RANDOM MINIMAX GAME TREES

LUC DEVROYE* AND OLIVIER KAMOUN*

Abstract. In this paper, we study random minimax trees of the incremental type. These are complete b-ary trees with n levels of edges, in which we associate independent identically distributed random variables with the edges. The value of a leaf is the sum of the edge values on the path to the root. The value of each internal node is obtained at alternating levels by taking the minimum or maximum value of the values of the children. We are interested in the behavior of the value of the root, V_n. For bounded generic edge random variable X, we show that V_n/n tends to a limit almost surely as $n \to \infty$. The limit is a highly nonlinear function of the distribution of X. For the case of a Bernoulli random variable with parameter p, the limit is a continuous function that is zero for p near zero, one for p near one, $1/2$ for p in an interval around $1/2$, and nonlinear inbetween. A comparison is made with a random minimax tree model studied by Pearl, in which the leaf values are independent.

Keywords and phrases. Game tree. Minimax tree. Probabilistic analysis. Asymptotic behavior. Random trees.

1. Introduction. Random MIN-MAX trees play a major role in the understanding of search strategies in game trees. They can be used to explain why certain search algorithms are preferable in some situations. For example, Pearl [Pearl80] considered a complete b-ary tree T with n levels of edges and associated with the leaves independent random variables all distributed as a given random variable X. To each internal node u with A_u its set of children, he associated a value $V(u)$ according to the standard MIN-MAX rule:

$$V(u) = \begin{cases} \max_{v \in A_u} \{V(v)\} & \text{if } u \text{ at even level in } T \\ \min_{v \in A_u} \{V(v)\} & \text{if } u \text{ at odd level in } T \end{cases} . \tag{1}$$

The parity of a level is with respect to distance from the root, not the leaves. The root is thus a maximizing node in all cases. The values of the nodes thus obtained should be thought of as values obtained by an evaluation function in a game. The trees obtained in this manner will be called Pearl trees. Pearl continued the study of this model in his book [Pearl84], in which several search alorithms are compared with respect to their average case behavior. The simplicity of Pearl's model leads to a beautiful and uncluttered analysis. Some results are quite striking: for example, if V_n is the value of a root node n levels removed from the leaf level, and if the generic random variable X is uniformly distributed on $[0, 1]$, then

$$\lim_{n\to\infty} V_{2n} = 1 - \xi_b \text{ almost surely ;}$$
$$\lim_{n\to\infty} V_{2n+1} = \xi_b \quad \text{almost surely ,}$$

* School of Computer Science, McGill University, Montreal, Canada H3A 2A7.

where ξ_b is the positive solution of $x^b + x - 1 = 0$. For $b = 2$, $\xi_b = (\sqrt{5}-1)/2 = 0.618\ldots$ is related the golden ratio. If X has distribution function F, one can obtain the correct limit result from this by applying the probability integral transform.

The limiting behavior of V_n depends upon whether the bottommost operation is a MIN or MAX. Hence the disturbing bi-asymptotic result mentioned above. If we had always started with a MIN operation from the leaf level up, then V_n would not have oscillated. Nau [Nau82a,Nau82b] pointed out this pathology in Pearl's model. It can all be traced back to the independence assumption for the leaf values. Incremental models are more responsive to this criticism. They effectively incorporate the notion that siblings in the tree should have highly correlated values: for an early survey, we refer to Newborn [Newb77]. Knuth and Moore [KM75] and Fuller [Newb77] have studied particular models of this kind. In this paper, we follow the model developed by Nau [Nau82a,Nau82b,Nau83]: with every edge in a b-ary complete tree T with n levels of edges, we associate an independently drawn random variable distributed as X. A leaf is given as value the sum of the edge values found on the path from the leaf to the root. Values of internal nodes are obtained by the rules laid out in (1). We take the liberty of calling our trees incremental trees, even though authors such as Nau [Nau82a,Nau82b] reserve this terminolgy for broader classes of trees.

The main question here is: how does the root's value (V_n) vary with n? The answer is both interesting and to some extent surprising. In discovering the answer, we will learn a lot about the rich structure of the tree. In any case, even for the simplest kind of edge random variable X, the analysis is much more involved than for the Pearl trees. Besides the general incremental tree with edge random variable X, we will also study the informative Bernoulli tree, which is an incremental tree in which X is a Bernoulli random variable with

$$\boldsymbol{P}\{X=1\} = 1 - \boldsymbol{P}\{X=0\} = p\ .$$

The first reassuring observation is that the oscillatory behavior of V_n seen for Pearl trees has disappeared. Our main result states that for all non-negative edge variables X, there exists a constant $\mathcal{V}(X)$ such that

$$\lim_{n\to\infty} \frac{V_n}{n} = \mathcal{V}(X) \text{ almost surely }.$$

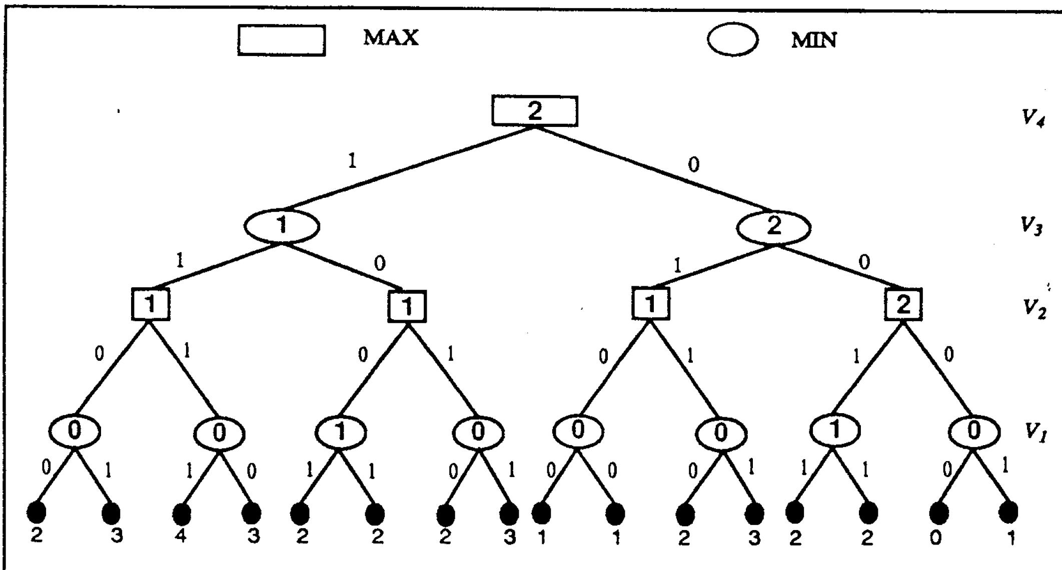
This result confirms that the root's value increases linearly with n, the number of levels. This was to be expected, since with $X \equiv 1$, we would have $V_n \equiv n$. We would of course like to know how $\mathcal{V}(X)$ is related to X. The relationship with the distribution of X is spectacularly nonlinear. For example, for Bernoulli trees with X Bernoulli (p), it is not true that $\mathcal{V}(X) \equiv p$, even though nearly all leaf values are very close to np. Consider now Pearl trees with n levels and in which leaf values are i.i.d. binomial (n,p) with $p \in (0,1)$ fixed. We have seen above that the root's value is either the ξ_b or $1-\xi_b$ quantile of the leaf distribution. For the binomial (n,p) distribution, both converge rapidly to np, and therefore, V_n/n is very close to p. This difference in behavior shows that incremental models are much more colorful and structurally interesting.

For the Bernoulli model in particular, we will give more details about the nonlinear behavior of $\mathcal{V}(X)$ as a function of p. For example, when p is near zero, $\mathcal{V}(X) = 0$. In this case, it will be shown that V_{2n} has a limit distribution function F_∞:

$$\lim_{n\to\infty} \boldsymbol{P}\{V_{2n} \le i\} = F_\infty(i)\ ,\ i \ge 0\ .$$

For $p > \alpha$, which is a fixed threshold depending upon b only, the behavior of V_n changes abruptly, as V_n starts to grow linearly with n: $\mathcal{V}(X) > 0$. The limit value $\mathcal{V}(X)$ varies nonlinearly with p. While the limit is continuous in p, it is constant and equal to $1/2$ on $[\beta, 1-\beta]$, an interval centered at $1/2$. As $b \to \infty$, this central interval extends to $(0,1)$, making the root's value basically independent of p.

We will also establish how close V_n is to $\boldsymbol{E}V_n$. The results in this paper are based upon the thesis of the second author [Kam92]. They are a first limited step towards the understanding of incremental models. Hopefully, the methods used below will be useful in the study of the performance of search algorithms on incremental trees. For more details and more elaborate simulations than those reported here, we refer to the thesis.



A SUM model.

