UNIVERSAL LIMIT LAWS FOR DEPTHS IN RANDOM TREES

Luc Devroye
School of Computer Science
McGill University

Abstract. Random binary search trees, b-ary search trees, median-of-(2k+1) trees, quadtrees, simplex trees, tries and digital search trees are special cases of random split trees. For these trees, we offer a universal law of large numbers and a limit law for the depth of the last inserted point, as well as a law of large numbers for the height.

Keywords and phrases. Binary search tree, data structures, expected time analysis, depth of a node, random tree, law of large numbers.

CR Categories: 3.74, 5.25, 5.5.

Research of the author was sponsored by NSERC Grant A3456 and by FCAR Grant 90-ER-0291. Address: School of Computer Science, McGill University, 3480 University Street, Montreal, Canada H3A 2K6. Email: luc@cs.mcgill.ca.

Random split trees.

We introduce a model for a random tree that is sufficiently general so that it encompasses many important families of random trees, such as random binary search trees, random m-ary search trees, random fringe-balanced trees, random median-of-(2k+1) trees, random quadtrees and random simplex trees. A skeleton tree T_b of branch factor b is an infinite rooted position tree, in which each node has b children, numbered 1 through b. A split tree (with branch factor $b > 0$, vertex capacity $s > 0$, and of cardinality $n \geq 0$) is a skeleton tree of branch factor b in which n balls are assigned to a collection of vertices, where each vertex may hold up to s balls. Nodes or vertices are denoted by u. $N(u)$ denotes the number of balls in the subtree rooted at u. $C(u)$ denotes the number of balls associated with vertex u. A vertex u is a leaf if $C(u) = N(u) > 0$, or equivalently if $C(u) > 0$ and $N(v) = 0$ for all b children v of u. A node u is useless if $N(u) = 0$. Two split trees with the same parameters b, s, n are equivalent if for all their vertices, the $N(u)$'s are identical. The trimmed split tree is the (finite) split tree from which all useless nodes are deleted.

We now introduce a random split tree with parameters b, s, s_0, s_1, $\mathcal{V}$ and n. The branch factor b, vertex capacity s and number of balls n are as for split trees. The additional integers s_0 and s_1 are needed to describe the ball distribution process and satisfy the inequalities

$$0 < s \ , \ 0 \leq s_0 \leq s \ , \ 0 \leq bs_1 \leq s + 1 - s_0 \ .$$

Finally, $\mathcal{V}$ is a prototype random vector $(V_1, \ldots, V_b)$ of probabilities: $\sum_i V_i = 1; V_i \geq 0$. A random split tree is a skeleton tree T_b in which each vertex u is given an independent copy of $\mathcal{V}$, and in which n balls are distributed in the manner described below over the vertices, where each vertex may hold up to s balls. The distribution is done in an incremental fashion as described below. Initially, there are no balls, so every node has a ball count $C(u) = 0$. Adding a ball to a tree rooted at u proceeds as follows. Let $(V_1, \ldots, V_b)$ be the probability vector associated with u.

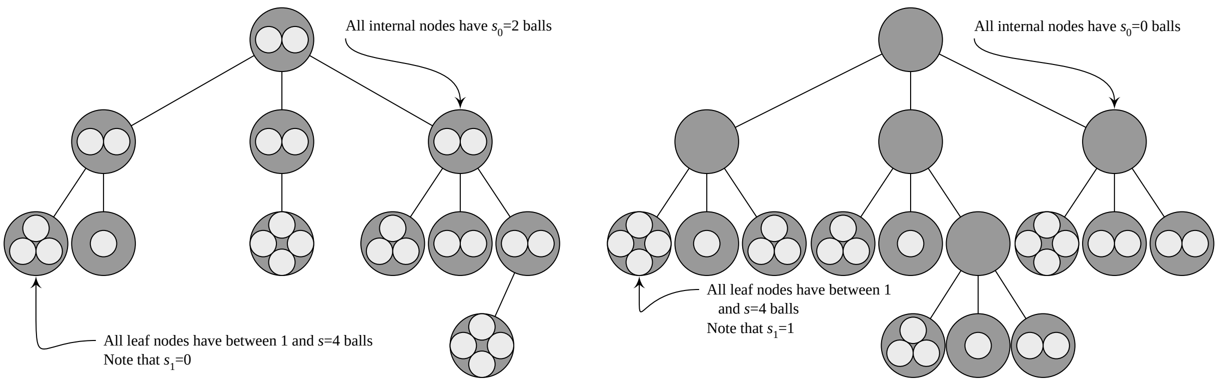
A. If u is not a leaf (so that $C(u) = s_0$), choose child i with probability V_i, increment $N(u)$ by 1, and recursively add the ball to the subtree rooted at child i.

B. If u is a leaf but $C(u) = N(u) < s$ (where s is the capacity of a vertex introduced in the previous paragraph), then add the ball to u and stop. $C(u)$ and $N(u)$ are both incremented by 1.

C. If u is a leaf but $C(u) = N(u) = s$, there is no room for the ball at u. In that case, we set $N(u) = s + 1$ and $C(u) = s_0$. We place $s_0 \leq s$ randomly selected balls at u, and send $s + 1 - s_0$ balls down to the children of u. This is done as follows.

We first give s_1 randomly selected balls to each child and adjust the ball counts for the children. The remaining $s+1-s_0-bs_1$ balls are sent down by choosing for each ball a child independently according to the probability vector $(V_1,\ldots,V_b)$ and applying the procedure "add a ball" to the tree rooted at the selected child. Note that this may have to be repeated several times if $s_0=0$, but only once if $s_0>0$ (because no child will reach the capacity s).

Note that every non-leaf node has $C(u)=s_0$ and every leaf has $0<C(u)\le s$. Two split trees, one with $(s,s_0,s_1)=(4,2,0)$ and the other with $(s,s_0,s_1)=(4,0,1)$ are shown below.



A <u>random trimmed split tree</u> with parameters b, s, s_0, s_1, $\mathcal{V}$ and n is a random split tree with parameters b, s, s_0, s_1, $\mathcal{V}$ and n from which we eliminate all useless nodes. Note that there is in general no simple relationship between the number of vertices and the number of balls (or points). The ball distribution method described above has several advantages, first and foremost among them the direct relationship with several species of trees that occur as natural data structures. The parameters s, s_0, s_1 add sufficient flexibility. We will see that all trees with fixed finite values of these parameters have the same asymptotic behavior for various shape parameters. When we refer to a random split tree, it is understood that we mean a random trimmed split tree.

The depth of a vertex is its distance from the root. The height of a tree is the maximal depth of any of the leaves. When the $(n+1)$-st ball is added to a random split tree of cardinality n (that is, a tree holding n balls), the depth of the vertex reached by the ball is denoted by D_{n+1}. If the leaf is split, the ball ends up, if $s_0>0$, at depth D_{n+1} or $D_{n+1}+1$, depending upon which ball we choose to follow in the splitting process. It is therefore convenient to study D_n.

Interestingly, the following property is valid for a tree with n balls rooted at u: if $n\le s$, all balls are in the root node, which is a leaf; if $n>s$, there are s_0 balls in the root node, and the cardinalities $(N_1,\ldots,N_b)$ of the b subtrees of the root are distributed as

$(s_1, \ldots, s_1)$ plus a multinomial $(n-s_0-bs_1, V_1, \ldots, V_b)$ random vector, where $(V_1, \ldots, V_b)$ is associated with u. This property is repeated recursively at every node. Roughly speaking, the subtrees rooted at the children have cardinalities close to $nV_1, \ldots nV_b$.

The behavior of several other parameters is easily deduced from that of D_n. For example, let D'_n be the average depth, i.e., the sum of the depths of the n balls divided by n. The incremental growing process described above shows that, if $s_0 > 0$,

$$\frac{n-an}{n}\mathbf{E}\{D_{an}\} \le \mathbf{E}\{D'_n\} \le \mathbf{E}\{D_n\} + 1$$

for any $a > 0$ such that an is integer. This implies that if $\mathbf{E}\{D_n\} \sim c\log n$ for some constant c, then $\mathbf{E}\{D'_n\} \sim c\log n$. For this reason, we will not investigate D'_n at length.

The purpose of this paper is to point out that within this general setting, the asymptotics—a law of large numbers and a limiting distribution—of D_n are easy to determine. Interestingly, one proof is offered that works for all trees mentioned above.

Throughout this paper, we assume that the components of $(V_1, \ldots, V_b)$ are identically distributed. Note that if they are not, a random permutation $(\sigma_1, \ldots, \sigma_b)$ of $(1, 2, \ldots, b)$ shows that we achieve this goal by taking $(V_{\sigma_1}, \ldots, V_{\sigma_b})$. This random permutation of the children does not affect the depth and height. If V_i has a distribution described by the probability measure μ_i, then the V_{σ_i}'s are identically distributed with common probability measure $(1/b)\sum_j \mu_j$. The latter is called the splitting distribution. A random variable V with the splitting distribution is called a *splitter*. Define $W = V_S$ where, given $(V_1, \ldots, V_b)$, $S = i$ with probability V_i. Observe that $\mathbf{E}V = \mathbf{E}W = 1/b$ in all cases. The law of large numbers and the limit law for D_n depend upon just two parameters,

$$\mu = \mathbf{E}\{\log(1/W)\} = b\mathbf{E}\{V\log(1/V)\}$$

and

$$\sigma^2 = \mathrm{Var}\{\log W\} = b\mathbf{E}\left\{V\log^2 V\right\} - \mu^2 \ .$$

We first state our main results without proof. Then we give a brief discussion of the random trees to which the results apply. The proofs are at the end of the paper.

The main result.

For all trees that follow the given model, if H_n denotes the height, that is, the maximal distance between the root and any leaf, we have $H_n = O(\log n)$ in probability. The behavior of H_n is related to that of the moment function

$$m(t) = \mathbf{E}\{V^t\} \ , \ t \ge 0 \ .$$

For later reference, we provide the key properties of this function.

LEMMA 1.

