to the endnodes of the respective links going out of the node of branching.

Rule C. The tree t_n is obtained by compressing the one constructed via Rule B; namely, each path of the latter is shortened (if possible) by deleting its no-branching nodes (i.e. those with outdegree 1), and putting together its remaining pieces.

Figure 1 shows the trees t_5 for the dyadic sequences $\omega(1) = (1, 0, 1, \cdots)$, $\omega(2) = (0, 0, 1, \cdots)$, $\omega(3) = (0, 1, 0, \cdots)$, $\omega(4) = (1, 0, 0, \cdots)$, $\omega(5) = (0, 0, 0, \cdots)$.

In the third tree, the shape of the paths from the root to the nodes $\omega(1)$, $\omega(4)$, and the mark 1 at the preceding node, imply that $\omega_1(1) = \omega_1(4) = 1$, $\omega_2(1) = \omega_2(4)$ and $\omega_3(1) = 1$, $\omega_3(4) = 0$. If this mark were some integer α, it would mean that $\omega_1(1) = \omega_1(4) = 1$, $\omega_2(1) = \omega_2(4)$, $\cdots$, $\omega_{2+(\alpha-1)}(1) = \omega_{2+(\alpha-1)}(4)$, and $\omega_{2+\alpha}(1) = 1$, $\omega_{2+\alpha}(4) = 0$. Generally, for this rule, t_n may have many marked nodes, and some of them may belong to the same paths.

Whatever the rule is, t_n is interpreted as a tree-type arrangement of n numbers (records) in a computer's memory. Upon request, any one of them can be found by using its consecutive digits (letters) as the pointers showing where to move

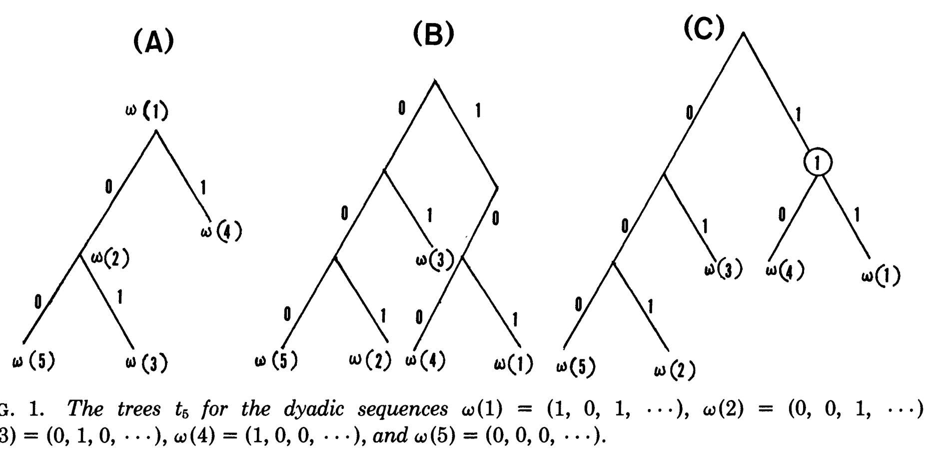

FIG. 1. *The trees* t_5 *for the dyadic sequences* $\omega(1) = (1, 0, 1, \cdots)$, $\omega(2) = (0, 0, 1, \cdots)$, $\omega(3) = (0, 1, 0, \cdots)$, $\omega(4) = (1, 0, 0, \cdots)$, *and* $\omega(5) = (0, 0, 0, \cdots)$.

next in the tree. For the rule C, a mark encountered in the search process indicates the respective amount of the next digits to be skipped over. The corresponding algorithms are named "Digital Tree Search", "Trie Search", and "Patricia" (abbr. "Practical Algorithm to Retrieve Information Coded in Alphanumeric"), [6].

For each rule, $L_{n\nu} = L_{n\nu}(\omega^\infty)$, the random length of the path leading to the node $\omega(\nu)$ is the amount of the digits of $\omega(\nu)$ to be checked before its position is found, $1 \le \nu \le n$. For the rule A, $L_{n\nu} = L_{\nu\nu} =_{\text{def}} L_\nu$, as the position of the node $\omega(\nu)$ in t is determined once and for all by comparing its digits with those of $\omega(1), \cdots, \omega(\nu-1)$. In case of the rules B and C, on the other hand, each $t_n = t_n(\omega(1), \cdots, \omega(n))$ is permutation invariant; hence $L_{n\nu}$, $1 \le \nu \le n$, are equidistributed. Another important characteristic of t_n is M_n, the length of the path from the root to $\omega(\mu)$, where $\omega(\mu)$ is randomly chosen with equal likelihood from $\{\omega(1), \cdots, \omega(n)\}$. Thus M_n might be viewed as a typical "retrieval time."

Throughout the paper we shall assume that the *generic* process ω satisfies the following condition: denote $\mathscr{F}_a^b$ the σ-field generated by $\omega_a, \cdots, \omega_b$, $1 \le a \le b$; there exist two positive constants $c_1 \le c_2$ and an integer $b_0 > 0$ such that for all $1 \le a \le a + b_0 \le b$,

$$c_1 P(\mathscr{A})P(\mathscr{B}) \le P(\mathscr{A}\mathscr{B} \le c_2 P(\mathscr{A})P(\mathscr{B}) \tag{1.1}$$

whenever $\mathscr{A} \in \mathscr{F}_1^a$, $\mathscr{B} \in \mathscr{F}_{a+b_0}^b$ (cf. [2], [3]). This condition implies not only ergodicity of ω, but also guarantees that it is strongly mixing [1]. For the rule C, we shall also assume that for each $n \ge 1$ and $s^n = (s_1, \cdots, s_n) \in S^n$

$$p(s^n) \le \rho p(s^{n-1}), \quad (p(s^\nu) = P(\omega_1 = s_1, \cdots, \omega_\nu = s_\nu)), \tag{1.2}$$

where $\rho \in (0, 1)$ and $p(s^0) = 1$.