2. Another construction of the incremental model. We can look at the incremental model with edge random variable X in an equivalent manner that is better suited for analysis. Again, we consider an n-level complete b-ary tree T. Let u be an internal node, and let A_u be the set of its children. For all $v \in A_u$, we associate with the edge (u, v) an independent drawing $E(u, v)$ of a given random variable X. Let F be the distribution function of X:

$$F(x) = \boldsymbol{P}\{X \le x\} .$$

With each node u we associate a value according to the following recurrence: if u is a leaf, then $V(u) = 0$. The level of a node is determined by its distance from the leaf level. For an internal node u we define

$$V(u) = \begin{cases} \max_{v \in A_u} \{V(v) + E(u, v)\} & \text{if } u \text{ at even level} \\ \min_{v \in A_u} \{V(v) + E(u, v)\} & \text{if } u \text{ at odd level} \end{cases} . \tag{2}$$

All nodes at path distance n from the leaf level are independent and identically distributed. A generic random variable of this kind is denoted by V_n. It is easy to see that this is the value of the root of a tree of height n which follows an incremental model with edge distribution F. Thus, $V_0 \equiv 0$. Clearly, we have the following distributional identities:

$$V_n \stackrel{\mathcal{L}}{=} \begin{cases} \max_{1 \le j \le b} \{V_{n-1,j} + X_j\} & \text{if } n \text{ is even} \\ \min_{1 \le j \le b} \{V_{n-1,j} + X_j\} & \text{if } n \text{ is odd} \end{cases} ,$$

where X_j denotes an independent copy of the random variable X, and $V_{n-1,j}$ denotes an independent copy of V_{n-1}. Let F_n be the distribution function of V_n:

$$F_n(x) = \boldsymbol{P}\{V_n \le x\} .$$

Clearly, we see that

$$F_0(i) = \begin{cases} 0 & \text{if } i < 0 \\ 1 & \text{if } i \ge 0 \end{cases} .$$

When X is a continuous random variable, the distribution function of $V_{2n-1,j} + X_j$ is the convolution of F and F_{2n-1}. Thus, we have the following relations:

$$\begin{aligned} F_{2n}(x) &= \left(\textstyle\int F_{2n-1}(x-t)dF(t)\right)^b , \\ F_{2n+1}(x) &= 1 - \left(1 - \textstyle\int F_{2n}(x-t)dF(t)\right)^b . \end{aligned} \tag{3}$$

When X is an integer-valued random variable, and f is the discrete probability density ($f(j) = \boldsymbol{P}\{X = j\}$),

$$\begin{aligned} F_{2n}(i) &= \left(\textstyle\sum_j f(j)\ F_{2n-1}(i-j)\right)^b , \\ F_{2n+1}(i) &= 1 - \left(1 - \textstyle\sum_j f(j)\ F_{2n}(i-j)\right)^b . \end{aligned} \tag{4}$$

3. Limit of $F_{2n}(0)$ as n tends to infinity. The behavior of incremental trees when the edge random variables are mostly zero is peculiar and forms the basis of further analysis in future sections. We assume that X is an integer-valued random variable. Let F be the distribution function of X and k be the smallest non-zero value taken by X. We define $p = \boldsymbol{P}\{X > 0\} = F(k)$. Then V_n is clearly stochastically bigger than kV_n' where V_n' is the root's value of the Bernoulli tree with parameter p. We prove the following theorem.

THEOREM 1. *For all b there exists $\alpha \in (0,1)$ such that, for $p \in [0,\alpha]$,*

$$\lim_{n\to\infty} F_{2n}(0) > 0,$$

and

$$\lim_{n\to\infty} F_{2n+1}(0) > 0.$$

And for $p > \alpha$

$$\lim_{n\to\infty} F_{2n}(0) = \lim_{n\to\infty} F_{2n+1}(0) = 0 .$$

Finally,

$$\alpha \leq 1 - b^{-1/(b+1)} \overset{b\to\infty}{\longrightarrow} 0.$$

Proof. We first prove Theorem 1 for X Bernoulli(p). Then using the remark given at the beginning of the chapter the theorem follows immediately for general discrete positive random variable. Let $p \in [0,1)$. Use recurrences (4) for $i = 0$:

$$\begin{aligned} F_0(0) &= 1, \\ F_{2n}(0) &= \Big((1-p)F_{2n-1}(0)\Big)^b, \\ F_{2n+1}(0) &= 1 - \Big(1-(1-p)F_{2n}(0)\Big)^b . \end{aligned}$$

Combining all this, we note that for $n \geq 1$,

$$F_{2n}(0) = (1-p)^b\left(1-\Big(1-(1-p)F_{2n-2}(0)\Big)^b\right)^b \overset{\text{def}}{=} G_0(F_{2n-2}(0)) ,$$

where $G_0(x) = (1-p)^b\Big(1-\big(1-(1-p)x\big)^b\Big)^b$. This is a simple functional iteration, the solution of which depends upon the behavior of the mapping G_0. $G_0(x)$ is an order b^2 polynomial that is a strictly increasing mapping: $[0,1] \to [0,1]$, since $G_0(0) = 0$ and $G_0(1) \leq 1$. Then $F_{2n}(0)$ is decreasing and converges to L_0, the greatest fixed point on $[0,1]$. Define the set of p such that $G_0(x)$ has a non-zero fixed point on $[0,1]$:

$$\Gamma = \{p \in [0,1] | L_0 > 0\}.$$

Define also

$$h(x,p) = \frac{G_0(x) - x}{x}.$$

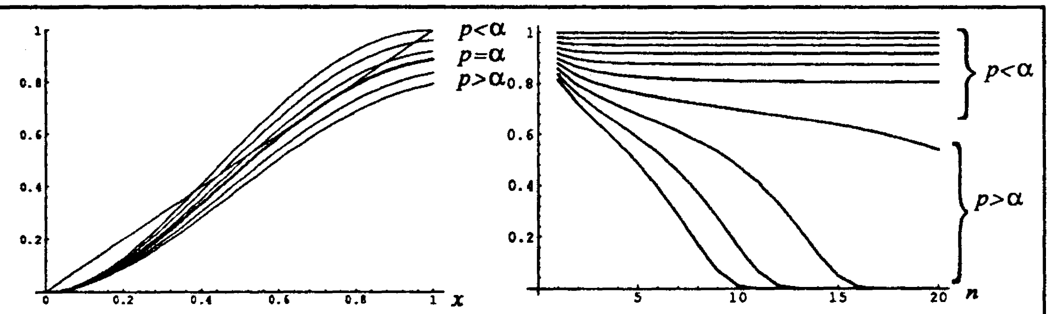
As $G_0(0) = 0$, h is a (b^2-1)-th order polynomial function of p and x. As $G_0'(0) = 0$, the derivative of $G_0(x) - x$ is equal to -1 at $x = 0$, and thus zero is a simple root of $G_0(x) - x$ and it is not a root of h. Thus we have,

$$\Gamma = \{p \in [0,1] | h(x,p) \text{ has a root in } [0,1]\}.$$

Since h is continuous, the inverse image of $\{0\}$ is a closed set of $\Re^2$, and Γ too is a closed set. Since G_0 is decreasing in p, h is also decreasing in p. We also have that $0 \in \Gamma$ since $h(1,0) = 0$. We will prove that there exists $\alpha \in \Re$ such that $\Gamma = [0,\alpha]$. We already know that Γ is a closed set containing zero. Thus, we just have to prove that Γ is convex. Assume that $p \in \Gamma$. Thus there exists $L_0 > 0$ such that $h(L_0, p) = 0$. Then for all $p' \in (0,p]$ we have,

$$h(L_0, p') \geq 0, \ h(1,p') \leq 0.$$

Thus $h(x,p')$ has a non-zero root in $[L_0, 1]$ and $p' \in \Gamma$. This implies that Γ is convex. Thus Theorem alphatheo is proved for $p \leq \alpha$.



$G_0(x)$ for b=2 and p between 0 an 0.1. $F_{2n}(0)$ for p=0 to p=0.1, b=2.

If $p > \alpha$ then G_0 has only zero as fixed point. Thus $F_{2n}(0)$ converges to zero. We get similar results with $F_{2n+1}(0)$. It tends to a positive limit if an only if $p \in [0,\alpha]$. These facts can be shown using

$$F_{2n+1}(0) = 1 - \left(1 - (1-p)F_{2n}(0)\right)^b .$$

We still have to prove that $\alpha > 0$. For $p = 0$, we have

$$G_0(x) = \left(1 - (1-x)^b\right)^b, \ G_0(1) = 1, \ G_0'(1) = 0 < 1.$$

As G_0 is differentiable and $G_0'(1) = 0$, there exists $0 < y < 1$ such that $G_0(y) > y$. And by continuity of G_0 in p, there exists an $\varepsilon > 0$ such that for all $p < \varepsilon$, $G_0(y) > y$. This implies that for $p < \varepsilon$, G_0 has a fixed point on $(0,1)$ and then we have $\alpha > \varepsilon > 0$. Thus the first part of Theorem 1 is proved.