A. *The function* m *decreases monotonically from* $m(0)=1$ *to* $\mathbf{P}\{V=1\}$ *as* $t\to\infty$.

B. m *is differentiable for all* $t>0$.

C. $\log m$ *is convex. In particular,* m'/m *is increasing on* $(0,\infty)$.

D. $m(t)^{1/t}\le m(s)^{1/s}$ *for* $t<s$. *Thus,* $\log(m(t))/t$ *is nondecreasing.*

E. *For* $t>0$, $m'(t)=\mathbf{E}\{V^t\log(V)\}$ *and* $m''(t)=\mathbf{E}\{V^t\log^2(V)\}$.

F. m'/m *takes every value between* $\mathbf{E}\{\log(V)\}$ *(as* $t\downarrow 0$*) and* $\log v_\infty$, *where* v_∞ *is the rightmost point in the support of* V.

G. *The solution of the equation* $m'(t)/m(t)=-1/c$ *is called* $t^*=t^*(c)$. *Then* t^* *is a monotonically increasing function of* c, *and a solution exists when*

$$-\frac{1}{\mathbf{E}\{\log(V)\}}<c<-\frac{1}{\log v_\infty}\ .$$

H. $t^*/c+\log m(t^*)$ *decreases in* c *(or* t^**). The value of* $t^*/c+\log m(t^*)$ *changes from* 0 *(at* $t^*=0$*) to* R *(possibly* $-\infty$*) as* $t^*\to\infty$, *where* $R=\lim_{t\to\infty}(\log m(t)-t\,m'(t)/m(t))$.

THEOREM 1. *Let* V *be a splitter for a random split tree. Assume that* $\mathbf{P}\{V=1\}=0$. *Then there exists a finite constant* c *such that*

$$\lim_{n\to\infty}\mathbf{P}\{H_n>c\log n\}=0\ .$$

If additionally $R<-\log b$, *where*

$$R=\lim_{t\to\infty}(\log m(t)-t\,m'(t)/m(t))\ ,$$

then the same is true for all $c>\gamma$ *and* $\gamma\in(0,\infty)$ *is a parameter only depending upon* b *and the distribution of* V, *and is defined by*

$$\gamma=\inf\{c:e^{t^*}(bm(t^*))^c<1\}=\inf\{c:t^*/c+\log(m(t^*))<-\log b\}\ ,$$

where t^* *is the unique solution of* $m'(t)/m(t)=-1/c$. *(See Lemma 1 below.)*

We note that under the conditions of Theorem 1, $H_n/\log n\to\gamma$ in probability. The lower bound that goes with the upper bound of Theorem 1 can be obtained by various methods, and its straightforward proof is not included here (as we focus mainly on depths in this paper). Galton-Watson processes (see Athreya and Ney, 1972) may be used directly (Devroye, 1987). One may also use extrema in branching random walks as exhibited in Devroye (1986) (see Biggins (1976, 1977), Hammersley (1974) and Kingman (1973) for branching random walks, and see Mahmoud (1992) for further applications). Pittel (1994) points out how one may use the Crump-Mode process in this respect.

Theorem 2. *Let D_n be the depth of the last node in a random split tree with n nodes and splitter V. If $\mu \neq 0$ and $\mathbf{P}\{V = 1\} = 0$, then*

$$\frac{D_n}{\log n} \to \frac{1}{\mu} \text{ in probability as } n \to \infty ,$$

and $\mathbf{E}\{D_n\}/\log n$ tends to the same limit. Furthermore, if $\sigma > 0$, then

$$\frac{D_n - (\log n)/\mu}{\sqrt{\sigma^2 (\log n)/\mu^3}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

where $\mathcal{N}(0,1)$ denotes the normal distribution and $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution.

The law of large numbers for D_n does not apply when $\mathbf{E}\{V \log V\} = 0$, i.e., when $\mathbf{P}\{V \in \{0,1\}\} = 1$. This degenerate case is excluded from further consideration. It suffices that V have a density (as in many examples that follow below). For the limit law, we need in addition $\mathrm{Var}\{\log W\} > 0$. This is equivalent to asking that V not be monoatomic. Of all the examples below, only special cases of tries—the symmetric tries and symmetric digital search trees—have a monoatomic splitter V ($V \equiv 1/b$). All other examples satisfy the latter condition.

Some properties of the beta distribution.

The beta distribution plays an important role in many important random split trees. We summarize some key properties. Define the beta (a, b) density

$$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} \ , \ 0 < x < 1 \ ,$$

where $a, b > 0$ are parameters, and $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$.

LEMMA 2. *If X is a beta (a,b) random variable, then*

$$\mathbf{E}\{\log(1/X)\} = \psi(a+b) - \psi(a)$$

where $\psi(u) = \Gamma'(u)/\Gamma(u)$ is the derivative of $\log \Gamma$ at u (also called the digamma function). Furthermore,

$$\mathbf{E}\{X \log(1/X)\} = \frac{a}{a+b}(\psi(a+1+b) - \psi(a+1)) \ .$$

Let ψ'—the trigamma function—be the derivative of ψ. Then

$$\mathbf{E}\{\log^2(X)\} = (\psi(a+b) - \psi(a))^2 + \psi'(a) - \psi'(a+b) \ .$$

Finally,

$$\mathbf{E}\{X \log^2(X)\} = \frac{a}{a+b}\left(\psi(a+1+b) - \psi(a+1)\right)^2 + \frac{a}{a+b}\left(\psi'(a+1) - \psi'(a+1+b)\right) \ .$$

For integrals such as those dealt with in Lemma 2, we refer to Sibuya (1979) or Gradshteyn and Ryzhik (1980, p. 538, p. 541). We recall that the digamma function basically behaves as the harmonic numbers (Abramowitz and Stegun, 1970, pp. 258–259): if γ is Euler's constant,

$$\psi(n) = -\gamma + \sum_{k=1}^{n-1} \frac{1}{k} \quad (n \geq 2) \ ; \ \psi(1) = -\gamma \ ; \ \psi(1/2) = -\gamma - 2\log(2) \ ;$$

$$\psi(z+1) = \psi(z) + \frac{1}{z} = -\gamma + \sum_{n=1}^{\infty} \frac{z}{n(n+z)} \ , \ z > -1 \ .$$

For the trigamma function, we have

$$\psi'(z) = \sum_{n=0}^{\infty} \frac{1}{(z+n)^2} \ .$$

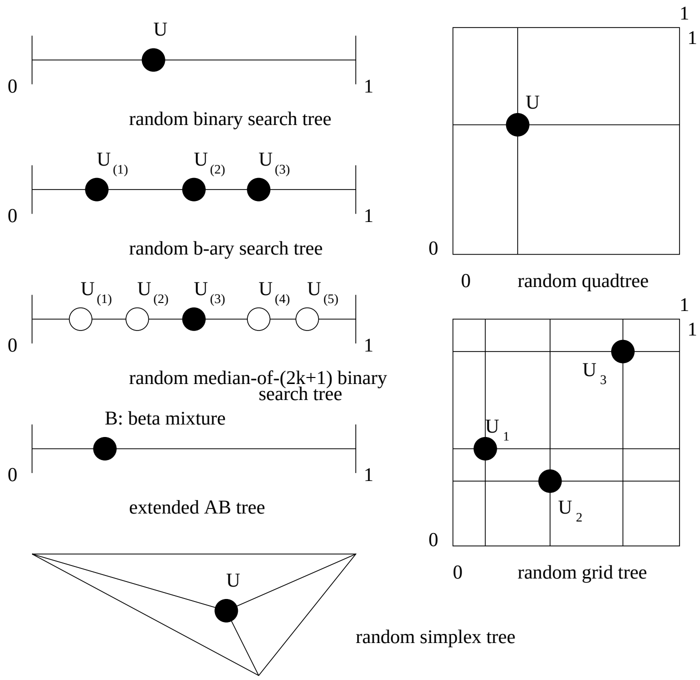
Also, $\psi'(z) = \psi'(z-1) - 1/(z-1)^2$, and $\psi'(1) = \pi^2/6$.

Examples: an overview.

We list a number of important special cases of random split trees. In this table, $U, U_1, U_2, \ldots$ are i.i.d. uniform $[0,1]$ random variables. Recall that s is the capacity of a node before it is split, s_0 is the number of balls left in a node after a split, s_1 is the minimum number of balls sent to any subtree, and b is the branch factor.

tree	$V \stackrel{\mathcal{L}}{=}$	s	s_0	s_1	b
random binary search tree	U, uniform $[0,1]$ beta (1,1)	1	1	0	2
random b-ary search tree	$U_{(b-1)}=\min(U_1,\ldots,U_{b-1})$ beta$(1,b-1)$	$b-1$	$b-1$	0	b
random quadtree	$\prod_{i=1}^{d} U_i$	1	1	0	2^d
random median-of-(2k+1) binary search tree	median $(U_1,\ldots,U_{2k+1})$ beta $(k+1,k+1)$	$2k$	1	k	2
random simplex tree	$\min_{1\le i\le d} U_i$ beta$(1,d)$	1	1	0	$d+1$
AB tree	symmetric beta (a,a)	1	0(1)	0	2
extended AB tree	uniform $\{$beta (a,b), beta $(b,a)\}$	1	0(1)	0	2
trie	uniform $\{p_1,\ldots,p_b\}$	1	0	0	b
digital search tree	uniform $\{p_1,\ldots,p_b\}$	1	1	0	b
random m-grid tree	$\prod_{i=1}^{d} U_i'$ $U_1',\ldots,U_d'$ i.i.d. beta$(1,m)$	m	m	0	$(m+1)^d$

This table shows that a large variety of trees may be dealt with in one sweep. The fixed parameters s, s_0 and s_1 are irrelevant for first term asymptotics and the law of large numbers for depths and heights. Only the distribution of the splitter V matters.



Various ways of splitting spaces are shown. All splits are applied recursively. In the figure, U refers to the uniform distribution, and B to a mixture of beta distributions.

Nonetheless, many trees cannot be molded into our framework, such as all trees whose depth does not grow logarithmically with n. For example, it is well-known that the uniform random binary tree has average depth and height of the order of $\sqrt{n}$ (Flajolet and Odlyzko, 1982; Vitter and Flajolet, 1990). Interestingly, Aldous (1993) has introduced a model that includes many (but not all) of the trees in the table above and the uniform random trees, as well as a continuum of trees that link them. Our work was inspired for a great deal by Aldous' paper.

