LIMITING DISTRIBUTION FOR THE DEPTH IN PATRICIA TRIES

May 17, 2000

Bonita Rais*
Dept. of Computer Science
Purdue University
W. Lafayette, IN 47907
U.S.A.

Philippe Jacquet†
INRIA
Rocquencourt
78153 Le Chesnay Cedex
France

Wojciech Szpankowski‡
Dept. of Computer Science
Purdue University
W. Lafayette, IN 47907
U.S.A.

Abstract

Digital tries occur in a variety of computer and communication algorithms including symbolic manipulations, compiling, comparison based searching and sorting, digital retrieval techniques, algorithms on strings, file systems, codes and communication protocols. We study the depth of the PATRICIA trie in a probabilistic framework. The PATRICIA trie is a digital tree in which nodes that would otherwise have only one branch have been collapsed into nodes having more than one branch. Because of this characteristic, the depth of the PATRICIA trie provides a measure on the compression of the keys stored in the trie. Here we consider n independent keys that are random strings of symbols from a V-ary alphabet. This model is known as the Bernoulli model. We show that the depth in the asymmetric case (i.e., symbols from the alphabet do *not* occur with the same probability) is asymptotically *normally* distributed. In the symmetric case, which surprisingly proved to be more difficult, we present the limiting generating function and the limiting distribution. In either case, our results lead us to the conclusion that the PATRICIA trie is with high probability a well-balanced tree.

*This research was in part supported by AFOSR grant 90-0107 and by NSF grant CCR-8900305.

†This research was primary supported by NATO Collaborative Grant 0057/89.

‡This author's research was supported in part by NATO Collaborative grant 00571/89, and in part by AFOSR grant 90-0107, by NSF grants CCR-8900305 and INT-8912631, and by grant R01 LM05118 from the National Library of Medicine.

1. INTRODUCTION

This paper establishes the limiting distribution for the depth of keys in a PATRICIA trie. A PATRICIA trie is a variation of the trie, a well-known tree structure, which is a frequently used data structure in many applications of computer science and telecommunications. These applications include dynamic hashing [5, 7], data compression [1], [2], pattern matching [1], and conflict resolution algorithms for broadcast communications [3, 13].

The *depth* of a leaf in a trie, also known as *depth of insertion* or *successful search time*, is the number of internal nodes on the path from the root of the trie to the leaf. It is of particular interest since it provides useful information in many applications. For example, when keys are stored in the leaves of the trie, the depth of a key gives an estimate of the search time for that key in searching and sorting algorithms [20]. Depth also gives the length of a conflict resolution session for tree-based communication protocols or, in compression algorithms, provides the length of a substring that may be copied or compressed [1].

The primary purpose of a trie is to store a set $\mathcal{S}$ of keys. Each key $X = x_1x_2x_3\ldots$ is a finite or infinite string of symbols taken from a finite alphabet $\mathcal{A} = \{\omega_1, \ldots, \omega_V\}$. The trie over $\mathcal{S}$ is built recursively as follows. For $|\mathcal{S}| = 0$, the trie is, of course, empty. For $|\mathcal{S}| = 1$, $trie(\mathcal{S})$ is a single node. If $|\mathcal{S}| > 1$, $\mathcal{S}$ is split into V subsets $\mathcal{S}_1, \mathcal{S}_2, .., \mathcal{S}_V$ so that a key is in $\mathcal{S}_j$ if its first symbol is ω_j. The tries $trie(\mathcal{S}_1), trie(\mathcal{S}_2), ..., trie(\mathcal{S}_V)$ are constructed in the same way except that at the k-th step, the splitting of sets is based on the k-th symbol. They are then connected from their respective roots to a single node to create $trie(\mathcal{S})$. A trie may have nodes with only one branch leading from it and it is this waste of space which the *PATRICIA trie* eliminates by collapsing one-way branches into a single node. Thus the depth of a key in a PATRICIA trie may be less than that of the same key in a regular trie.

Consider the following example. Let $\mathcal{A} = \{0, 1, 2\}$ and $\mathcal{S} = \{A, B, C, D, E, F\}$ as defined in Figure 1. Then, the PATRICIA trie built over the set $\mathcal{S}$ is shown in Figure 1. We can also vary both the trie and PATRICIA trie to a more general structure by allowing a leaf to hold at most b keys [7, 12]. This is the case in algorithms for extendible hashing in which the capacity of a page or other storage unit is b.

Tries have been analyzed by many authors under various probabilistic models, most having independent keys [7, 12, 25, 16, 18, 22]. Frequently the symbols of $\mathcal{A}$ are also independent with $\Pr\{x_j = \omega_i\} = p_i$ for any $j = 1, 2, \ldots$ where $\sum_{i=1}^{V} p_i = 1$, and we do adopt these assumptions in this paper. Such a model is known as the *Bernoulli model* provided the number of keys n is fixed. If $p_1 = p_2 = \cdots = p_V = 1/V$, then the distribution of symbols is *symmetric*, otherwise it is *asymmetric*. Studies of the binary symmetric model

$A = 000$
$B = 010$
$C = 012$
$D = 100$
$E = 200$
$F = 221$

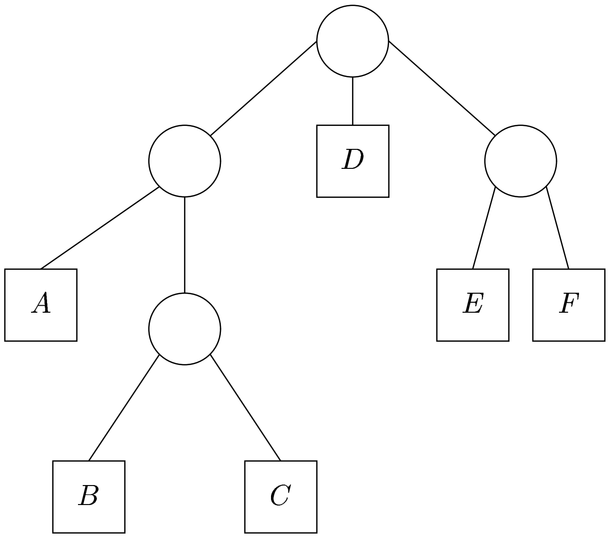

Figure 1: Example of a 3-ary digital trie with $n = 6$.

have been carried out by Knuth [20], Flajolet [7], and Kirschenhofer and Prodinger [16]. The variance of the depth was also obtained in [16] (see also [18]). The limiting distribution for the depth of a regular trie was obtained independently by Jacquet and Régnier [12] (limiting distribution), Pittel [22] (limiting distribution), and Szpankowski [25] (all moments for the asymmetric independent model). The limiting distribution of depth in tries using a Markovian dependency among symbols is presented by Jacquet and Szpankowski in [14]. Pittel [21] has proved convergence *in probability* for the depth for a more general dependency among symbols (i.e., mixing stationary sequences).

PATRICIA tries have not been studied as extensively but the moments of the successful search for the asymmetric model (see also [16], [20] for the binary symmetric case), and moments of the unsuccessful search for binary symmetric model have been obtained in Szpankowski [26], and the variance of the external path length by Kirschenhofer *et al.* [18]. Pittel [22] provided the leading term in the *almost sure* convergence for the depth and the height. Also, Devroye [4] has obtained results for depth and height of PATRICIA tries under a model in which the keys are random variables with a continuous density f on [0,1]. In this paper we obtain the convergence *in distribution* of the depth in the Bernoulli model. From the probabilistic viewpoint, this is the best and the strongest possible result regarding typical behavior of the depth in the PATRICIA.

Assuming independence among keys as well as symbols, our aim is to analyze the limiting distribution of the depth for a PATRICIA trie. In order to accomplish this we will use the

Poisson transform and study the *Poisson model* in which the number of keys in the trie follows a Poisson distribution with parameter n. After deriving results for this model, we will make use of the inverse Poisson transform to get the results for our Bernoulli model. In either model, the distribution of the depth in the asymmetric case is asymptotically a Gaussian-like distribution. In our analysis of this, we use properties of the Mellin transform and follow the method suggested by Jacquet and Régnier in [12]. However, in the symmetric case where we obtain very different results, another approach is necessary.