For general $b \geq 2$, we derive an upper bound for α. We have :

$$\begin{aligned} G_0(x) &= q^b\left(1-\left(1-qx\right)^b\right)^b \\ &\leq q^b(1-(1-qx)^b) \\ &\leq q^b(1-(1-bqx)) \\ &= bq^{b+1}x. \end{aligned}$$

Thus if $bq^{b+1} < 1$, $G_0(x) < x$ for all $x > 0$. This implies that it cannot have a non-zero fixed point and thus $p > \alpha$. Thus we have

$$\alpha \leq 1 - \sqrt[b+1]{\frac{1}{b}}.$$

This implies that α tends to zero when b tends to infinity. □

REMARK 1. For $b = 2$,

$$\alpha = 1 - \sqrt[3]{\left(\frac{27}{32}\right)} \approx 0.05506.$$

4. The main theorem of convergence. THEOREM 2. *For the incremental tree with bounded edge variable X, $\boldsymbol{E}V_n/n$ converges to a finite limit $\mathcal{V}(X)$ as n tends to infinity.*

The proof will be spread over the next two sections. Assume without loss of generality that for some finite μ, $0 \leq X \leq \mu$. It is helpful to introduce an associated tree T' in which we introduce new node values $V'(u)$. However, at every node, we force $V'(u) \geq V(u)$. First fix the integers $N \geq 1$ and $k \geq 1$. The leaves of T' have value zero. At any level i that is not a multiple of N, we follow the standard rules (2) as for an incremental tree with edge variable X. If i is a multiple of N, say $i = lN$, then we set for any node u at level i

$$W(u) = \begin{cases} \max_{v \in A_u} \{V'(v) + E(u,v)\} & \text{if } i \text{ is even} \\ \min_{v \in A_u} \{V'(v) + E(u,v)\} & \text{if } i \text{ is odd} \end{cases},$$

and,

$$V'(u) = \begin{cases} l\boldsymbol{E}V_N + (2l-1)k\mu & \text{if } W(u) \leq l\boldsymbol{E}V_N + (2l-1)k\mu \\ \infty & \text{if } W(u) > l\boldsymbol{E}V_N + (2l-1)k\mu \end{cases}.$$

Note that many nodes may have the value infinity. We call T' the (k, N) associated tree. Let V_n' be the random variable defined as the root value of such a tree with n levels of edges. Note that for all nodes at levels that are multiple of N, the values of $V'(u)$ are either ∞ or a given fixed finite value that is the same at that level. This device already hints at the behavior of incremental trees: all nodes at a given level have approximately equal values—no serious imbalances occur. The entire paper rests on three technical inequalities that deserve of a section of their own.

5. The fundamental inequalities. The first result states that it is "doubly exponentially" unlikely that the root of a Pearl tree with Bernoulli leaf values takes the value one, if the Bernoulli parameter is small enough.

LEMMA 1. *Let T be a b-ary Pearl tree with Bernoulli (q) leaf values, where*

$$q \leq \xi \stackrel{\text{def}}{=} \frac{1}{2} b^{-b} .$$

Then, regardless of whether we begin with either MIN *or* MAX *nodes, and regardless of the parity of n,*

$$\mathbf{P}\{V_n = 1\} \leq b 2^{-b^{\lfloor n/2 \rfloor}} .$$

Proof. V_n is maximal if we begin with a MAX level. For a leaf value V_0, we have $\mathbf{P}\{V_0 = 1\} = q$. Define $p_n = \mathbf{P}\{V_n = 1\}$. Then the following recursion holds: $p_0 = q$, and

$$p_{2n} = \left(1 - (1 - p_{2n-2})^b\right)^b \leq (b p_{2n-2})^b .$$

This yields

$$p_{2n} \leq b^{b^{n+1}} q^{b^n} .$$

For p_{2n+1} we get

$$p_{2n+1} \leq 1 - \left(1 - b^{b^{n+1}} q^{b^n}\right)^b \leq b b^{b^{n+1}} q^{b^n} .$$

Let q such that $b^b q \leq \frac{1}{2}$. Then

$$p_{2n} \leq 2^{-b^n}, \text{ and } p_{2n+1} \leq b 2^{-b^n},$$

and in general, regardless of whether we start with a MIN or a MAX level,

$$p_n \leq b 2^{-b^{\lfloor n/2 \rfloor}} .$$

□

The second technical result establishes quite simply that V_n is close to $\boldsymbol{E}V_n$ in all circumstances.

LEMMA 2. *Let $0 \le X \le \mu$ in a random incremental tree. For all $\varepsilon > 0$,*

$$\boldsymbol{P}\{|V_n - \boldsymbol{E}V_n| \ge \varepsilon\} \le 2e^{-2\varepsilon^2/n\mu^2} .$$

Proof. At the i^{th} level of edges, starting from the topmost level, we find b^i independent edge values. These are collected in a random vector U_i. Clearly then, $V(u) = f(U_1, \ldots, U_n)$ for some function f. Furthermore, if U_i is replaced by a different vector U_i', $V(u)$ changes by at most μ. Thus, we can apply the McDiarmid's inequality (1989)[McDi89]: for all $\varepsilon > 0$,

$$\boldsymbol{P}\{|V_n - \boldsymbol{E}V_n| \ge \varepsilon\} \le 2e^{-2\varepsilon^2/n\mu^2} . \tag{5}$$

□

Our third inequality is fundamental in proving that V_n increases about linearly in n. It states that a node at level lN has a value not much larger than l times the value of a node at level N. The explicit non-asymptotic nature of the bound will be helpful as well.

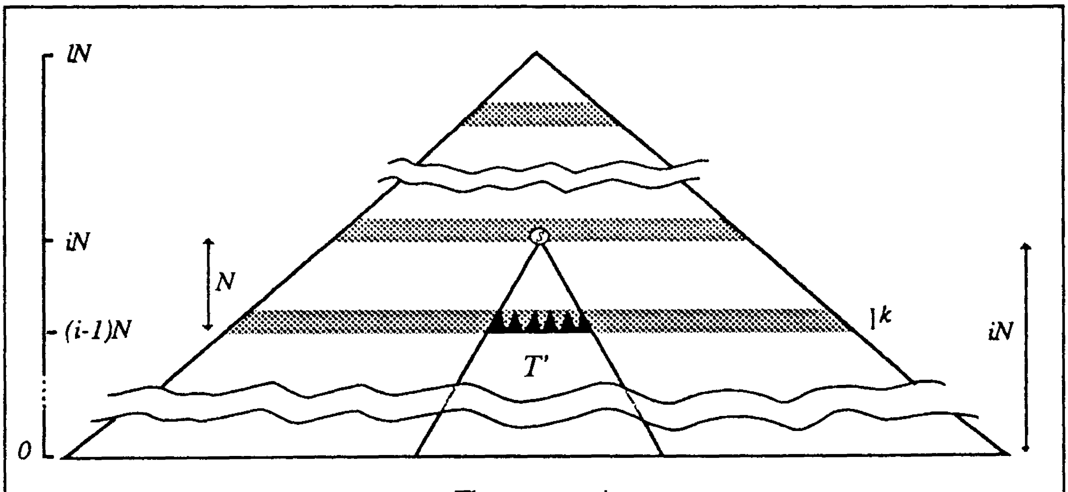
LEMMA 3. *Let $0 \le X \le \mu$ in a random incremental tree. For N large enough and for all $l > 0$,*

$$\boldsymbol{P}\{V_{lN} \ge l\boldsymbol{E}V_N + (2l-1)k\mu\} \le b^{N-N^{2/3}} b\, 2^{-b^{\frac{N^{2/3}}{2}-1}} + 2e^{-2N^{1/3}} \stackrel{\text{def}}{=} R(b, N),$$

where $k = \lceil N^{2/3} \rceil$. Finally for all $\varepsilon > 0$ there exists an N such that for all $n > N$,

$$\boldsymbol{P}\{|V_n - l\boldsymbol{E}V_N| \ge (2l-1)k\mu + N\mu\} \le \varepsilon,$$

where $l = \lfloor n/N \rfloor$ and $k = \lceil N^{2/3} \rceil$.



The construction

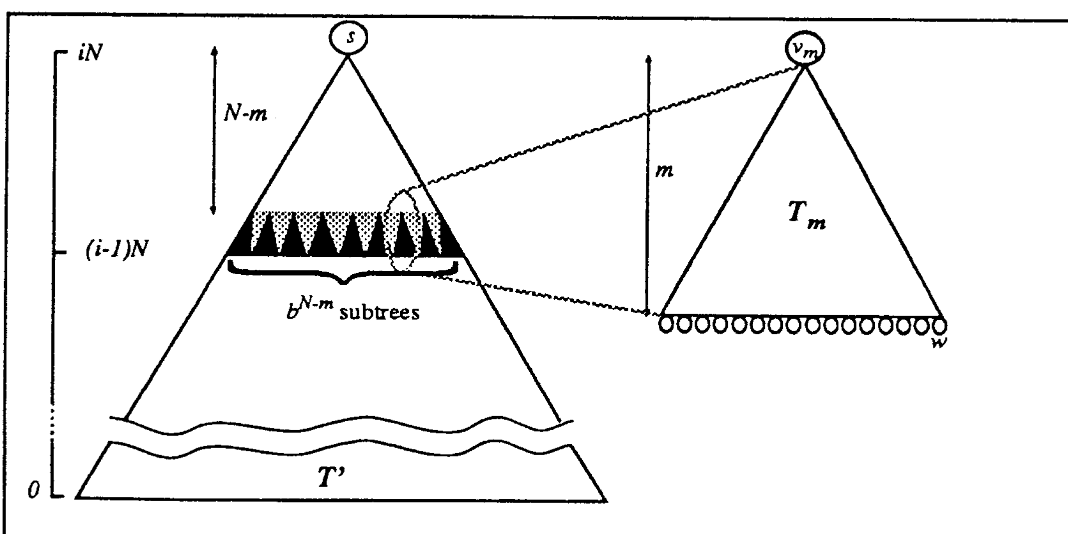
Proof. Assume that the tree has $n = lN$ levels. We consider the (k, N) associated tree with $k = \lceil N^{2/3} \rceil$. Let N be so large that $R(b, N) \le (1/2)b^{-b}$. We prove by induction that for such N and for all $i \ge 1$,

$$\boldsymbol{P}\{V'_{iN} = \infty\} \le R(b, N). \tag{6}$$

For $i = 1$, we obtain

$$\begin{aligned} \boldsymbol{P}\{V'_N = \infty\} &= \boldsymbol{P}\{V_N > \boldsymbol{E}V_N + k\mu\}, \\ &\le 2e^{-2(k\mu)^2/N\mu^2} \qquad \text{(according to (5))} \\ &\le 2e^{-2N^{1/3}} \\ &\le R(b, N). \end{aligned}$$

Now we assume that $\boldsymbol{P}\left\{V'_{(i-1)N} = \infty\right\} \le R(b, N)$. The nodes at level iN are i.i.d. distributed as V'_{iN}. Let T' be an associated tree with iN levels and s its root node. Then look at the $m < N$ levels of this tree from depth $N - m$ to depth N. This part consists of b^{N-m} m-level subtrees as shown in the next figure. Let T_m be one of these subtrees and let v_m be its root. Thus $V'(v_m)$ is distributed as $V'_{(i-1)N+m}$. The leaves w of T_m are nodes of T' at level $(i-1)N$. Thus their values $(V'(w))$ are bi-valued i.i.d. random variables distributed as $V'_{(i-1)N}$.