To formulate the result, introduce

$$h = \lim_{n\to\infty} n^{-1} E(\ln(1/p(\omega^n))), \tag{1.3}$$

$$h_1 = \lim_{n\to\infty} n^{-1} \max\{\ln(1/p(s^n)) \colon p(s^n) > 0\}, \tag{1.4}$$

$$h_2 = \lim_{n\to\infty} n^{-1} \min\{\ln(1/p(s^n)) \colon p(s^n) > 0\}, \tag{1.5}$$

$$h_3 = \lim_{n\to\infty} (2n)^{-1} \ln(1/E(p(\omega^n))), \tag{1.6}$$

$\omega^n = (\omega_1, \cdots, \omega_n)$. The limit in (1.3) exists for any stationary process ω (Kolmogorov-Sinai theorem, [1]), and it is called the entropy per letter. Existence of the limits in (1.4)–(1.6) is proven below (Lemma 1) under the condition (1.1). It can be seen that $0 \le h_3 \le h_2 \le 2h_3 \le h \le h_1$. Assume from now on that

$$h_1 < \infty, \quad h_3 > 0. \tag{1.7}$$

THEOREM. *Under the conditions* (1.1), (1.7) *and, for the rule C, also* (1.2), *we have*: (a)

$$(P)\lim_{n\to\infty} L_n/\ln n = (P)\lim_{n\to\infty} M_n/\ln n = 1/h; \tag{1.8}$$

(b) *almost surely (a.s.)*

(1.9) $$\liminf_n L_n/\ln n = \sigma_1, \quad \limsup_n L_n/\ln n = \sigma_2,$$

where $\sigma_1 = 1/h_1$, $\sigma_2 = 1/h_2$ *(Rules* A, C), *and* $\sigma_2 = 1/h_3$ *(Rule* B).

NOTES. (1) Typically, $\sigma_1 < 1/h < \sigma_2$, so this theorem reveals an interesting property of $L_n/\ln n$: if n is large then, with high probability, $L_n/\ln n$ is close to $1/h$; on the other hand, almost surely $L_n/\ln n$ oscillates between σ_1 and σ_2 as $n \to \infty$. Our proofs reveal also that the height of t_n is a.s. equivalent to $\sigma_2 \ln n$.

(2) Suppose that the generic process ω is a sequence of independent trials with a distribution $\{p(s): s \in S\}$ for each trial. A little reflection shows that $h_t = \ln(1/p_t)$, $1 \le t \le 3$, where

$$p_{1,2} = \min, \max\{p(s): s \in S\}, \quad p_3 = (\textstyle\sum_s p^2(s))^{1/2},$$

and, of course, $h = \sum_s p(s)\ln(1/p(s))$.

In particular, if $p(s) \equiv 1/m$, then $h_1 = h_2 = \ln m$ and $h_3 = \ln m/2$. Thus, $L_n/\ln n \to 1/\ln m$ a.s. (Rules A, C) and $2^{-1}\limsup_n L_n/\ln n = \liminf_n L_n/\ln n = 1/\ln m$ a.s. (Rule B). The case of equal probabilities was previously studied by Konheim and Newman ($m = 2$) [7] who proved that $E(M_n) = \log_2 n + O(1)$, $\lim_{n\to\infty} P(M_n \le (1 - \varepsilon)\log_2 n) = 0$, $\varepsilon > 0$, (Rule A), and also by Knuth (arbitrary m) [6] (Section 6.3) who showed (for all the rules) that $E(L_n) = E(M_n) = \log_m n + O(1)$. From [4], [6] (Section 5.2.2, analysis of radix exchange sorting), it follows also that, for the Rule B, $E(L_n) = E(M_n) = \ln n/h + O(1)$ in case $m = 2$ and general $p(1)$, $p(2)$. The bulk of Knuth's proofs concerns the study of the term $O(1)$ in the above estimates of $E(L_n)$, $E(M_n)$. (The author has obtained results on the limiting *distributions* of L_n and the lengths of the longest and shortest path in t_n (Rule B) for a general distribution $\{p(s): s \in S\}$, [9].)

(3) Suppose that ω is a stationary Markov chain. It satisfies (1.1), (1.7) if its transition matrix $[p(s, s')]$, $s, s' \in S$, is irreducible and aperiodic; the condition (1.2) is valid if $\max\{p(s, s'): s, s' \in S\} < 1$. Since $p(s^\nu) = \pi(s_1)p(s_1, s_2) \cdots p(s_{\nu-1}, s_\nu)$, where $\{\pi(s): s \in S\}$ is the stationary distribution, it follows from a (nonprobabilistic) result of Romanovski [13]:

(1.10) $$h_{1,2} = \max, \min(|\mathscr{C}|^{-1}\ell(\mathscr{C}));$$

here max, min are taken over all *simple* cycles $\mathscr{C} = (s_1, \cdots s_\nu, s_1)$ on S, ($\ell(\mathscr{C}) = \sum_{\mu=1}^{\nu} \ln(1/p(s_\mu, s_{\mu+1}))$, $s_{\nu+1} = s_1$, $|\mathscr{C}| = \nu$, $1 \le \nu \le m$), such that $\ell(\mathscr{C}) < \infty$. From the formula for $p(s^\nu)$, it follows also that $h_3 = -2^{-1}\ln q$, where q is the spectral radius of the matrix $[p^2(s, s')]$. Finally, it is known that $h = \sum_{s,s'} \pi(s)p(s, s')\ln(1/p(s, s'))$.

(4) A more general case of an r-dependent stationary Markov chain can be treated similarly.

(5) Another well known search-insertion algorithm is based on comparisons between the numbers, rather than their digits. For asymptotical results, the reader is referred to [6] (Section 6.2.2), [8], [10], and [11].

2. Proofs.

LEMMA 1. *There exist the (possibly infinite) limits* h_1, h_2, h_3, (*see* (1.4)–(1.6)).