The paper is organized as follows. The next section will give all necessary definitions and tools not yet defined. It will also give a statement of the main results and consequences of our findings. In the last section we will prove all the results given in the second section.

2. MAIN RESULTS

Before making precise statements of our results, it is necessary to give some definitions and notation. We let the random variable D_n be the depth of a randomly chosen key in a PATRICIA trie holding n keys. Then, $\Pr\{D_n = k\}$ is the probability that the depth of a key is k when the PATRICIA trie holds n keys, that is, k is the length of the path from the root to a randomly selected key. Then, $D_n(u)$ is the *ordinary generating function*, and $D(z,u)$ is the *Poisson generating function* where

$$D_n(u) = \sum_{k=0}^{\infty} \Pr\{D_n = k\} u^k$$

$$D(z,u) = \sum_{n=0}^{\infty} D_n(u) \frac{z^n}{n!} e^{-z}.$$

The first function, $D_n(u)$, is used in the Bernoulli model where the number of keys is fixed at n, and the probability of generating the i-th symbol from the alphabet $\mathcal{A}$ is equal to p_i for $1 \le i \le V$. When z is real, $D(z,u)$ is the generating function for the Poisson model in which the number of keys in the trie follows a Poisson distribution with parameter z. These functions are well-defined for any complex numbers z and u such that $|u| < 1$. However, in our analysis we need to extend analytically the functions to $|u| < 1 + \delta$ for some $\delta > 0$. When we replace u by e^t where t is a complex number, we obtain the *characteristic function* of the respective distribution.

We summarize the main results of our study in the following theorems. The first theorem gives a complete characterization of the asymptotic behavior of the depth in a PATRICIA trie under the Bernoulli model with an asymmetric alphabet.

THEOREM 1. *Consider the* **asymmetric** *model of PATRICIA tries described above. Then,*

(i) *For large n the average depth ED_n of a PATRICIA trie is*

$$ED_n = \frac{1}{H}\log n + O(1)$$

and the variance $varD_n$ of the depth is

$$varD_n = \frac{H^2 - H_2}{H^3}\log n + O(1)$$

where $H = -\sum_{i=1}^{V} p_i \log p_i$ is the entropy of the alphabet, and $H_2 = \sum_{i=1}^{V} p_i \log^2 p_i$.

(ii) *The random variable $\left(\frac{D_n - ED_n}{\sqrt{varD_n}}\right)$ is asymptotically normal with mean zero and variance one, that is,*

$$\lim_{n\to\infty} Pr\{D_n \le ED_n + x\sqrt{varD_n}\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt.$$

In addition, for all integer m the following convergence in moments holds

$$E\left(\frac{D_n - ED_n}{\sqrt{varD_n}}\right)^m \to \begin{cases} \frac{m!}{2^{m/2}(m/2)!} & m \;\; even \\ 0 & m \;\; odd \end{cases}$$

where the right-hand side of the above presents moments of the normal distribution. ∎

The second theorem provides the limiting generating function as well as the limiting distribution for the symmetric case. Proofs for both theorems are presented in the next section.

THEOREM 2. *Consider the* **symmetric** *model of PATRICIA tries described above.*
(i) `Limiting Generating Function`. *The limiting generating function $D_n(u)$ for the depth in a PATRICIA trie for large n is*

$$D_n(u) = u^{\log_V n}\exp\{\frac{\alpha(u)}{\log V} - \frac{\log u}{2} + \frac{1}{\log V}\sum_{k\neq 0} g^*(s_k,u)n^{-s_k}\} + O(\frac{1}{\sqrt{n}}) \tag{1}$$

where $s_k = 2\pi ik/\log V$, $k \neq 0$,

$$\alpha(u) = (1-u)\int_0^\infty \frac{e^{-z}\log z}{e^{-z} + u(1-e^{-z})}dz \tag{2}$$

and

$$g^*(s,u) = \int_0^\infty \log[e^{-z} + u(1-e^{-z})]z^{s-1}dz \tag{3}$$

where $\alpha(u)$ is defined for $u \notin (-\infty, 0]$.

(ii) `Limiting Distribution`. *Let* $k_0 = \lfloor \log_V n \rfloor$ *and* $q = e^{-n/V^{k_0}}$. *Toss a* k_0*-long sequence of biased coins with the probability of heads (success) for the* j*-th coin equal to* q^{V^j}. *Let* X_1 *be the number of successes. Toss an infinite sequence of biased coins with probability of heads (success) for the* j*-th coin equal to* $q^{V^{-j}}$. *Let* X_2 *be the number of failures. Then asymptotically*

$$\Pr\{D_n - \log_V n \le x\} = \Pr\{X_2 - X_1 \le x\} + O(1/\sqrt{n}) \ . \tag{4}$$

and X_1 *and* X_2 *are independent. Moreover, this characterization leads to*

$$\lim_{n\to\infty} \Pr\{D_n - \log_V n \le x\} = f^P(V^{-x})$$

where

$$f^P(y) = \sum_{m=0}^{\infty} B_m(yV^{m-1}) e^{-yV^{m-1}} \tag{5}$$

and

$$B_m(z) = e^{-z} \sum_{\substack{J\subset\{1,2,\ldots\} \\ |J|=m}} \prod_{j\in J} (e^{V^{-j}(V-1)z} - 1)$$

In particular, $B_0(z) = e^{-z}$ *and* $B_1(z) = e^{-z}\sum_{j=1}^{\infty}(e^{V^{-j}(V-1)z} - 1)$. ■

Remarks.

(i) *Symmetric PATRICIA*. Although the limiting distribution is computed here for the first time, it was shown previously [20] that for large n the average ED_n depth of a PATRICIA trie is

$$ED_n = \log_V n + \frac{\gamma}{\log V} + \log_V(1 - 1/V) + \frac{1}{2} + P_1(\log n)$$

and the variance $varD_n$ of the depth is constant, more precisely [16], [26]

$$varD_n = \frac{\pi^2}{6\log^2 V} + \frac{1}{12} - \frac{2}{\log V} \log \prod_{j=1}^{\infty}\left(1 + \frac{1}{V^j}\right) + P_2(\log_V n)$$

where $\gamma = 0.577$ is the Euler constant, and $P_1(\log_V n)$ and $P_2(\log_V n)$ are fluctuating functions of very small amplitude. We also obtain these results from the limiting generating function of (1). These formulas follow from our limiting distribution (4). Finally, we notice also that

$$\lim_{n\to\infty} \Pr\{D_n - \log_V n \le x\} = F_1(x) \cdot F_2(x)$$

where

$$F_1(x) = \exp(-V^{-x}) \quad , \quad F_2(x) = \sum_{m=0}^{\infty} B_m(V^{-x+m-1}) \exp(-V^{-x}(V^{m-1} - 1)) \ .$$

The function $F_1(x)$ is a distribution function. In fact, it is known as the standard extreme distribution (i.e., the so called *double exponential* distribution $\exp(-e^{-x})$). Let Z be a random variable distributed according to $F_1(x)$. Then, $EZ = \gamma/\log V - 1/2$ and $var Z = \pi^2/(6\log^2 V) + 1/12$, where the terms $1/2$ and $1/12$ are Sheppard's corrections for continuity (cf. [15]) of the average value and the variance respectively. In fact, as discussed in Remark (iii) below, $\log_V n + Z$ is distributed as the depth in regular tries.

(ii) *Aldous' representation for the symmetric case.* For completeness, we include an observation of D. Aldous who points out that the problem of depth in the symmetric PATRICIA trie can be alternatively described in the following way. Consider an infinite row of boxes, and box i receives a Poisson number of balls with mean $V^{-i}(1-1/V)$. Let D_n^T be the number of the rightmost box containing a ball and C be the number of empty boxes to the left of box D_n^T. In terms of tries, D_n^T gives the depth of a key in a regular symmetric trie and C gives the number of non-branching nodes on the path from the root to the key (in the Poisson model). Thus the depth in a PATRICIA trie D_n^P is given by $D_n^T - C$. Unfortunately, the random variables D_n^T and C are *dependent*, and therefore this representation is rather useless for the limiting distribution analysis. However, it is interesting to note that in the symmetric binary case, $\Pr\{C = d\} = 2^{-d}$ as is shown by Knuth in *American Math Monthly* (vol. 94, p.189). This distribution of C was independently discovered by Pittel and Rubin [23].