We assign to each leaf node w of T_m a value $V''(w)$ as follows:

$$V''(w) = \begin{cases} \infty & \text{if } V'(w) = \infty \\ 0 & \text{if } V'(w) < \infty \end{cases}.$$

And to each internal node u of T_m we assign a value $V''(u)$ using the MIN-MAX rules 1:

$$V''(u) = \begin{cases} \max_{v \in A_u} \{V''(v)\} & \text{if } u \text{ is a MAX node of } T' \\ \min_{v \in A_u} \{V''(v)\} & \text{if } u \text{ is a MIN node of } T' \end{cases}.$$

Then $V''(v_m)$ is distributed as the root of a m-level b-ary Pearl tree, where the leaves take value ∞ with probability $q = \boldsymbol{P}\{V''(w) = \infty\}$. The bottom level is a MIN or a MAX according to the parity of $(i-1)N$. Thus as $q = \boldsymbol{P}\{V''(w) = \infty\} = \boldsymbol{P}\left\{V'_{(i-1)N} = \infty\right\} \leq R(b, N) \leq (1/2)b^{-b}$, by Lemma 1 about Pearl trees,

$$\boldsymbol{P}\{V''(v_m) = \infty\} \leq b2^{-b^{\lfloor m/2 \rfloor}}.$$

Let u be an internal node of T_m. As it is not at a level that is a multiple of N in T', $V'(u)$ is computed with the standard rules of the incremental model. Thus $V'(u)$ is infinity if and only if $V''(u)$ is infinity. If $V''(v_m) = 0$, then

$$V'(v_m) \leq (i-1)\boldsymbol{E}V_N + (2i-3)k\mu + m\mu. \tag{7}$$

Thus,

$$\boldsymbol{P}\{V'(v_m) = \infty\} \leq b2^{-b^{\lfloor m/2 \rfloor}}.$$

Furthermore we have

$$\begin{aligned} Q_m &\stackrel{\text{def}}{=} \boldsymbol{P}\{V'(v_m) = \infty \text{ for at least one node } v_m \text{ at depth } N-m \text{ from the top of } T'\} \\ &\leq b^{N-m} b2^{-b^{\lfloor m/2 \rfloor}}. \end{aligned}$$

Now we take $m = k = \lceil N^{2/3} \rceil$. If there is no infinity node at depth $N-k$ in T', then each $V'(v_m)$ is less than $(i-1)\boldsymbol{E}V_N + (2i-3)k\mu + k\mu$, and $V'(s)$ is stochastically less than $V'(v_m) + V_N$. Thus

$$\begin{aligned} \boldsymbol{P}\{V'_{iN} = \infty\} &= \boldsymbol{P}\{V'(s) > i\boldsymbol{E}V_N + (2i-1)k\mu\} \\ &\leq Q_k + \boldsymbol{P}\{(i-1)\boldsymbol{E}V_N + (2i-3)k\mu + k\mu + V_N \\ &\qquad > i\boldsymbol{E}V_N + (2i-1)k\mu\} \\ &\leq Q_k + \boldsymbol{P}\{V_N > \boldsymbol{E}V_N + k\mu\} \\ &\leq b^{N-N^{2/3}} b\, 2^{-b^{\frac{N^{2/3}}{2}-1}} + 2e^{-2N^{1/3}} \\ &= R(b, N) \\ &\leq \frac{1}{2} b^{-b}. \end{aligned}$$

Thus the induction proof of (6) is finished and we have

$$\boldsymbol{P}\{V_{lN} \geq l\boldsymbol{E}V_N + (2l-1)k\mu\} \leq R(b, N).$$

Next, we are left with a minor cleanup to handle the case when n is not a multiple of N. Now we consider b-ary incremental trees with n levels, n not a multiple of N, and we set $l = \lfloor \frac{n}{N} \rfloor$. Using 7 with $m = n - lN < N$ we get

$$\boldsymbol{P}\{V_n \geq l\boldsymbol{E}V_N + (2l-1)k\mu + m\mu\} \leq \boldsymbol{P}\{V''(v_m) = \infty\} \leq b2^{-b^{\lfloor m/2 \rfloor}}.$$

Thus if $N > m \geq N^{1/4}$ we have,

$$P\{V_n \geq lEV_N + (2l-1)k\mu + N\mu\} \leq b2^{-b\lfloor N^{1/4}\rfloor/2}. \tag{8}$$

If $m \leq N^{1/4}$ the probability that $V_n \geq lEV_N + (2l-1)k\mu + N\mu$ is less than the probability that there is at least an infinity node at level lN of the associated tree. Thus,

$$\begin{aligned} P\{V_n \geq lEV_N + (2l-1)k\mu + N\mu\} &\leq b^m R(b,N) \\ &\leq b^{N^{1/4}} R(b,N). \end{aligned} \tag{9}$$

Finally using (8) and (9) we have for all n,

$$P\{V_n \geq lEV_N + (2l-1)k\mu + N\mu\} \leq b2^{-b\lfloor N^{1/4}\rfloor} + b^{N^{1/4}} R(b,N).$$

The right hand side tends to zero when N tends to infinity. Thus, for all $\varepsilon > 0$, there exists an N such that for $n > N$,

$$P\{V_n \geq lEV_N + (2l-1)k\mu + N\mu\} \leq \varepsilon,$$

where $k = \lceil N^{2/3}\rceil$ and $l = \lfloor n/N \rfloor$. □

6. Convergence: proof of Theorem 2. We show that

$$\limsup_{n\to\infty} \frac{EV_n}{n} \leq \liminf_{n\to\infty} \frac{EV_n}{n}, \tag{10}$$

by showing that for given $\varepsilon > 0$,

$$\limsup_{n\to\infty} \frac{EV_n}{n} \leq \frac{EV_N}{N} + 3\varepsilon$$

for all N large enough. Then, by definition of the limit infimum we can find an N so large that

$$\frac{EV_N}{N} \leq \liminf_{n\to\infty} \frac{EV_n}{n} + \varepsilon,$$

so that we may conclude (10) by the arbitrary nature of ε.

We use the notation of the preceding part: u is the root of an n-level complete b-ary incremental tree with edge variable X, and $V'(u)$ is its value for the (k,N) associated tree. Define $l = \lfloor n/N \rfloor$. If $V'(u) \leq N\mu + lEV_N + (2l-1)k\mu$ then

$$\begin{aligned} V(u) &\leq N\mu + \left\lfloor \frac{n}{N} \right\rfloor EV_N + \left(2\left\lfloor \frac{n}{N}\right\rfloor - 1\right)k\mu \\ &\leq N\mu + \frac{n}{N}EV_N + 2\frac{nk}{N}\mu, \end{aligned}$$

so that (recalling $k = \lceil N^{2/3} \rceil$),

$$
\begin{aligned}
\frac{V(u)}{n} &\le \frac{N}{n}\mu + \frac{\boldsymbol{E}V_N}{N} + 2\frac{k}{N}\mu \\
&\le \frac{N}{n}\mu + \frac{\boldsymbol{E}V_N}{N} + 2\left(\frac{N^{2/3}+1}{N}\right)\mu \\
&\le \frac{\boldsymbol{E}V_N}{N} + 2\varepsilon \qquad (11)
\end{aligned}
$$

for N large enough and $n \ge N/\mu\varepsilon$. Using Lemma 3, we can find N large enough such that

$$\boldsymbol{P}\left\{V_n \ge l\boldsymbol{E}V_N + (2l-1)k\mu + N\mu\right\} < \varepsilon$$

for all p and for $n > N$. Thus we have,

$$
\begin{aligned}
\frac{\boldsymbol{E}V_n}{n} &\le \boldsymbol{P}\left\{\frac{V_n}{n} > \frac{\boldsymbol{E}V_N}{N} + 2\varepsilon\right\} + \frac{\boldsymbol{E}V_N}{N} + 2\varepsilon \\
&\le \boldsymbol{P}\left\{V_n \le l\boldsymbol{E}V_N + (2l-1)k\mu + N\mu\right\} + \frac{\boldsymbol{E}V_N}{N} + 2\varepsilon \\
&\le \frac{\boldsymbol{E}V_N}{N} + 3\varepsilon
\end{aligned}
$$

for N large enough and $n \ge N/\mu\varepsilon$. This implies that

$$\limsup_{n\to\infty} \frac{\boldsymbol{E}V_n}{n} \le \frac{\boldsymbol{E}V_N}{N} + 3\varepsilon$$

as required. Thus $\boldsymbol{E}V_n/n$ has a limit $\mathcal{V}(X)$ when n tends to infinity. □

7. A law of large numbers. THEOREM 3. *If $X \in [0,\mu]$ and $\boldsymbol{P}\{X > 0\} > \alpha$, where α is defined in Theorem 1, we have*

$$\lim_{n\to\infty} \frac{\boldsymbol{E}V_n}{n} = \mathcal{V}(X) > 0$$

and

$$\lim_{n\to\infty} \frac{V_n}{\boldsymbol{E}V_n} = \lim_{n\to\infty} \frac{V_n}{n\mathcal{V}(X)} = 1$$

almost surely as n tends to infinity.