The parameters μ and σ^2 are computed for the trees mentioned above. Most of the limit laws and laws of large numbers are known, but the unified approach of this paper explains things in a stronger way. More details are provided in the nine subsequent subsections, in which each tree is briefly discussed separately. The symbols $\mathcal{H}$ and $\mathcal{H}_2$ are properly defined in the section on tries.

tree	σ^2	$1/\mu$ (limit of $D_n/\log n$)
random binary search tree	$1/4$	2
random b-ary search tree	$\sum_{i=2}^{b} \frac{1}{i^2}$	$\frac{1}{\sum_{i=2}^{b} \frac{1}{i}}$
random quadtree	$d^2/4$	$2/d$
random median-of-(2k+1) binary search tree	$\sum_{j=k+2}^{2k+2} \frac{1}{j^2}$	$\frac{1}{\sum_{i=k+2}^{2k+2} \frac{1}{i}}$
random simplex tree	$\sum_{i=2}^{d+1} \frac{1}{i^2}$	$\frac{1}{\sum_{i=2}^{d+1} \frac{1}{i}}$
AB tree	$(\psi(2a+1)-\psi(a+1))^2 + (\psi'(a+1)-\psi'(2a+1))$	$\frac{1}{\psi(2a+1)-\psi(a+1)} = \frac{1}{\sum_{n=1}^{\infty} \frac{a}{(n+a+1)(n+2a+1)}}$
extended AB tree	$\frac{a}{a+b}\left(\psi(a+1+b)-\psi(a+1)\right)^2 + \frac{a}{a+b}\left(\psi'(a+1)-\psi'(a+1+b)\right) + \frac{b}{a+b}\left(\psi(a+1+b)-\psi(b+1)\right)^2 + \frac{b}{a+b}\left(\psi'(b+1)-\psi'(a+1+b)\right)$	$\frac{a+b}{a(\psi(a+1+b)-\psi(a+1)) + b(\psi(a+1+b)-\psi(b+1))}$
b-ary trie	$\mathcal{H}_2 - \mathcal{H}^2$	$1/\mathcal{H}$
b-ary digital search tree	$\mathcal{H}_2 - \mathcal{H}^2$	$1/\mathcal{H}$
random m-grid tree	$d\sum_{j=2}^{m+1} \frac{1}{j^2} + d(d-1)\left(\sum_{j=2}^{m+1} \frac{1}{j}\right)^2$	$\frac{1}{d\sum_{i=2}^{m+1} \frac{1}{i}}$

Example 1: the random binary search tree.

In a random binary search tree with n nodes, the following operation is applied independently and recursively: a random node is chosen from the n nodes at hand, and it is made the root. The nodes with a smaller label travel to the left subtree of the root, and the others to the right subtree. The size of the left subtree is distributed as $\lfloor nU \rfloor$, where U is uniform $[0,1]$. The size of the right subtree has a similar distribution. Equivalently, attach to each of the balls an independent copy of a uniform $[0,1]$ random variable, to get $U_1, U_2, \ldots, U_n$. Put U_1 in the root and partition the others into left and right subsets by comparison with U_1. Repeating the splitting process at each node creates a random split tree. In a third equivalent representation, that of the random split tree, we may associate with each node an independent random split vector (V_1, V_2) distributed as $(U_1, 1-U_1)$.

It is known that $D_n/\log n \to 2$ in probability (Lynch, 1965; Knuth, 1973; Devroye, 1988). The limit law for D_n was derived by Devroye (1988): $(D_n - 2\log n)/\sqrt{2\log n} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$. Robson (1979), Pittel (1984) and Devroye (1985, 1987) showed that $H_n/\log n \to 4.31107\ldots$ in probability. All of these results are contained in Theorems 1 and 2 as $m(t) = 1/(t+1)$, $\mu = 1/2$ and $\sigma^2 = 1/4$.

An important variant of the binary search tree related to the standard occurs when no balls are stored in internal nodes; so, $s_0 = 0$. This leads to a binary search tree in which all balls are at leaves. This too is a random split tree, and it follows the same limit laws as the ordinary random binary search tree.

Example 2: the random b-ary search tree.

Let n balls be given and associate with each ball an independent uniform $[0,1]$ random variable. In a random b-ary search tree with n nodes, the following operation is applied independently and recursively: $b-1$ random balls are chosen from the n balls at hand, and are associated with the root. The other balls, if there are any, are partitioned into b sets by membership in the intervals induced by the $b-1$ balls. If $(N_1, \ldots, N_b)$ are the number of balls in the intervals (with $\sum_i N_i = n-b+1$ of course), then this vector is multinomial $(n-b+1, V_1, \ldots, V_b)$, where the V_i's are the lengths of the intervals (or spacings; see Pyke, 1975). The split vector of V_i's is thus distributed as the collection of b spacings induced by $b-1$ i.i.d. uniform $[0,1]$ random variables on $[0,1]$. In particular, $V = V_1$ is distributed as a beta $(1, b-1)$ random variable.

We easily compute $\mu = \sum_{i=2}^{b} 1/i$ and $\sigma^2 = \sum_{i=2}^{b} 1/i^2$. This yields

$$\frac{D_n}{\log n} \to \frac{1}{\sum_{i=2}^{b} \frac{1}{i}} \quad \text{in probability}$$

(a result of Mahmoud and Pittel, 1984) and

$$\frac{D_n - (1/\mu)\log n}{\sqrt{(\sigma^2/\mu^3)\log n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \ .$$

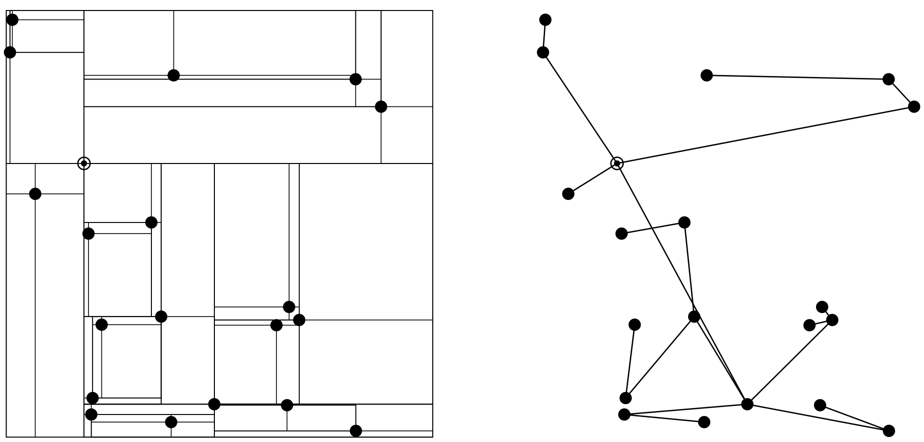
As an example, if $b = 3$, $\mu = 5/6$, $\sigma^2 = 78/125$, and

$$\frac{D_n - (6/5)\log n}{\sqrt{(78/125)\log n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \ .$$

We also know that $H_n/\log n \to c$ in probability for a function c of b given in Devroye (1990) and indicated in Theorem 1.

Example 3: the random quadtree.

The point quadtree in R^d (Finkel and Bentley, 1974; see Samet (1990b) for a survey) generalizes the binary search tree. One ball is put in each node of a tree with branch factor 2^d; each ball has associated with it a d-vector for the point it represents; each subtree of a node corresponds to one of the quadrants formed by considering the ball's d-vector as the new origin. Insertion into point quadtrees is as for binary search trees.

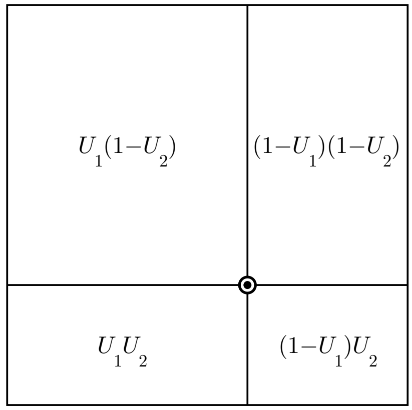


At the left, a partition of the plane by a random quadtree is shown. The circled point is the root. It partitions the space into four quadrants, and the splitting rule is recursively applied to each quadrant. At the right, the same points are shown together with the edges in the quadtree.

We assume that a random quadtree is constructed on the basis of an i.i.d. sequence drawn from the uniform distribution on $[0,1]^d$. In that case, it is convenient to index the split vector by a bit sequence of length d: $(b_1 \ldots b_d)$. The vector $(V_{00\ldots00}, \ldots, V_{11\ldots11})$ has components that may be written as

$$V_{(b_1 \ldots b_d)} = \prod_{j=1}^{d} U_j^{b_j} (1 - U_j)^{1-b_j}$$

where $(U_1, \ldots, U_d)$ is the point in the node where the split takes place.



The split induced by the root of the quadtree is shown. The random variables U_1, U_2 are i.i.d. uniform $[0,1]$. The areas of the four rectangles are all distributed like products of two independent uniform $[0,1]$ random variables.

The height H_n of a random quadtree is in probability asymptotic to $(c/d)\log n$, where $c = 4.31107\ldots$ is the constant in the height of the random binary search tree (Devroye, 1987). This also follows from Theorem 1 as $m(t) = \mathbf{E}\{V^t\} = 1/(t+1)^d$. Write $V = V_{11\ldots11} = \prod_{j=1}^d U_j$. Then it takes just a moment to verify that

$$\begin{aligned}
\mu = \mathbf{E}\{\log W\} &= 2^d\mathbf{E}\{V\log(1/V)\} \\
&= 2^d \sum_{j=1}^d \mathbf{E}\left\{\prod_{k=1}^d U_k \log(1/U_j)\right\} \\
&= 2^d d\,\mathbf{E}\left\{\prod_{k=2}^d U_k\right\}\mathbf{E}\left\{U_1\log(1/U_1)\right\} \\
&= 2d\,\mathbf{E}\left\{U_1\log(1/U_1)\right\} \\
&= \frac{d}{2}\ .
\end{aligned}$$

From this, we see that

$$\frac{D_n}{\log n} \to \frac{2}{d} \quad \text{in probability}\ ,$$

a result first noted by Devroye and Laforest, 1990. See also Flajolet, Gonnet, Puech and Robson (1991). The computations of the variance are a bit more tedious. We have

$$\begin{aligned}
\sigma^2 + \mu^2 &= 2^d\mathbf{E}\{V\log^2(1/V)\} \\
&= 2^d\mathbf{E}\left\{\prod_{k=1}^d U_k\left(\sum_{j=1}^d \log(1/U_j)\right)^2\right\}
\end{aligned}$$

$$
\begin{aligned}
&= 2^d d\,\mathbf{E}\left\{\prod_{k=2}^{d} U_k\right\}\mathbf{E}\left\{U_1 \log^2(1/U_1)\right\} \\
&\qquad + 2^d\, d(d-1)\mathbf{E}\left\{\prod_{k=3}^{d} U_k\right\}\mathbf{E}^2\left\{U_1 \log(1/U_1)\right\} \\
&= 2d\,\mathbf{E}\left\{U_1 \log^2(1/U_1)\right\} + 4d(d-1)\mathbf{E}^2\left\{U_1 \log(1/U_1)\right\} \\
&= \frac{d^2}{4} + \frac{d}{4}\ .
\end{aligned}
$$

Hence, $\sigma^2 = d/4$. This yields the limit law

$$
\frac{D_n - (2/d)\log n}{\sqrt{(2/d^2)\log n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)\ ,
$$

valid for any $d \geq 1$. This result was obtained via complex analysis by Flajolet and Lafforgue (1994).

