A PROBABILISTIC MIN-MAX TREE

by
Olivier KAMOUN

School of Computer Science
McGill University
Montréal, Canada

November 1992

A Thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements of the degree of
Master of Science

ABSTRACT. MIN-MAX trees have been studied for thirty years as models of game trees in artificial intelligence. Judea Pearl introduced a popular probabilistic model that assigns random independent and identically distributed values to the leaves. Among the dependent models, incremental models assume that terminal values are computed as sums of edge values on the path from the root to a leaf. We study a special case called the SUM model where the edge values follow a Bernoulli distribution with mean p. Let V_n be the root's value of a complete b-ary, n-level SUM tree. We prove that $\mathbf{E}V_n/n$ tends to a uniformly continuous function $\mathcal{V}(p)$. Surprisingly, $\mathcal{V}(p)$ is very nonlinear and has some flat parts. More formally, for all b, there exist $\alpha, \beta \in (0,1)$ such that,

$$\begin{cases} \text{if } p \in [0, \alpha] & : \mathbf{E}V_n \text{ has a finite limit} \\ \text{if } p \in [1-\alpha, 1] & : n - \mathbf{E}V_n \text{ has a finite limit} \\ \text{if } p \in [\beta, 1-\beta] & : \mathbf{E}V_n/n \text{ tends to } 1/2 \end{cases}.$$

Finally β and α tend to zero when b tends to infinity.

RÉSUMÉ. Depuis trente ans les arbres MIN-MAX sont souvent étudiés pour modéliser les arbres de jeux en intelligence artificielle. Judea Pearl a introduit le modèle probabiliste généralement utilisé qui attribue des valeurs aléatoires, indépendantes et identiquement distribuées aux feuilles. Parmi les modèles dépendants introduits, les modèles incrémentiels supposent que les valeurs terminales sont égales à la somme des valeurs des branches du chemin reliant la racine à la feuille. Nous étudions un cas particulier de ces modèles, appelé SUM-modèle où les branches suivent une distribution de Bernoulli de moyenne p. Soit V_n la racine d'un SUM-arbre b-aire à n niveaux. Nous montrons que $\mathbf{E}V_n/n$ converge vers une fonction $\mathcal{V}(p)$ uniformément continue en p quand n tend vers l'infini. $\mathcal{V}(p)$ possède la carastéristique surprenante de présenter des paliers. Ceci s'explique par l'existence pour tout b de α, $\beta \in (0,1)$ tel que

$$\begin{cases} \text{si } p \in [0, \alpha] & : \mathbf{E}V_n \text{ a une limite finie} \\ \text{si } p \in [1-\alpha, 1] & : n - \mathbf{E}V_n \text{ a une limite finie} \\ \text{si } p \in [\beta, 1-\beta] & : \mathbf{E}V_n/n \text{ tend vers } 1/2 \end{cases}$$

De plus, α et β tendent vers zero quand b tend vers l'infini.

Acknowledgments

I would like to thank the technical staff and the departmental secretaries for their assistance and warm welcome which especially helped me when I arrived from France. I especially thank my advisor, Luc Devroye, for giving me my first taste of scientific research, for his financial support and for the good working environment in which we have worked during the year. And finally, thanks to Montreal which is a great city in which to study, to have fun and to be frozen.

TABLE OF CONTENTS.

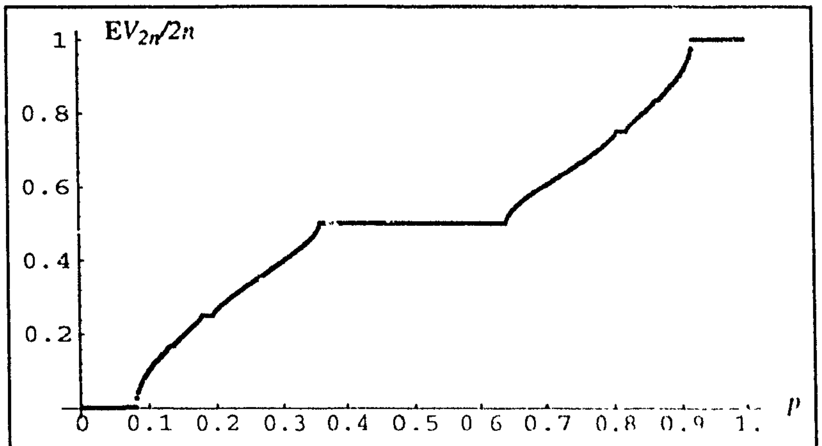
INTRODUCTION

In this thesis, we study the asymptotic properties of random MIN-MAX trees with a given number of levels. It is well known that the efficiency of a game tree search algorithm depends very much on the underlying probabilistic model. For this reason, we are studying a model that is more realistic than the early models considered in the literature. The model most frequently studied is the one in which the leaf nodes are independent identically distributed random variables. Judea Pearl [Pearl80][Pearl84] thoroughly studied the asymptotic behavior of these trees Chapter II of this thesis presents some of these results. The independence assumption has been challenged as it does not seem to model real games very well. As a result, other models have been proposed. Some of these are presented in the first chapter They have been studied to compute the efficiency of the alpha-beta pruning, for example. The model we study belongs to a class called incremental models. In these models a leaf value is computed as the sum of the edge values of the path from the root to the leaf. We study a particular case introduced by Nau [Nau82a] where the edge values are independent Bernoulli random variables taking the value one with probability p and zero with probability $1 - p$. We call this the SUM model

The asymptotic behavior of these models is largely unknown. To compound matters, numerical simulations are only possible for small trees with less than twenty levels. In this thesis we give a method to compute the distribution function of the value of the root of such a tree in polynomial time as a function of the depth. With this method we simulate trees with up to one thousand levels. We also prove the following results on the asymptotic behavior of V_n, the value of the root of a b-ary SUM tree with n levels:

- For all p, EV_n/n converges as n tends to infinity, and the limit is uniformly continuous in p.
- There is a range for p close to zero where V_n tends to a bona fide distribution with a finite expected value.
- When b is fixed and greater than 2, there is a range for p around 1/2 where the expected value EV_{2n} is asymptotic to n and this range tends to fill the interval (0, 1) when b tends to infinity

These results can be summarized by the graph below. It shows $EV_{2n}/2n$ as a function of p for a tree with ten children per node and with $n = 200$.



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

In chapter III, we study the asymptotic behavior of the root for the SUM model for small p. The proofs combine probability theoretical arguments with analysis of some recurrences. In chapter IV we prove the convergence of EV_n/n for general p as well as the existence of the flat part around 1/2.

The nonlinear behavior of the limit of EV_n/n as a function of p may come as a surprise to some. The work presented here could be used elsewhere in the study of various search algorithm for the SUM model. It could also form the basis of future research on SUM models for general edge value distributions, and for generalized SUM models in which the number of children per node is random, as in a Galton-Watson branching process.

I

MIN-MAX TREES AND GAME-PLAYING PROGRAMS

Since the beginning of computers we have tried to compare human intelligence to the computer's capacity through game-playing programs. Games like chess and checkers have always been used as prototypes in this comparison. At the center of these game-playing programs are the MIN-MAX trees. In this chapter, we give some basic definitions and discuss various probabilistic models. For more details, the reader is referred to Pearl [Pearl84].

1. Game-playing programs.

The games we consider are two-player perfect information games such as chess, checkers or GO match There are two adversaries that we call MAX and MIN We are on MAX's side. Play alternates between the players. For each position there is a finite number of possible moves defined by the rules. The game must end after a finite number of turns. To model these games we use trees called game trees or MIN-MAX trees. Each node represents a position and each edge a move. Thus a node has a child for each possible move. The root node represents the current position. A terminal node is a position in which the player cannot move any further.

First we consider WIN-LOSS trees Each terminal node is a WIN, LOSS or DRAW position according to the game rules Such nodes are called WIN, LOSS or DRAW nodes. The problem is to know if from the current position one can force a win. If you can force a win, we call the corresponding node a WIN node When it is your turn you can force a win if one of the possible moves lead you to a win node And if it is your opponent's turn you can force a win if all the moves lead to a WIN node. We call S_u the function which assigns to the node u its WIN, LOSS or DRAW status. If u is a terminal node, S_u is the status of the corresponding terminal position according to the rules. If u is a nonterminal MAX node we have,

$$S_u = \begin{cases} \text{WIN} & \text{if one of } u\text{'s successors is a WIN node} \\ \text{LOSS} & \text{if all of } u\text{'s successors are LOSS nodes} \\ \text{DRAW} & \text{if at least one of } u\text{'s successors is a DRAW and none is a WIN.} \end{cases}$$

If u is a nonterminal MIN node we have,

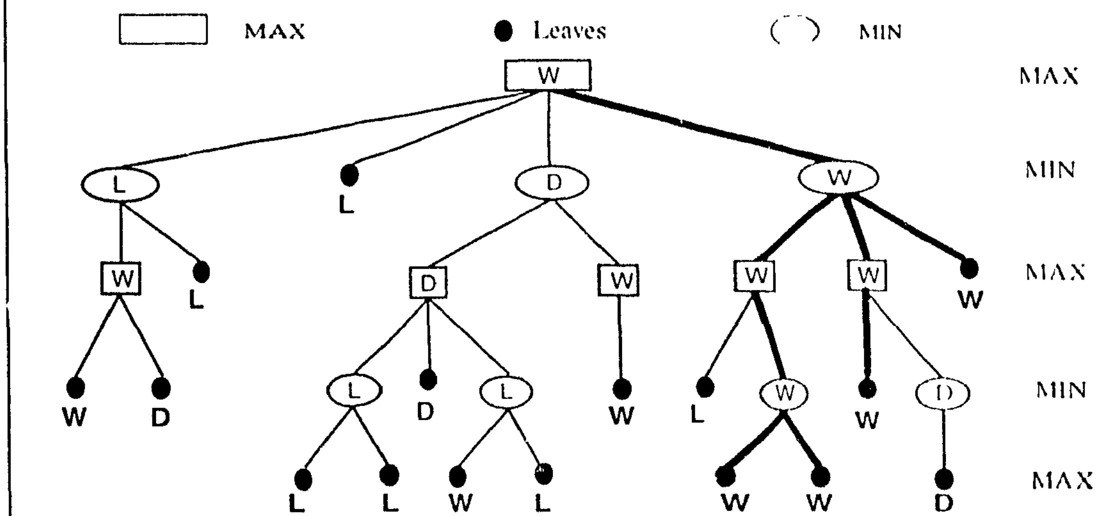
$$S_u = \begin{cases} \text{WIN} & \text{if all of } u\text{'s successors are WIN nodes} \\ \text{LOSS} & \text{if one of } u\text{'s successors is a LOSS node} \\ \text{DRAW} & \text{if at least one of } u\text{'s successors is a DRAW and none is a LOSS.} \end{cases}$$

If we give the value 1 to a WIN node, 0 to a DRAW node and -1 to a LOSS node then we have when u is a nonterminal MAX node with A_u the set of its children,

$$S_u = \max_{v \in A_u} S_v.$$

And for a nonterminal MIN node,

$$S_u = \min_{v \in A_u} S_v.$$



A WIN-LOSS-DRAW game tree evaluated with MIN MAX rules
The solution tree T^+ in boldface

The above figure shows such a tree evaluated with the MIN-MAX rules. The label below the leaves represents the leaf values, and the label inside the nodes represents the MIN-MAX status. The root label of this example is WIN: this means that for each possible move of MIN, MAX has a strategy to force the win. This strategy is represented by a sub tree T^+ called a solution tree. It starts at the root of T. At the MIN levels it has all the children of T and at the MAX levels it has only one child which is the best move for MAX.