Proof. We first prove this lemma for Bernoulli trees with parameter p. Let u be a node of T in an n-level complete b-ary incremental tree with parameter p. We associate with u the value $V'(u)$ related to $V(u)$ by monotonicity: $V'(u) \le V(u)$. The idea is to cut the tree into pieces of N levels each, and for every second piece, we force all edge values to be zero. On those pieces, we use results about Pearl trees. The exact definition of

$V'(u)$ is given below. We denote by V'_n the value of the root of the n-level model. We will show that $\liminf_{n\to\infty} \boldsymbol{E}V'_n/n > 0$.

Let N be a large fixed positive integer. For all nodes u at level i, we determine $V'(u)$ from $V'(v), v \in A_u$ as follows for $l = \lceil n/(2N) \rceil$:

1) If $(2l-2)N < i < (2l-1)N$, then $V'(u)$ is determined from $V'(v), v \in A_u$ as in the incremental tree with parameter p.
2) If $i = (2l-1)N$, then first $W(u)$ is determined from $V'(v), v \in A_u$, by MIN-MAX rules as in the incremental tree with parameter p, and we set

$$V'(u) = \begin{cases} -\infty & \text{if } W(u) < l \\ l & \text{if } W(u) \geq l \end{cases}.$$

(Thus, at this level, $V'(u)$ is bi-valued!)

3) If $(2l-1)N < i \leq 2lN$, then the edge values are considered to be zero and thus $V'(u)$ is determined by the MIN-MAX rules

$$V'(u) = \begin{cases} \max_{v\in A_u} \{V'(v)\} & \text{if } u \text{ at even level in } T \\ \min_{v\in A_u} \{V'(v)\} & \text{if } u \text{ at odd level in } T \end{cases} \tag{12}$$

It is easy to verify by induction that $V'(u) \leq V(u)$ for every node. Now we prove by induction that if $p > \alpha$, for all $\varepsilon > 0$ we can find N such that for all integer $l > 0$ we have

$$\boldsymbol{P}\{V'_{lN} = -\infty\} \leq \min\left(\varepsilon, \frac{1}{2}b^{-b}\right). \tag{13}$$

For $l = 1$, this is true since

$$\boldsymbol{P}\{V'_N = -\infty\} = \boldsymbol{P}\{W_N = 0\} = F_N(0)$$

where F_N is the distribution function of the value of the root of an N-level incremental tree with parameter p, and $W_N = W(u)$ is the value of the root of this tree. (Recall that for $p > \alpha$, $F_N(0) \to 0$ as $N \to \infty$.) Thus we choose N so large that

$$F_N(0) < \frac{1}{2}\min\left(\varepsilon, \frac{1}{2}b^{-b}\right).$$

For the induction we have to distinguish between two cases. First we consider a node u at a level $2lN$. We assume (13) to be true for all $l' < 2l$. All nodes v at level $(2l-1)N$ have a value $V'(v)$ equal to l or $-\infty$. We consider the N-level subtree T_N rooted at the node u and in which the leaf values are the $V'(v)$ from level $(2l-1)N$ of T. Also $V'(u)$ is distributed as the root of a Pearl tree where the leaves have value l or $-\infty$. By the induction hypothesis the value $-\infty$ occurs with probability $q \leq (1/2)b^{-b}$. Thus by Lemma 1 we have,

$$\boldsymbol{P}\{V'(u) = -\infty\} \leq b2^{-b^{\lfloor N/2 \rfloor}} \leq \min\left(\varepsilon, \frac{1}{2}b^{-b}\right), \quad \text{for } N \text{ large enough.}$$

For all $m \geq 0$,

$$\text{(14)} \qquad \boldsymbol{P}\{V_{2lN+m} < l\} \leq \min\left(\varepsilon, \frac{1}{2}b^{-b}\right).$$

This concludes the first part. Let us now consider a node u at level $(2l-1)N$. According to the hypothesis, at level $(2l-2)N$, there are nodes with value $(l-1)$ and nodes with value $-\infty$. The probability that at least one node v at level $(2l-2)N$ has value $V'(v) = -\infty$ is less than $b^N b 2^{-b^{\lfloor N/2 \rfloor}}$. If the b^N nodes at level $(2l-2)N$ have the value $l-1$, then $\boldsymbol{P}\{V'(u) = -\infty\} = \boldsymbol{P}\{W_N = 0\} = F_N(0)$. Thus if we choose N such that $b^{N+1}2^{-b^{\lfloor N/2 \rfloor}} \leq \frac{1}{2}\min(\varepsilon, \frac{1}{2}b^{-b})$,

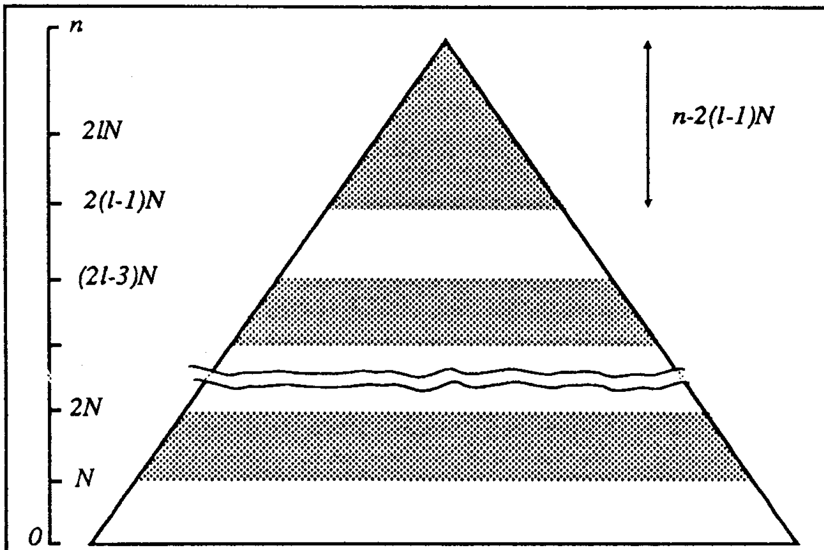
$$\begin{aligned}
\boldsymbol{P}\{V'(u) = -\infty\} &\leq b^{N+1}2^{-b^{\lfloor N/2 \rfloor}} + F_N(0) \\
&< \frac{1}{2}\min\left(\varepsilon, \frac{1}{2}b^{-b}\right) + \frac{1}{2}\min\left(\varepsilon, \frac{1}{2}b^{-b}\right) \\
&= \min\left(\varepsilon, \frac{1}{2}b^{-b}\right).
\end{aligned}$$

Thus the induction is shown and we have for all integer l,

$$\boldsymbol{P}\{V'_{lN} = -\infty\} \leq \varepsilon.$$

Thus for all $l > 0$,

$$\boldsymbol{P}\{V_{2lN} \geq l\} \geq \boldsymbol{P}\{V'_{2lN} = l\} \geq 1 - \varepsilon.$$



Inside the gray parts we assume the edge values to be zero.

We now generalize the result for incremental trees with a number of levels that is not a multiple of N. Let u be the root of T, an n-level b-ary incremental tree with parameter $p > \alpha$. Let $l = \lfloor n/2N \rfloor$. Using (14), with $m = n - 2lN$, we have

$$\boldsymbol{P}\{V(u) < l\} \le \min\left(\varepsilon, \frac{1}{2}b^{-b}\right).$$

Thus,

$$\boldsymbol{E}V_n \ge l(1-\varepsilon).$$

As a consequence,

$$\frac{\boldsymbol{E}V_n}{n} \ge (1-\varepsilon)\left(\frac{n}{2N} - 1\right)\frac{1}{n} \ge \frac{1-\varepsilon}{2N} - o(1),$$

so that

$$\liminf_{n\to\infty} \frac{\boldsymbol{E}V_n}{n} \ge \frac{1-\varepsilon}{2N},$$

and finally

$$\lim_{n\to\infty} \frac{\boldsymbol{E}V_n}{n} > 0.$$

We now consider X to be a positive random variable bounded by μ such that

$$p \stackrel{\text{def}}{=} \boldsymbol{P}\{X > 0\} > \alpha.$$

Then there exists $\delta > 0$ such that

$$\boldsymbol{P}\{X \ge \delta\} > \alpha.$$

Then

$$\mathcal{V}(X) \ge \delta\mathcal{V}(p) > 0.$$

The theorem follows immediately. In particular, whenever X is continuous, $\boldsymbol{P}\{X > 0\} = 1$, and thus $\mathcal{V}(X) > 0$.