Example 4: the random median-of-(2k+1) binary search tree.

Bell (1965) and Walker and Wood (1976) introduced the following method for constructing a binary search tree. Take $2k+1$ points at random from the set of n points on which a total order is defined, where k is integer. The median of these points serves as the root of a binary tree. The remaining points are thrown back into the collection of points and are sent to the subtrees. Following Poblete and Munro (1985), we may look at this tree by considering internal nodes and external nodes, where internal nodes hold one data point and external nodes are bags of capacity $2k$. Insertion proceeds as usual. As soon as an external node overflows (i.e., when it would grow to size $2k+1$), its bag is split about the median, leaving two new external nodes (bags) of size k each, and an internal node holding the median. After the insertion process is completed, we may wish to expand the bags into balanced trees. Using the branching process method of proof (Devroye, 1986b, 1987, 1990; see also Mahmoud, 1992) the almost sure limit of $H_n/\log n$ for all k may be obtained (Devroye, 1993). For another possible proof method, see Pittel (1992). The depth D_n of the last node when the fringe heuristic is used has been studied by the theory of Markov processes or urn models in a series of papers, notably by Poblete and Munro (1985), Aldous, Flannery and Palacios (1988). See also Gonnet and Baeza-Yates (1991, p. 109). Poblete and Munro (1985) showed that

$$
\frac{D_n}{\log n} \to \frac{1}{\sum_{i=k+2}^{2k+2} \frac{1}{i}}
$$

in probability. It should be clear by now that this tree is a random split tree with $s = 2k$, $s_0 = 1$, $s_1 = k$, $b = 2$ and split vector (V_1, V_2) distributed as $(B, 1-B)$, where B is beta

$(k+1, k+1)$. That is, B is distributed as the median of $2k+1$ i.i.d. uniform $[0,1]$ random variables. This representation is obtained by associating with each point in the data an independent uniform $[0,1]$ random variable. Clearly,

$$\mu = \sum_{i=k+2}^{2k+2} \frac{1}{i} \ .$$

Also, if X is beta (a,a), and a is integer-valued, Lemma 2 and the properties of the digamma and trigamma functions imply

$$\begin{aligned} \mathbf{E}\{X \log^2(X)\} &= \frac{1}{2}\left(\psi(2a+1) - \psi(a+1)\right)^2 + \frac{1}{2}\left(\psi'(a+1) - \psi'(2a+1)\right) \ . \\ &= \frac{1}{2}\left(\left(\sum_{j=a+1}^{2a} \frac{1}{j}\right)^2 + \sum_{j=a+1}^{2a} \frac{1}{j^2}\right) \ . \end{aligned}$$

Thus,

$$\sigma^2 = \sum_{j=k+2}^{2k+2} \frac{1}{j^2} \ .$$

Therefore, we obtain a limit law for all k. As an example, for $k=1$, we obtain $\mu = 1/3 + 1/4 = 7/12$, $\sigma^2 = 1/9 + 1/16 = 25/144$, and thus

$$\frac{D_n - (12/7)\log n}{\sqrt{(300/343)\log n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \ .$$

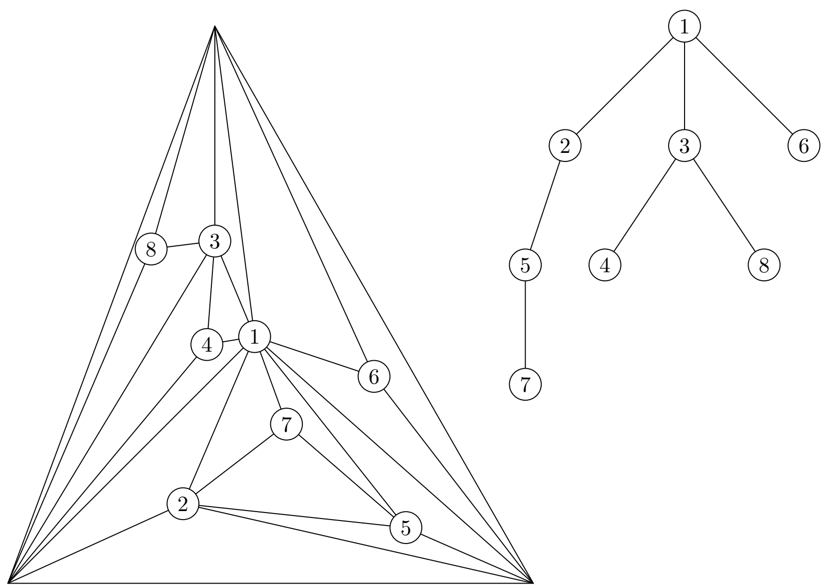
We rediscover results for the number of comparisons C_n for median-of-(2k+1) quicksort. As $\mathbf{E}C_n = \mathbf{E}\{nD_n\}$, where D_n is the depth of the n-th point when inserted in a median-of-(2k+1) binary search tree holding $n-1$ points. From the above results,

$$\lim_{n\to\infty} \frac{\mathbf{E}C_n}{n \log n} = \frac{1}{\frac{1}{k+2} + \frac{1}{k+3} + \cdots + \frac{1}{2k+2}} \ .$$

Thus, for median-of-5 quicksort, the expected number of comparisons grows as $60/47 n \log n$. By generating function methodology (Vitter and Flajolet, 1990; Kemp, 1984; Sedgewick, 1983) or via urn models (Aldous et al, 1988), results of this nature are harder to obtain. However, our method does not allow one to compute anything but the main asymptotic term in $\mathbf{E}C_n$.

Example 5: random simplex trees.

Triangulating polygons and objects in the plane is an important problem in computational geometry. An $O(n \log n)$ algorithm for triangulating n points was found by Avis and El Gindy (1987). Arkin, Held, Mitchell and Skiena (1994) obtained a simple fast $O(n \log n)$ expected time algorithm for triangulating any collection of n planar points in general position. We look more specifically at their triangulation and its d-dimensional extension to simplices, and ask what the tree generated by this partitioning looks like if the points are uniformly distributed in the unit simplex. Given are n vectors $X_1, \ldots, X_n$ taking values in a fixed simplex S of $\mathbb{R}^d$. It is assumed that this is an i.i.d. sequence with a uniform distribution on S for the purposes of analysis. X_1 is associated with the root of a $d+1$-ary tree. It splits S into $d+1$ new simplices by connecting X_1 with the $d+1$ vertices of S. Associate with each of these simplices the subset of $X_2, \ldots, X_n$ consisting of those points that fall in the simplex. Each nonempty subset is sent to a child of the root, and the splitting is applied recursively to each child. As every split takes linear time in the number of points processed, it is clear that the expected time is proportional to $n\mathbf{E}D_n$, where D_n is the expected depth of a random node in the tree. The partition consists of $dn+1$ simplices, each associated with an external node of the tree. There are precisely n nodes in the tree and each node contains one point.



A triangle is triangulated by a cloud of random points. The corresponding ternary tree is shown on the right.

If $|S|$ denotes the size of a simplex S, then the following crucial property is valid.

LEMMA 3. *If simplex S is split into $d+1$ simplices $S_1, \ldots, S_{d+1}$ by a point X distributed uniformly in S, then $(|S_1|, \ldots, |S_{d+1}|)$ is jointly distributed as $(|S|V_1, \ldots, |S|V_{d+1})$, where $V_1, \ldots, V_{d+1}$ are the spacings of $[0,1]$ induced by d i.i.d. uniform $[0,1]$ random variables.*

PROOF. It is known that X is distributed as $\sum_{i=1}^{d+1} V_i t_i$, where $t_1, \ldots, t_{d+1}$ are the vertices of S (see Rubinstein (1982), Smith (1984) or Devroye (1986a)). But for a simplex S, we know that

$$|S| = \frac{1}{d!} \det \begin{pmatrix} t_1 & t_2 & t_3 & \cdots & t_{d+1} \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}$$

Apply this formula to S_1 by replacing t_1 by X:

$$\begin{aligned} |S_1| &= \frac{1}{d!} \det \begin{pmatrix} X & t_2 & t_3 & \cdots & t_{d+1} \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \\ &= \frac{1}{d!} \det \begin{pmatrix} \sum_i V_i t_i & t_2 & t_3 & \cdots & t_{d+1} \\ \sum_i V_i & 1 & 1 & \cdots & 1 \end{pmatrix} \\ &= \frac{V_1}{d!} \det \begin{pmatrix} t_1 & t_2 & t_3 & \cdots & t_{d+1} \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \\ &= V_1 |S| \ . \end{aligned}$$

The statement then follows trivially. □

It is immediate that the random simplex tree is a random split tree with split vector distributed as the spacings defined by d i.i.d. uniform $[0,1]$ random variables on $[0,1]$, $s_0 = 1$, $s = 1$, $s_1 = 0$ and $b = d+1$. Therefore, by Theorem 2, D_n behaves precisely as for the random $d+1$-ary tree discussed earlier. Thus, $\mu = \sum_{i=2}^{d+1} 1/i$ and $\sigma^2 = \sum_{i=2}^{d+1} 1/i^2$. This yields

$$\frac{D_n}{\log n} \to \frac{1}{\sum_{i=2}^{d+1} \frac{1}{i}} \quad \text{in probability}$$

and

$$\frac{D_n - (1/\mu)\log n}{\sqrt{(\sigma^2/\mu^3)\log n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \ .$$

As an example, if $d = 2$, then $\mu = 5/6$, $\sigma^2 = 78/125$, and

$$\frac{D_n - (6/5)\log n}{\sqrt{(78/125)\log n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \ .$$

For $d = 3$, $\mu = 1/2 + 1/3 + 1/4 = 13/12$. Thus, $D_n/\log n \to 12/13$ in probability. We also know that $H_n/\log n \to c$ in probability for a function c of d that may be computed via the recipe described in Theorem 1.