We can also consider trees with numerical terminal values. These values give a certain weight to each terminal position, and the goal of MAX is to get the biggest value. To compute the root's value we use the MIN-MAX rules. Let u be a node of T. If u is terminal then V_u is its value given as part of the data. If u is a nonterminal MAX node with A_u the set of its children,

$$V_u = \max_{v \in A_u} V_v.$$

For a nonterminal MIN node,

$$V_u = \min_{v \in A_u} V_v.$$

The root's value is the guaranteed value MAX can get regardless how well MIN plays. The SUM model we study belongs to this class of trees.

2. Game strategies.

If we compute the root value and the winning MAX strategy of the entire game tree and if this value is WIN we are sure to win against any player. For a game such as chess or checkers it is virtually impossible to compute the root label of the game tree. For chess the size of the game tree is estimated at about 10^{120} nodes (10^{40} for checkers). If 3 billion nodes can be generated each second it will take 10^{101} (10^{21} for checkers) centuries to generate it. Even if you find the solution tree by chance, you will need to store about 10^{20} moves for a single certified checkers strategy. Thus we cannot expect to find a solution tree which guarantees the optimal strategy.

Another strategy uses evaluation functions. An evaluation function is a function that assigns a value to any position according to static characteristics of the position. The greater the function, the better the position. One just computes the values of the children of the current position and chooses the greatest one. This strategy gives rather poor results as it is hard to get an accurate evaluation function based on static characteristics.

The strategy generally used is the bounded look-ahead strategy, which is a mixture of the preceding concepts. The idea is to generate the game tree up to a fixed number of levels, to compute the value of the leaf nodes with a static evaluation function, and to compute the value of the root and the solution tree with the min-max rules. One picks the move which guarantees the best value. The leaf values can ideally be considered as the probability for MAX to win. And the root's value is the probability for MAX to win if MIN plays optimally.

A practical problem we won't consider here, is related to the efficient computation of the root's value. Alpha-beta pruning for example (see, e.g.,[KM75], [Pearl84], [Newb77]) allows one to find the root's value without generating the entire tree. For other computationally efficient methods, we refer to [Pearl84], [CM83].

3. Probabilistic models.

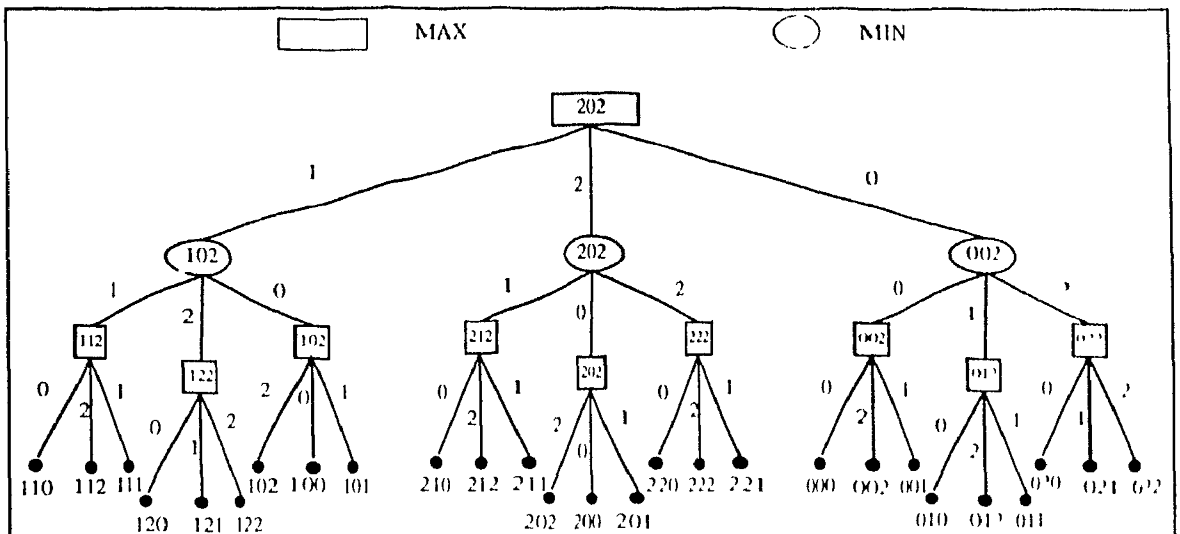
In order to study the probabilistic behavior of an algorithm we need a random model for the data. In general, we may introduce a random variable B, possibly depending upon the depth, that gives the number of children of a node. One may also introduce random leaf values. We restrict ourselves to complete b-ary trees with n levels.

The standard probabilistic model assumes that the leaf values are independent identically distributed (see chapter II). Many argued that the independence hypothesis was not realistic. In chapters III and IV, a model with dependent leaf values in introduced.

4. Incremental models.

Incremental models, such as the model of Nau[Nau82], introduce dependent leaf values. With the incremental model we assign a random value to each edge, and define the value of a leaf as the sum of the edge values of the path from the root to the node. In the few examples from the literature given below, the trees are b-ary.

Model 1: This model, proposed by Knuth and Moore[KM75], assumes that the edge values at level k form a uniform random permutation of $\{0, 1/b^k, 2/b^k, \ldots, (b-1)/b^k\}$. This implies that all leaf nodes have distinct values. This model weighs branches near to the root more than branches deep in the tree.



Model 1 totaly dependent (proposed by Knuth and Moore)

Model 2: Proposed by Fuller[Newb77], this model assumes that the edge values at level k are randomly assigned a distinct value from the set $\{0, 1, 2, \ldots, b-1\}$. In models 1 and 2 the root's value is not random. In fact, as all the set values are taken, we know that if we are in a MAX node we pick the edge with the biggest value, i.e., $b-1$ for model 2 and $(b-1)/b^k$ for model 1 and in a MIN node, we choose the edge with the smallest value, i.e., 0. If we consider a tree with depth $2n$ and if the top level is a MAX node, the root's value is always $n(b-1)$ for model 2 and $\sum_{k \in \{0, n/2-1\}} (b-1)/b^{k}$ for model 1. Nevertheless, these models remain interesting in the study of alpha-beta pruning because the pruning depends only on the random order of the nodes.

Model 3: In this model proposed by Nau [Nau82], the edge values are assigned from $\{-1, 1\}$ without the hypothesis of distinction. In this model the root's value is random. Up to a translation and rescaling it is equivalent to a special case of the model studied in this thesis.

All the above models have been used to study the pruning factor of the alpha-beta method. The results of these studies are reviewed by Newborn [Newb77]. It has been shown that the independent model brings about a pathological situation (a pathology occurs if by increasing the depth of the search in the game tree we decrease the quality of the decision). Pathologies do not appear with incremental models. They also seem to be absent in real games like chess or checkers. In such games strong positions have often strong children. This creates some strong "sections" in the tree where there are a lot of strong nodes. That is why the dependent model seems to fit better this kind of game.

The model we will study is related to model 3. We consider an n-level b-ary tree, and we assign to each edge e an independent drawing of a basic random variable X, denoted by X_e. Then if v is a leaf, and $\mathcal{P}_v$ the path from the root to v, we assign to v the value

$$\sum_{e \in \mathcal{P}_v} X_e.$$

This defines the SUM model with random edge variable X. The root value is computed with the MIN-MAX rules. The SUM model in which X is Bernoulli with parameter p is called the SUM model with parameter p. It will be studied in Chapters III, IV and V.

II ASYMPTOTIC BEHAVIOR OF THE STANDARD PROBABILISTIC MODEL

In this chapter we consider complete b-ary trees with independent and identically distributed leaf values. We present a WIN-LOSS model and a model with arbitrary distribution of terminal values, and we give a short summary of the results given in [Pearl84].

1. The WIN-LOSS model.

Let G_n be an n-level b-ary tree and we assign to each leaf node a WIN status with probability P_0 and a LOSS status with probability $1 - P_0$. The bottom level is a MIN level. We denote by P_{2n} the probability that MAX can force a WIN with a game tree G_{2n}, in which MAX has to play first and MIN plays last. We denote by P_{2n-1} the probability that MAX can force a WIN with a game tree G_{2n-1} noting that MIN will play the next and the last turn. With the MIN-MAX rules we can assign values to all the nodes of the tree. A MAX node is a WIN position if one of its b children is a WIN node. We recall that such a tree is called a b-ary Pearl tree with parameter P_0. We have

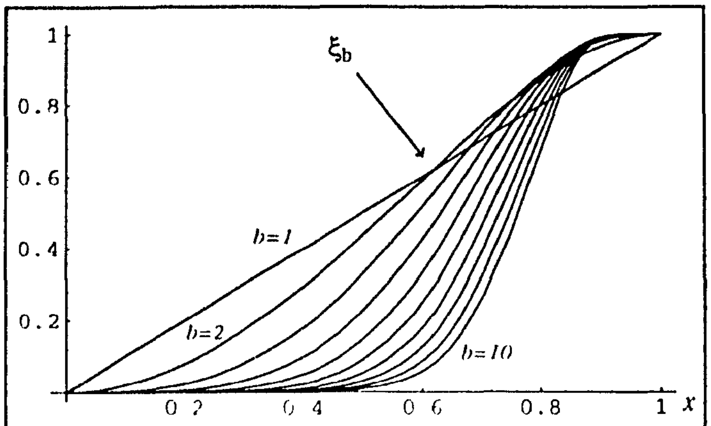
$$P_{2n} = 1 - (1 - P_{2n-1})^b.$$

A MIN node is a WIN position if all b children are WIN nodes. Thus we have,

$$P_{2n-1} = P_{2n-2}^b.$$

Using the two preceding equations we obtain the following recurrence for P_{2n}:

$$P_{2n} = 1 - (1 - P_{2(n-1)}^b)^b \stackrel{\text{def}}{=} f(P_{2(n-1)}).$$



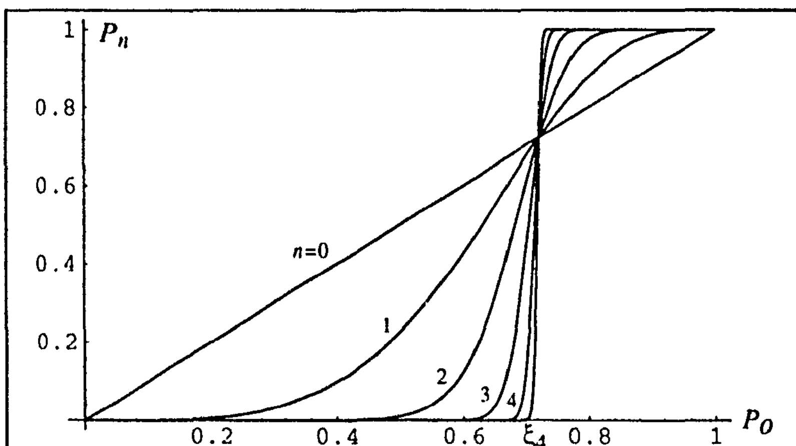
The iteration function $f(x)$, for different b between 2 and 10.