For all $p > \alpha$ there exist $c > 0$ and $n_0 > 1$ such that for all $n > n_0$, $\boldsymbol{E}V_n > cn$. According to Lemma 2, for any p, and any $\varepsilon > 0$,

$$\boldsymbol{P}\{|V_n - \boldsymbol{E}V_n| \ge \varepsilon\} \le 2e^{-2\varepsilon^2/n\mu^2}.$$

Thus for $n > n_0$,

$$\begin{aligned}
\boldsymbol{P}\left\{\left|\frac{V_n}{\boldsymbol{E}V_n} - 1\right| \ge \varepsilon\right\} &= \boldsymbol{P}\{|V_n - \boldsymbol{E}V_n| \ge \varepsilon\boldsymbol{E}V_n\} \\
&\le \boldsymbol{P}\{|V_n - \boldsymbol{E}V_n| \ge \varepsilon cn\} \\
&\le 2^{-2\varepsilon^2 c^2 n/\mu^2}.
\end{aligned}$$

Thus by the Borel-Cantelli Lemma, $V_n/\boldsymbol{E}V_n \to 1$ when $n \to \infty$ almost surely. □

8. Robustness, continuity, and embedding. A natural question is to ask what happens for two incremental trees that are very much alike. Clearly, if the incremental model is to be widely accepted, it should have a certain robustness property with respect to small changes in the edge random variable X. In this section, a simple method is introduced for analyzing this sort of situation. For starters, we give the key technical result.

LEMMA 4. *Consider a complete b-ary tree with n levels of edges. With each edge e of this tree we associate a uniform $[0,1]$ random variable U_e. Denote by $\mathcal{P}_n$ the collection of all b^n paths from the root to the leaves. Let $\varphi = \varphi(\varepsilon)$ be defined as follows:*

$$\varphi = \inf\left\{x : 1 > x \geq \varepsilon, b\left(\frac{1-\varepsilon}{1-x}\right)^{1-x}\left(\frac{\varepsilon}{x}\right)^{x} \leq 1\right\}.$$

Then, for all $\varepsilon > 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \boldsymbol{E}\left\{\max_{P\in\mathcal{P}_n} \sum_{e\in P} I_{[U_e\leq\varepsilon]}\right\} \leq \varphi(\varepsilon).$$

Furthermore, $\varphi(\varepsilon) \to 0$ as $\varepsilon \to 0$. For $\varepsilon \geq 1/b$, $\varphi(\varepsilon) = 1$, while for $\varepsilon < 1/b$, $\varphi(\varepsilon) < 1$.

REMARK: BRANCHING RANDOM WALKS. We need the explicit bound of the previous Lemma in what follows. However, note that by the theory of maxima in branching random walks, it easily follows that

$$\lim_{n\to\infty} \frac{1}{n} \boldsymbol{E}\left\{\max_{P\in\mathcal{P}_n} \sum_{e\in P} I_{[U_e\leq\varepsilon]}\right\} = \varphi(\varepsilon).$$

See for example the work by Hammersley [Ha74], Kingman [Ki75] or Biggins [Bi76,Bi77].

Proof. For $\varepsilon \geq 1/b$ the statement is trivial. We assume $\varepsilon < 1/b$. For every $P \in \mathcal{P}_n$, $B = \sum_{e\in P} I_{[U_e\leq\varepsilon]}$ is binomial (n,ε) distributed. Thus, by Bonferroni's inequality, for $1 > x \geq \varepsilon$,

$$\begin{aligned} \boldsymbol{P}\left\{\max_{P\in\mathcal{P}_n} \sum_{e\in P} I_{[U_e\leq\varepsilon]} \geq xn\right\} &\leq \sum_{P\in\mathcal{P}_n} \boldsymbol{P}\{B \geq xn\} \\ &\leq b^n\left(\left(\frac{1-\varepsilon}{1-x}\right)^{1-x}\left(\frac{\varepsilon}{x}\right)^{x}\right)^n, \end{aligned}$$

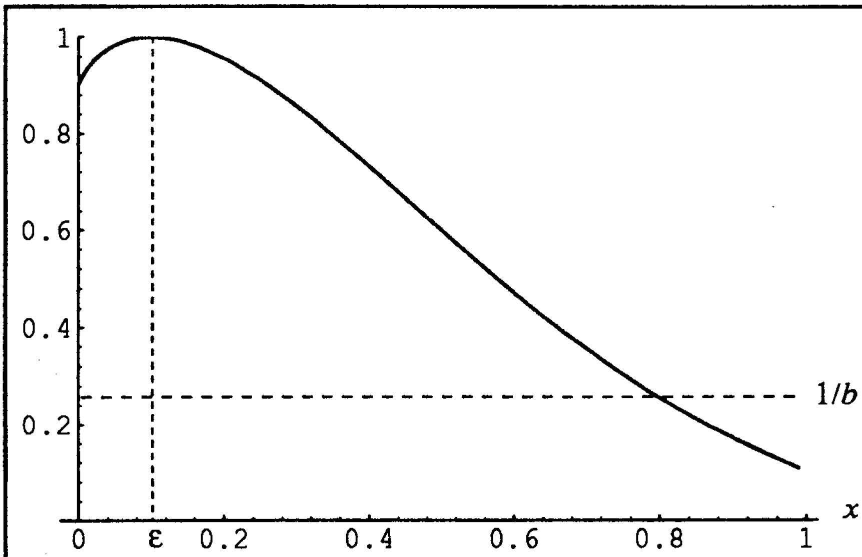
where we use Chernoff's bound for the tail of a binomial distribution (see for example Hoeffding, [Hoeff63], Theorem 1). We denote

$$H(\varepsilon,x) = \left(\frac{1-\varepsilon}{1-x}\right)^{1-x}\left(\frac{\varepsilon}{x}\right)^{x}.$$

Thus $\varphi(\varepsilon)$ is the smallest solution greater than ε and smaller than 1 of

$$H(\varepsilon, x) = 1/b.$$

It is a simple analytical exercise to show that $H(\varepsilon, x)$ is monotonically decreasing from 1 at $x = \varepsilon$ to ε at $x = 1$ (see figure below).



$H(\varepsilon,x)$, for $\varepsilon=0.1$

We see that $\varphi(\varepsilon)$ is well-defined and that for $\varepsilon < 1/b$, $\varepsilon < \varphi(\varepsilon) < 1$. Furthermore,

$$\varphi(\varepsilon) \to 0 \text{ as } \varepsilon \to 0$$

because

$$H(\varepsilon, f(\varepsilon)) \sim \varepsilon^{f(\varepsilon)} \to 0 \text{ as } \varepsilon \to 0$$

for any increasing function $f(\varepsilon)$ with $f(\varepsilon)\log(1/\varepsilon) \to \infty$, and $f(\varepsilon) \to 0$, as $\varepsilon \to 0$ ($f(\varepsilon) = 1/\sqrt{\log(1/\varepsilon)}$ will do). For ε small enough, $\varphi(\varepsilon) \le f(\varepsilon) \to 0$. We have for all $\delta > 0$,

$$\begin{aligned} \boldsymbol{E}\left\{\max_{P\in\mathcal{P}_n} \frac{\sum_{e\in P} I_{[U_e\le\varepsilon]}}{n}\right\} &\le \boldsymbol{P}\left\{\max_{P\in\mathcal{P}_n} \sum_{e\in P} I_{[U_e\le\varepsilon]} \ge (\varphi(\varepsilon)+\delta)n\right\} + \varphi(\varepsilon) + \delta \\ &\le o(1) + \varphi(\varepsilon) + \delta. \end{aligned}$$

By the arbitrary nature of δ, Lemma 4 follows. □

We can use this Lemma in a variety of ways.

EXAMPLE 1. Assume that we have two incremental trees with random edge variables X and Y respectively, where $\boldsymbol{P}\{X \ne Y\} = p$, and $0 \le X, Y \le \mu$.

The case of interest is when p is small. Then we may consider the complete b-ary tree with n levels of edges, in which we we give every edge e a value Z_e according to the rule

$$Z_e = \begin{cases} 0 & \text{if } X = Y \\ \mu & \text{if } X \neq Y \end{cases} .$$

Call $V_n(X)$ and $V_n(Y)$ the root values in both incremental trees. By Lemma 4,

$$|V_n(X) - V_n(Y)| \leq \max_{P \in \mathcal{P}_n} \sum_{e \in P} Z_e \leq (\varphi(p) + o(1))n\mu$$

almost surely. By a trivial argument,

$$|\mathcal{V}(X) - \mathcal{V}(Y)| \leq \varphi(p)\mu .$$

EXAMPLE 2. Assume that X and Y are Bernoulli edge variables with parameters p and q respectively. We may couple these on a common probability space. For example, we could consider a uniform $[0,1]$ random variable U associated with each edge in a complete b-ary tree. Set $X = I_{U \leq p}$ and $Y = I_{U \leq q}$. Clearly, $X \neq Y$ with probability $|p-q|$. Therefore, by Example 1,

$$|\mathcal{V}(X) - \mathcal{V}(Y)| \leq \varphi(|p-q|) .$$

The properties of φ insure that $\mathcal{V}(X)$ is a uniformly continuous function of p for Bernoulli trees.

EXAMPLE 3. For general random variables X and Y, having distribution functions F and G respectively, we may construct a common probability space based upon a uniform $[0,1]$ random variable U once again. Note that X is distributed as $F^{\mathrm{inv}}(U)$ and Y as $G^{\mathrm{inv}}(U)$. However, there are other kinds of couplings as well. Associate with each edge e in a complete b-ary tree a random variable Z_e, where

$$Z_e = \begin{cases} \mu & \text{if } |X - Y| > \delta ; \\ \delta & \text{if } |X - Y| \leq \delta , \end{cases}$$

where $\delta > 0$, and X and Y are coupled as above. A little thought shows that

$$\begin{aligned} |\mathcal{V}(X) - \mathcal{V}(Y)| &\leq \max_{P \in \mathcal{P}_n} \sum_{e \in P} Z_e + n\delta \\ &\leq n\mu(\varphi(\boldsymbol{P}\{|X-Y| > \delta\}) + o(1)) + n\delta \end{aligned}$$

almost surely, where we use Lemma 4. Therefore, collecting things,

$$|\mathcal{V}(X) - \mathcal{V}(Y)| \leq \inf_{\delta \geq 0} \left[\mu\varphi \left(\inf_{\text{all couplings of } X \text{ and } Y} \boldsymbol{P}\{|X-Y| > \delta\} \right) + \delta \right] .$$

This indeed shows the robustness of $\mathcal{V}(X)$ with respect to small changes in the distribution of X.