PROOF. (a) Introduce $f(n) = \max\{\ln(1/p(s^n)): p(s^n) > 0\}$. Let

$$n = j + k + b_0, \quad j \geq 0, \quad k \geq 0;$$

according to (1.1) and stationarity of ω,

$$f(j + k + b_0) \leq \max\{\ln(1/p(s^j)) + \ln(1/p(\tilde{s}^k)): p(s^j) > 0, p(\tilde{s}^k) > 0\} + \ln c_1^{-1} \leq f(j) + f(k) + c, \tag{2.1}$$

where $\tilde{s}^k = (s_{n-k+1}, \cdots, s_n)$, $c = \ln c_1^{-1}$. Let $n_t \to \infty$ be such that

$$\liminf_n f(n)/n = \lim_{t\to\infty} f(n_t)/n_t.$$

For $n \geq b_0$ and a fixed t, write $n - b_0 = qn_t + r$, where $q \geq 0$ and $0 \leq r < n_t$; clearly, $q = q(n) \to \infty$ as $n \to \infty$. Repeatedly using (2.1), we have

$$f(n) = f(qn_t + r + b_0) \leq qf(n_t) + f(r + b_0) + qc.$$

Hence

$$\limsup_n f(n)/n \leq f(n_t)/n_t + c/n_t,$$

or, letting $t \to \infty$,

$$\limsup_n f(n)/n \leq \liminf_n f(n)/n.$$

Existence of $h_1 = \lim_{n\to\infty} f(n)/n$ is proven. The case of h_2 is similar.

(b) Introduce the process $\{\bar{\omega}_n\}_{n=1}^{\infty} = \{(\omega_n(1), \omega_n(2))\}_{n=1}^{\infty}$; $\bar{\omega}$ is obviously stationary, with the state space $S^2 = S \times S$. Observe that

$$E(p(\omega^n)) = \textstyle\sum_{s^n} p^2(s^n) = P(\omega_\nu(1) = \omega_\nu(2), 1 \leq \nu \leq n).$$

Also, by independence of $\omega(1)$ and $\omega(2)$, the process $\bar{\omega}$ satisfies the condition (1.1) with $\mathscr{F}_1^a$, $\mathscr{F}_{a+b_0}^b$ replaced respectively by $\mathscr{F}_1^a \times \mathscr{F}_1^a$ and $\mathscr{F}_{a+b_0}^b \times \mathscr{F}_{a+b_0}^b$, and $\bar{c}_1 = c_1^2$, $\bar{c}_2 = c_2^2$. Arguing as in (2.1), we obtain: the function $g(n) = \ln P(\omega_\nu(1) = \omega_\nu(2), 1 \leq \nu \leq n)$ satisfies

$$g(j + k + b_0) \leq g(j) + g(k) + \bar{c}, \quad \bar{c} = \ln \bar{c}_2.$$

Again, this relation implies existence of $h_3 = -\lim_{n\to\infty} g(n)/2n$.

LEMMA 2. *There exist* $\hat{w}, \tilde{w} \in \Omega$ *such that*

$$h_1 = \lim_{n\to\infty} n^{-1}\ln(1/p(\hat{w}^n)), \quad h_2 = \lim_{n\to\infty} n^{-1}\ln(1/p(\tilde{w}^n)).$$

PROOF. Consider, for example, h_2. By Lemma 1, there exists a sequence $\{\tilde{s}^n\}_{n=1}^{\infty}$, $\tilde{s}^n \in S^n$, for which $h_2 = \lim_{n\to\infty} n^{-1}\ln(1/p(\tilde{s}^n))$. By induction, one can prove existence of a sequence $t(\nu) \in S^{b_0}$, $\nu \geq 1$, such that

$$p(\tilde{w}^{n_\nu}) \geq (m^{-b_0}c_1)^{\nu-1} p(\tilde{s}^1) \cdots p(\tilde{s}^\nu), \quad \nu \geq 1. \tag{2.2}$$

Here c_1 is the constant from (1.1), and

$$\tilde{w}^{n_1} = \tilde{s}^1,$$

$$\tilde{w}^{n_\nu} = (\tilde{s}^1, t(1), \tilde{s}^2, \cdots, t(\nu - 1), \tilde{s}^\nu), \quad \nu \geq 2,$$

so that $n_\nu = \nu(\nu + 1)/2 + b_0(\nu - 1)$. Introduce an infinite sequence $\tilde{w} = (\tilde{s}^1, t(1), \tilde{s}^2, t(2), \cdots)$. By definition of $\{\tilde{s}^\nu\}_{\nu=1}^\infty$ and (2.2),

$$\begin{aligned}
\limsup_\nu n_\nu^{-1}\ln(1/p(\tilde{w}^{n_\nu})) &\leq \lim_{\nu\to\infty} n_\nu^{-1}[-(\nu - 1)\ln(m^{-b_0}c_1) + \textstyle\sum_{\mu=1}^\nu \ln(1/p(\tilde{s}^\mu))] \\
&= \lim_{\nu\to\infty}[\nu(\nu + 1)/2]^{-1}(\textstyle\sum_{\mu=1}^\nu \mu(\mu^{-1}\ln(1/p(\tilde{s}^\mu)))) = h_2.
\end{aligned}$$

(Cesaro-type averages of a convergent sequence converge to the same limit.) Since obviously $\liminf_\nu n_\nu^{-1}\ln(1/p(\tilde{w}^{n_\nu})) \geq h_2$, we have

$$\lim_{n\to\infty} n_\nu^{-1}\ln(1/p(\tilde{w}^{n_\nu})) = h_2. \tag{2.3}$$

As $\lim_{\nu\to\infty}(n_{\nu+1}/n_\nu) = 1$ and $\ln(1/p(\tilde{w}^n))$ is a nondecreasing function of n, the relation (2.3) entails in a usual way that $\lim_{n\to\infty} n^{-1}\ln(1/p(\tilde{w}^n)) = h_2$, too.

REMARK. Since $h_1, h_2 < \infty$, (see (1.7)), $p(\hat{w}^k) > 0$, $p(\tilde{w}^k) > 0$ for each $k \geq 1$. A sequence $w \in \Omega$ is called feasible if $p(w^k) > 0$, $k \geq 1$. Thus, $\hat{w}$, $\tilde{w}$, are feasible. Since $h_2 > 0$ as well, $p(w^k) \in (0, 1)$, $k \geq 1$, for every feasible w.

We consider L_n separately for the three rules.