(iii) *Comparison with regular tries.* When considering either the case of a symmetric or asymmetric alphabet, we can make the following observations. Although the expected depth of either the regular or PATRICIA trie is $\log n/H + O(1)$, the constant is not the same. Examination of this constant shows that the expected depth of a regular trie is greater than that of a PATRICIA trie. More importantly, the variance of the depth for a PATRICIA is smaller than for regular tries. This leads us to conclude that the PATRICIA trie is a better balanced trie than the regular trie.

We can offer further support of this claim in the symmetric case. In particular, as shown in [26], the difference in the variance for small alphabet is significant. For example, for binary regular tries we have $var D_n^T = 3.507\ldots$ while for binary PATRICIA $var D_n^P = 1.000\ldots$. In fact, as proved in [17] the variance $var D_n^P = 1.000000000000\ldots$ (twelve zeros). We note also that the difference becomes smaller for larger values of V, as expected. We can also compare the limiting distribution for the depth in a PATRICIA trie with that for a regular trie. From [12, 22] we know that the limiting distribution for a regular trie under

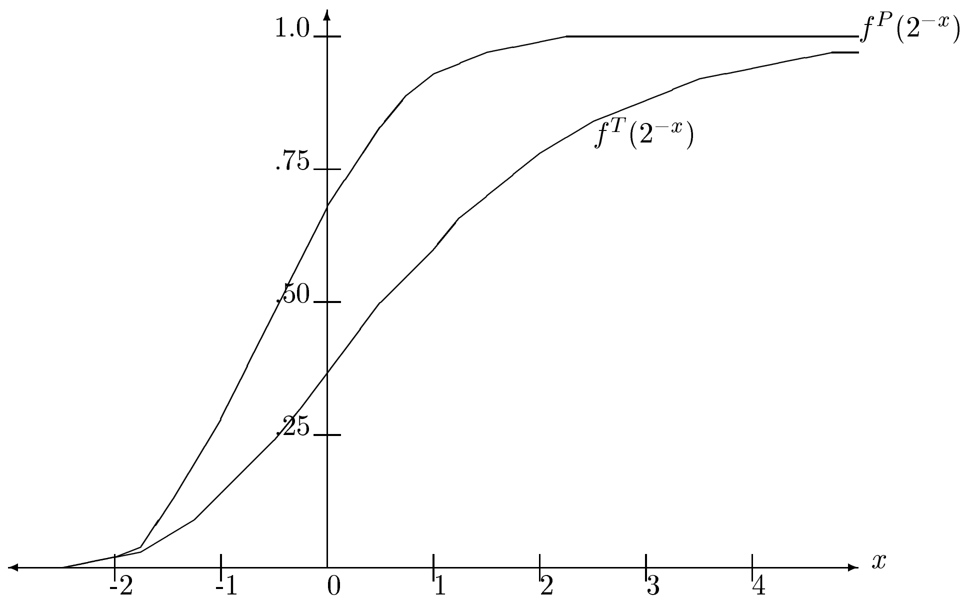

Figure 2: Comparison of distributions of tries and PATRICIA tries for $V = 2$.

the symmetric alphabet is given by

$$\lim_{n\to\infty} \Pr\{D_n - \log_V n \leq x\} = e^{-V^{-x}} \ . \tag{6}$$

A simple proof of this is given in Pittel [22] which proceeds as follows. First observe that $\Pr\{D_n \leq k\} = (1 - V^{-k})^{n-1}$. By letting $k = x + \log_V n$ and taking the limit as $n \to \infty$, Pittel obtains the limiting distribution $f^T(V^{-x})$ as given in (6), where $f^T(x) = e^{-x}$.

As easy to see from Figure 2 – which compares $f^T(2^{-x})$ and $f^P(2^{-x})$ – the probability of the depth of a randomly chosen key being at most $\log_V n + x$ is greater in a PATRICIA trie than in a regular trie. Since the mean depth is $\log_V n + O(1)$ for both structures, this supports the conclusion that the PATRICIA trie is better balanced than the regular trie.

(iv) *How well is the PATRICIA balanced* ? A tree built over a V-ary alphabet is well-balanced if (a) the average depth of a key is $\log_V n + O(1)$, (b) the variance of the depth is significantly smaller than the average depth, and (c) large derivations from the average value are very unlikely. Many algorithms using trees need balanced trees (e.g., the extendible hashing algorithm [5]) to run efficiently, so often times these algorithms include a costly rebalancing step. This rebalancing operation is customarily justified by the worst-case analysis. But our average case analysis shows that this costly operation seems to be unnecessary since, with high probability, the tree is already balanced, that is, a random shape of a PATRICIA trie resembles the shape of the well-balanced structure of a complete tree [2]. In the symmetric case, we know that the expected depth of a PATRICIA trie is $\log_V n + O(1)$ and its variance is $O(1)$, thus we can expect that the PATRICIA trie is well-balanced. In the asymmetric case, we show that the limiting distribution for the depth is normal with mean $\frac{\log n}{H} + O(1)$ and variance $\frac{H^2 - H_2}{H^3} \log n + O(1)$. The coefficient $1/H$ in the mean shows that the more asymmetric the distribution of the symbols is, the more skew the PATRICIA trie is. However, the standard deviation is $O(\sqrt{\log n})$, so the PATRICIA trie is still, on average, balanced. Efficiently preprocessing the asymmetric alphabet to obtain a more symmetric alphabet will improve the balance of the PATRICIA trie.

(v) *Poisson model.* In the proofs of our theorems, we will also establish similar convergence results for the Poisson model, in which the number of keys is not fixed but rather a random variable distributed according to Poisson law. That is, in the asymmetric case, the depth of a key in a PATRICIA trie with Poisson number of records, once normalized and centered, is asymptotically normal.

3. ANALYSIS

The primary focus of this section is the proof of our results. As mentioned earlier, different approaches are necessary to compute the limiting distributions for the depth of a PATRICIA trie in the symmetric and asymmetric cases. Before giving details of our analysis, we briefly identify tools that are useful in manipulating the generating functions defined in the previous section in both the symmetric and asymmetric cases. Then we will prove Theorems 1 and 2 in the following subsections.

An important tool that will enable us to obtain asymptotic results is the *Mellin transform*, an integral transform from complex analysis which is defined as follows [8] (cf. also [10]). Let $F(x)$ be a piecewise continuous function on the interval $[0,\infty)$. If $F(x) = O(x^\alpha)$ for $x \to 0$, and $F(x) = O(x^\beta)$ for $x \to \infty$, then the Mellin transform of $F(x)$, denoted $F^*(s)$, is defined for any complex number s in the strip $-\alpha < \Re(s) < -\beta$ and

$$F^*(s) = \int_0^\infty F(x) x^{s-1} dx.$$

The importance of the Mellin transform is that it provides information concerning the asymptotic behavior of a function $F(x)$ around 0 and ∞ through the poles of $F^*(s)$. In fact, the asymptotic expansion of $F(x)$ is obtained directly from the poles of its transform [8]:

$$F(x) \sim \pm \sum_{\lambda \in \mathcal{H}} Res\{F^*(s) x^{-s}, s = \lambda\}$$

where $Res\{f(s),\, s=\lambda\}$ is the residue of $f(s)$ at $s=\lambda$ and $\mathcal{H}$ is the set of poles of $F^*(s)$ in the left (right) half-plane giving the asymptotic expansion as $x \to 0$ ($x \to \infty$). This is obtained using the inverse Mellin transform [10]

$$F(x) = \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} F^*(s) x^{-s} ds \ ,$$

where $c \in (-\alpha, -\beta)$, and residue theory. Provided $F^*(s)$ is small at $\pm i\infty$ and has only isolated singularities, for the case in which $x \to \infty$ we can close the contour to the right and derive (cf. [8], [10])

$$F(x) = -\sum_{\lambda} Res\{F^*(s)x^{-s}, s=\lambda\} + O(x^{-M})$$

where the sum is taken over all poles λ such that $-\beta \leq \Re(\lambda) \leq M$ for any M.