This is a simple functional iteration. f has three fixed points on [0,1], one at zero, one at one and one at $\xi_b \in (0,1)$. We can show that if

$$\lim_{n\to\infty} P_{2n} = \begin{cases} 1 & \text{if } P_0 > \xi_b \\ \xi_b & \text{if } P_0 = \xi_b \\ 0 & \text{if } P_0 < \xi_b \end{cases}.$$

Furthermore,

$$\lim_{n\to\infty} P_{2n-1} = \left(\lim_{n\to\infty} P_{2n}\right)^b = \begin{cases} 1 & \text{if } P_0 > \xi_b \\ \xi_b^b & \text{if } P_0 = \xi_b \\ 0 & \text{if } P_0 < \xi_b \end{cases}.$$



Probability of winning an n cycle game P_{2n} as a function of P_0, with $b=4$.

For $b = 2$ we find $\xi_2 = (\sqrt{5} - 1)/2 \approx .61803$. The convergence of P_n is super-exponential. The next lemma gives an explicit upper bound for the probability of a WIN at the root of a Pearl tree. Pearl [Pearl84] has obtained similar inequalities with sharper a threshold for ξ, but for our purposes, the explicit bound derived here suffices.

Sometimes, it is convenient to work with node values. In that case, a WIN is assigned the value 1 and a LOSS the value 0 The value of the root of an n-level Pearl tree is denoted by V_n.

LEMMA 1. *Let T be a b-ary Pearl tree with parameter $q \le \xi \stackrel{\text{def}}{=} \frac{1}{2} b^{-b}$, and let V_n be the value of a node n levels away from the leaves. Then regardless of whether we begin with* MIN *or* MAX *nodes, and regardless of the parity of n,*

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

PROOF. V_n is maximal if we begin with a MAX level. Let V_0 be the value of a leaf (corresponding to a tree with n levels of edges and with a MAX level at the bottom), so that

$$\mathbf{P}\{V_0 = 1\} = q.$$

Define $P_n = \mathbf{P}\{V_n = 1\}$. Then the following recursion holds: $P_0 = q$, and

$$P_{2n} = \left(1 - (1 - P_{2n-2})^b\right)^b \le (bP_{2n-2})^b.$$

Therefore,

$$\begin{aligned}
\log P_{2n} &\le b \log b + b \log P_{2n-2} \\
&\le (b + b^2) \log b + b^2 \log P_{2n-4} \\
&\vdots \\
&\le (b + b^2 + b^3 + \cdots + b^n) \log b + b^n \log P_0 \\
&\le b^{n+1} \log b + b^n \log q.
\end{aligned}$$

Thus we have,

$$P_{2n} \le b^{b^{n+1}} q^{b^n}.$$

And for P_{2n+1} we get

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

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

$$P_{2n} \le 2^{-b^n}, \text{ and } P_{2n+1} \le b 2^{-b^n},$$

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

$$P_n \le b 2^{-b^{\lfloor n/2 \rfloor}}. \square$$

2. Limit results for trees with arbitrary distribution of the leaf values.

We consider a b-ary tree of depth $2n$, and we assign to each terminal node an independent value drawn from the distribution of X. We denote by V_{2n} the root value of such a tree. It is easy to see that since we are using only maxima and minima, the properties of V_k remain invariant under monotone transformations of the X's. In fact, if every X is replaced by $F(X)$, where F is strictly monotone, then V_k is replaced by $F(V_k)$. If X has a density, then we can let F be the distribution function of X, so that $F(X)$ is uniform on $[0, 1]$. We denote by F_{2n} the distribution function of V_{2n}. Then we have the following recurrence:

$$F_{2n} = \left(1 - (1 - F_{2n-2})^b\right)^b,$$

and

$$F_0 = F.$$

One can show that F_{2n} converges to a step-function F_∞ as n tends to infinity. We have

$$F_\infty(v) = \begin{cases} 0 & \text{if } F(v) < 1 - \xi_b \\ 1 - \xi_b & \text{if } F(v) = 1 - \xi_b \\ 1 & \text{if } F(v) > 1 - \xi_b \end{cases},$$

where ξ_b is the solution of

$$1 - \xi = (1 - \xi^b)^b.$$

Thus the root's value is almost certain to be very close to $v^* \stackrel{\text{def}}{=} F^{\text{inv}}(1 - \xi_b)$. Note the ξ_b is the same as for the WIN-LOSS model discussed earlier.

If the leaf values are discrete taking values in $\{x_1, x_2, \ldots, x_m\}$, and if for all i, $F(x_i) \neq 1 - \xi_b$, the root's value converges to a limit which is the smallest x_i satisfying

$$F(x_{i-1}) < 1 - \xi_b < F(x_i).$$

All these results are due to Pearl [Pearl80, Pearl84]

III

THE SUM MODEL: DEFINITION AND DISTRIBUTION

The weakness of the classical model stems from the fact that it provides independent distributed values for the leaf nodes. It is more realistic to assume dependence between close nodes. Incremental models, and in particular the SUM model, provide a strong dependence between close nodes.

1. Definition and notation.

Consider an n-level b-ary tree in which each node at depths 0 through $n-1$ has b children, and each node at depth n is a leaf. 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:

$$F(x) = \mathbf{P}\{X \leq x\} .$$

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,

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

To make things simpler, we will speak of MAX and MIN nodes. 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:

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

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:

$$F_n(x) = \mathbf{P}\{V \leq x\} .$$

Clearly, we see that

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

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:

$$F_{2n}(x) = \left(\int F_{2n-1}(x - t)dF(t)\right)^b,$$

$$F_{2n+1}(x) = 1 - \left(1 - \int F_{2n}(x - t)dF(t)\right)^b.$$

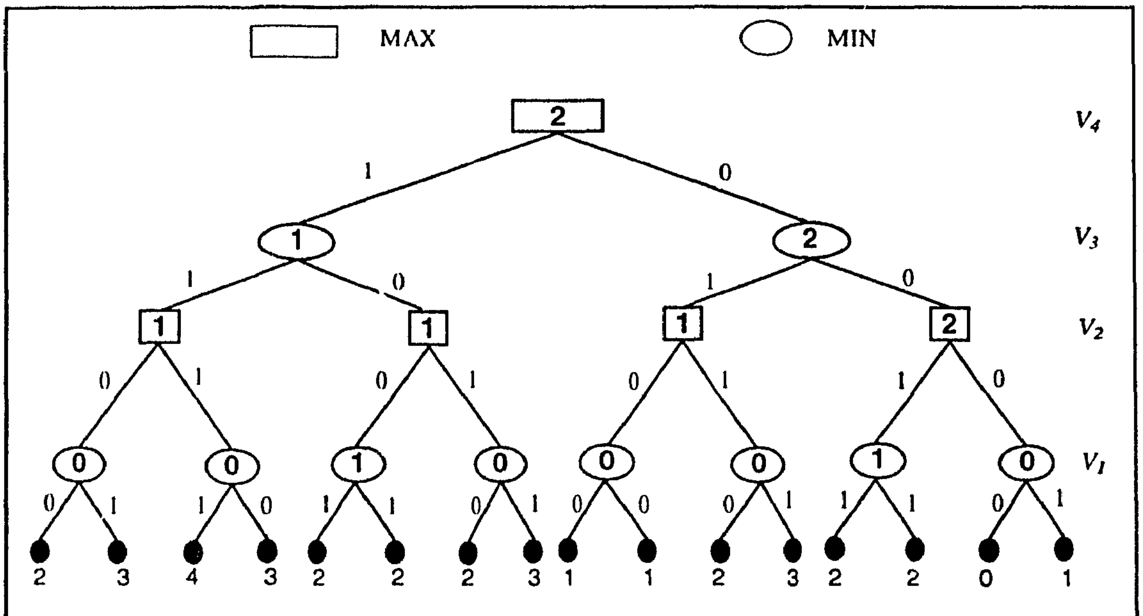
To understand the process, we will study the model in which X is a Bernoulli random variable:

$$\mathrm{P}\{X = 1\} = p \in (0, 1),$$

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

We call this model the SUM model with parameter p. The recurrence becomes

$$\begin{aligned} F_{2n}(i) &= \left(pF_{2n-1}(i - 1) + (1 - p)F_{2n-1}(i)\right)^b, \\ F_{2n+1}(i) &= 1 - \left(p(1 - F_{2n}(i - 1)) + (1 - p)(1 - F_{2n}(i))\right)^b. \end{aligned} \tag{1}$$

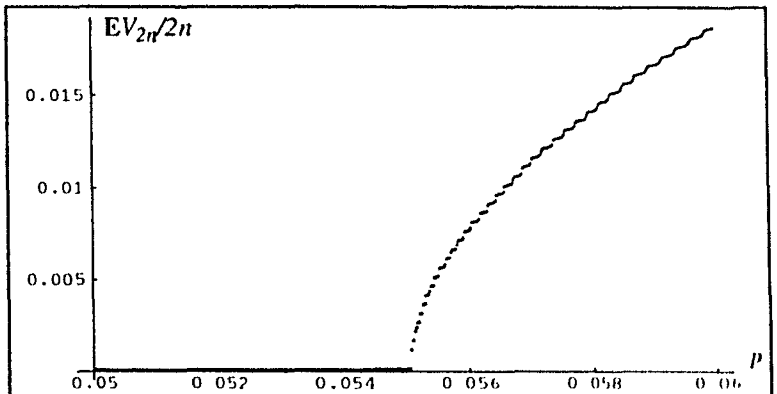


A SUM model.

We may consider this model also as a classical MIN-MAX tree with the values of leaf nodes equal to the sum of the edge values of the path from the root to the leaf. These leaf values then follow a binomial distribution with $2n$ trials and success probability p. Notice, however, that the binomial leaf values are heavily dependent.