Rule A. Each feasible $w \in \Omega$ determines an infinite path in the tree t. Denote the path and its k-long initial segment by w and w^k respectively, $k \geq 0$. Let

$$X_n(w) = \max\{k : w^k \text{ is contained in } t_n\},$$

so $X_1(w) = 0$. $X_n(w)$ is $\mathscr{F}^n$-measurable, where $\mathscr{F}^n = \sigma(\omega(1), \cdots, \omega(n))$. Since $\mathscr{F}^n$ and $\omega(n + 1)$ are independent, it follows from the description of Rule A that on $(X_n(w) = k)$, $k \geq 0$,

$$P(X_{n+1}(w) = k + 1 \mid \mathscr{F}^n) = 1 - P(X_{n+1}(w) = k \mid \mathscr{F}^n) = p(w^{k+1}).$$

Hence, $\{X_n(w)\}$ is a Markov chain with

$$P(X_{n+1}(w) = k + 1 \mid X_n(w) = k) = p(w^{k+1}),$$

and remaining (conditional) probability mass at $X_{n+1}(w) = k$. It follows then that

$$\begin{aligned}
T_k(w) &=_{\text{def}} \min\{n : X_n(w) = k\} \\
&= 1 + \textstyle\sum_{r=1}^k (T_r(w) - T_{r-1}(w)) = 1 + \sum_{r=1}^k G(p(w^r)),
\end{aligned} \tag{2.4}$$

where $\{G(p(w^r))\}$ is a sequence of independent geometrically distributed random variables with parameters $\{p(w^r)\}$. In particular, if $Y = G(p)$, then

$$\begin{aligned}
&\text{(a) } P(Y = j) = p(1 - p)^{j-1}, \quad (j \geq 1), \quad &&\text{(b) } E(Y) = 1/p, \\
&\text{(c) } E(1/Y) = p\ln(1/p)/(1 - p), &&\text{(d) } E(x^Y) = xp/(1 - x(1 - p)).
\end{aligned} \tag{2.5}$$

LEMMA 3. *Let w be such that there exists*

$$\lim_{n\to\infty} n^{-1}\ln(1/p(w^n)) = h(w) > 0. \tag{2.6}$$

Then a.s.

$$\lim_{n\to\infty} X_n(w)/\ln n = 1/h(w). \tag{2.7}$$

PROOF OF LEMMA 3. Since $X_n(w) = k$ for $n = T_k(w)$, (2.7) follows—by easy monotonicity arguments—from

$$\lim_{k\to\infty} \ln T_k(w)/k = h(w) \quad \text{a.s.} \tag{2.8}$$

To prove (2.8), observe first that $\lim_{k\to\infty}[\exp(-h(w))p(w^k)^{1/k}] = 1$ by (2.6). Then

$$P[\ln T_k > (1+\varepsilon)kh(w)] = P[T_k > \exp((1+\varepsilon)kh(w))] \tag{2.9}$$

tends to 0 exponentially fast via Chebyshev's inequality and (2.4), (2.5b), while

$$\begin{aligned} P[\ln T_k < (1-\varepsilon)kh(w)] &= P[T_k < \exp((1-\varepsilon)kh(w))] \\ &= P[1 + \textstyle\sum_{r=1}^{k} G(p(w^r)) < \exp((1-\varepsilon)kh(w))] \\ &< P[1/G(p(w^k)) > \exp(-(1-\varepsilon)kh(w))] \end{aligned} \tag{2.10}$$

tends to 0 exponentially fast via Chebyshev's inequality and (2.5c). Hence, (2.8) follows from (2.9), (2.10) by the Borel-Cantelli lemma.

DEFINITION. A node of t is called feasible if a path, say w^k, connecting it with root has $p(w^k) > 0$.

Clearly, the nodes of all trees t_n are feasible a.s. Denote ℓ_n and $\mathscr{L}_n$ the length of the shortest and the longest path from the root to the feasible external nodes of the tree t_{n-1}. Then, $\ell_n \le L_n \le \mathscr{L}_n$ a.s.

LEMMA 4. $\limsup_n \ell_n/\ln n \le 1/h_1$, $\liminf_n \mathscr{L}_n/\ln n \ge 1/h_2$ *a.s.*

PROOF. The statement follows from Lemmas 2, 3 and obvious inequalities

$$\ell_n \le X_{n-1}(\hat{w}) + 1, \quad X_{n-1}(\tilde{w}) + 1 \le \mathscr{L}_n.$$

LEMMA 5. $\liminf_n \ell_n/\ln n \ge 1/h_1$, $\limsup_n \mathscr{L}_n/\ln n \le 1/h_2$ *a.s.*

COROLLARY 1. $\lim_{n\to\infty} \ell_n/\ln n = 1/h_1$, $\lim_{n\to\infty} \mathscr{L}_n/\ln n = 1/h_2$ *a.s.*

COROLLARY 2. $\liminf_n L_n/\ln n = 1/h_1$, $\limsup_n L_n/\ln n = 1/h_2$ *a.s.*

PROOF OF COROLLARY 2. A.s. $L_n = \ell_n$ whenever $\ell_{n+1} > \ell_n$, which happens infinitely often a.s., since $\ell_n \to \infty$ a.s. Then a.s. $\liminf_n L_n/\ln n = \lim_{n\to\infty} \ell_n/\ln n = 1/h_1$. The case of the second limit is similar.

PROOF OF LEMMA 5. Since $\ell_n = 1 + \min_w X_{n-1}(w)$,

$$P(\ell_n \le r) \le \sum_{w^r} P(X_{n-1}(w) \le r - 1) = \sum_{w^r} P(T_r(w) \ge n). \tag{2.11}$$

Here, (see (2.4), (2.5d)),

$$P(T_r(w) \ge n) \le x^{-n}E(x^{T_r(w)}) = x^{-n+1} \prod_{\mu=1}^{r} xp(w^\mu)/(1 - xq(w^\mu)), \tag{2.12}$$

for all $x \in [1, 1/q(w^r))$, $(q(w^\mu) = 1 - p(w^\mu))$. Similarly,

$$P(\mathscr{L}_n > r) \le \sum_{w^r} P(X_{n-1}(w) > r - 1) = \sum_{w^r} P(T_r(w) < n), \tag{2.13}$$

where

$$P(T_r(w) < n) \le x^{-n}E(x^{T_r(w)}), \quad x \in (0, 1). \tag{2.14}$$

(a) Given $\varepsilon > 0$, by Lemma 2,

$$p(w^\mu) \ge c\alpha^\mu, \quad \alpha = \exp(-(1 + \varepsilon^2)h_1), \quad c = c(\varepsilon) \in (0, 1). \tag{2.15}$$