A second tool of great importance that will allow us to extract results for $D_n(u)$ from the results for $D(z,u)$, is the Poisson generating function. It is derived from Cauchy's integral formula which says that [10]

$$D_n(u) = \frac{n!}{2\pi i} \oint D(z,u) e^z \frac{dz}{z^{n+1}} \tag{7}$$

where the integration is taken over a circle of arbitrary radius centered at the origin. The following important result is derived from this formula when the radius of the circle is chosen to be n.

Depoissonization Lemma. *Let S_θ be a cone $S_\theta = \{z : |\arg z| < \theta,\ 0 < \theta < \pi/2\}$. If for $z \in S_\theta$ and $z \to \infty$ the following holds for all u in a compact set $\mathcal{U}$*

$$|D(z,u)| < \beta_1 |z|^\epsilon \tag{8}$$

for some β_1, $\epsilon > 0$, and for $z \notin S_\theta$ and all $u \in \mathcal{U}$

$$|D(z,u)e^z| \leq \beta_2 |z|^\epsilon e^{\alpha|z|} \tag{9}$$

for some $0 < \alpha < 1$ and a constant $\beta_2 > 0$, then for large n uniformly in $u \in \mathcal{U}$ the generating function $D_n(u)$ satisfies

$$D_n(u) = D(n,u) + O(n^{\epsilon - 1/2}) \tag{10}$$

with $\epsilon < 1/2$.

Proof. We proceed along the lines of [13] with some necessary modifications. We apply Cauchy's formula (7). By the change of variable $z = ne^{it}$ we have

$$D_n(u) = \frac{n!}{n^n e^{-n}\sqrt{2\pi n}}\sqrt{\frac{n}{2\pi}}\int_{-\theta}^{\theta} D(ne^{it}, u)\exp(n(e^{it} - it - 1))dt + O(n^{\varepsilon} e^{-(1-\alpha)n}) \ , \quad (11)$$

where the last term of the above is a simple consequence of our condition (9). We note that by Stirling's formula $n!/(n^n e^{-n}\sqrt{2\pi n}) = 1 + O(1/n)$. Hence, (11) can be reduced to $D_n(u) = (1 + O(1/n))I_n(u) + O(n^{\varepsilon} e^{-(1-\alpha)n})$, where

$$I_n(u) = \sqrt{\frac{n}{2\pi}}\int_{-\theta}^{\theta} D(ne^{it}, u)\exp(n(e^{it} - it - 1))dt \ .$$

Now, we evaluate the integral $I_n(u)$. After introducing a new change of variable $x = t\sqrt{n}$ we obtain

$$I_n(u) = \frac{1}{\sqrt{2\pi}}\int_{-\theta}^{\theta} D(ne^{ix/\sqrt{n}}, u)\exp(n(e^{ix/\sqrt{n}} - ix/\sqrt{n} - 1))dx \ .$$

The following expansions are easy to derive

$$\exp(n(e^{ix/\sqrt{n}} - ix/\sqrt{n} - 1)) = e^{-x^2/2}\left(1 - \frac{ix^3}{6\sqrt{n}} + O(1/n)\right) \ ,$$

and

$$D(ne^{ix/\sqrt{n}}, u) = D(n, u) + n(e^{ix/\sqrt{n}} - 1)D'(n, u) + O(1/n)$$

where $D'(n, u)$ is the first derivative of $D(z, u)$ with respect to z at $z = n$. But using again the Cauchy's formula in the form (cf. [10])

$$D'(n, u) = \frac{1}{\pi i}\oint \frac{D(z, u)}{(z - n)^2}dz$$

and our first condition (8), we finally prove that $D'(n, u) = O(n^{\varepsilon - 1})$ uniformly in $u \in \mathcal{U}$. The leads to (10) by noting that $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}dx = 1$ (cf. [6]). ∎

This lemma gives us the conditions necessary to transform our Poisson model results into those for the Bernoulli model, so we call it *depoissonization lemma* (in fact, (10) can be called the inverse Poisson transform – see also [9] and [11]).

3.1 Asymmetric Case

In this section we will adopt the approach of Jacquet and Régnier [12], making necessary changes required by the PATRICIA trie. At first, we give a rough plan of our analysis that will lead to the proof of our main results. To get the limiting distribution for depth in a PATRICIA trie under our Bernoulli model, we will begin by deriving its probability generating function, $D_n(u)$. Unfortunately, it is not easy to derive the limiting distribution directly. However, we use the Poisson transform to compute the generating function $D(z,u)$ for the Poisson model in which the number of keys follows a Poisson distribution with parameter z. This model is easier to analyze, since the generating function for the Poisson model has a closed form. This is not so in the Bernoulli model.

Since we are interested in the asymptotic behavior of the Poisson probability generating function $D(z,u)$, we make use of the Mellin transform. We also replace u by e^t where t is complex. This will guarantee that each generating function in our sequence is analytic. The limit function of a sequence of analytic function is again analytic, so all its derivatives are well-defined. In this way, we will also get convergence in moments.

We then show the following limit where $\mu(z)$ is the mean and $\sigma(z)$ is the standard deviation for the Poisson model with parameter z, and $\tau = i\nu$ for any real number ν:

$$\lim_{z\to\infty} e^{-\tau\mu(z)/\sigma(z)} D(z, e^{\tau/\sigma(z)}) = e^{\tau^2/2} \ , \tag{12}$$

which is a modification of Goncharov's theorem (cf. [19] Chap. 1.2.10, Ex. 13). By proving (12) we shall show that the depth in the Poisson model is asymptotically normal with mean $\mu(z)$ and variance $\sigma^2(z)$. The next step is to extract from (12) information about the Bernoulli model, that is to *depoissonizate* the above formula. We do this by applying the *depoissonization* lemma to (12) and we obtain

$$\lim_{n\to\infty} e^{-\tau\mu_n/\sigma_n} D_n(e^{\tau/\sigma_n}) = e^{\tau^2/2}$$

for all $\tau = i\nu$ and $-\infty < \nu < \infty$, where $D_n(u)$ is the generating function of the depth D_n. But, the above is exactly the Goncharov's theorem (cf. [19], Chap. 1.2.10, Ex. 13), which states that a sequence of random variables D_n with mean μ_n and standard deviation σ_n approaches a normal distribution if the above holds. This way we shall prove Theorem 1.

Having this plan in mind, we present below some more details. To simplify our notation, hereafter we will assume a binary alphabet, noting that our derivation extends easily to any finite alphabet. We denote the probabilities of the symbols ω_1 and ω_2 as p and q, where $p + q = 1$.

Let $\mathcal{J}_n$ be the set of all possible PATRICIA tries of n keys from the alphabet $\mathcal{A}$, and let $\mathcal{T}$ be a particular trie from $\mathcal{J}_n$. If $S^k_{\mathcal{T}}$ denotes the number of keys at depth k in $\mathcal{T}$, then the generating function associated with $\mathcal{T}$ is given by

$$S_{\mathcal{T}}(u) = \sum_{k=0}^{\infty} S^k_{\mathcal{T}} u^k \ .$$

Note that the sum is actually finite since the maximum depth in a PATRICIA trie with n keys is $n-1$. Clearly, the following statements are true when the left subtree $\mathcal{T}_\alpha$ and right subtree $\mathcal{T}_\beta$ hold k and $n-k$ keys, respectively, and $\delta_{j,k}$ is the Kronecker delta (i.e., $\delta_{j,k}=1$, if $j=k$, $\delta_{j,k}=0$ otherwise)

$$S_{\mathcal{T}}(u) = n, \quad n \le b$$

$$S_{\mathcal{T}}(u) = u\{S_{\mathcal{T}_\alpha}(u) + S_{\mathcal{T}_\beta}(u)\} + (1-u)(\delta_{k,n} + \delta_{n-k,n})S_{\mathcal{T}}(u), \quad n > b.$$

We note that the above recurrence holds for a particular tree $\mathcal{T}$ in $\mathcal{J}_n$. The leading factor u must be present since the depth of a key in either $\mathcal{T}_\alpha$ or in $\mathcal{T}_\beta$ is one less than its depth in the trie $\mathcal{T}$. The second term avoids one-way branching. For example, if $k=0$, then the right branch is a one-way branch and $S_{\mathcal{T}}(u) = uS_{\mathcal{T}}(u) + (1-u)S_{\mathcal{T}}(u)$, which means that the subtree begins at the root of $\mathcal{T}$.