2. Asymptotic behavior for small p.

In this part V_n denotes the root value of the n-level b-ary SUM model with parameter p. The graph of $EV_n/2$ for $b = 2$ and large n has two "flat" parts, one for small p and one for p near 1. In these flat parts small variations of p do not influence the limit of EV_n/n. We observe the same kind of behavior for other b. All of this is captured in Theorem 1.



$EV_{2n}/2n$ for n=1000 and b=2.

As $n \to \infty$, the behavior of V_n depends very much on b and p. We consider first small values of p. Below some threshold value α which itself is a function of b, we see that V_n tends to a limit distribution that depends upon p in the sense that for all $i \geq 0$,

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

and

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

where F_∞ and H_∞ are bona fide distribution functions that put positive mass on all nonnegative integers. The limit distributions will be described below. It is noteworthy that such a limit result is only possible because the zeros overwhelm the ones for small values of p. For $p > \alpha$, we will show that $V_n \to \infty$ in probability. Thus, at $p = \alpha$, there is an abrupt change in the asymptotic behavior of V_n. In this section, we show the following:

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

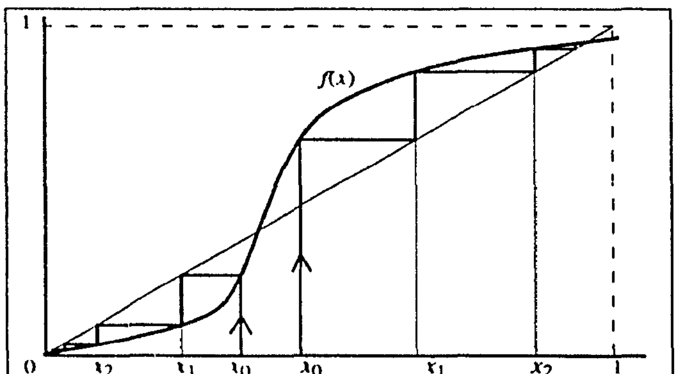
$$\alpha \leq 1 - b^{-1/(b+1)} \stackrel{b\to\infty}{\longrightarrow} 0.$$

The remainder of this section contains the proof of this result. We first show a small useful result about recurrences.

LEMMA 2. *Let x_n be a sequence defined by*

$$x_n = f(x_{n-1})\,,\ \text{and}\ x_0 = a \in [0,1],$$

where f is a continuous increasing function from $[0,1]$ to $[0,1]$. Then f has at least one fixed point. If $f(a) < a$ then x_n converges to the greatest fixed point smaller than a. If $f(a) > a$ then x_n converges to the smallest fixed point greater than a.



Two sequences reccurently define.

PROOF. $f : [0,1] \to [0,1]$ implies $f(0) \geq 0$ and $f(1) \leq 1$ and then f has at least one fixed point on $[0,1]$. Since f is increasing we have by induction,

$$\text{if } x_1 \geq x_0\,,\ x_n \geq x_{n-1} \text{ for all } n > 0,$$
$$\text{if } x_1 \leq x_0\,,\ x_n \leq x_{n-1} \text{ for all } n > 0.$$

Thus x_n is monotone and bounded, and it converges. Let l be the limit. By continuity of f we have $f(l) = l$. To see that this limit is L the smallest fixed point greater than a, we argue by contradiction. Assume that x_n is increasing and it converges to a limit l not equal to L. Thus l is a fixed point and $l > L$ and there exists N such that $x_N < L$ and $x_{N+1} > L$. This implies that $f(x_N) = x_{N+1} > L$. As f is increasing and $x_N < L$ we have $f(x_N) \leq f(L) = L$. Contradiction! For x_n decreasing the proof is similar. □

Part 1: Limit of $F_n(0)$.

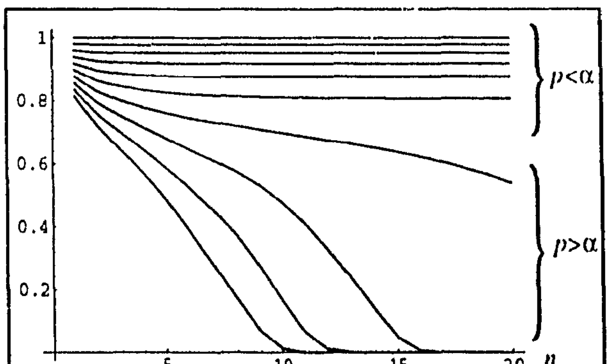
We first prove Theorem 1 for $i = 0$.

LEMMA 3. *For all b there exists $\alpha \in [0, 1)$ such that for $p \in [0, \alpha]$,*

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

If $p > \alpha$,

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



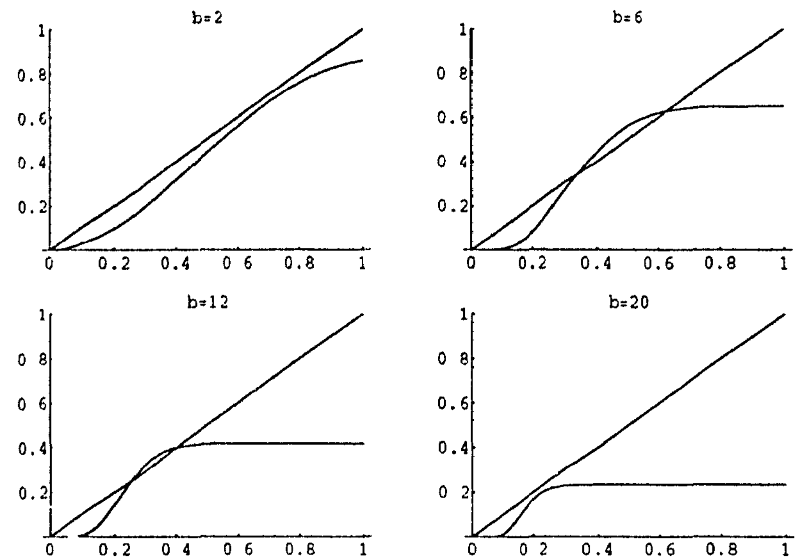
$F_{2n}(0)$ for p=0 to p=0.1, b=2.

PROOF. We assume $p \in [0, 1)$. We consider the recurrence (1) for $i = 0$:

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

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

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



$G_0(x)$ for p=0 07, and for different b.

where $G_0(x) = (1-p)^b\left(1-(1-(1-p)x)^b\right)^b$. 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 $G_0(1) \leq 1$. G_0 satisfies the hypotheses of Lemma 2. Thus $F_{2n}(0)$ converges and as $F_0(0) = 1$, it 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

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

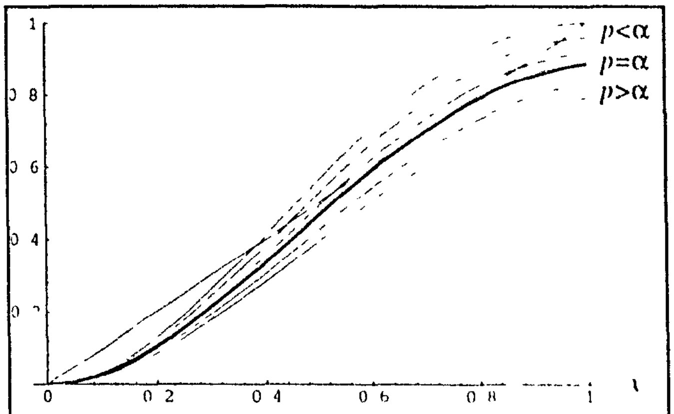
As $G_0(0) = 0$, h is a (b^2-1)-th order polynomial function of p and x. As $G_0'(0) = 0$, the derivative of $G_0(x) - x$ is equal to -1 at $x = 0$, and thus zero is a simple root of $G_0(x) - x$ and it is not a root of h. Thus we have,

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

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 Lemma 3 is proved for $p \leq \alpha$.



$G_0(x)$ for b=2 and p between 0 an 0 1

If $p > \alpha$ then G_0 has only zero as fixed point. Thus according to Lemma 2, $F_{2n}(0)$ converges to zero. If $b = 2$, G_0 becomes a fourth order polynomial and we can get the exact value of α. For general $b > 2$ we will derive a theoretical upper bound of α. 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

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

Part 2: Value of α.

We first prove that $\alpha > 0$. Then we give the exact value of α when $b = 2$ and an upper bound for general b.

LEMMA 4. *For all $b \geq 2$,*

$$1 - \sqrt[b+1]{\frac{1}{b}} \geq \alpha > 0.$$

And for $b = 2$,

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

PROOF. We first 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 Lemma 4 is proved.

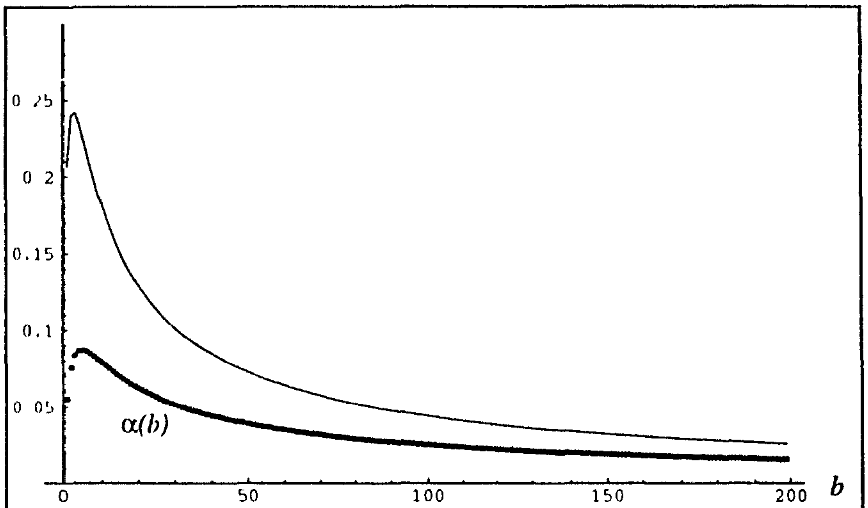
For general $b \geq 2$, we derive an upper bound for α, and an upper bound for $F_{2n}(0)$ when p is greater than α's upper bound. We have :

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

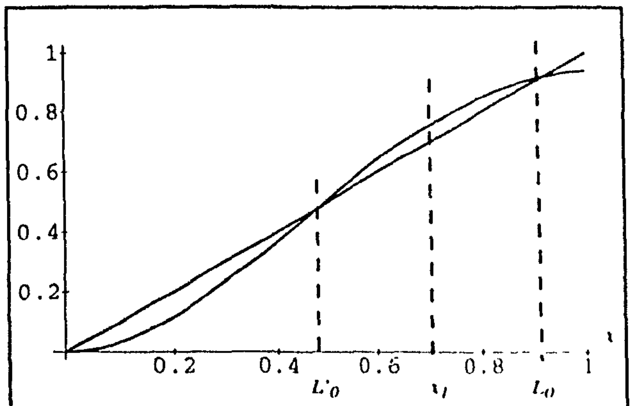
Thus if $bq^{b+1} < 1$, $G_0(x) < x$ for all $x > 0$. This implies that it cannot have a non-zero fixed point and thus $p > \alpha$. Thus we have

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

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



The value of α and its upper bound .