Since

$$1/q(w^r) = (1 - p(w^r))^{-1} > 1 + p(w^r) \ge 1 + c\alpha^r,$$

we may, and shall, choose $x = 1 + c\alpha^r$ in (2.12). For $x > 1$, the function $p/(1 - x(1 - p))$ decreases when p increases. Thus in (2.12)

$$\begin{aligned}\prod_{\mu=1}^{r} p(w^\mu)/(1 - xq(w^\mu)) &\le \prod_{\mu=1}^{r} [(1 + c\alpha^r) - c\alpha^{r-\mu}]^{-1}\\ &\le \prod_{\mu=1}^{r} (1 - c\alpha^{r-\mu})^{-1}\\ &\le \prod_{\mu=0}^{\infty} (1 - c\alpha^\mu)^{-1} \doteq \text{const} < \infty.\end{aligned} \tag{2.16}$$

Also,

$$\begin{aligned}x^{r-n+1} &= (1 + c\alpha^r)^{r-n+1} = O((1 + c\alpha^r)^{-n})\\ &= O(\exp[-cn\alpha^r + O(n\alpha^{2r})]).\end{aligned} \tag{2.17}$$

Take $r = [(1 - \varepsilon)\ln n/h_1]$. Since, (see (2.15)),

$$\alpha n^{\varepsilon + O(\varepsilon^2)} \le n\alpha^r \le n^{\varepsilon + O(\varepsilon^2)},$$

choosing ε small enough, we obtain from (2.12), (2.16), (2.17):

$$P(T_r(w) \ge n) \le \exp(-cn^{\varepsilon/2}), \quad n \ge n(\varepsilon).$$

Thus, (see (2.11),

$$P(\ell_n \le r) \le m^r \exp(-cn^{\varepsilon/2}) = \exp(-cn^{\varepsilon/2} + O(\ln n)),$$

and, by the Borel-Cantelli lemma, $\liminf_n \ell_n/\ln n \ge 1/h_1$ a.s.

(b) Fix $\varepsilon \in (0, 1)$ and estimate $P(T_r(w) < n)$ for

$$r = [(1 + \varepsilon)\ln n/h_2]. \tag{2.18}$$

By Lemma 2 again, for all $\mu \ge 1$,

$$p(w^\mu) \le c\beta^\mu, \quad \beta = \exp(-(1 - \varepsilon^2)h_2), \quad c = c(\varepsilon). \tag{2.19}$$

Choose $x = (1 + c/n)^{-1}$ in (2.14). For $x < 1$, $p/(1 - x(1 - p))$ increases with p.

Thus, in (2.14), (see (2.12)),

$$x^{-n}E(x^{T_r(w)}) \le x^{-n+1}\left(\prod_{\mu=1}^{r} x\beta^{\mu}\right)\exp\left(-\sum_{\mu=1}^{r} \ln(1/n + \beta^{\mu})\right). \tag{2.20}$$

Here

$$\begin{aligned}\sum_{\mu=1}^{r} \ln(1/n + \beta^{\mu}) &\ge \int_1^{r+1} \ln(1/n + \beta^{\mu})\, d\mu = (\ln 1/\beta)^{-1} \int_{\beta^{r+1}}^{\beta} \ln(1/n + y)/y\, dy \\ &= (\ln 1/\beta)^{-1}\left[\int_{\beta^{r+1}}^{1/n} + \int_{1/n}^{\beta}\right] = (\ln 1/\beta)^{-1}[I_1 + I_2];\end{aligned}$$

(by (2.18), (2.19), $\beta^{r+1} < 1/n < \beta$ for $\varepsilon \le \varepsilon_0$ and $n \ge n(\varepsilon)$). Further,

$$I_1 \ge \ln(1/n) \int_{\beta^{r+1}}^{1/n} 1/y\, dy = \ln^2 n - (r+1)\ln n \ln(1/\beta),$$

$$I_2 \ge \int_{1/n}^{\beta} \ln y/y\, dy = 2^{-1}[\ln^2\beta - \ln^2 n],$$

and, after some work,

$$\begin{aligned}\sum_{\mu=1}^{r} \ln(1/n + \beta^{\mu}) \ge{}& (2\ln(1/\beta))^{-1}(\ln n - r\ln(1/\beta))^2 \\ &+ \ln(1/\beta)(1 - r^2)/2 - \ln n.\end{aligned}$$

Then, ($x^{-n+r+1} = O(1)$ for $x = (1 + c/n)^{-1}$), by (2.14), (2.20) we have

$$\begin{aligned}P(T_r(w) < n) &\le \exp[-(r^2 + r)\ln(1/\beta)/2 - (2\ln(1/\beta))^{-1}(\ln n - r\ln(1/\beta))^2 \\ &\qquad + r^2\ln(1/\beta)/2 + O(\ln n)] \\ &= \exp[-(2\ln(1/\beta))^{-1}(\ln n - r\ln(1/\beta))^2 + O(\ln n)].\end{aligned} \tag{2.21}$$

Here, by (2.18), (2.19), $(\ln n - r\ln(1/\beta))^2$ is of order $\ln^2 n$, if ε is small; so combination of (2.13), (2.21) yields

$$P(\mathscr{L}_n > r) \le \exp(-c\ln^2 n), \quad c = c(\varepsilon) > 0.$$

As in (a), it implies that $\limsup_n \mathscr{L}_n/\ln n \le 1/h_2$ a.s.

REMARK. Denote H_n the height of the tree t_n. Since $H_n = \mathscr{L}_{n+1} - 1$ a.s., it follows from Corollary 1 that $\lim_{n\to\infty} H_n/\ln n = 1/h_2$ a.s.

To complete the study of Rule A, it remains to prove

LEMMA 6. $(P)\lim_{n\to\infty} L_n/\ln n = (P)\lim_{n\to\infty} M_n/\ln n = 1/h$.