Averaging $S_{\mathcal{T}}(u)$ over all tries $\mathcal{T}$ in $\mathcal{J}_n$, we derive a new generating function $S_n(u) = E(S_{\mathcal{T}}(u))$. Of course, for $n \le b$ we have $S_n(u) = n$. Otherwise,

$$S_n(u) = u\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}[S_k(u) + S_{n-k}(u)] - (u-1)(p^n+q^n)S_n(u) \ . \tag{13}$$

Here the sum in the first term ranges over k from 0 to n. We know that in a PATRICIA trie, there are no one-way branches, thus we must subtract those terms from the sum. However, it is possible that all n keys begin with the same symbol from $\mathcal{A}$. This occurs with probability p^n+q^n, and so we add this term. The last term in (13) is what makes the analysis of the depth of the PATRICIA trie different from that of the regular trie.

Define $S(z,u) = \sum_{n=0}^{\infty} S_n(u)\frac{z^n}{n!}e^{-z}$, and note that $S(z,u)$ is the generating function of the depth in the Poisson model. Since $S_n(1)=n$, we see that for any z we have $S(z,1)=z$. Using the relation in (13), we obtain the following functional equation

$$S(z,u) = u[S(pz,u) + S(qz,u)] + (1-u)e^{-z}ze_{b-1}(z) +$$

$$(1-u)[S(pz,u)e^{-qz} + S(qz,u)e^{-pz}] + (u-1)e^{-z}[pze_{b-1}(pz) + qze_{b-1}(qz)]$$

where $e_m(x) = 1 + x + \cdots + x^m/m!$.

The generating function $D_n(u) = \frac{S_n(u)}{n}$ is the *probability* generating function [22] for the depth of a leaf in a PATRICIA since the coefficient of u^k is the probability that a randomly chosen key in a randomly chosen trie $\mathcal{T}$ in $\mathcal{J}_n$ is at depth k. Thus the Poisson generating function for depth of a leaf is $D(z,u) = \frac{S(z,u)}{z}$ and $D(\alpha z, 1) = 1$ for all α and z. (Note, in fact, $D(z,u) = e^{-z}\sum_{n\ge 0} D_{n+1}(u)\frac{z^n}{n!}$.) Consequently we have

$$D(z,u) = upD(pz,u) + uqD(qz,u) + (1-u)e^{-z}e_{b-1}(z) +$$

$$(1-u)[D(pz,u)pe^{-qz}+D(qz,u)qe^{-pz}]+(u-1)e^{-z}[pe_{b-1}(pz)+qe_{b-1}(qz)]\ . \tag{14}$$

We have in (14) the functional equation corresponding to the Poisson generating function for the depth in a PATRICIA trie. The first three terms give the functional equation for the regular trie [12]. Since we are interested in its asymptotic behavior, we would like to solve it. This, however, is too difficult, so we will use the Mellin transform to calculate the asymptotics of $D(z,u)$. Since the strip on which the Mellin transform of $D(z,u)$ is defined is empty, let $D^*(s,u)$ be the Mellin transform of $D(z,u)-1$. Note that it is defined for all s in the strip where $-1<\Re(s)<0$. (In fact, it is defined on the larger strip $-b<\Re(s)<0$, since as $z\to 0$, $D(z,u)-1=O(z^b)$, and $z\to\infty$, $D(z,u)-1=O(z^\epsilon)$ for some $\epsilon>0$; see Appendix. Subsequent integral computations require the smaller strip.) Computing $D^*(s,u)$ requires the evaluation of many integrals but ultimately we arrive at the following

$$D^*(s,u)=\frac{(1-u)G^*(s,u)}{1-u(p^{1-s}+q^{1-s})} \tag{15}$$

where

$$\begin{aligned}G^*(s,u)=&\int_0^\infty \left(pe^{-qz}[D(pz,u)-1]+qe^{-pz}[D(qz,u)-1]\right)z^{s-1}dz\\ &+\frac{\Gamma(s+b)}{s(b-1)!}-\sum_{j=0}^{b-1}\Gamma(s+j)\frac{p^{j+1}+q^{j+1}}{j!}-\Gamma(s)(pq^{-s}+qp^{-s})\end{aligned} \tag{16}$$

Although (15) looks very much like that for regular tries in [12], it is, in fact, very different. For regular tries $G^*(s,u)$ is exactly the second term of (16), but in (15), we see that $D^*(s,u)$ is only implicitly given since $G^*(s,u)$ contains an integral depending on $D(pz,u)$ and $D(qz,u)$. The analysis of $D(z,u)$ is clearly more difficult in the case of PATRICIA tries than in the case of regular tries.

Now, in order to prove (12), we let $u=e^t$, where t is complex. We want to evaluate the asymptotics of $D(z,e^t)$ as t goes to 0. We can recover $D(z,e^t)$ from $D^*(s,u)$ by evaluating the integral

$$D(z,u)=\frac{1}{2i\pi}\int_{-1/2-i\infty}^{-1/2+i\infty} z^{-s}D^*(s,u)ds$$

This is the inverse Mellin transform [10]. We will use Cauchy's residue theorem to evaluate this integral but first we must find the poles of the integrand. These correspond to the roots of

$$e^t(p^{1-s}+q^{1-s})=1\ . \tag{17}$$

Now, following [12] we analyze the roots of (17) lying in the strip $\Re(s)\le 1$. We denote them as $s_k(t)$. Let also $R_k(t)$ be the residues of $\frac{1}{1-e^t(p^{1-s}+q^{1-s})}$ at these points for $k=0$,

$\pm 1, \pm 2, \ldots$. Then we can write $D(z, e^t)$ as follows:

$$D(z,e^t) = R_0(t)G^*(s_0(t),e^t)(1-e^t)z^{-s_0(t)} + (1-e^t)\sum_{k\neq 0} R_k(t)G^*(s_k(t),e^t)z^{-s_k(t)} + O(z^{-M}), \tag{18}$$

for arbitrary $M > 0$ Now we compute the components of (18) and we begin with $s_0(t)$. Since (17) and the following equation

$$e^{-t} = p^{1-s} + q^{1-s} \tag{19}$$

are equivalent, we will solve above for $s_0(t)$. First, expand both sides using a Taylor's series up to terms of degree two. We then have

$$1 - t + t^2/2 + O(t^3) = 1 + Hs_0(t) + H_2 s_0(t)^2/2 + O(s_0(t)^3) \tag{20}$$

where $H = -(p\log p + q\log q)$ and $H_2 = p\log^2 p + q\log^2 q$. Since $s_0(0) = 0$, we can write

$$s_0(t) = at + bt^2 + O(t^3). \tag{21}$$

Substitute (21) into (19). Equating coefficients and solving for a and b we see that $s_0(t) = -\frac{t}{H} + \frac{1}{2}\left(\frac{1}{H} - \frac{H_2}{H^3}\right)t^2 + O(t^3)$. We also note that its residue $R_0(t) = -1/H + O(t)$.

Now, we are ready to show that (18) can be written as

$$D(z,e^t) = z^{-s_0(t)}(1 + O(t|z|^{-At^2})) \tag{22}$$

for some constant A. We begin with the first term of (18). The behavior of $R_0(t)$ and $(1-e^t)$ when $t \to 0$ has already been determined, so we continue by examining $G^*(s_0(t), e^t)$, which is given here in an alternate form than (16).

$$G^*(s,e^t) = \int_0^\infty \left(pe^{-qz}[D(pz,e^t)-1] + qe^{-pz}[D(qz,e^t)-1]\right) z^{s-1}dz$$
$$+ \sum_{j=1}^{b-1}\Gamma(s+j)\frac{1-(p^{j+1}+q^{j+1})}{j!} - \Gamma(s)(pq^{-s}+qp^{-s})\ . \tag{23}$$

Near zero, $\Gamma(z) = z^{-1} - \gamma + O(z)$ and $n^{-z} = 1 - z\log n + O(z^2)$. Thus the last term of (23) behaves as $-H/t + O(1)$ and the middle term behaves as a constant as $t \to 0$. Finally for the integral in (23) we note that, since $D(\alpha z, e^t)$ is continuous and $D(\alpha z, 1) = 1$, as $t \to 0$ we have $D(\alpha z, e^t) - 1 \to 0$ as $t \to 0$. Therefore, the integral will converge to 0 provided that it converges uniformly. This can be shown to be true [10]. Thus, as $t \to 0$, $G^*(s_0(t), e^t) \to H/t$ and $R_0(t)G^*(s_0(t), e^t)(1 - e^t) \to 1$.