$G_0(x)$ for p=0.03 and b=2

Assume $b = 2$. We have

$$G_0(x) = (1-p)^4 x^2 (2-(1-p)x)^2.$$

Zero is a fixed point for G_0, thus, in order to find the other fixed points we study

$$h(x) \stackrel{\text{def}}{=} \frac{G_0(x) - x}{x}.$$

We note that

$$h(x) = (1-p)^4 x (2 - x(1-p))^2 - 1$$

and $h(0) = -1$, $h(1) = (1-p)^2(1+p)^2 - 1 < 0$. Observe that

$$h'(x) = (1-p)^4 (3(1-p)^2 x^2 - 8(1-p)x + 4),$$

the roots of which are

$$x_1 = \frac{2}{3(1-p)},\ x_2 = \frac{2}{1-p}.$$

The second root is greater than 1 for all p, and is therefore of no consequence. A little thought shows that $x_1 < 1$ for $p < 1/3$ and that

$$h(x_1) = \left(\frac{32}{27(1-p)}\right)^3 - 1.$$

This implies that $h(x_1) > 0$ if and only if

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

If $p \in (0, \alpha)$, h has two roots on (0, 1), and G_0 has two fixed points L_0', L_0 satisfying

$$0 < L_0' < x_1 < L_0 < 1. \square$$

Remark: We have also shown that

$$F_{2n}(0) \leq (bq^{b+1})^n.$$

Part 3: A study of the variation of the fundamental recurrence.

Using (1) we get the fundamental recurrence:

$$F_{2n}(i) = G(F_{2n-2}(i-1), F_{2n-2}(i-2), F_{2n-2}(i)) = G(u,v,x)\ ,\ n \geq 1,$$
$$F_0(i) = \begin{cases} 0 & \text{if } i < 0 \\ 1 & \text{if } i \geq 0 \end{cases}, \tag{2}$$

where

$$\sqrt[b]{G(u,v,x)} \stackrel{\text{def}}{=} p\left(1 - \left(p(1-v) + (1-p)(1-u)\right)^b\right)$$
$$+ (1-p)\left(1 - \left(p(1-u) + (1-p)(1-x)\right)^b\right).$$

According to the definition of G_0 we have,

$$G_0(x) = G(0,0,x).$$

LEMMA 5. *We assume $p \in (0,1)$. G is a strictly increasing function of u, v and x for $(u,v,x) \in [0,1]^3$. For all $(u,v) \in [0,1]^2$, $G(u,v,x)$ has a unique inflection point $z(u,v) \in [0,1]$, where $z(u,v)$ is a decreasing function of u and v. Finally $G(u,v,x)$ has at most three fixed points in x on* $[0,1]$ *and at most two fix points on* $[z(u,v),1]$.

PROOF. The increasing or decreasing nature of $G(u,v,x)$ is the same as that of

$$H(u,v,x) \stackrel{\text{def}}{=} 1 - p\left(p(1-v) + (1-p)(1-u)\right)^b - (1-p)\left(p(1-u) + (1-p)(1-x)\right)^b.$$

Since $(u,\ v,\ x) \in [0,1]$, and $p \in (0,1)$ the values inside the powers always stay positive, and these values are strictly decreasing with u, v, x. Thus H is strictly increasing in u, v and x when $p \in (0,1)$. Defining

$$A = 1 - p\,(p(1-v) + (1-p)(1-u))^b\,,$$
$$B = p(1-u) + (1-p)(1-x),$$

we have

$$B' = \frac{dB}{dx} = -(1-p)$$

and

$$H = A - (1-p)B^b.$$

Thus,

$$G' = \frac{dG}{dx} = bH'H^{b-1},$$

and

$$G'' = \frac{d^2G}{dx^2} = bH^{b-2}\left(H''H + (b-1)H'^2\right),$$

where

$$H' = (1-p)^2 bB^{b-1},$$

and

$$H'' = -(1-p)^3 b(b-1)B^{b-2}.$$

Therefore

$$G'' = (1-p)^3 b^2 (b-1) H^{b-2} B^{b-2} \Big((1-p)(1+b)B^b - A \Big).$$

The sign of G'' is determined by the sign of $(1-p)(1+b)B^b - A$. Since B is a strictly decreasing function of x, we have equality for at most one x. Thus G has at most one inflection point, and at most three fixed points.

The second derivative G'' vanishes for

$$x = 1 - \frac{1}{(1-p)} \left(\sqrt[b]{\frac{A}{(1-p)(1+b)}} - p(1-u) \right) \stackrel{\text{def}}{=} z(u,v).$$

As A is an increasing function of u and v, $z(u,v)$ is a decreasing function of u, v. The second derivative of G is negative on $[z(u,v), 1]$ and this implies that G has at most two fixed points on $[z(u,v), 1]$. □

LEMMA 6. *If $p = \alpha$, $G_0(x)$ has two fixed points on $[0,1]$: $0, L_\alpha$. L_α is a double fixed point. For $p < \alpha$, $G_0(x)$ has exactly three simple fixed points on $x \in [0,1]$: 0, L_0', L_0. And for all $(u,v) \in [0,1]^2$, $G(u,v,x)$ has exactly one simple fixed point on $[L_0, 1]$ denoted $L(u,v)$. Futhermore for $p \in (0,\alpha]$, $L(u,v)$ is a continuous function of (u,v) on $[0,1]^2$ when $(u,v) \neq (0,0)$ and $(u,v) \neq (1,1)$.*

PROOF.

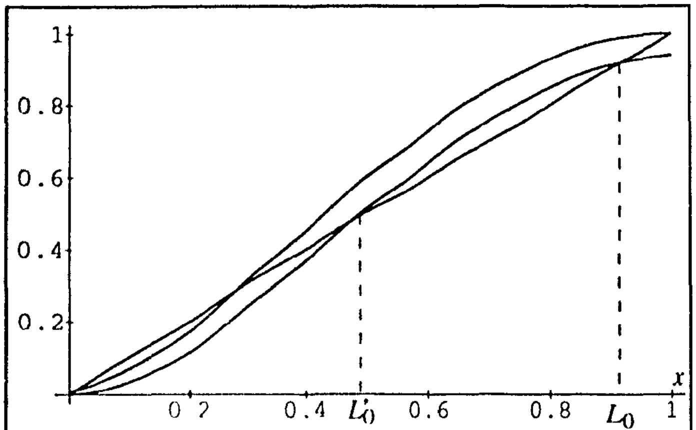
Assume $p = \alpha$, then by the definition of α, $G_0(x)$ has a fixed point on $(0,1)$. We denote by L_α the largest one. We prove by contradiction that it is the only one This would imply that it is a fixed point of multiplicity 2. Assume that L is a fixed point not equal to 0 or L_α Then G_0 has three simple fixed points, i.e., $0, L, L_\alpha$. And as $G_0(x)$ has a unique inflection point on $[0,1]$ this implies that for some $x \in (L, L_\alpha)$ we have $G_0(x) > x$ Then using the continuity of G_0 in p we can show that there exists ε such that for $p = \alpha + \varepsilon$, $G_0(x) > x$. As $G_0(1) < 1$, this implies the existence of a non-zero fixed point, contradicting the definition of α. Hence, L_α is the only fixed point of $G_0(x)$ on $(0,1)$ when $p = \alpha$ and it is a double fixed point

Assume $0 < p < \alpha$. We will show that $G_0(x)$ has three simple fixed points on $[0,1]$. We have

$$G_0 = (1-p)^b \left(1 - (1-(1-p)x)^b\right)^b,$$

and

$$G_0(0) = 0, \ \frac{dG_0}{dx}(0) = 0 < 1, \ G_0(1) < 1.$$



$G(0,0,x)$ and $G(1,1,x)$ for p=0.03.

We already know that zero is a fixed point and according to the preceding Lemma, G_0 has at most three fixed points. Since $G_0'(0) = 0$ and $G_0(0) = 0$, there exists $\varepsilon < L_\alpha$ such that

$$G_0(\varepsilon) < \varepsilon. \tag{3}$$

Since for all $x > 0$, G_0 is strictly decreasing with p, we have

$$G_0(L_\alpha) > L_\alpha \text{ for all } p \in (0, \alpha). \tag{4}$$

Using (3) and (4), $G_0(x)$ has a fixed point $L_0' \in (\varepsilon, L_\alpha)$ and a fixed point $L_0 \in (L_\alpha, 1]$, and these are simple fixed points. This also implies $z(0,0) < L_0$ as $G'(L_0') > 1$ and $G'(L_0) < 1$.

Next, we study the fixed points of $G(u, v, x)$ in x for $0 < p \leq \alpha$. We already know that for all $(u, v) \in [0, 1]^2$ with $(u, v) \neq (0, 0)$ and $(u, v) \neq (1, 1)$,

$$G(u, v, L_0) > L_0, \quad G(u, v, 1) < 1. \tag{5}$$

As $z(u, v)$ is strictly decreasing with u, v, we have $z(u, v) < L_0$. Thus $G(u, v, x)$ does not have an inflection point on $(L_0, 1]$, and it has at most two fixed points on this interval. (5) implies that $G(u, v, x)$ has an odd number of fixed point on $(L_0, 1]$. Thus it has exactly one simple fixed point on $(L_0, 1]$ denoted by $L(u, v)$. $L(u, v)$ is also the only fixed point on $(L_0', 1]$. We have

$$L = L(u, v) = \max\{x \in [0, 1] \mid G(u, v, x) = x\}.$$

For $p = 0$, $G(u, v, x)$ has three simple fixed points which are independent of u and v. The largest is 1.

In order to prove the continuity of L we use the theorem of implicit functions [Schw67] (p. 278). L is defined as the greatest solution of $G(u, v, x) - x = 0$ on $[0, 1]$. Recall that G is polynomial. According to this theorem the implicit function exists and is continuous at the point (u, v) if the derivative of $G(u, v, x) - x$ on x is non-zero at the point $(u, v, L(u, v))$. This derivative is zero if and only if L is not a simple root. We just proved that for $0 < p \leq \alpha$ and $(u, v) \neq (0, 0)$, L is a simple fixed point. Thus if $p < \alpha$ the implicit function exists and is continuous and Lemma 6 is proved. □

Part 4: Convergence of $F_{2n}(i)$ for $0 < p \leq \alpha$.