PROOF. It suffices to consider L_n, as $P(M_n = r) = n^{-1}\sum_{\nu=1}^{n} P(L_\nu = r)$. Fix

$\varepsilon \in (0, 1)$ and a positive integer μ_0. Introduce

$$A_n = [\omega^\infty : | L_n/\ln n - 1/h | \geq \varepsilon/h],$$

$$B_{n\mu_0} = [\omega^\infty : | \mu^{-1}\ln(1/p(\omega^\mu(n))) - h | \leq \varepsilon^2 h,\ \mu \geq \mu_0].$$

We have then

$$P(A_n) \leq P(A_n B_{n\mu_0}) + P(\mu_0, \varepsilon),$$

$$P(\mu_0, \varepsilon) = P(\sup_{\mu \geq \mu_0} | \mu^{-1}\ln(1/p(\omega^\mu)) - h | \geq \varepsilon^2 h),$$

where $\omega = (\omega_1, \omega_2, \cdots)$ is the generic process. By Shannon-McMillan-Breiman Theorem [1], $\lim_{\mu_0 \to \infty} P(\mu_0, \varepsilon) = 0$ for each $\varepsilon \geq 0$. Further, (compare with (2.11), (2.13)),

$$P(A_n B_{n\mu_0}) \leq \textstyle\sum_r P_{nr}, \quad P_{nr} = \sum_{w^r} P(T_r(w) = n),$$

where in the first sum $| r/\ln n - 1/h | \geq \varepsilon/h$, and in the second sum $\max\{| \mu^{-1}\ln(1/p(w^\mu)) - h | : \mu_0 \leq \mu \leq r\} \leq \varepsilon^2 h$. It suffices to show that $\sum_r P_{nr} \to 0$ as $n \to \infty$. Observe that, for *each* $x \in (0, 1/q(w^r))$,

$$P(T_r(w) = n) \leq x^{-n} E(x^{T_r(w)}),$$

(see (2.12)). By the definition of w^r, for $1 \leq \mu \leq r$,

$$c_1 \alpha_1^\mu \leq p(w^\mu) \leq c_2 \alpha_2^\mu, \quad \alpha_{1,2} = \exp(-(1 \pm \varepsilon^2)h).$$

Using the lower (resp. upper) estimate of $p(w^\mu)$ for $r \leq r_1 = [(1-\varepsilon)\ln n/h]$, (resp. for $r > r_2 = [(1+\varepsilon)\ln n/h]$) and arguing as in Part (a) (resp. Part (b)) of the proof of Lemma 5, we get: for small enough $\varepsilon > 0$,

$$\textstyle\sum_r P_{nr} = \sum_{r \leq r_1} P_{nr} + \sum_{r > r_2} P_{nr} \leq \exp(-c^1 n^{\varepsilon/2}) + \exp(-c'' \ln^2 n), \quad c', c'' > 0.$$

Rule B. Introduce ℓ_n, $\mathscr{L}_n$, H_n defined exactly as in the case of Rule A.

LEMMA 7. (a) $\ell_n/\ln n \to 1/h_1$ *a.s.*, (b) $\mathscr{L}_n/\ln n \to 1/h_3$ *a.s.* (*As in the case of rule A, it follows from* (b) *that* $H_n/\ln n \to 1/h_3$ *a.s.*)

PROOF OF LEMMA 7. (a) According to the rule, if, for some r and n, $\ell_n \leq r$ then there is w^r such that $| \{\nu : 1 \leq \nu \leq n-1,\ \omega^r(\nu) = w^r\} | \leq 1$. Then denoting $p(r) = \min\{p(w^r)\}$, $q(r) = 1 - p(r)$,

$$\begin{aligned} P(\ell_n \leq r) &\leq \textstyle\sum_{w^r} [q(w^r)^{n-1} + (n-1)q(w^r)^{n-2}p(w^r)] \\ &\leq m^r[q(r)^{n-1} + (n-1)q(r)^{n-2}p(r)] \leq n m^r q(r)^{n-1} \end{aligned} \tag{2.22}$$

for large enough r, since $q^{n-1} + (n-1)q^{n-2}p$ (where $p = 1 - q$) increases with q on (0, 1) and (see (1.4), (1.7)), $p(r) \to 0$. On the other hand, if $\ell_n > r$ then $| \{\nu : 1 \leq \nu \leq n-1,\ \omega^r(\nu) = w^r\} | \leq 1$, where w^r minimizes $p(w^r)$, i.e. $p(w^r) = p(r)$. Then

$$P(\ell_n > r) \leq 1 - q(r)^{n-1}. \tag{2.23}$$

Since $r^{-1}\ln(1/p(r)) \to 1/h_1$, $(r \to \infty)$, the estimates (2.22), (2.23) imply that, for

$\varepsilon \le \varepsilon_0$, $n \ge n(\varepsilon)$,

$$P(\ell_n \le (1-\varepsilon)\ln n/h_1) \le \exp(-n^{\varepsilon/2}), \quad P(\ell_n \ge (1+\varepsilon)\ln n/h_1) \le n^{-\varepsilon/2}.$$

It remains to apply the Borel-Cantelli lemma in conjunction, for the second estimate, with the fact that ℓ_n is nondecreasing and ln n is a slowly varying function, (cf., e.g., [5]).

(b) First, if $\mathscr{L}_n > r$ then $|\{(\nu, \mu): 1 \le \nu, \mu \le n-1, \omega^r(\nu) = \omega^r(\mu)\}| \ge 1$. Hence

$$P(\mathscr{L}_n > r) \le \binom{n-1}{2} \sum_{w^r} p(w^r)^2 = \binom{n-1}{2} E(p(\omega^r)). \tag{2.24}$$

Since $h_3 = \lim_{r\to\infty}(2r)^{-1}\ln(1/E(p(\omega^r)))$, (see Lemma 1), it follows from (2.24) that $P(\mathscr{L}_n > (1+\varepsilon)\ln n/h_3) \le n^{-\varepsilon}$, and $\lim\sup_n \mathscr{L}_n/\ln n \le 1/h_3$ a.s.

How to estimate $\mathscr{L}_n$ from below? Notice that, for a given path w in t, there is at least one pair $\omega(\nu)$, $\omega(\mu)$, $1 \le \nu, \mu \le n-1$, such that $\omega^r(\nu) = \omega^r(\mu)$, $r = X_{n-1}(w) - 1$. Hence $\mathscr{L}_n > X_{n-1}(w)$ and choosing $w = \tilde{w}$, by Lemmas 2, 3, we have that $\lim\inf_n \mathscr{L}_n/\ln n \ge 1/h_2$ a.s. This, together with the upper estimate, shows that $\lim_{n\to\infty} \mathscr{L}_n/\ln n = 1/h_3$ a.s. if $h_3 = h_2$.