Example 6: extended AB trees and simulation.

When generating random trees that resemble botanical trees, a number of mathematical models have been proposed. We refer to Viennot's survey (1980) or the book by Prusinkiewicz and Lindenmayer (1990) for further references. Stripped from geometrical considerations, most trees are binary. The main parameter one needs to control is the height of the tree as a function of the number of nodes. Alternately, one may wish to control the average distance from a node to the root. For this, it is necessary to have a family of random trees in which these parameters can take any large value. In the context of this paper, if we had a family of splitting trees—a continuum of trees, really—with parameter α and for which $D_n/\log n \to c(\alpha)$ in probability, we would be saved if the domain of values of $c(\alpha)$ would be $(\log 2, \infty)$ as α varies over a given range.

In 1993, Aldous introduced a family of random split trees with $s = 1$, $s_1 = 0$ and $s_0 = 0$. With this set-up, all balls are put in leaves, and internal nodes have no balls. Aldous splits with the aid of $(V, 1 - V)$, where V is beta (a, a) for some parameter $a > 0$. By varying a and even extending it beyond its natural range, Aldous creates a one-parameter family that may be used to model certain splitting processes in biology. He also studies the depths of nodes in these trees and obtains laws of large numbers for their heights. We define an AB tree (for *Aldous beta*) in a similar fashion but take $s_0 = s = 1$ and $s_1 = 0$. This change is only cosmetic, as it will not affect any asymptotic result. For $a = 1$, we obtain the random binary search tree. As $a \downarrow 0$, the tree becomes more elongated, and the amount of stretching may be controlled by a. As $a \to \infty$, every split is nearly fifty-fifty, and the height of the tree is in probability asymptotic to $\log_2 n$. The laws of large numbers for depths and heights are essentially those obtained by Aldous for his model.

We feel that a lot is gained by considering two-parameter families for modeling biological phenomena and simulating botanical trees. This may be achieved by extending the AB trees and taking B an equal mixture of a beta (a, b) and a beta (b, a) density. The splitter V remains symmetric, but as a and b diverge so that $a/(a + b) \to p \in (0, 1)$, we see that in the limit V is p or $1 - p$ with equal probability. This creates trees of height about $\log_{1/\min(p,1-p)} n$. The AB trees are obtained at $b = a$. We call this versatile family of trees extended AB trees. We will report on the drawing of realistic-looking trees via extended AB trees elsewhere.

From Lemma 2, the parameters are easily obtained:

$$\mu = \frac{a(\psi(a+1+b) - \psi(a+1)) + b(\psi(a+1+b) - \psi(b+1))}{a+b} ,$$

and

$$\sigma^2 = \frac{a}{a+b}\left(\psi(a+1+b)-\psi(a+1)\right)^2 + \frac{a}{a+b}\left(\psi'(a+1)-\psi'(a+1+b)\right) + \frac{b}{a+b}\left(\psi(a+1+b)-\psi(b+1)\right)^2 + \frac{b}{a+b}\left(\psi'(b+1)-\psi'(a+1+b)\right) .$$

If we set $b=1$, and a is integer, the limit for $\mathbf{E}D_n/\log n$ is

$$\frac{a+1}{\sum_{i=1}^{a}\frac{1}{i}} .$$

This grows unbounded like $a/\log a$ as $a\to\infty$. As $a\downarrow 0$, the limit is $(a+1)/(a/(a+1)+\psi(a+2)-1)\sim 1/a$. This grows unbounded as well. In the AB trees, we have $a=b$, and thus,

$$\frac{D_n}{\log n} \to \frac{1}{\psi(2a+1)-\psi(a+1)} = \frac{1}{\sum_{n=1}^{\infty}\frac{a}{(n+a+1)(n+2a+1)}} \text{ in probability.}$$

This result matches that of Aldous (1993) where the limit is written as an integral. It is easy to verify that as $a\downarrow 0$, the limit is asymptotic to $1/(a(\pi^2/6-1))$. At $a=1$, we have a limit of 2 as in the random binary search tree. As $a\to\infty$, the limit approaches $1/\log 2$, and the splits because nearly all perfectly balanced. The variance is given by

$$\sigma^2 = \left(\psi(2a+1)-\psi(a+1)\right)^2 + \left(\psi'(a+1)-\psi'(2a+1)\right) .$$

Finally, consider extended AB trees in which as $a,b\to\infty$, $a=p(a+b)$ with $p\in(0,1)$ fixed. The limit then behaves as

$$\frac{1}{-p\log p-(1-p)\log(1-p)} = \frac{1}{\mathcal{H}(p)} ,$$

where $\mathcal{H}(p)$ is the entropy of a Bernoulli (p) random variable. Here we rediscover a known property of the entropy $\mathcal{H}=-\sum_i p_i\log(p_i)$ of a discrete distribution $(p_1,p_2,\ldots)$: split a set of size n into subsets of sizes close to $np_1, np_2, np_3, \ldots$. Associate each subset with a child of the root, and repeat this process until no further splitting is possible (note that there is no randomness involved in this splitting). If one grabs a random node in the resulting tree, its depth is in probability equal to $(1/\mathcal{H})\log n$. This is exactly like the behavior of random nodes in tries (Fredkin, 1960) in which the symbols have probabilities $p_1,p_2,p_3,\ldots$; see Pittel (1985, 1986), Szpankowski (1988), and the next section.

Example 8: tries and digital search trees.

Tries are b-ary trees for storing infinite strings. Assume that the data consists of n infinite strings of $\{1, \ldots, b\}$-valued symbols, called $X_1, \ldots, X_n$. Each string carves out an infinite path in the infinite complete b-ary tree. For a node u, let $N(u)$ denote the number of strings that pass through node u. Now, eliminate all nodes with $N(u) = 0$ and eliminate all those with $N(u) = 1$ whose parent also has $N(u) = 1$. The resulting tree has n leaves with $N(u) = 1$, and every non-leaf v has $N(v) > 1$. Invented in 1960 by Fredkin, this structure is called a trie. Assume that all the symbols are drawn independently, and that each symbol takes the value i with probability p_i. Define the entropy $\mathcal{H}$ by

$$\mathcal{H} = -\sum_{i=1}^{b} p_i \log p_i \ ,$$

and the second order entropy by

$$\mathcal{H}_2 = \sum_{i=1}^{b} p_i \log^2 p_i \ ,$$

The trie may be viewed as a random split tree with $s = 1$, $s_0 = s_1 = 0$, in which a node u at which $N(u) = n$ is not split if $n = 1$, and in which a split occurs when $n > 1$; in the latter case, the sizes of the subtrees are distributed jointly as a multinomial $(n, p_1, \ldots, p_b)$ random variable. If V is p_S, where S is uniformly distributed on $\{1, 2, \ldots, b\}$, then

$$b\mathbf{E}\{V \log(1/V)\} = \mathcal{H} \ .$$

Therefore, from Theorem 2,

$$\frac{D_n}{\log n} \to \frac{1}{\mathcal{H}} \text{ in probability } .$$

Also,

$$\frac{D_n - \log n/\mathcal{H}}{\sqrt{(\mathcal{H}_2 - \mathcal{H}^2)\log n/\mathcal{H}^3}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \ .$$

The law of large numbers is due to Pittel (1985). The limit law was discovered independently by Jacquet and Régnier (1986) and Pittel (1986). See Szpankowski (1988) for additional results and references. Our result unifies the analysis of tries and binary search trees. The normal limit law as stated above is not valid if $\mathcal{H}_2 = \mathcal{H}^2$. This occurs if and only if $p_1 = p_2 = \cdots = p_b = 1/b$. In the random split tree, this situation corresponds to a monoatomic distribution for the splitter ($V \equiv 1/b$), which has zero variance.

The digital search tree of Coffman and Eve (1970) is like a trie. It is best described by its incremental construction. It has n nodes, one per string. Nodes are added one by one, starting with X_1 and ending with X_n. The node associated with X_n is the first node u in the infinite path of string X_n that has $N(u) = 0$ before X_n is inserted. This is a random split tree in which a node with $N(u) = n$ spawns subtrees when $n > 1$ of sizes that are jointly distributed as a multinomial $(n-1, p_1, \ldots, p_b)$ random variable. The limit laws given above for tries remain valid here, without change. These were known; see Pittel (1985) for the law of large numbers, and Pittel (1986) and Louchard (1987) for the normal limit law. Again, when all p_i's are equal, the normal limit law as stated above is not valid. Another proof method is needed to deal with that situation. Also, Theorem 1 only states that for nondegenerate tries and digital search trees, $H_n = O(\log n)$ in probability. Theorem 1 cannot be used to get a finer result. However, the behavior of the height is well-known (Pittel, 1985).

Example 9: the random grid tree.