EXAMPLE 4. Given an arbitrary unbounded positive edge variable X with finite moment generating function ($\boldsymbol{E}\exp(tX) < \infty$ for some $t > 0$). Define a coupled random variable $Y = \min(X, \mu)$, where μ is large but fixed. The root values of the trees based upon Y and X may be linked in a simple manner. Using arguments not unlike those above, one can establish that $\boldsymbol{E}V_n/n \to \mathcal{V}(X) < \infty$. Thus, the boundedness condition in our main convergence theorem is not required after all.

9. Asymptotic behavior of the Bernoulli tree with parameter p. We first give a result stronger then Theorem 1 on the asymptotic behavior of V_n for Bernoulli trees when p is small. For more details, we refer to [Kam92].

LEMMA 5. *For $p \in [0, \alpha]$, there exist bona fide distribution functions F_∞ and H_∞ with finite expected values that put positive mass on all the nonnegative integers, such that*

$$\lim_{n\to\infty} F_{2n}(i) = F_\infty(i)$$

and

$$\lim_{n\to\infty} F_{2n+1}(i) = H_\infty(i) \ .$$

Furthermore, for $p > \alpha$, we have for all fixed $i \geq 0$,

$$\lim_{n\to\infty} F_{2n}(i) = \lim_{n\to\infty} F_{2n+1}(i) = 0 \ ,$$

and $V_n \to \infty$ almost surely when $n \to \infty$. Finally,

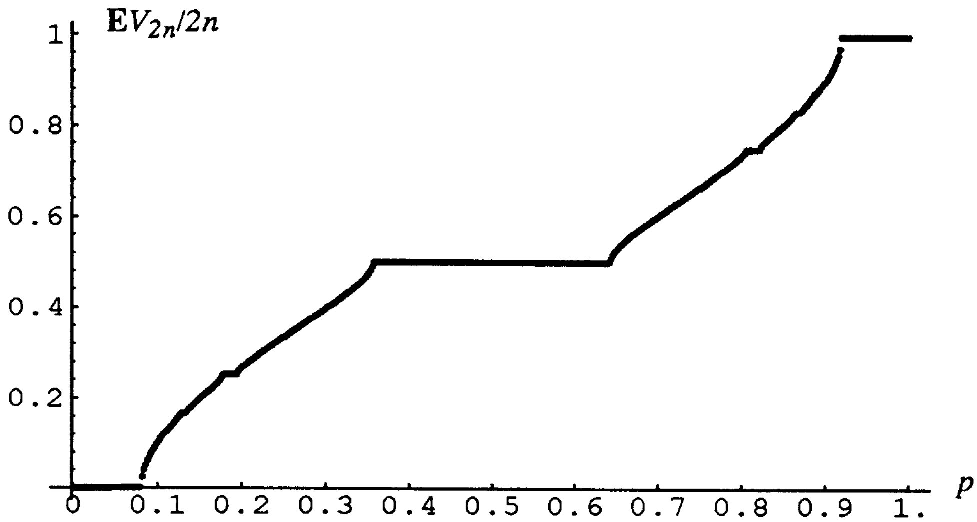
$$\alpha \leq 1 - b^{-1/(b+1)} \overset{b\to\infty}{\longrightarrow} 0.$$

We recall that if X is Bernoulli (p) and Y is Bernoulli (q), then

$$|\mathcal{V}(X) - \mathcal{V}(Y)| \leq \varphi(|p - q|) \ .$$

Then using Theorem 3, the following lemma is trivial.

LEMMA 6. *In a Bernoulli tree with parameter p, $\boldsymbol{E}V_n/n$ converges to a finite limit $\mathcal{V}(p)$, where $\mathcal{V}$ is a uniformly continuous function of p. Furthermore if $\alpha < p < 1 - \alpha$ then $0 < \mathcal{V}(p) < 1$ and $V_n/\boldsymbol{E}V_n \to 1$ almost surely when $n \to \infty$.*



$EV_{2n}/2n$ for $b=10$, $n=200$.

The above figure shows $\mathcal{V}(p)$ for $b = 10$, the flat parts close to 0 and 1 are explained by Theorem 1. Computations show that $\mathcal{V}(X)$ is flat around $p = 1/2$ as well whenever $b > 2$. In this respect, we offer the following theorem.

THEOREM 4. *For all b, $\mathcal{V}(1/2) = 1/2$ and for all p,*

$$\mathcal{V}(1-p) = 1 - \mathcal{V}(p).$$

Furthermore there exists $\beta \in (0, 1/2]$ such that

$$\mathcal{V}(p) \begin{cases} = \frac{1}{2} & \text{if } p \in [\beta, 1-\beta] \\ < \frac{1}{2} & \text{if } p \in [0, \beta) \\ > \frac{1}{2} & \text{if } p \in (1-\beta, 1] \end{cases} .$$

When $\beta < 1/2$, the range $[\beta, 1-\beta]$ is called the flat part around $p = 1/2$. For $p \in (0, 1/2)$, let $L(p, b)$ be the largest root of $1-(1-px^b)^b = x$ on $[0, 1]$. If $L(p, b) \neq 0$, then

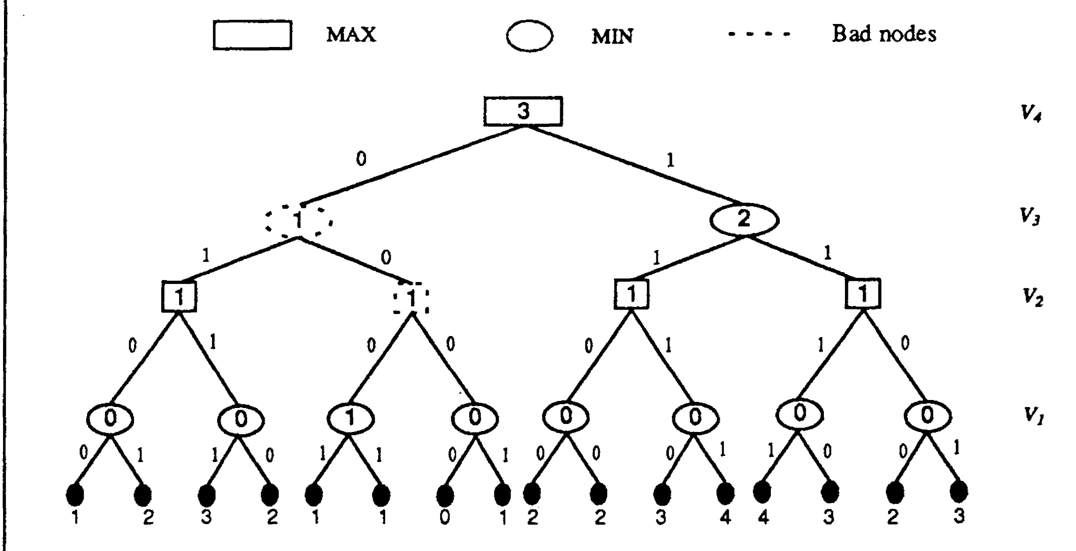
$$\left\lfloor \frac{n}{2} \right\rfloor - \sqrt{n \log n} \leq EV_n \leq \left\lfloor \frac{n}{2} \right\rfloor + \sqrt{n \log n}\ , \ (n \geq \sqrt{2/L^b(p,b)})\ ,$$

$\beta \geq p$, and $V_n/n \to 1/2$ almost surely. Furthermore, β tends to zero as b tends to infinity. Thus the flat part exists and tends to the full range as $b \to \infty$. Finally, for $b \geq 8$, we have $0 < \beta < 1/2$.

Proof. We first show that

$$\left\lfloor \frac{n}{2} \right\rfloor - \sqrt{n \log n} \le \boldsymbol{E} V_n .$$

The symmetric inequality is easily obtained by considering a tree with MAX nodes at the bottom (see [Kam92]). We consider a random $2n$-level b-ary incremental tree with parameter p. The nodes in the tree are marked good or bad. The leaves are all good. Consider a node at an odd level $2n+1$ with b children at level $2n$. Such a node corresponds to a MIN node in the tree. We mark it good only if all the children are good; otherwise, it is marked bad. For a node at level $2n$ (a MAX node) with b children, we mark it good if there exists at least one good child whose edge value is one. Thus, the root u is good if and only if there is a path from the root to the bottom level where all the MAX nodes provide at least one "1".



Tree with good and bad nodes.