Let $h_2 > h_3$. For a fixed r, $\mathscr{L}_n \le r$ iff, for *each* w^{r-1}, $|\{\nu: 1 \le \nu \le n-1, \omega^{r-1}(\nu) = w^{r-1}\}| \le 1$. Thus

$$P(\mathscr{L}_n = r) = \sum \prod_{\nu=1}^{n-1} p(w^{r-1}(\nu)),$$

where the sum is taken over all samples $\{w^{r-1}(1), \cdots, w^{r-1}(n-1)\}$ from S^{r-1} with order and without replacement. Equivalently, this sum is $(n-1)!$ times the coefficient of x^{n-1} in $\prod_{w^{r-1}} (1 + xp(w^{r-1}))$. So,

$$P(\mathscr{L}_n \le r) \le (n-1)!x^{-(n-1)} \prod_{w^{r-1}} [1 + xp(w^{r-1})], \quad x > 0. \tag{2.25}$$

Choose $x = n-1$. Then

$$\begin{aligned} &\ln\Big(\prod_{w^{r-1}} [1 + xp(w^{r-1})]\Big) \\ &= \sum_{w^{r-1}} [(n-1)p(w^{r-1}) - (n-1)^2 p^2(w^{r-1})/2 + O(n^3 p^3(w^{r-1}))] \\ &= (n-1) - (n-1)^2/2 \cdot \sum_{w^{r-1}} p(w^{r-1})^2 + O(n^3 \cdot \sum_{w^{r-1}} p(w^{r-1})^3). \end{aligned} \tag{2.26}$$

Take $r = [(1-\varepsilon)\ln n/h_3]$. Then, by definition of h_2, h_3,

$$\begin{aligned} (n-1)^2 \sum_{w^{r-1}} p(w^{r-1})^2 &= \exp[2\varepsilon \ln n + o(\ln n)], \\ (n-1)\max\{p(w^{r-1})\} &= \exp[\ln n - rh_2 + o(\ln n)] \\ &= \exp[\ln n(1 - (1-\varepsilon)h_2/h_3) + o(\ln n)]. \end{aligned} \tag{2.27}$$

Subsequently,

$$\begin{aligned} &n^3 \cdot \sum_{w^{r-1}} p(w^{r-1})^3 \\ &\le n \max\{p(w^{r-1})\} \cdot n^2 \sum_{w^{r-1}} p(w^{r-1})^2 \\ &= \exp\{\ln n[1 - h_2/h_3 + \varepsilon(2 + h_2/h_3)] + o(\ln n)\} \to 0, \quad n \to \infty, \end{aligned} \tag{2.28}$$

provided that $\varepsilon < (h_2/h_3 - 1)/(2 + h_2/h_3)$ (remember that $h_2 > h_3$). Combining

(2.25)–(2.28) and the Stirling formula, we obtain, for sufficiently small $\varepsilon > 0$ and $n \geq n(\varepsilon)$,

$$P(\mathscr{L}_n \leq r) \leq (n-1)![(n-1)/e]^{-(n-1)}\exp(-n^{\varepsilon}) = O(n^{1/2}\exp(-n^{\varepsilon})).$$

Thus, $\liminf_n \mathscr{L}_n/\ln n \geq 1/h_3$ a.s.

COROLLARY 3. *$\liminf_n L_n/\ln n = 1/h_1$, $\limsup_n L_n/\ln n = 1/h_3$ a.s.*

PROOF. From the description of Rule B, it follows that a.s. (a) $\ell_n \leq L_n \leq \mathscr{L}_n$, (b) $L_n = \ell_n$ if $\ell_{n+1} > \ell_n$, and (c) $L_n = \mathscr{L}_{n+1} - 1$ if $\mathscr{L}_{n+1} > \mathscr{L}_n$. These relations and Lemma 7 imply the statement.

The next lemma completes the study of rule B.

LEMMA 8. *$(P)\lim_{n\to\infty} L_n/\ln n = (P)\lim_{n\to\infty} M_n/\ln n = 1/h$.*

PROOF. It suffices to consider L_n, since $L_n \equiv_{\mathscr{D}} M_n$. Fix $\varepsilon > 0$, $\mu_0 \geq 1$, $\delta \in (1/h_3, \infty)$. Using again the Shannon-McMillan-Breiman Theorem, we can write

$$P(|L_n/\ln n - 1/h| \geq \varepsilon/h) \leq \sum_{r\in D_n} P_{nr} + P(L_n \geq \delta \ln n) + P(\mu_0, \varepsilon),$$

where $\lim_{\mu_0\to\infty} P(\mu_0, \varepsilon) = 0$, and

$$P_{nr} = P(L_n = r,\ |r^{-1}\ln[1/p(\omega^r(n))] - h| \leq \varepsilon^2 h \text{ if } \mu_0 \leq r),$$

$$D_n = \{r \geq 1 : |r/\ln n - 1/h| \geq \varepsilon/h \text{ and } r \leq \delta \ln n\}.$$

Since $P(L_n \geq \delta \ln n) \to 0$, (see Corollary 3), we have to prove only that $\sum_{r\in D_n} P_{nr} = o(1)$. To this end, notice that if $L_n = r$ then $\omega^r(n) \neq \omega^r(\nu)$, $1 \leq \nu \leq n-1$, but $\omega^{r-1}(n) = \omega^{r-1}(\nu)$ for at least one such ν. This shows that

$$P_{nr} \leq (n-2)\sum_{w^r} p(w^r)p(w^{r-1})(1-p(w^r))^{n-3}, \quad n \geq 3, \tag{2.29}$$

where w^r satisfies

$$|r^{-1}\ln(1/p(w^r)) - h| \leq \varepsilon^2 h, \quad \text{if} \quad \mu_0 \leq r. \tag{2.30}$$

Since $\sum_{w^r} p(w^{r-1}) = m$, it follows after simple estimations from (2.29), (2.30) that, for $\varepsilon \leq \varepsilon_0$, $P_{nr} \leq \exp(-n^{\varepsilon/2})$ if $r \leq [(1-\varepsilon)\ln n/h]$ and $P_{nr} \leq n^{-\varepsilon/2}$ if $r \geq [(1+\varepsilon)\ln n/h]$. Thus,

$$\sum_{r\in D_n} P_{nr} \leq \text{const} \ln n[\exp(-n^{\varepsilon/2}) + n^{-\varepsilon/2}] = o(1).$$

Finally, *Rule C*. Let $\{t'_n\}$ be the correspondent sequence of trees; t'_n is obtained by compressing the tree t_n constructed according to Rule B, (see Introduction for details). Let L'_n, M'_n, H'_n be defined for t'_n as the counterparts of L_n, M_n, H_n for t_n.