The quadtree is easily generalized as follows: consider a collection of m $\mathbb{R}^d$-valued points drawn from the data, and partition the space into $(m+1)^d$ hyperrectangles by the d perpendicular hyperplanes centered at each of the m points. In a quadtree, $m = 1$. This generates a tree, the m-grid tree, with fan-out $(m+1)^d$, and with up to m points per node. If the data consist of n independent random vectors uniformly distributed over $\mathbb{R}^d$, the tree thus constructed becomes a random split tree with split vector $(V_1, \ldots, V_b)$ in which each V_i is distributed as $V = \prod_{j=1}^d B_j$, and $B_1, \ldots B_d$ are independent beta $(1, m)$ random variables. While not exactly the same, the random grid tree borrows ideas from the celebrated grid file data structure (Nievergelt, Hinterberger and Sevcik, 1984; Nievergelt and Hinrichs, 1993). We note the following:

$$\begin{aligned}
\mu = \mathbf{E}\{\log W\} &= (m+1)^d \mathbf{E}\{V \log(1/V)\} \\
&= (m+1)^d \sum_{j=1}^d \mathbf{E}\left\{\prod_{k=1}^d B_k \log(1/B_j)\right\} \\
&= (m+1)^d d\, \mathbf{E}\left\{\prod_{k=2}^d B_k\right\} \mathbf{E}\left\{B_1 \log(1/B_1)\right\} \\
&= (m+1) d\, \mathbf{E}\left\{B_1 \log(1/B_1)\right\} \\
&= d \sum_{j=2}^{m+1} \frac{1}{j} \, .
\end{aligned}$$

From this, we see that

$$\frac{D_n}{\log n} \to \frac{1}{d \sum_{j=2}^{m+1} \frac{1}{j}} \quad \text{in probability} \, .$$

Also,

$$\sigma^2 + \mu^2 = (m+1)^d \mathbf{E}\{V \log^2(1/V)\}$$

$$
\begin{aligned}
&= (m+1)^d \mathbf{E}\left\{ \prod_{k=1}^d B_k \left(\sum_{j=1}^d \log(1/B_j) \right)^2 \right\} \\
&= (m+1)^d d\, \mathbf{E}\left\{ \prod_{k=2}^d B_k \right\} \mathbf{E}\left\{ B_1 \log^2(1/B_1) \right\} \\
&\qquad + (m+1)^d\, 2d(d-1) \mathbf{E}\left\{ \prod_{k=3}^d B_k \right\} \mathbf{E}^2 \left\{ B_1 \log(1/B_1) \right\} \\
&= (m+1)d\, \mathbf{E}\left\{ B_1 \log^2(1/B_1) \right\} + 2(m+1)^2 d(d-1) \mathbf{E}^2 \left\{ B_1 \log(1/B_1) \right\} \\
&= d \sum_{j=2}^{m+1} \frac{1}{j^2} + d \left(\sum_{j=2}^{m+1} \frac{1}{j} \right)^2 + 2d(d-1) \left(\sum_{j=2}^{m+1} \frac{1}{j} \right)^2 .
\end{aligned}
$$

Hence,

$$
\sigma^2 = d \sum_{j=2}^{m+1} \frac{1}{j^2} + d(d-1) \left(\sum_{j=2}^{m+1} \frac{1}{j} \right)^2 .
$$

This yields the limit law obtained earlier for the quadtree when $m = 1$. For $m = 2$, we have

$$
\frac{D_n - (6/5d) \log n}{\sqrt{\left(\frac{180d-102}{125d^2} \right) \log n}} \stackrel{\mathcal{L}}{\to} \mathcal{N}(0,1) ,
$$

for any $d \ge 1$. For $d = 1$, this coincides with the result obtained earlier for the random 3-ary search tree.

Proof of Theorem 1: the height.

We show that $\mathbf{P}\{H_n \ge (c+3\epsilon) \log n\} \to 0$ for all $\epsilon > 0$ and $c > \gamma$. Define $\delta = s_1$, $k' = \lfloor \epsilon \log n \rfloor$, and $l = k'(\delta + 1)$. If n is the number of balls stored in a random split tree, then the cardinalities of the subtrees at distance k from the root are bounded from above by quantities of the form

$$
\begin{aligned}
Z_k &\stackrel{\text{def}}{=} \text{binomial}\left(n, \prod_{i=1}^k V^{(i)} \right) + \text{binomial}\left(\delta, \prod_{i=2}^k V^{(i)} \right) \\
&\qquad + \text{binomial}\left(\delta, \prod_{i=3}^k V^{(i)} \right) + \cdots + \text{binomial}\left(\delta, V^{(k)} \right) + \delta \\
&\le \text{binomial}\left(n, \prod_{i=1}^k V^{(i)} \right) + \text{binomial}\left(\delta(k-k'+1), \prod_{i=k-k'+1}^k V^{(i)} \right) + k'\delta
\end{aligned}
$$

where $V^{(1)}, \ldots, V^{(k)}$ is a sequence of i.i.d. random variables distributed as V, and the inequality is in a stochastic sense only. For each of the b^k paths down to a node at distance k, a different sequence is obtained. Thus, by Boole's inequality and the fact that all splitters are identically distributed, we have for integer $k, l > 0$,

$$\mathbf{P}\{H_n \geq k + 3l\} \leq b^k \mathbf{P}\{Z_k \geq 3l\} .$$

We further argue as follows:

$$\begin{aligned}\mathbf{P}\{Z_k \geq 3l\} \leq \mathbf{P}&\left\{\text{binomial}\left(n, \prod_{i=1}^{k} V^{(i)}\right) \geq l\right\}\\ &+ \mathbf{P}\left\{\text{binomial}\left(\delta(k-k'+1), \prod_{i=k-k'+1}^{k} V^{(i)}\right) \geq l\right\}\\ &+ \mathbf{P}\left\{k'\delta \geq l\right\} .\end{aligned}$$

The last term is zero by choice of l. Conditioned on the $V^{(i)}$'s, the first term is easily bounded using Markov's inequality and Chernoff's bounding method. Let $t > 0$ be picked later. Then, if $Z = \prod_{i=1}^{k} V^{(i)}$,

$$\begin{aligned}\mathbf{P}\{\text{binomial}(n, Z) \geq l | Z\} &\leq \mathbf{E}\left\{(1 - Z + Ze^t)^n | Z\right\} e^{-tl}\\ &\leq \mathbf{E}\left\{e^{(e^t-1)nZ} | Z\right\} e^{-tl}\\ &\leq \mathbf{E}\left\{e^{l-nZ+l\log(nZ)} | Z\right\} \quad (\text{take } e^t = l/(nZ)) .\end{aligned}$$

Take the expectation with respect to Z. The inequality is then further developed by noting the following: for $z \in (0,1)$, and $t > 0$,

$$\begin{aligned}\mathbf{P}\{\text{binomial}(n, Z) \geq l\} &\leq e^{l-z+l\log(z)} + \mathbf{P}\{nZ > z\}\\ &\leq (ez)^l + (n/z)^t \mathbf{E}\{Z^t\} .\end{aligned}$$

Similarly, for $z' \in (0,1)$, $t' > 0$ and $Z' = \prod_{i=k-k'+1}^{k} V^{(i)}$,

$$\begin{aligned}\mathbf{P}\{\text{binomial}(\delta(k-k'+1), Z') \geq l\} &\leq e^{l-z'+l\log(z')} + \mathbf{P}\{\delta(k-k'+1)Z' > z'\}\\ &\leq (ez')^l + (\delta(k-k'+1)/z')^{t'} \mathbf{E}\{Z'^{t'}\} .\end{aligned}$$

Take $z = z' = b^{-2k/l}/e$ and note that $ze \sim b^{-2c/\epsilon(\delta+1)}$. Then, combining the previous bounds,

$$\begin{aligned}b^k \mathbf{P}\{Z_k \geq 3l\} &\leq 2b^{-k} + b^k (ne)^t b^{2kt/l} m(t)^k + b^k (\delta(k-k'+1)e)^{t'} b^{2kt'/l} m(t')^{k'}\\ &\stackrel{\text{def}}{=} I + II + III ,\end{aligned}$$

where $m(t)$ is the t-th moment of V. Clearly, $I = o(1)$. Choose t' large enough so that $b\, m(t')^{\epsilon/c} < 1$. This is possible, as $\mathbf{P}\{V = 1\} = 0$ and thus $m(t') \to 0$ as $t' \to \infty$ (Lemma 1). With this choice of t', $III = o(1)$. To treat II, fix t and observe that $II = o(1)$ if

$$b^k n^t m(t)^k \to 0 ,$$

which occurs if

$$(bm(t))^c e^t < 1 \text{ or, equivalently, } c\log(bm(t)) + t < 0 \ .$$

Here we distinguish between the two statements in the theorem. For the first statement, take t so large that $bm(t) < 1$. Then take c large enough to ensure that $(bm(t))^c e^t < 1$. For the second statement, we must be a bit more careful. The minimal value of $c\log(bm) + t$ is obtained at the solution $t^* = t^*(c)$ of the equation $m'(t)/m(t) = -1/c$. From Lemma 1, a solution exists when

$$-\frac{1}{\mathbf{E}\{\log(V)\}} < c < -\frac{1}{\log v_\infty} \ .$$

Replace t by t^* (as we are allowed to choose any positive t) and let $c > \gamma$. As $R < -\log b$, $c\log(bm(t^*)) + t^* < 0$ and thus $II = o(1)$. This concludes the proof of Theorem 1.

Proof of Theorem 2: the depth.

The following lemma will be useful for bounding tail probabilities.

LEMMA 4. *If X is binomial (n, Z) (written $B_{n,Z}$) where $Z \in [0, 1]$ is a random variable, then for $0 < a < n$,*

$$\mathbf{P}\{X \geq a\} \leq \mathbf{P}\{Z > a/(2n)\} + \left(\frac{e}{4}\right)^{a/2} \ .$$

Similarly,

$$\mathbf{P}\{X \leq a\} \leq \mathbf{P}\{Z < 2a/n\} + \left(\frac{2}{e}\right)^{a} \ .$$

PROOF. If X is binomial (n, p),then , for $1 > u \geq p$ by Chernoff's bound (Chernoff, 1952; Okamoto, 1958),

$$\mathbf{P}\{X \geq nu\} \leq \left(\left(\frac{p}{u}\right)^u \left(\frac{1-p}{1-u}\right)^{1-u}\right)^n \ .$$

Interestingly, the same bound applies for $\mathbf{P}\{X \leq nu\}$ if $0 < u \leq p$. In particular,

$$\mathbf{P}\{X \geq 2np\} \leq \left(\left(\frac{1}{2}\right)^{2p} \left(\frac{1-p}{1-2p}\right)^{1-2p}\right)^n \leq \left(\left(\frac{1}{2}\right)^{2p} e^p\right)^n = \left(\frac{e}{4}\right)^{np} \ .$$

Also,

$$\mathbf{P}\{X \leq np/2\} \leq \left((2)^{p/2} \left(1 - \frac{p/2}{1-p/2}\right)^{1-p/2}\right)^n \leq \left((2)^{p/2} e^{-p/2}\right)^n = \left(\sqrt{\frac{2}{e}}\right)^{np} \ .$$

Applying this, we have

$$\begin{aligned}
\mathbf{P}\{X > a\} &\le \mathbf{P}\{Z > a/(2n)\} + \mathbf{P}\{B_{n,a/(2n)} > a\} \\
&\le \mathbf{P}\{Z > a/(2n)\} + \left(\frac{e}{4}\right)^{a/2} .
\end{aligned}$$

Similarly, assuming without loss of generality that $n \ge 2$,

$$\begin{aligned}
\mathbf{P}\{X \le a\} &\le \mathbf{P}\{Z < 2a/n\} + \mathbf{P}\{B_{n,2a/n} \le a\} \\
&\le \mathbf{P}\{Z < 2a/n\} + \left(\frac{2}{e}\right)^{a} . \square
\end{aligned}$$

The convergence of $\mathbf{E}\{D_n\}$ follows from the weak convergence,

$$\mathbf{E}\{D_n\} = \int_0^1 \mathbf{P}\{D_n > t\}\, dt ,$$

and the tail probabilities for $\mathbf{P}\{D_n > t\}$ developed below. The details are omitted.