If the value $V(u)$ of the root of the tree is good, then $V(u) \ge n$. Also, for a node v at level $2n+1$ we have $V(v) \ge n$. Let p_n denote the probability that a node at level n is marked as good. Then, by the previous discussion,

$$\boldsymbol{P}\{V_{2n} \ge n\} \ge p_n .$$

Furthermore, we have a simple recursion:

$$p_0 = 1 ,$$

and

$$p_{2n} = f(p_{2n-2}) ,$$

where

$$f(x) \stackrel{\text{def}}{=} 1 - \left(1 - px^b\right)^b .$$

The function f is continuous and increases monotonically from 0 to $f(1) = 1 - (1-p)^b$. We note therefore that p_{2i} decreases monotonically in i to a limit which is either zero or a positive number. The limit is the largest root on $[0,1)$ of the equation $f(x) = x$. Let us call this limit $L(p,b)$. Thus, the following interesting inequalities are true:

$$\begin{aligned} &\inf_n \mathbf{P}\{V_{2n} \geq n\} \geq L(p,b), \\ &\inf_n \mathbf{P}\{V_{2n+1} \geq n\} \geq L^b(p,b). \end{aligned}$$

Therefore,

$$\inf_n \boldsymbol{P}\left\{V_n \geq \left\lfloor \frac{n}{2} \right\rfloor\right\} \geq L^b(p,b).$$

Continuing this discussion, we consider the set of all p for which $L(p,b) > 0$. We know by McDiarmid's inequality that

$$\boldsymbol{P}\left\{|V_n - \boldsymbol{E}V_n| > \sqrt{n \log n}\right\} \leq \frac{2}{n^2}.$$

Therefore, if $2/n^2 < L(p,b)$, we see that

$$\boldsymbol{E}V_n \geq \left\lfloor \frac{n}{2} \right\rfloor - \sqrt{n \log n} .$$

For the remainder of the proof, we refer to [Kam92]. □

REMARK. Numerical computations show that the distribution of the root's value is even more concentrated than this Lemma shows.

10. Concluding remarks. Although we have discovered many properties of the limit $\mathcal{V}(X)$, we have been unable to provide a precise analytic formulation of this limit function. Many inequalities are available however. If $\mathcal{V}(p)$ denotes the limit for a Bernoulli tree with parameter p, we see that for any positive random variable X, since $X \geq aI_{X \geq a}$,

$$\mathcal{V}(X) \geq \sup_a a\mathcal{V}(\boldsymbol{P}\{X \geq a\}) .$$

(We hope the reader will permit this abusive notation.) In particular, if a is the median of a continuous random variable X, this inequality shows that

$$\mathcal{V}(X) \geq \frac{\text{median } (X)}{2} .$$

If $X \leq \mu$, we see that $\mathcal{V}(\mu - X) = \mu - \mathcal{V}(X) \geq a\mathcal{V}(\boldsymbol{P}\{\mu - X \geq a\})$. Thus,

$$\mathcal{V}(X) \leq \mu - \sup_a a\mathcal{V}(\boldsymbol{P}\{X \leq \mu - a\}) .$$

For example, when X is uniform on $[0,1]$, and we take $a = 1/2$, we obtain the inequalities

$$\frac{1}{4} \leq \mathcal{V}(X) \leq \frac{3}{4} .$$

This was of course predictable, as we know that $\mathcal{V}(X) = 1/2$ by a relatively simple symmetry argument.

Another inequality follows from $\mathcal{V}(1-\alpha) = 1$: we obtain for any positive X,

$$\mathcal{V}(X) \geq X_{(\alpha)} ,$$

where $X_{(\alpha)}$ is the α quantile of X.

Note that $\mathcal{V}(a + bX) = a + b\mathcal{V}(X)$. Assume that X is bounded and positive, and that Y is independent of X, bounded, and positive. Then, one would be tempted to infer that

$$\mathcal{V}(X+Y) = \mathcal{V}(X) + \mathcal{V}(Y) ,$$

where $\mathcal{V}(X+Y)$ refers to the incremental tree with edge values distributed as $X+Y$. This is clearly false: just take X and Y Bernoulli with parameter α. Then $X+Y$ is stochastically greater than a Bernoulli with parameter $1-(1-\alpha)^2 = 2\alpha - \alpha^2$. From our results,

$$\mathcal{V}(2\alpha - \alpha^2) > 0 ,$$

while

$$\mathcal{V}(X) = \mathcal{V}(Y) = 0 .$$

The lack of linearity seen throughout our analysis and experiments makes it difficult to get a good grip on the limit function except in special cases.

In the Bernoulli model, if the ones were represented in the minimax path in proportion to their frequency in the tree, we would have $\mathcal{V}(p) = p$. Clearly, we do not have this—the tree's behavior depends upon more than just frequencies or averages. Also, in general, we do not have $\mathcal{V}(X) = \boldsymbol{E}X$. The Bernoulli model shows that $\mathcal{V}(X)/\boldsymbol{E}X$ can be zero!

A comparison with the Pearl tree with the same distribution on the leaf values is helpful. In the Bernoulli model, we associate with the leaves i.i.d. binomial (n,p) random variables. Pearl's result implies that the root value V_n of his tree satisfies

$$\frac{V_n}{n} \to p = \boldsymbol{E}X$$

almost surely: the ones are proportionally represented in the minimax path.

For game tree searching strategies, the incremental model is very promising, as it incorporates different behaviors for different values of the edge parameters such as p in the Bernoulli model. The incremental trees are teeming with different sorts of life, perhaps modeling both very easy and very hard search problems for different choices of edge distributions. For p near zero, the zeroes overwhelm the ones, even more than their proportions would suggest.

The natural continuation of this study is the consideration of search heuristics. Here we note that algorithms that expand an edge have access to the edge's value. The purpose is to find strategies that expand few nodes (relatively speaking) in an incremental tree, yet lead to an n-level path of value close to V_n in some sense. For example, if we were to take a random child and if the opponent were infinitely smart (knowing the entire tree), we would end up with a path of total value less than about np (as this would be obtained against a random opponent), and less than about $n\mathcal{V}(p)$ (as we ourselves are not infinitely smart). The true path value thus concentrates around $nS(p)$, where $0 \leq S(p) \leq \min(p, \mathcal{V}(p))$. For more interesting heuristics based upon backtracking and pruning, the model promises to be exciting.

Finally, one may wonder what happens with random edge variables that have large infinite tails. This may occur for example for the normal distribution. More specifically, we may consider edge variables with a symmetric stable distribution with parameter $\alpha \in (0, 2]$ (these have characteristic function $\exp(-|t|^\alpha)$). Consider first a Pearl tree, in which we associate with each leaf independently a sum of n independent symmetric stable random variables, to be able to make a fair comparison with our model. We can prove that $V_n/\sqrt{n}$ oscillates on alternating levels between a positive value c and its negative counterpart $-c$, when the stable distribution is normal ($\alpha = 2$), and when new levels are added at the bottom of the tree. For the Cauchy distribution ($\alpha = 1$), $|V_n/n|$ tends to a positive constant c almost surely, while the sign of V_n alternates on different levels. For $\alpha < 1$, we have distributions with very big tails, and $|V_n/n| \to \infty$ almost surely. For these distributions, the swings in the oscillatory pendulum are unbearably big. In the incremental model, there are distributions for which $\limsup_{n\to\infty} V_n/n = \infty$ almost surely if we add new levels at the root.

REFERENCES

[Bi76] J. D. Biggins, "The first and last-birth problems for a multitype age-dependent branching process," *Advances in Applied Probability*, vol. 8, pp. 446-459, 1976.

[Bi77] J. D. Biggins "Chernoff's theorem in the branching random walk," *Journal of Applied Probability*, vol. 14, pp. 630-636, 1977.

[CM83] M.S. Campbell and T.A. Marsland, "A comparison of minimax tree search algorithms," in: *Artificial Intelligence*, 20: 347-367, 1983.

[CL91] E.G. Coffman, Jr. and G.S. Lueker, *Probabilistic Analysis of Packing and Partitioning Algorithms*, Wiley-Interscience Series, New York, 1991.

[Ha74] J. M. Hammersley, "Postulates for subadditive processes," *Annals of Probability*, vol. 1, pp. 652-680, 1974.

[Hoeff63] W. Hoeffding: "Probability inequalities for some series of bounded random variables," in: *Journal of the American Statistical Association*, 58:13-29,1963.

[Kam92] O. Kamoun, "A Probabilistc Min-Max Tree," M.Sc. Thesis, School of Computer Science, McGill University, Montreal, 1992.

[Ki75] J. F. C. Kingman, "Subadditive ergodic theory," *Annals of Probability*, vol. 1, pp. 883-909, 1973.

[KM75] D.E. Knuth and R.W. Moore, "An analysis of alpha-beta pruning," in: *Artificial Intelligence*, 6: 293-326, 1975.

[McDi89] C. McDiarmid, "On the method of bounded differences," in: *Surveys in Combinatorics 1989*, vol. 141, pp. 148–188, London Mathematical Society Lecture Notes Series, Cambridge University Press, Cambridge, 1989.

[Nau82a] D.S. Nau, "The last player theorem," in: *Artificial Intelligence*, 18: 53-65, 1982.

[Nau82b] D.S. Nau, "An investigation of the causes of pathology in games," in: *Artificial Intelligence*, 19: 257-278, 1982.

[Nau83] D.S. Nau, "Pathology on game trees revisited, and an alternative to minimaxing," in: *Artificial Intelligence*, 21: 221-244, 1983.

[Newb77] M. Newborn, "The efficiency of the alpha-beta search of trees with branch-dependent terminal node score," in: *Artificial Intelligence*, 8: 137-153, 1977.

[Pearl80] J. Pearl, "Asymptotic properties of Minimax trees and game-searching procedures," in: *Artificial Intelligence*, 14: 113-138, 1980.

[Pearl84] J. Pearl, *Heuristics*, Addison-Wesley Publishing Company, Reading, MA, 1984.

[Schw67] L. Schwartz, *Cours d'Analyse*, Hermann, Paris, 1967.