LEMMA 7. *For all $p \in (0, \alpha]$, there are numbers $F_\infty(i) \in (0, 1)$ such that*

$$\lim_{n \to \infty} F_{2n}(i) = F_\infty(i), \text{all } i \geq 0.$$

These values can be computed by the recurrence

$$F_\infty(0) = L_0,$$
$$F_\infty(1) = L(L_0, 0),$$
$$F_\infty(i) = L\left(F_\infty(i-1), F_\infty(i-2)\right) \text{ for all } i > 1$$

PROOF. We prove that $F_{2n}(i)$ converges to a limit $F_\infty(i) \in (0, 1)$ by induction on i. For $i = 0$, according to Lemma 3, $F_{2n}(0)$ is a decreasing sequence with n and it converges to $F_\infty(0) = L_0 \in (0, 1)$ when n tends to infinity. Thus, for all ε there exists a N such that for all $n > N$

$$F_\infty(0) \leq F_{2n}(0) \leq F_\infty(0) + \varepsilon < 1$$

We define the two sequences x_{2n} and y_{2n} by

$$y_{2N} \stackrel{\text{def}}{=} x_{2N} \stackrel{\text{def}}{=} F_{2N}(1),$$
$$y_{2n} \stackrel{\text{def}}{=} G(F_\infty(0), 0, y_{2n-2}) \quad n > N$$
$$x_{2n} \stackrel{\text{def}}{=} G(F_\infty(0) + \varepsilon, 0, x_{2n-2}), \quad n > N.$$

As $G(u, v, x)$ is increasing with u and v,

$$y_{2n} \leq F_{2n}(1) \leq x_{2n}.$$

According to Lemma 6, $G(F_\infty(0), 0, x)$ and $G(F_\infty(0)+\varepsilon, 0, x)$ have only one fixed point ($L(F_\infty(0), 0)$ and $L(F_\infty(0) + \varepsilon, 0)$ respectively) on $[L_0, 1]$. Thus, since G is increasing and y_{2N} and x_{2n} are greater than L_0, using Lemma 2 we have,

$$\lim_{n \to \infty} y_{2n} = L(F_\infty(0), 0)$$
$$\lim_{n \to \infty} x_{2n} = L(F_\infty(0) + \varepsilon, 0).$$

Using the continuity of L, for all ε there exists an N such that for all $n > N$,

$$|F_{2n}(1) - F_\infty(1)| \leq \varepsilon.$$

Thus $F_{2n}(1)$ tends to $F_\infty(1)$ when n tends to infinity. For $p > 0$ we have $F_\infty(1) > F_\infty(0)$.

Suppose that $F_{2n}(i-1)$ and $F_{2n}(i-2)$ converge to $F_\infty(i-1)$ and $F_\infty(i-2)$ respectively. We prove that $F_{2n}(i)$ converges to $F_\infty(i) = L(F_\infty(i-1), F_\infty(i-2))$. Since L is a continuous function from $(0, 1) \times [0, 1]$ to $(0, 1)$, we see that for all $\delta > 0$, there exists $\varepsilon_2 > 0$ such that for all

$(w, w') \in ((0,1)\times[0,1])^2$ with $\|w-w'\| < \varepsilon_2$, we have $|L(w)-L(w')| < \delta$. We set $\varepsilon = \min\{\varepsilon_2/4, \delta\}$. Hence

$$\begin{aligned} |L(F_\infty(i-1)+\varepsilon, F_\infty(i-2)+\varepsilon) - F_\infty(i)| &< \delta, \\ |L(F_\infty(i-1)-\varepsilon, F_\infty(i-2)-\varepsilon) - F_\infty(i)| &< \delta, \end{aligned} \tag{6}$$

where $F_\infty(i) = L(F_\infty(i-1), F_\infty(i-2))$.

Using the convergence of $F_{2n}(i')$ for $i' < i$, there exists a finite N such that for $n > N$,

$$\begin{aligned} |F_{2n}(i-1) - F_\infty(i-1)| &\le \varepsilon \le \delta, \\ |F_{2n}(i-2) - F_\infty(i-2)| &\le \varepsilon \le \delta. \end{aligned} \tag{7}$$

Let x_{2n} and y_{2n} be sequences defined for $n > N$ by the recurrences

$$\begin{aligned} y_{2N} &\stackrel{\text{def}}{=} x_{2N} \stackrel{\text{def}}{=} F_{2N}(i), \\ y_{2n} &\stackrel{\text{def}}{=} G(F_\infty(i-1)-\varepsilon, F_\infty(i-2)-\varepsilon, y_{2n-2}),\ n > N \\ x_{2n} &\stackrel{\text{def}}{=} G(F_\infty(i-1)+\varepsilon, F_\infty(i-2)+\varepsilon, x_{2n-2}),\ n > N. \end{aligned}$$

Then as G is increasing and using (7) we know that for all $n > N$,

$$y_{2n} \le F_{2n}(i) \le x_{2n}\,. \tag{8}$$

Since, for all n, $F_n(i) \ge F_n(1)$, we have $F_\infty(i-1) \ge F_\infty(1) > L_0$.We can choose δ small enough to have

$$F_{2N}(i) \ge F_{2N}(i-1) \ge F_\infty(i-1) - \delta \ge L_0.$$

Thus, since the starting values of x_{2n} and y_{2n} are $F_{2N}(i) \ge L_0$, using Lemma 6 and Lemma 3, we have

$$\lim_{n\to\infty} x_{2n} = L(F_\infty(1)+\varepsilon, F_\infty(2)+\varepsilon)$$

and

$$\lim_{n\to\infty} y_{2n} = L(F_\infty(1)-\varepsilon, F_\infty(2)-\varepsilon). \tag{9}$$

Using (7),(8) and (9), there exists a finite N' such that for $n > N'$,

$$F_\infty(i) - 2\delta < F_{2n}(i) < F_\infty(i) + 2\delta\,.$$

By the arbitrary nature of δ we conclude that for $p < \alpha$,

$$\lim_{n\to\infty} F_{2n}(i) = L(F_\infty(i-1), F_\infty(i-2)) > 0\,.$$

Using

$$F_{2n+1}(i) = 1 - \Big(p(1-F_{2n}(i-1)) + (1-p)(1-F_{2n}(i))\Big)^b,$$

one can show the convergence of $F_{2n+1}(i)$ when n tends to infinity to a non-zero limit. □

Part 5: The limit of $F_\infty(i)$ when i tends to infinity.

Assume $0 < p \leq \alpha$. In order to have a bona fide distribution we have to prove that $F_\infty(i)$ tends to 1 when i tends to infinity. As G is a continuous function we can take the limit in the fundamental recurrence on both sides of the equality. Then we get the following recurrence:

$$F_\infty(0) = L_0,$$
$$F_\infty(1) = L(L_0, 0),$$
$$F_\infty(i) = G(F_\infty(i-1), F_\infty(i-2), F_\infty(i)).$$

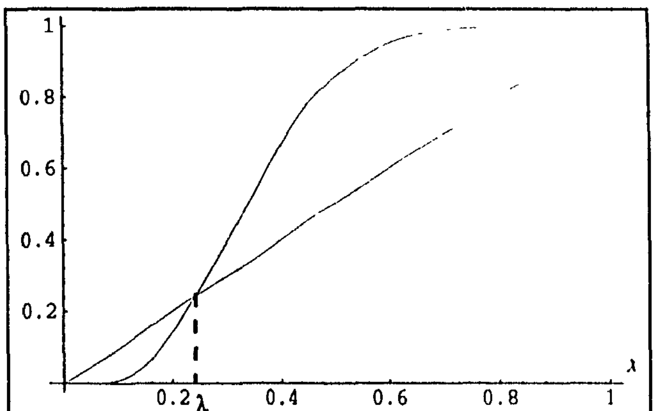
Thus

$$F_\infty(i) \geq G(F_\infty(i-2), F_\infty(i-2), F_\infty(i-2))$$

since

$$F_\infty(i-2) \leq F_\infty(i-1) \leq F_\infty(i) \tag{10}$$

We define $J(x) = G(x,x,x) = \left(1-(1-x)^b\right)^b$. J has 0 and 1 as fixed points, and a third one on $(0,1)$ denoted λ.



$J(x)$ for b=5.

We define the sequence y_i:

$$y_0 = L_0$$
$$y_i = J(y_{i-1}), \text{ for } i \geq 1. \tag{11}$$

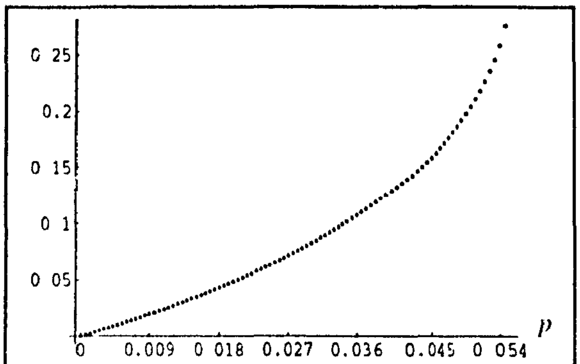
Thus using (10)

$$1 \geq F_\infty(2i+1) \geq F_\infty(2i) \geq y_i. \tag{12}$$

Since G is strictly increasing in (u,v) when $p > 0$, we have $J(x) > G_0(x)$ and thus $y_0 > \lambda$. Then according to Lemma 2, y_i converges to 1. Thus, from (12),

$$\lim_{i\to\infty} F_\infty(i) = 1. \square$$

Part 6: The expected value when $p \leq \alpha$.



EV_{2n} for $2n$=400, b=2.

LEMMA 8. *If $p \leq \alpha$, EV_∞ is finite.*

PROOF. Define y_i as in (11). Using

$$EV_\infty = \sum_{i \geq 0} (1 - F_\infty(i)) .$$

(10), and (12) we see that EV_∞ is finite if

$$\sum_{i \geq 0} (1 - y_i) < \infty.$$

But

$$\begin{aligned} 1 - y_{i+1} &= 1 - \left(1 - (1 - y_i)^b\right)^b, \\ &\leq b(1 - y_i)^b, \\ &\leq \frac{1 - y_i}{2} \end{aligned}$$

if

$$1 - y_i \leq \frac{1}{(2b)^{1/(b-1)}}.$$

Let I be the smallest integer $1 - y_I \leq \frac{1}{(2b)^{1/(b-1)}}$. (This exists because $y_i \to 1$ as $i \to \infty$.) Then for $i > I$,

$$1 - y_i \leq \frac{1 - y_I}{2^{i-I}}.$$

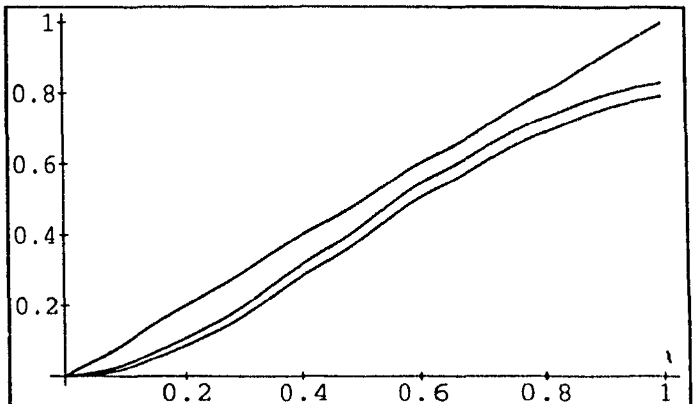
Therefore,

$$\sum_{i \geq I} (1 - y_i) \leq 2(1 - y_I).$$

The remainder of the proof is trivial. □

Part 7: The limit of $F_{2n}(i)$ for $p > \alpha$.