Finally, for the Poisson model it remains only to show that the sum in (18) is small when t is small. The proof of this is similar to that which appears in [12] and [14]. It relies on showing that $\sum_{k\neq 0} |R_k(t)G^*(s_k(t), e^t)| = O(1)$ and that $\Re(s_k(t)) \geq s_0(\Re(t))$. This implies that for some $A > 0$

$$|D(z, e^t)| \leq z^{-\Re(s_0(t))}(1 + O(t|z|^{-At^2})$$

which behaves as $z^{-\Re(s_0(t))}$ as $t \to 0$. Therefore, the sum in (18) contributes $z^{-s_0(t)}o(1)$, giving (22). Writing (22) in a more convenient form, we have

$$D(z, e^t) = \exp\left\{\frac{t}{H}\log z - \frac{1}{2}\left[\frac{1}{H} - \frac{H_2}{H^3}\right]t^2 \log z + O(t^3)\log z\right\}(1 + O(t|z|^{At^2})). \qquad (24)$$

This directly leads to (12), that is,

$$e^{-t\mu(z)/\sigma(z)} D(z, e^{t/\sigma(z)}) = e^{-t^2/2}(1 + o(1)).$$

Hence, we see that the mean of the distribution of the Poisson model with parameter z is $\mu(z) = \frac{\log z}{H} + O(1)$ and its variance is $\sigma^2(z) = \left[\frac{1}{H} - \frac{H_2}{H^3}\right]\log z + O(1)$.

Finally, to prove our main result for the Bernoulli model, we need to use our depoissonization lemma. But this requires to verify hypotheses (8) and (9). This is rather technical and appears in the appendix. Then, we can compute $D_n(e^t)$ from (24) and (10), and by Goncharov's theorem, we prove that the limiting distributions of the Bernoulli model is normal with mean $\mu_n = \frac{\log n}{H} + O(1)$ and variance $\sigma_n^2 = \left[\frac{1}{H} - \frac{H_2}{H^3}\right]\log n + O(1)$. This completes the proof of Theorem 1.

3.2 Symmetric Case

Notice that in the preceding analysis, when $p = q = 1/2$, the variance $varD_n$ becomes $O(1)$ due to the fact that in this case $H_2 = H^2 = \log^2 V$. Hence, from (24) we conclude that Goncharov's theorem cannot hold, and we need a little different analysis. More precisely, the Mellin transform (15) in this case becomes

$$D^*(s, u) = \frac{G^*(s, u)}{1 - ue^{s\log V}} \ .$$

The poles are all on the axis defined by $\Re(s\log V + \log u) = 0$. Therefore, by the Mellin inverse formula, we get

$$D(z, e^t) = \frac{z^{t/\log V}}{\log V}\left[G^*\left(-\frac{t}{\log V}, e^t\right) + \sum_{k\neq 0} G^*\left(-\frac{t + 2ik\pi}{\log V}, e^t\right) z^{2ik\pi/\log V}\right] + O(z^{-M}) \qquad (25)$$

with M as large as we want. Then from the depoissonization lemma, we have

$$D_n(e^t) = D(n, e^t) + O(n^{\epsilon - 1/2}).$$

This form for the limiting distribution in the symmetric case is unsatisfying since it gives little information except that the distribution is periodic with period $\log V$. Thus we have to search for an alternative representation. From now on, we will consider the case where $b = 1$.

We begin another approach by again considering $S_n(u)$. The recurrence relation for $S_n(u)$ is obtained from (22) by setting $p = q = 1/2$:

$$S_n(u) = u \sum_{k=0}^{n} \binom{n}{k} 2^{-n} \{S_k(u) + S_{n-k}(u)\} + (1-u)2^{-n} 2 S_n(u) \tag{26}$$

We cannot solve the recurrence in (26) directly, so we define $S(z, u) = \sum_{n=0}^{\infty} S_n(u) \frac{z^n}{n!}$ as we did before. This gives another recurrence

$$S(z, u) = 2S(z/2, u)\{ue^{z/2} - u + 1\}$$

similar to that derived in the asymmetric model. Finally, since $D_n(u) = \frac{S_n(u)}{n}$ (we set $D_0(u) = 0$), by defining $D(z, u) = \frac{S(z,u)e^{-z}}{z}$ we obtain the Poisson generating function for the depth in a PATRICIA trie:

$$D(z, u) = D(z/2, u)\{e^{-z/2} + u(1 - e^{-z/2})\} \tag{27}$$

Iterating it and knowing that $D(0, u) = 1$, we are able to express $D(z, u)$ as an infinite product (cf. [13])

$$D(z, u) = \prod_{k=1}^{\infty} \{e^{-z/2^k} + u(1 - e^{-z/2^k})\} \tag{28}$$

We will need this form of the generating function in order to prove our theorem. It should be noted here that the function given in (28) is the generating function of $\sum_{k \geq 1} Y_k$, where the Y_k are independent random variables that take on the values 0 or 1 with $\Pr\{Y_k = 1\} = 1 - e^{-z/2^k}$.

We start with proving Theorem 2(i) concerning the limiting generating function of the depth. Although (28) provides the generating function, it is difficult to extract information concerning the distribution. However, since we are primarily interested in deriving a limiting law, we can make use of the Mellin transform. We let $l(z, u) = \log(D(z, u))$ and compute its Mellin transform. Note that $l(z, u)$ can be written as

$$l(z, u) = \sum_{k=1}^{\infty} \log[e^{-z/2^k} + u(1 - e^{-z/2^k})].$$

Let $g(z,u) = \log[e^{-z} + u(1-e^{-z})]$. Using a special property of the Mellin transform concerning harmonic sums (cf. [8]), $l^*(s,u)$ can be computed as the product of $g^*(s,u)$ and $\sum_{k=1}^{\infty}(2^{-k})^{-s}$. First we determine the strip on which $g^*(s,u)$ is defined. Notice that as $z \to 0$, $g(z,u) = O(z)$ and as $z \to \infty$, $g(z,u) = O(1)$. Thus, the Mellin transform of $g(z,u)$, and therefore of $l(z,u)$, is defined in the strip $-1 < \Re(s) < 0$. Furthermore, using (3) to compute $g^*(s,u)$, we obtain

$$g^*(s,u) = \frac{-\log u}{s} + \alpha(u) + O(s) \tag{29}$$

where $\alpha(u)$ is defined as in (2). Finally, we have

$$l^*(s,u) = \frac{2^s}{1-2^s} g^*(s,u) \tag{30}$$

Now we can use $l^*(s,u)$ to determine the asymptotic expansion of $l(z,u)$. By definition, the inverse Mellin transform is given by

$$l(z,u) = \frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} l^*(s,u) z^{-s} ds$$

This integral can be computed using Cauchy's theorem on residues. Since we want the expansion to hold for large values of z, we close the contour to the right, with the left boundary as the line $\Re(s) = -1/2$. Our next step then is to identify the poles of the integrand $l^*(s,u)z^{-s}$ with respect to s and determine their residues.

Clearly the only poles are those of $l^*(s,u)$. There are poles of multiplicity one at $s_k = 2\pi i k/\log 2$ for all integers k since $1 - 2^{s_k} = 0$. But $s_0 = 0$ is actually a double pole since it also is a pole of $g^*(s,u)$. The residues at the single poles $s_k, k \neq 0$, are easily calculated and are equal to $\frac{-g^*(s_k,u)}{\log 2} z^{-s_k}$. The residue associated with s_0 requires more work.