For the proof of Theorem 2, we consider an infinite random path in the tree, $u_0, u_1, u_2, \ldots$, where u_0 is the root, and given u_i, and the split vector $(V_1, \ldots, V_b)$ for u_i, u_{i+1} is the j-child of u_i with probability V_j. Put n balls in the tree as in the construction of a random split tree, and let u^* be the unique leaf on the infinite path. Then D_n is less than or equal to the distance between u^* and the root.

We first show that for all $c > 1/\mu$, $\mathbf{P}\{D_n > c\log n\} \to 0$. Take $k = \lfloor c \log n \rfloor$. Let u_0 be the root, and let u_0, u_1, … be the path of nodes followed by the inserted point from the root down. We have, if $\beta = (s_0+1)k$,

$$[D_n > k + l] \subseteq [N(u_k) > \beta] \cup [H_\beta > l] ,$$

where H_β is the height of a random split tree with β balls. But

$$\mathbf{P}\{N(u_k) > \beta\} \le \mathbf{P}\left\{ ks_0 + \text{binomial}\left(n, \prod_{i=0}^{k-1} W_i\right) > \beta \right\}$$

where $W_0, W_1, \ldots$ are i.i.d. random variables distributed as $W = V_S$, and $S = i$ with probability V_i. Here we made use of the fact that a binomial (N, p) in which N is binomial (n, q) is distributed as a binomial (n, pq). By Lemma 4, we see that

$$\begin{aligned}
\mathbf{P}\left\{ \text{binomial}\left(n, \prod_{i=0}^{k-1} W_i\right) > \beta - ks_0 \right\} &\le \mathbf{P}\left\{ \prod_{i=0}^{k-1} W_i > \frac{\beta - ks_0}{2n} \right\} + \left(\frac{e}{4}\right)^{\frac{\beta - ks_0}{2}} \\
&= \mathbf{P}\left\{ \sum_{i=0}^{k-1} \log W_i > \log\left(\frac{\beta - ks_0}{2n}\right) \right\} + \left(\frac{e}{4}\right)^{\frac{k}{2}} \\
&= I + II .
\end{aligned}$$

Clearly, $II = o(1)$. Also, $I = o(1)$ by the law of large numbers, as

$$\frac{\sum_{i=0}^{k-1} \log W_i}{k\mathbf{E}\{\log W\}} \to 1$$

almost surely. Recall that $\mu = \mathbf{E}\{\log(1/W)\}$. Also, we used the fact that

$$\liminf_{n\to\infty} \frac{\log\left(\frac{\beta - ks_0}{2n}\right) + k\mu}{k} = -\frac{1}{c} + \mu > 0$$

since $c > 1/\mu$. To wrap up the proof of the first part, we must show that $\mathbf{P}\{H_\beta > l\} \to 0$, where l is our choice. Let us pick $l = \lfloor 2\gamma \log \beta \rfloor = \lfloor 2\gamma \log((s_0+1)k) \rfloor$ where $\gamma > 0$ is as in Theorem 1. Then $\mathbf{P}\{H_\beta > l\} \to 0$, because

$$\lim_{n\to\infty} \mathbf{P}\{H_\beta > 2\gamma \log \beta\} = 0 \ ,$$

As $l \sim \log\log n$, the first part of the law of large numbers is proved.

Next, we show that for all $c < 1/\mu$, $\mathbf{P}\{D_n < c\log n\} \to 0$. Take $k = \lfloor c \log n \rfloor$. Then if $N(.)$ refers to the tree with $n-1$ balls,

$$[D_n < k] \subseteq [N(u_k) = 0] \ .$$

But

$$\mathbf{P}\{N(u_k) = 0\} \le \mathbf{P}\left\{ -ks + \text{binomial}\left(n-1, \prod_{i=0}^{k-1} W_i\right) \le 0 \right\}$$

where $W_0, W_1, \ldots$ are as in the earlier part of this proof. By Lemma 4, we see that

$$\begin{aligned}
\mathbf{P}\left\{\text{binomial}\left(n-1, \prod_{i=0}^{k-1} W_i\right) \le ks\right\} &\le \mathbf{P}\left\{\prod_{i=0}^{k-1} W_i \le \frac{2ks}{n-1}\right\} + \left(\frac{2}{e}\right)^{ks} \\
&= \mathbf{P}\left\{\sum_{i=0}^{k-1} \log W_i \le \log\left(\frac{2ks}{n-1}\right)\right\} + \left(\frac{2}{e}\right)^{ks} \\
&= I + II \ .
\end{aligned}$$

Obviously, $II = o(1)$. $I = o(1)$ by the law of large numbers, as

$$\frac{\sum_{i=0}^{k-1} \log W_i}{k\mathbf{E}\{\log W\}} \to 1$$

almost surely. and

$$\limsup_{n\to\infty} \frac{\log\left(\frac{2ks}{n}\right) - k\mathbf{E}\{\log W\}}{k} = -\frac{1}{c} - \mathbf{E}\{\log W\} = -\frac{1}{c} + \mu < 0 \ .$$

This concludes the proof of the lower bound for the law of large numbers.

The limit law is obtained by using the same upper and lower bounds introduced in the proof of the law of large numbers. Additionally, we will use the fact that

$$\frac{\sum_{i=0}^{k-1} \log W_i + k\mu}{\sqrt{k\sigma^2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \ .$$

Consider first the probability $\mathbf{P}\{D_n > k\}$, where $k = \lfloor (1/\mu)\log n + u\sqrt{\log n} \rfloor$ and $u \in \mathbb{R}$. We have, if $\beta = (s_0+1)k$,

$$[D_n > k + l] \subseteq [N(u_k) > \beta] \cup [H_\beta > l] \ ,$$

where H_β is the height of a random split tree with β balls. Arguing as in the first part of this proof,

$$\begin{aligned}
\mathbf{P}\{N(u_k) > \beta\} &\le \mathbf{P}\left\{ \text{binomial}\left(n, \prod_{i=0}^{k-1} W_i\right) > \beta - ks_0 \right\} \\
&= \mathbf{P}\left\{ \sum_{i=0}^{k-1} \log W_i > \log\left(\frac{\beta - ks_0}{2n}\right) \right\} + \left(\frac{e}{4}\right)^{\frac{k}{2}} \\
&= \mathbf{P}\left\{ \frac{\sum_{i=0}^{k-1} \log W_i + k\mu}{\sqrt{k\sigma^2}} > \frac{\log\left(\frac{\beta - ks_0}{2n}\right) + k\mu}{\sqrt{k\sigma^2}} \right\} + o(1) \\
&= \mathbf{P}\left\{ \mathcal{N}(0,1) > \frac{\log\left(\frac{\beta - ks_0}{2n}\right) + k\mu}{\sqrt{k\sigma^2}} \right\} + o(1) \\
&= \mathbf{P}\left\{ \mathcal{N}(0,1) > \frac{u\mu^{3/2}}{\sigma} \right\} + o(1) \ .
\end{aligned}$$

Recall that for $l = \lfloor 2\gamma \log \beta \rfloor = \lfloor 2\gamma \log((s_0+1)k) \rfloor$, we obtain $\mathbf{P}\{H_\beta > l\} \to 0$. As $l \sim \log\log n$, the first part of the limit law is proved:

$$\mathbf{P}\{D_n > k\} \le \mathbf{P}\{N > u\mu^{3/2}/\sigma\} + o(1) \ .$$

Using the arguments for the lower bound, we may prove in a similar fashion that

$$\mathbf{P}\{D_n < k\} \le \mathbf{P}\{N < u\mu^{3/2}/\sigma\} + o(1) \ .$$

Taken together, this proves that

$$\lim_{n\to\infty} \mathbf{P}\{D_n < k\} = \mathbf{P}\{N < u\mu^{3/2}/\sigma\} \ ,$$

which was to be shown.

Other possible universal models for random split trees.

We could have developed this theory based on other models. In a random split tree, the subtree sizes are multinomial $(n, V_1, \ldots, V_b)$, where $(V_1, \ldots, V_b)$ in turn is a random split vector. This introduces two levels of randomization. A more rigid and perhaps less universal model would fix an integer δ and require that the subtree sizes $N(u_1), \ldots, N(u_b)$ for the children $u_1, \ldots, u_b$ of a node u satisfy:

$$\max_{1 \le i \le b} |N(u_i) - nV_i| \le \delta \ .$$

Some of the trees discussed earlier fall into this framework. For example, in a random binary search tree, it is well-known that the left and right subtrees of the root have cardinalities (N_1, N_2) that are jointly distributed as $(\lfloor nU \rfloor, \lfloor n(1-U) \rfloor)$, where U is uniform $[0,1]$ and n is the cardinality of the tree. Setting $(V_1, V_2) = (U, 1-U)$, we thus have $\max_i |N_i - nV_i| \le 1$. Theorems 1 and 2 have straightforward equivalent versions (with the same dependence upon μ, σ^2 and $m(t)$). We should note that for generating extended AB trees for the purpose of simulation, the model of this section is more convenient. Here the split vectors V_1 and $V_2 = 1 - V_1$ are mixtures of beta random variables, but no multinomial sampling is necessary, as we use $(N_1, N_2) = (\lfloor nV_1 \rfloor, \lfloor nV_2 \rfloor)$ to determine subtree sizes at the root of a subtree of cardinality n. This way, each node will receive one ball.