LEMMA 9. (a) *$\lim_{n\to\infty} H'_n/\ln n = 1/h_2$, a.s.*; (b) *$\limsup_n L'_n/\ln n = 1/h_2$, $\liminf_n L'_n/\ln n = 1/h_1$ a.s.*

LEMMA 10. $(P)\lim_{n\to\infty} L_n'/\ln n = (P)\lim_{n\to\infty} M_n'/\ln n = 1/h$.

Since the proofs of these lemmas do not contain significant new elements, in comparison with the Rules A, B, we restrict ourselves to just several remarks.

First of all, the condition (1.2)—the ρ-condition—implies that each feasible node of the infinite tree t has at least two feasible direct descendants. This shows that (a.s.) $L_n' \geq \ell_n - 1$ and $L_n' = \ell_n$ whenever $\ell_{n+1} > \ell_n$, hence a.s. $\lim \inf_n L_n'/\ln n = 1/h_1$ (Lemma 7). The estimate $\lim \inf_n H_n'/\ln n \geq 1/h_2$ (a.s.) is proven (predictably) by considering the path $\tilde{w}$ (see Lemma 2) and using the ρ-condition. To get the estimate $\lim \sup_n H_n'/\ln n \leq 1/h_2$ (a.s.), we fix k, $r \geq 1$ and observe that if $H_n' \geq k + r - 1$ then there are some sequences $\omega(\nu_1), \cdots, \omega(\nu_k)$, $1 \leq \nu_1 < \nu_2 < \cdots < \nu_k \leq n$, such that $\omega^r(\nu_1) = \cdots = \omega^r(\nu_k)$. Subsequently, $\lim \sup_n H_n'/\ln n \leq 1/h(k)$, where $h(k) = \lim_{r\to\infty}(kr)^{-1}\ln(1/E(p(\omega^r)^{k-1}))$, and it remains to use the fact that $h(k) \to h_2$ as $k \to \infty$.

To prove Lemma 10, it suffices to show that $P(\tilde{L}_n/\ln n \leq (1 - \varepsilon)/h) = o(1)$, where $\tilde{L}_n$ is the length of the initial segment of the path ω^n in t_n such that all its nodes, except the last one, are of branching type. (This portion of the path ω^n will not change after compressing the tree t_n into the tree t_n', whence $L_n' \geq \tilde{L}_n$.) The proof uses again the ρ-condition.

Acknowledgement. I wish to thank the participants of the probability seminar at the Ohio State University for a useful discussion of the results of this paper. I am very grateful to the editor and a referee for many critical comments and specific suggestions, which helped me to improve the presentation of the paper. In particular, I am indebted to the referee for his detailed suggestions to use the random times $T_k(w)$ in the proofs of Lemmas 3, 4—these times were used previously in the proofs of Lemmas 5, 6—and to make a short cut in the proof of Lemma 5.

REFERENCES

[1] BILLINGSLEY, P. (1965). *Ergodic theory and information.* Wiley, New York.

[2] BLUM, J. R., HANSON, D. L. and KOOPMANS, L. H. (1963). On the strong law of large numbers for a class of stochastic processes. *Z. Wahrsch. verw. Gebiete* **2** 1–11.

[3] EDGAR, G. A., MILLET, A. and SUCHESTON, L. (1982). On compactness and optimality of stopping times. "Martingale Theory in Harmonic Analysis and Banach Spaces," Proceedings, Cleveland 1981, Springer-Verlag, 36–61.

[4] VAN EMDEN, M. H. (1970). Increasing the efficiency of the Quicksort. *Comm. ACM* **13** 563–567.

[5] KINGMAN, J. F. C. (1973). Subadditive ergodic theory. *Ann. Probab.* **1** 883–899.

[6] KNUTH, D. E. (1973). The art of computer programming 3. Addison-Wesley, New York.

[7] KONHEIM, A. G. and NEWMAN, D. J. (1973). A note on growing binary trees. *Discrete Math.* **4** 57–63.

[8] MAHMOUD, H. and PITTEL, B. (1984). On the most probable shape of a search tree grown from a random permutation. *SIAM J. Algebraic Discrete Methods* **5** 69–81.

[9] PITTEL, B. (1984). Paths in a random digital tree: limiting distributions. *J. Appl. Probab.* (submitted).

[10] PITTEL, B. (1984). On growing random binary trees. *J. Math. Anal. Appl.* **103** 461–480.

[11] ROBSON, J. M. (1979). The height of binary search trees. *Austral. Comput. J.* **11** 151–153.

[12] ROMANOVSKIJ, J. V. (1967). Optimization of stationary control in a discrete deterministic process. *Kibernetika* (*Kiev*) **2** 66–78. (In Russian).

Added in proof.

[13] DEVROYE, L. (1982). A note on the average depth of tries. *Computing* **28** 367–371.

[14] DEVROYE, L. A probabilistic analysis of the height of tries and of the complexity of triesort. (Unpublished).

[15] FLAJOLET, P. and STEYAERT, J. M. (1982). A branching process arising in dynamic hashing, trie searching and polynomial factorization. *Proceedings of the Ninth ICALP Colloquium, Lecture Notes in Computer Science* **140** 239–251.

[16] MENDELSON, H. (1982). Analysis of extendible hashing. *IEEE Trans. Software Engrg.* **SE-8** 611–619.

[17] REGNIER, M. (1982). On the average height of trees in digital search and dynamic hashing. *Inform. Process. Letters* **13** 64–66.

[18] YAO, A. C. (1980). A note on the analysis of extendible hashing. *Inform. Process. Letters* **11** 84–86.

DEPARTMENT OF MATHEMATICS
THE OHIO STATE UNIVERSITY
COLUMBUS, OHIO 43210