We begin this computation by expanding all factors of $l^*(s,u)z^{-s}$. The expansion of $g^*(s,u)$ is shown in (29). The other factors are then written as

$$\frac{2^s}{1-2^s} = \frac{-1}{s\log 2} - \frac{1}{2} + O(s) \tag{31}$$

and

$$z^{-s} = 1 - s\log z + O(s^2) \tag{32}$$

Multiplying (29), (31) and (32) and taking the coefficient of $1/s$ gives us the residue at s_0, namely $-\log z \frac{\log u}{\log 2} - [\frac{\alpha(u)}{\log 2} - \frac{\log u}{2}]$. We, therefore, can write

$$l(z,u) = \log z \frac{\log u}{\log 2} + \frac{\alpha(u)}{\log 2} - \frac{\log u}{2} + \frac{1}{\log 2}\sum_{k\neq 0} g^*(s_k,u) z^{-s_k} + O(z^{-M}) \ .$$

Finally $D(z,u) = e^{l(z,u)}$. This gives us the Poisson generating function, so we have a result for the Poisson model. From this we obtain the results for the Bernoulli model, the generating function $D_n(u)$ of (1) by applying the *depoissonization* lemma from the previous section. Of course, the hypotheses of the lemma must first be verified, but this is a simple variation of the proof that appears in the appendix. Thus, proof of part (i) of Theorem 2 is complete.

We have not yet obtained the limiting distribution which is our ultimate goal. Ordinarily this can be found from $D_n(u)/(1-u)$ using Cauchy's integral formula [10]. However, we cannot use this technique here since the expression we have obtained for $D_n(u)$ does not appear to be analytic within any circle about the origin, a necessary condition of Cauchy's formula, due to the presence of $\log u$. Therefore another approach is necessary to obtain the limiting distribution for the depth in a PATRICIA trie. We can, however, compute all moments from this limiting generating function since all of its derivatives exist at $u = 1$.

Now we turn our attention to the limiting distribution and the proof of part (ii) of Theorem 2. To show that $D_n - \log_2 n$ can be written as the difference of the random variables X_1 and X_2 as defined in Theorem 2, let $G_{X_1}(u)$ and $G_{X_2}(u)$ be their respective generating functions. Recall from probability theory [6] that $G_{X_2-X_1}(u) = G_{X_2}(u)G_{X_1}(1/u)$. Clearly then,

$$G_{X_2-X_1}(u) = \prod_{j=1}^{\infty}[q^{2^{-j}} + u(1-q^{2^{-j}})] \prod_{j=0}^{k_0}[u^{-1}q^{2^j} + (1-q^{2^j})] \tag{33}$$

By applying the depoissonization lemma, we have $D_n(u) = D(n,u) + O(1/\sqrt{n})$. Using this in (28), dividing both sides of the result by u^{k_0}, and replacing $e^{-n/2^{k_0}}$ by q, we have exactly (33). This proves that asymptotically $D_n - \log_2 n = X_2 - X_1$.

To show the second part of Theorem 2, consider the functional equation of (27) and define a new function $F(z,u) = D(z,u)/(1-u)$. The generating function $F(z,u)$ is then

$$F(z,u) = uF(z,u) + e^{-z/2}D(z/2,u)$$

For now we assume that $u < 1$. Iterating repeatedly and again using the fact that $D(0,u) = 1$,

$$F(z,u) = \sum_{k=0}^{\infty} u^k e^{-z/2^{k+1}} D(z/2^{k+1},u) \tag{34}$$

Define $B_m(z)$ so that $D(z,u) = \sum_{m=0}^{\infty} B_m(z)u^m$. Then, substituting this into (34),

$$F(z,u) = \sum_{k=0}^{\infty} u^k \sum_{m=0}^{k} \{B_m(z/2^{k+1-m})e^{-z/2^{k+1-m}}\} \tag{35}$$

So, we can obtain the limiting distribution if we can compute $B_m(z)$ for $m \geq 0$. To do this, consider (28). The coefficient of u^m in the expansion of this product is exactly $B_m(z)$. Clearly then, with some algebraic manipulation, $B_0(z) = e^{-z}$ and

$$B_1(z) = e^{-z} \sum_{j_1=1}^{\infty} (e^{z/2^{j_1}} - 1)$$

$$B_2(z) = e^{-z} \sum_{j_1=1}^{\infty} \{(e^{z/2^{j_1}} - 1) \sum_{j_2=1, j_2 \neq j_1}^{\infty} (e^{z/2^{j_2}} - 1)\}$$

Other $B_m(z)$ for $m > 2$ are similar to $B_2(z)$, having m sums with the condition that no two j_i's are equal.

Again using the *depoissonization* lemma, $F(n,u) = D_n(u)/(1-u) + O(n^{\epsilon - 1/2})$, making

$$\Pr\{D_n \leq k\} = \sum_{m=0}^{k} B_m(n/2^{k+1-m}) e^{-n/2^{k+1-m}}$$

Let $k = \log_2 n + x$. Then

$$\Pr\{D_n - \log_2 n \leq x\} = \sum_{m=0}^{\log_2 n + x} B_m(2^{-(x+1-m)}) e^{-2^{-(x+1-m)}} \tag{36}$$

The proof of Theorem 2 is now complete.

APPENDIX

In this Appendix we prove that conditions of the Depoissonization Lemma hold for our problem, and we can use the inverse Poisson transform to prove the limiting distribution of PATRICIA for the Bernoulli model. Although the proof is written for the asymmetric case, it also holds in the symmetric case. We start with the following proposition.

Proposition 1. *For each $\epsilon > 0$, there exists a neighborhood of 1, $\mathcal{U}(1)$, such that for all u in $\mathcal{U}(1)$, z in S_θ and $|z| > 1$ the following holds: $|D(z,u)| < |z|^\epsilon$.*

Proof. Let us define ρ such that $\rho > 1$ and $\rho(p^{1+\epsilon} + q^{1+\epsilon}) < 1 - \epsilon'$, for some $\epsilon' > 0$. Suppose also that $p > q$. Let us choose A such that $A > 1/q$ and such that for $z \in S_\theta$ and $|z| \geq A$ the following holds

$$(1+\rho)\{|pz|^\epsilon |e^{-pz}| + |qz|^\epsilon |e^{-qz}| + |e^{-z}\{e_{b-1}(z) - pe_{b-1}(pz) - qe_{b-1}(qz)\}|\} < \epsilon' |z|^\epsilon.$$

Let us define a sequence of domains $R_0 = \{z : \ z \in S_\theta \ , \ B \leq |z| \leq A\}$ with $1 < B < A$, and for m natural $R_m = \{z : \ z \in S_\theta \ , \ 1 < |z| < A/\rho^m\}$. An interesting fact is that $z \in R_m - R_{m-1}$ implies $qz \in R_{m-1}$ and $pz \in R_{m-1}$. We prove our proposition by recursion on domains $R_m \cap S_\theta$. Since $R_0 \cap S_\theta$ is compact, $D(z, 1) = 1$ and $|z|^\epsilon > 1$, there is a neighborhood $\mathcal{U}(1)$ of 1 such that for all $u \in \mathcal{U}(1)$ and $z \in R_0 \cap S_\theta$, we have $|D(z, u)| < |z|^\epsilon$. We can restrict u such that $|u| < \rho$ (by redefining $\mathcal{U}(1)$, if necessary).

Now let us suppose that for all $z \in R_m \cap S_\theta$ and for all $u \in V(1)$, Proposition 1 holds, that is, $|D(z, u)| < |z|^\epsilon$. We will prove that the proposition is true for all $z \in R_{m+1} \cap S_\theta$. Let $z \in (R_{m+1} - R_m) \cap S_\theta$ and u in $\mathcal{U}(1)$. Then, by (14)

$$
\begin{aligned}
D(z, u) &= upD(pz, u) + uqD(qz, u) + (1 - u)\{D(pz, u)pe^{-qz} + D(qz, u)qe^{-pz}\} \\
&+ (1 - u)\{e_{b-1}(z) - pe_{b-1}(pz) - qe_{b-1}(qz)\}e^{-z} \ .
\end{aligned}
$$

Since pz and qz are in $R_m \cap S_\theta$, we can make use of the fact that $|D(pz, u)| < |pz|^\epsilon$ and $|D(qz, u)| < |qz|^\epsilon$. So,

$$
\begin{aligned}
|D(z, u)| &< \rho(p^{1+\epsilon} + q^{1+\epsilon})|z|^\epsilon + (1 + \rho)\{|pz|^\epsilon p|e^{-qz}| + |qz|^\epsilon q|e^{-pz}|\} \\
&+ (1 + \rho)|\{e_{b-1}(z) - pe_{b-1}(pz) - qe_{b-1}(qz)\}e^{-z}| \ .
\end{aligned}
$$

According to the fact that $|z| > A$ and $z \in S_\theta$, we can make use of the hypothesis about A:

$$
|D(z, u)| < (1 - \epsilon')|z|^\epsilon + \epsilon'|z|^\epsilon = |z|^\epsilon
$$

and this completes the induction step. ∎

To verify the second condition of the Depoissonization Lemma, we now need only to check that $D(z, u)$ outside cone S_θ does not grow faster than exponential. We prove this below.