Acknowledgment.

I thank Paul Kruszewski and two anonymous referees for great feedback on the manuscript.

References.

M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Tables*, Dover Publications, New York, NY, 1970.

A. V. Aho, J. E. Hopcroft, and J. D. Ullman, *Data Structures and Algorithms*, Addison-Wesley, Reading, MA, 1983.

D. Aldous, "Probability distributions on cladograms," Technical Report, Institute of Mathematics and Applications, University of Minnesota, 1993.

D. Aldous, B. Flannery, and J. L. Palacios, "Two applications of urn processes: the fringe analysis of search trees and the simulation of quasi-stationary distributions of Markov chains," *Probability in the Engineering and Information Sciences*, vol. 2, pp. 293–307, 1988.

E. Arkin, M. Held, J. Mitchell , and S. Skiena, “Hamiltonian triangulations for fast rendering,” in: *Algorithms–ESA’94*, ed. J. van Leeuwen, vol. 855, pp. 36–47, Lecture Notes in Computer Science, Springer-Verlag, 1994.

K. B. Athreya and P. E. Ney, *Branching Processes*, Springer-Verlag, Berlin, 1972.

D. Avis and H. El Gindy, “Triangulating point sets in space,” *Discrete Computational Geometry*, vol. 2, pp. 99–111, 1987.

C. J. Bell, “An investigation into the principles of the classification and analysis of data of an automatic digital computer,” Doctoral Dissertation, Leeds University, 1965.

J. D. Biggins, “The first and last-birth problems for a multitype age-dependent,” *Advances in Applied Probability*, vol. 8, pp. 446–459, 1976.

J. D. Biggins, “Chernoff’s theorem in the branching random walk,” *Journal of Applied Probability*, vol. 14, pp. 630–636, 1977.

H. Chernoff, “A measure of asymptotic efficiency of tests of a hypothesis based on the sum of observations ,” *Annals of Mathematical Statistics*, vol. 23, pp. 493–507, 1952.

E. G. Coffman and J. Eve, “File structures using hashing functions,” *Communications of the ACM*, vol. 13, pp. 427–436, 1970.

L. Devroye, *Non-Uniform Random Variate Generation*, Springer-Verlag, New York, 1986a.

L. Devroye, “A note on the height of binary search trees,” *Journal of the ACM*, vol. 33, pp. 489–498, 1986b.

L. Devroye, “Branching processes in the analysis of the heights of trees,” *Acta Informatica*, vol. 24, pp. 277–298, 1987.

L. Devroye, “Applications of the theory of records in the study of random trees,” *Acta Informatica*, vol. 26, pp. 123–130, 1988.

L. Devroye, “On the height of random m-ary search trees,” *Random Structures and Algorithms*, vol. 1, pp. 191–203, 1990.

L. Devroye, “On the expected height of fringe-balanced trees,” *Acta Informatica*, vol. 30, pp. 459–466, 1993.

L. Devroye and L. Laforest, “An analysis of random d-dimensional quadtrees,” *SIAM Journal on Computing*, vol. 19, pp. 821–832, 1990.

S. W. Dharmadhikari and K. Jogdeo, “Bounds on moments of certain random variables,” *Annals of Mathematical Statistics*, vol. 40, pp. 1506–1508, 1969.

R. A. Finkel and J. L. Bentley, "Quad trees: a data structure for retrieval on composite keys," *Acta Informatica*, vol. 4, pp. 1–9, 1974.

P. Flajolet, G. Gonnet, C. Puech, and J. M. Robson, "The analysis of multidimensional searching in quad-trees," in: *Proceedings of the Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 100–109, SIAM, Philadelphia, 1991.

P. Flajolet and T. Lafforgue, "Search costs in quadtrees and singularity perturbation analysis," *Discrete and Computational Geometry*, vol. 12, pp. 151–175, 1994.

P. Flajolet and A. Odlyzko, "The average height of binary trees and other simple trees," *Journal of Computer and System Sciences*, vol. 25, pp. 171–213, 1982.

P. Flajolet and R. Sedgewick, "Digital search trees revisited," *SIAM Journal on Computing*, vol. 15, pp. 748–767, 1986.

E. H. Fredkin, "Trie memory," *Communications of the ACM*, vol. 3, pp. 490–500, 1960.

D. K. Fuk and S. V. Nagaev, "Probability inequalities for sums of independent random variables," *Theory of Probability and its Applications*, vol. 16, pp. 643–660, 1971.

G. H. Gonnet and R. Baeza-Yates, *Handbook of Algorithms and Data Structures*, Addison-Wesley, Workingham, England, 1991.

I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, New York, 1980.

G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, Oxford University Press, 1992.

J. M. Hammersley, "Postulates for subadditive processes," *Annals of Probability*, vol. 2, pp. 652–680, 1974.

P. Jacquet and M. Régnier, "Trie partitioning process: limiting distributions," in: *Lecture Notes in Computer Science*, vol. 214, pp. 196–210, 1986.

R. Kemp, *Fundamentals of the Average Case Analysis of Particular Algorithms*, B.G.Teubner, Stuttgart, 1984.

J. F. C. Kingman, "Subadditive ergodic theory," *Annals of Probability*, vol. 1, pp. 883–909, 1973.

D. E. Knuth, *The Art of Computer Programming, Vol. 3 : Sorting and Searching*, Addison-Wesley, Reading, MA, 1973.

G. Louchard, "Exact and asymptotic distributions in digital and binary search trees," *Theoretical Informatics and Applications*, vol. 21, pp. 479–496, 1987.

W. C. Lynch, “More combinatorial problems on certain trees,” *Computer Journal*, vol. 7, pp. 299–302, 1965.

H. M. Mahmoud, “On the average internal path length of m-ary search trees,” *Acta Informatica*, vol. 23, pp. 111–117, 1986.

H. M. Mahmoud, *Evolution of Random Search Trees*, John Wiley, New York, 1992.

H. M. Mahmoud and B. Pittel, “On the joint distribution of the insertion path length and the number of comparisons in search trees,” *Discrete Applied Mathematics*, vol. 20, pp. 243–251, 1988.

H. Mahmoud and B. Pittel, “On the most probable shape of a search tree grown from a random permutation,” *SIAM Journal on Algebraic and Discrete Methods*, vol. 5, pp. 69–81, 1984.

J. Marcinkiewicz and A. Zygmund, “Sur les fonctions indépendantes,” *Fundamentales de Mathématiques*, vol. 29, pp. 60–90, 1937.

S. V. Nagaev and N. F. Pinelis, “Some inequalities for sums of independent random variables,” *Theory of Probability and its Applications*, vol. 22, pp. 248–256, 1977.

J. Nievergelt and K. H. Hinrichs, *Algorithms and Data Structures with Applications to Graphics and Geometry*, Prentice-Hall, Englewood Cliffs, NJ, 1993.

J. Nievergelt, H. Hinterberger, and K. C. Sevcik, “The grid file: an adaptable, symmetric multikey file structure,” *ACM Transactions on Database Systems*, vol. 9, pp. 38–71, 1984.

M. Okamoto, “Some inequalities relating to the partial sum of binomial probabilities,” *Annals of Mathematical Statistics*, vol. 10, pp. 29–35, 1958.

V. V. Petrov, *Sums of Independent Random Variables*, Springer-Verlag, Berlin, 1975.

B. Pittel, “On growing random binary trees,” *Journal of Mathematical Analysis and Applications*, vol. 103, pp. 461–480, 1984.

B. Pittel, “Asymptotical growth of a class of random trees,” *Annals of Probability*, vol. 13, pp. 414–427, 1985.

B. Pittel, “Paths in a random digital tree: limiting distributions,” *Advances in Applied Probability*, vol. 18, pp. 139–155, 1986.

B. Pittel, “Note on the heights of random recursive trees and random m-ary search trees,” *Random Structures and Algorithms*, vol. 5, pp. 337–347, 1994.

P. V. Poblete and J. I. Munro, "The analysis of a fringe heuristic for binary search trees," *Journal of Algorithms*, vol. 6, pp. 336–350, 1985.

P. Prusinkiewicz and A. Lindenmayer, *The Algorithmic Beauty of Plants*, Springer-Verlag, New York, 1990.

R. Pyke, "Spacings," *Journal of the Royal Statistical Society Series B*, vol. 7, pp. 395–445, 1965.

J. M. Robson, "The height of binary search trees," *The Australian Computer Journal*, vol. 11, pp. 151–153, 1979.

R. Y. Rubinstein, "Generating random vectors uniformly distributed inside and on the surface of different regions," *European Journal of Operations Research*, vol. 10, pp. 205–209, 1982.

H. Samet, *Applications of Spatial Data Structures*, Addison-Wesley, Reading, MA, 1990a.

H. Samet, *The Design and Analysis of Spatial Data Structures* , Addison-Wesley, Reading, MA, 1990b.

R. Sedgewick, "Mathematical analysis of combinatorial algorithms," in: *Probability Theory and Computer Science*, ed. G. Louchard and G. Latouche, pp. 123–205, Academic Press, London, 1983.

M. Sibuya, "Generalized hypergeometric, digamma and trigamma distributions," *Annals of the Institute of Statistical Mathematics*, vol. 31, pp. 373–390, 1979.

R. L. Smith, "Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions," *Operations Research*, vol. 32, pp. 1296–1308, 1984.

W. Szpankowski, "Some results on V-ary asymmetric tries," *Journal of Algorithms*, vol. 9, pp. 224–244, 1988.

X. G. Viennot, "Trees everywhere," in: *CAAP 90*, ed. A. Arnold, vol. 431, pp. 18–41, Lecture Notes in Computer Science, Springer-Verlag, Berlin, 1990.

J. S. Vitter and P. Flajolet, "Average-case analysis of algorithms and data structures," in: *Handbook of Theoretical Computer Science, Volume A: Algorithms and Complexity*, ed. J. van Leeuwen, pp. 431–524, MIT Press, Amsterdam, 1990.

A. Walker and D. Wood, "Locally balanced binary trees," *Computer Journal*, vol. 19, pp. 322–325, 1976.