We know that G is increasing in u, v and x, if $u, v, x \in [0,1]$. We studied $G(0,0,x)$ during the study of $F_n(0)$, and we showed that by definition of α, if $p > \alpha$, $G(0,0,x)$ has zero as unique fixed point on $[0,1]$. Thus for all $\delta \in (0,1)$, there exists $\varepsilon > 0$, such that for all $x \in [\delta, 1]$, $G(0,0,x) < x - \varepsilon$. Also there exists $\varepsilon_1 > 0$ such that for all $x \in [\delta, 1]$ and for all $(u,v) \in [0,\varepsilon_1]^2$, $G(u,v,x) < x - \varepsilon/2$. Since $G(u,v,0) \geq 0$, we see that for all $(u,v) \in [0,\varepsilon_1]^2$, $G(u,v,x)$ has a unique fixed point $L(u,v) \in [0,\delta]$.



$G(0,0,x)$ and $G(0.1, 0.1, x)$ for p=0.1.

We prove the convergence of $F_{2n}(i)$ to zero by induction on i. According to Lemma 2, $F_{2n}(0)$ converges to zero. The following proof is valid for $i = 1$ noting $F_n(-1) = 0$. We begin the induction with $i \geq 1$. Suppose that $F_{2n}(i-1)$ and $F_{2n}(i-2)$ converge to zero. Then we know that for all $\varepsilon_1 > 0$ there exists N such that for all $n > N$,

$$F_{2n}(i-1) \leq \varepsilon_1, \; F_{2n}(i-2) \leq \varepsilon_1 .$$

Let x_{2n} be a sequence defined by the recurrence

$$x_{2n} \stackrel{\text{def}}{=} G(\varepsilon_1, \varepsilon_1, x_{2n-2}),$$
$$x_{2N} \stackrel{\text{def}}{=} F_{2N}(i) .$$

As G is increasing in u and v we have for $n > N$, $0 \leq F_{2n}(i) \leq x_{2n}$. The sequence x_{2n} converges to $L(\varepsilon_1, \varepsilon_1)$. Thus there exists $N_1 > N$ such that for all $n > N_1$,

$$|x_{2n} - L(\varepsilon, \varepsilon)| \leq \delta.$$

Thus for all δ there exists N_1 such that for all $n > N_1$

$$F_{2n}(i) \leq 2\delta.$$

By the arbitrary nature of δ,

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

This implies $V_{2n} \longrightarrow \infty$ in probability. □

3. p close to 1.

This part generalizes the results of the preceding part for the case p close to one by a symmetrical argument.

<u>THEOREM 2.</u> *For $p \in [0, 1-\alpha]$, and i fixed,*

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

If $p \in [1-\alpha, 1]$,

$$\lim_{n\to\infty} F_{2n}(2n-i) < 1 \text{ for all } i > 0.$$

PROOF. We recall the definition of V_n in the SUM model:

$$V_0 = 0;$$
$$V_n \overset{\mathcal{L}}{=} \begin{cases} \max_{1\le j\le b}\{V_{n-1,j}+X_j\} & \text{if } n \text{ is even} \\ \min_{1\le j\le b}\{V_{n-1,j}+X_j\} & \text{if } n \text{ is odd} \end{cases},$$

where the X_j's denote independent Bernoulli (p) random variables corresponding to adjacent edges. The $V_{n-1,j}$'s are independent copies of V_{n-1}. Let us define $Y_j \overset{\mathcal{L}}{=} 1 - X_j$ and $V_n'' \overset{\mathcal{L}}{=} n - V_n$. Then we have the following distributional identities:

$$V_{2n}'' \overset{\mathcal{L}}{=} 2n - \max_{1\le j\le b}\left\{(2n-1) - V_{2n-1,j}'' + 1 - Y_j\right\},$$
$$V_{2n+1}'' \overset{\mathcal{L}}{=} 2n + 1 - \min_{1\le j\le b}\left\{2n - V_{2n,j}'' + 1 - Y_j\right\}.$$

From this, we obtain

$$V_n'' \overset{\mathcal{L}}{=} \begin{cases} \min_{1\le j\le b}\left\{V_{n-1,j}''+Y_j\right\} & \text{if } n \text{ is even} \\ \max_{1\le j\le b}\left\{V_{n-1,j}''+Y_j\right\} & \text{if } n \text{ is odd} \end{cases}.$$

Here the Y_j's are i.i.d. Bernoulli with parameter $1-p$. Thus for $n \ge 2$, V_{2n+1}'' follows the same recurrence as V_{2n}. We denote by F_n'' the distribution function of V_n''. The recurrence for $F_{2n+1}''(i)$ is

$$F_1''(0) = 1 - p^b,$$
$$F_1''(1) = 1,$$
$$F_{2n+1}''(i) = G(F_{2n-1}''(i-1), F_{2n-1}''(i-2), F_{2n-1}''(i)).$$

If $p < 1-\alpha$ then $q > \alpha$ and the recurrence function only has zero as fixed point and $F_{2n+1}'(i)$ converges to 0. If $p > 1-\alpha$ one can show that the proof for convergence remains valid, and $F_n''(i)$ converges to a non-zero limit. □

IV

THE SUM MODEL: EXPECTED VALUE

1. The main theorem.

In this chapter we prove that the expected value of the root divided by n converges and that limit is a continuous function of p. To prove this we first show that the distribution of the root value is highly concentrated around the expected value.

THEOREM 3. *For every p, $\mathrm{E}V_n/n$ converges to a finite limit $\mathcal{V}(p)$, and $\mathcal{V}$ is a uniformly continuous function of p. Furthermore if $\alpha < p < 1-\alpha$ then $0 < \mathcal{V}(p) < 1$ and $V_n/\mathrm{E}V_n \longrightarrow 1$ almost surely when $n \longrightarrow \infty$.*

2. Construction and notation.

We recall the definition of V_n, the root value of a complete b-ary tree with n levels of edges in which we associate with each edge 1 or 0 with probability p and $1-p$ respectively. The values of the nodes are found by the following recursive rule: all leaves have value 0, and for every node u with A_u as its set of children we have

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

where $E(u,v)$ is the value of the edge (u,v). This defines the n-level b-ary SUM tree with parameter p. V_n is the random variable defined as the root value of such a tree. Then we get the following distributional identities: $V_0 \stackrel{\mathcal{L}}{=} 0$,

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

where X_j are i.i.d. Bernoulli random variables with parameter p, and $\{V_{n-1,j}\}$ are i.i.d. copies of V_{n-1}.

We obtain a b-ary Pearl tree with parameter $q \in (0,1)$ if in the previous construction we set $E(u,v) = 0$, and $V_0 = 1$ with probability q and $V_0 = 0$ with probability $1-q$. Finally, in our proof, we need an associated tree. We fix the integers $N \ge 1$ and $k \ge 1$, and we consider a b-ary SUM tree with parameter p. A node u in the associated tree has value $V'(u)$. The associated tree has the property that for every node u, $V'(u) \ge V(u)$. The leaves have value zero. At any level i that is not a multiple of N, we follow the standard rules as for a b-ary SUM tree with parameter p. 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,

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

Note that many nodes may have the value infinity. We call this 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' are either ∞ or a given fixed finite value.

3. The fundamental inequalities.

The following lemma is the main part of the proof of convergence. The proof uses results on Pearl trees. Consider a SUM tree with lN levels, and cut it into pieces of N levels each. In each part the lowest k levels are used as a filter. The following Lemma forms the basis of the entire chapter.

LEMMA 9. *Let V_n be the root value of a n-level b-ary* SUM *tree with parameter $p \in [0,1]$. For all $\varepsilon > 0$ we have,*

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

Furthermore, for N large enough and for all $l > 0$,

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

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

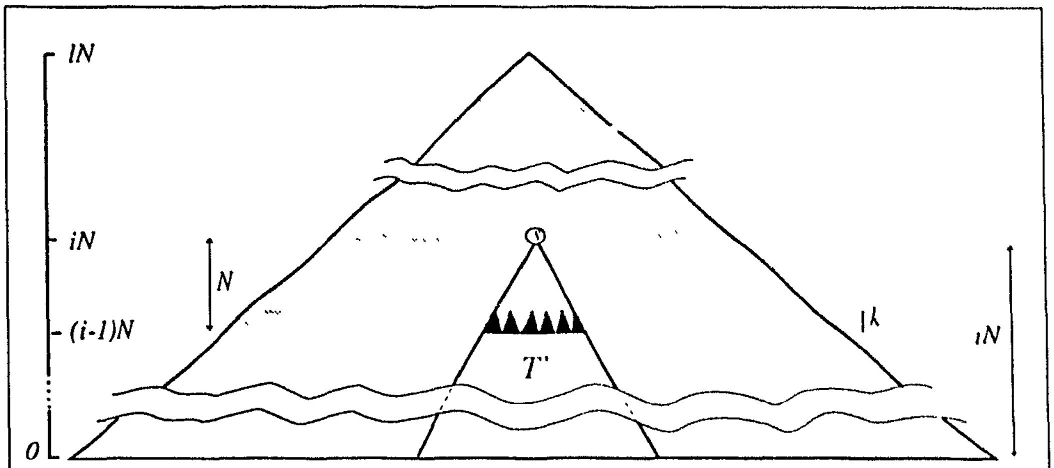
$$\mathbf{P}\{|V_n - l\mathbf{E}V_N| \ge (2l-1)k + N\} \le \varepsilon,$$

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

PROOF. Consider an n-level b-ary SUM tree with root u. 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 1. Thus, we can apply the McDiarmid's inequality (1989)[McDi89]: for all $\varepsilon > 0$,

$$\mathbf{P}\{|V_n - \mathbf{E}V_n| \ge \varepsilon\} \le 2e^{-2\varepsilon^2/n}. \quad (13)$$

The first bound is proved.