Proposition 2. *There exists $\alpha < 1$ and a neighborhood $\mathcal{U}(1)$ of 1 such that for all $u \in \mathcal{U}(1)$, $z \notin S_\theta$ and $|z| > 1$ implies that $|D(z, u)e^z| \le |z|^\epsilon e^{\alpha|z|}$.*

Proof. Essentially we have $\cos\theta < \alpha < 1$ because $|e^z| = e^{\Re(z)} \le e^{\cos\theta|z|}$ for z not in S_θ. Let S_θ^c be the complementary set of S_θ in the complex plane. Let us choose $A > 0$ such that $A > 1/q$ and such that for $z \in S_\theta^c$ and $|z| \ge A$ the following holds

$$
(1 + \rho)\{|pz|^\epsilon pe^{p\alpha|z|} + |qz|^\epsilon qe^{q\alpha|z|} + |e_{b-1}(z) - pe_{b-1}(pz) - qe_{b-1}(qz)|\} < \epsilon'|z|^\epsilon e^{\alpha|z|}.
$$

Using the domains R_m as defined in the previous proof, we can establish Proposition 2 by mathematical induction on domains $R_m \cap S_\theta^c$. Since the above sets are compact, $D(z, 1)e^z = e^z$, and $|e^z| < |z|^\epsilon e^{\alpha|z|}$, there exists a neighborhood $\mathcal{U}(1)$ such that for all u in $\mathcal{U}(1)$ and $z \in R_0 \cap S_\theta^c$, the following holds: $|D(z, u)e^z| \le |z|^\epsilon e^{\alpha|z|}$.

Now, let us suppose that the property is true on $R_m \cap S_\theta^c$, and we will prove that it also holds for $m+1$. Let $z \in (R_{m+1} - R_m) \cap S_\theta^c$ with u in $\mathcal{U}(1)$. Then by (14) we have

$$\begin{aligned} D(z,u)e^z &= upD(pz,u)e^z + uqD(qz,u)e^z + (1-u)\{D(pz,u)pe^{pz} + D(qz,u)qe^{qz}\} \\ &+ (1-u)\{e_{b-1}(z) - pe_{b-1}(pz) - qe_{b-1}(qz)\} \ . \end{aligned}$$

Therefore, taking into account the mathematical induction hypotheses, that is, $|D(az,u)e^{-az}| < |az|^\epsilon e^{\alpha a|z|}$ with a either p or q, we finally obtain

$$\begin{aligned} |D(z,u)e^z| &\leq \{\rho p^{1+\epsilon} e^{\alpha p|z|}|e^{qz}||z|^\epsilon + \rho q^{1+\epsilon} e^{\alpha q|z|}|e^{pz}||z|^\epsilon\} + \epsilon'|z|^\epsilon e^{\alpha|z|} \\ &\leq (\rho p^{1+\epsilon} + \rho q^{1+\epsilon})|z|^\epsilon e^{\alpha|z|} + \epsilon'|z|^\epsilon e^{\alpha|z|} \leq |z|^\epsilon e^{\alpha|z|} \end{aligned}$$

and this completes the proof of Proposition 2, and also verification of hypotheses (8) and (9) in the Depoissonization Lemma. ■

ACKNOWLEDGEMENT

We are very grateful to Dr. Philippe Flajolet for a suggestion that led to part (ii) of our Theorem 2. We also would like to thank an anonymous referee for suggestions that improved the presentation of this paper.

References

[1] A. Apostolico, The Myriad Virtue of Suffix Trees, in Springer NATO ASI Ser. F12, 85–96 (1985).

[2] A. Aho, J. Hopcroft and J. Ullman, *Data Structures and Algorithms*, Addison-Wesley (1983).

[3] J. Capetanakis, Tree Algorithms for Packet Broadcast Channels, *IEEE Trans. on Inf. Theory*, 25,pp. 605-615 (1980).

[4] L. Devroye, A Study of Trie-like Structures Under the Density Model, *Annals of Applied Probability*, 1992 (in press).

[5] R. Fagin, J. Nievergelt, N. Pippenger and H. Strong, Extendible Hashing; A Fast Access Method for Dynamic Files, *ACM TODS*4, pp. 315-344 (1979).

[6] W. Feller, *An Introduction to Probability Theory and its Applications*, John Wiley & Sons, New York (1988).

[7] P. Flajolet, On the Performance Evaluation of Extendible Hashing and Trie Searching, *Acta Informatica* 20, pp. 345-369 (1983).

[8] P. Flajolet, M. Régnier and R. Sedgewick, Some Uses of the Mellin Transform Techniques on the Analysis of Algorithms, *Combinatorial Algorithms on Words*, Springer NATO ASI Ser. F12, 241-254 (1985).

[9] G. Gonnet and J. Munro, The Analysis of Linear Probing Sort by the Use of a New Mathematical Transform, *Journal of Algorithms*, 5, pp. 451–470 (1984).

[10] P. Henrici, *Applied and Computational Complex Analysis*, John Wiley & Sons, New York (1977).

[11] M. Hofri, *Probabilistic Analysis of Algorithms*, Springer-Verlag (1987).

[12] P. Jacquet and M. Régnier, Limiting Distributions for Trie Parameters, *Lecture Notes in Computer Science*, 214, 196-210 (1986).

[13] P. Jacquet and W. Szpankowski, Ultimate Characterization of the Burst Response of an Interval Searching Algorithm: A Study of a Functional Equation, *SIAM J. Comp.*, 18, pp. 777-791 (1989).

[14] P. Jacquet and W. Szpankowski, Analysis of Digital Tries with Markovian Dependency, *IEEE Transactions on Information Theory*, 37, pp. 1470-1475 (1991).

[15] M. Kendall and A. Stuart, *The Advanced Theory of Statistics*, 4th Ed. Vol. 1, London, Charles Griffin and Co. Ltd., (1977).

[16] P. Kirschenhofer and H. Prodinger, Further Results on Digital Search Trees, *Theoretical Computer Science*, 58, pp. 143-154 (1988).

[17] P. Kirschenhofer, H. Prodinger and J. Schoissengeier, Zur Auswertung gewisser Reihen mit Hilfe modularer Functionen, in: E. Hlawka, ed., *Zahlentheoretische Analysis 2, Lecture Notes in Mathematics 1262,* Springer, Berlin, 108-110 (1987).

[18] P. Kirschenhofer, H. Prodinger and W. Szpankowski, On the Balance Property of Patricia Tries: External Path Length Viewpoint, *Theoretical Computer Science,* 68, pp. 1-17 (1989).

[19] D. Knuth, *The Art of Computer Programming*, vol. I, Addison-Wesley (1973).

[20] D. Knuth, *The Art of Computer Programming. Sorting and Searching.* vol. III. Addison-Wesley (1973).

[21] B. Pittel, Asymptotic Growth of a Class of Random Trees, *The Annals of Probability*, 18, 414-427 (1985).

[22] B. Pittel, Paths in a Random Digital Tree: Limiting Distributions, *Adv. Appl. Prob.*, 18, 139-155 (1986).

[23] B. Pittel and H. Rubin, How Many Random Questions Are Necessary to Identify n Distinct Objects?, *J. Combinatorial Theory, Ser. A*, 55, 292-312, 1990.

[24] A. Renyi, *Probability Theory*, North-Holland, Amsterdam (1970).

[25] W. Szpankowski, Some Results on V-ary Asymmetric Tries, *Journal of Algorithms*, 9, 224-244 (1988).

[26] W. Szpankowski, Patricia Tries Again Revisited, *Journal of the ACM,* 37, 691-711 (1991).