The construction

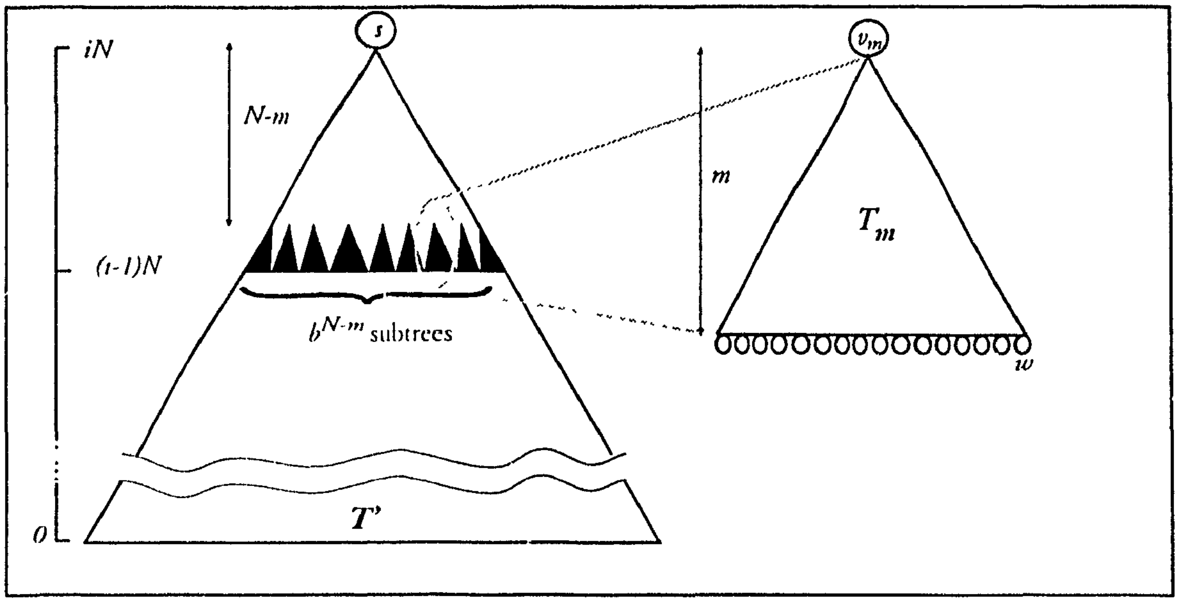
We now assume that the SUM tree has $n = lN$ levels. We consider the (k, N) associated tree with $k = \lceil N^{2/3} \rceil$. Let N be so large that $R(b, N) \leq (1/2)b^{-b}$. We will prove by induction that for such N we have for all $i \geq 1$,

$$\mathbf{P}\left\{V'_{iN} = \infty\right\} \leq R(b, N). \tag{14}$$

For $i = 1$, we obtain

$$\begin{aligned} \mathbf{P}\left\{V'_N = \infty\right\} &= \mathbf{P}\left\{V_N > \mathbf{E}V_N + k\right\}, \\ &\leq 2e^{-2k^2/N} \qquad \text{(according to (13))} \\ &\leq 2e^{-2N^{1/3}} \\ &\leq R(b, N). \end{aligned}$$

Now we assume that $\mathrm{P}\left\{V'_{(i-1)N} = \infty\right\} \leq R(b, N)$. 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. 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 of T_m are nodes of T' at level $(i-1)N$. Thus their values are i.i.d. distributed as $V'_{(i-1)N}$. Let w be a leaf of T_m. $V'(w)$ is distributed as $V'_{(i-1)N}$ independently of the other leaves.



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:

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

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

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

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 SUM model. Thus $V'(u)$ is infinity if and only if $V''(u)$ is infinity. If $V''(v_m) = 0$ then

$$V'(v_m) \leq (i-1)\mathrm{E}V_N + (2i-3)k + m. \tag{15}$$

Thus,

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

Furthermore we have

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

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 $(i-1)\mathbf{E}V_N + (2i-3)k + k$, and $V'(s)$ is stochastically less than $V'(v_m) + V_N$. Thus

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

Thus the induction proof of (14) is finished and we have

$$\mathbf{P}\{V_{lN} \ge l\mathbf{E}V_N + (2l-1)k\} \le R(b,N).$$

Now we consider b-ary SUM trees with n levels, n not a multiple of N, and we set $l = \lfloor \frac{n}{N} \rfloor$. Using (13) with $m = n - lN < N$ we get

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

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

$$\mathbf{P}\{V_n \ge l\mathbf{E}V_N + (2l-1)k + N\} \le b2^{-b^{\lfloor N^{1/4} \rfloor/2}}. \tag{16}$$

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

$$\begin{aligned}
\mathbf{P}\{V_n \ge l\mathbf{E}V_N + (2l-1)k + N\} &\le b^m R(b,N) \\
&\le b^{N^{1/4}} R(b,N).
\end{aligned} \tag{17}$$

Finally using (16) and (17) we have for all n,

$$\mathbf{P}\{V_n \ge l\mathbf{E}V_N + (2l-1)k + N\} \le b2^{-b^{\lfloor N^{1/4} \rfloor}} + b^{N^{1/4}} R(b,N).$$

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

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

where $k = \lceil N^{2/3} \rceil$ and $l = \lfloor n/N \rfloor$. Using the same method, one can easily show that we also have

$$\mathbf{P}\{V_n \le l\mathbf{E}V_N - (2l-1)k - N\} \le \varepsilon$$

under the same conditions. □

4. Convergence.

LEMMA 10. *For every p, $\mathbf{E}V_n/n$ has a limit $\mathcal{V}(p)$ when n tends to infinity. Furthermore for all $\varepsilon_1, \varepsilon_2 > 0$, there exists an N such that for $n > N/\varepsilon_1$,*

$$\mathbf{P}\left\{\left|\frac{V_n}{n} - \frac{\mathbf{E}V_N}{N}\right| \geq 2\varepsilon_1\right\} \leq \varepsilon_2.$$

Finally for all $\varepsilon_1, \varepsilon_2$ there exists a N such that for $n > N/\varepsilon_1$,

$$\mathbf{P}\left\{\left|\frac{V_n}{n} - \mathcal{V}(p)\right| \geq 3\varepsilon_1\right\} \leq \varepsilon_2.$$

PROOF. We show that

$$\limsup_{n\to\infty} \frac{\mathbf{E}V_n}{n} \leq \liminf_{n\to\infty} \frac{\mathbf{E}V_n}{n}, \tag{18}$$

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

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

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

$$\frac{\mathbf{E}V_N}{N} \leq \liminf_{n\to\infty} \frac{\mathbf{E}V_n}{n} + \varepsilon,$$

so that we may conclude (18) by the arbitrary nature of ε.

We use the notations of the preceding part: u is the root of an n-level b-ary SUM tree with parameter p, and $V'(u)$ is its value for the (k, N) associated tree. Let set $l = \lfloor n/N \rfloor$. If $V'(u) \leq N + l\mathbf{E}V_N + (2l-1)k$ then

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

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

$$\begin{aligned} \frac{V(u)}{n} &\leq \frac{N}{n} + \frac{\mathbf{E}V_N}{N} + 2\frac{k}{N} \\ &\leq \frac{N}{n} + \frac{\mathbf{E}V_N}{N} + 2\left(\frac{N^{2/3}+1}{N}\right) \\ &\leq \frac{\mathbf{E}V_N}{N} + 2\varepsilon \end{aligned} \tag{19}$$

for N large enough and $n \geq N/\varepsilon$. Thus using Lemma 9, we can find N large enough such that

$$\mathbf{P}\{V_n \geq l\mathbf{E}V_N + (2l-1)k + N\} < \varepsilon$$

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

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

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

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

as required. Thus $\mathbf{E}V_n/n$ has a limit $\mathcal{V}(p)$ when n tends to infinity. The second part of Lemma 10 follows easily from the above argument. □

5. An embedding lemma.

We consider a b-ary tree with n levels of edges and with each edge e of this tree we associate a uniform $[0,1]$ random variable U_e. Let $I_{[U\leq\varepsilon]}$ be the indicator function taking the value one if $U \leq \varepsilon$. We denote by $\mathcal{P}_n$ the collection of all b^n paths from the root to the leaves.

LEMMA 11. *There exists a positive function φ such that for all $\varepsilon > 0$,*

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

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

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 $1 > x \geq \varepsilon$,

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

where we use the Chernoff's bound for the tail of a binomial distribution (see for example Hoeffding, [Hoeff63], Theorem 1). Let $\varphi = \varphi(\varepsilon)$ be defined as follows:

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

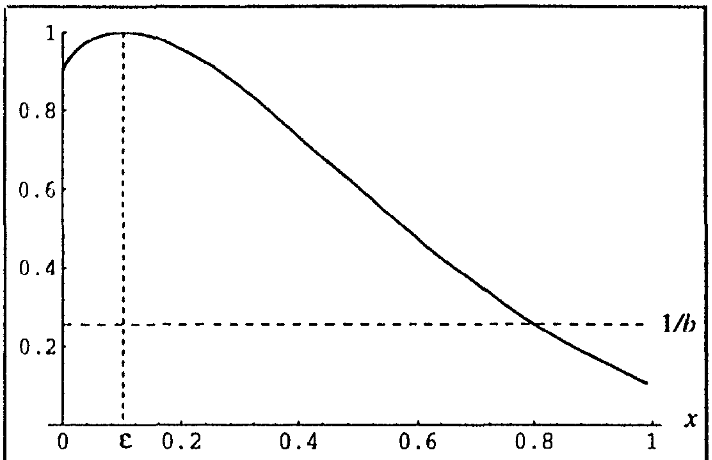
We denote

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

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 ε=0.1

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

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

because

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

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

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

By the arbitrary nature of δ, Lemma 11 follows. □

6. Continuity.

LEMMA 12. *Let $V_n(p)$ be the root value of an n-level b-ary SUM tree with parameter p. Then*

$$\lim_{\epsilon\to 0} \sup_{p,q\,|p-q|\le\epsilon} \limsup_{n\to\infty} \left| \frac{\mathbf{E}V_n(p)}{n} - \frac{\mathbf{E}V_n(q)}{n} \right| = 0,$$

and thus $\mathbf{E}V_n(p)/n$ is uniformly continuous in p for all n. If $\mathcal{V}(p)$ is the limit of $\mathbf{E}V_n/n$ when n tends to infinity, then $\mathcal{V}(p)$ is uniformly continuous in p as well.

PROOF. We use an embedding argument, associating with each edge in a b-ary tree an independent copy of a uniform [0, 1] random variable U. To obtain $V_n(p)$, we associate with each edge the value $I_{[U\le p]}$, where I is the indicator function. In this manner, $V_n(p)$ and $V_n(q)$, although both random quantities, are heavily coupled. Also, if $q > p$, then $V_n(q) \ge V_n(p)$. Next, let P be a path from the root to a terminal node, let e be a typical edge, and let U_e be the uniform [0, 1] random variable associated with that edge. Let $\mathcal{P}_n$ be the collection of all b^n paths from the root to a leaf in a b-ary tree of height n. The embedding construction shows immediately the following: if $p+\epsilon > q > p$, then

$$0 \le \mathbf{E}V_n(p) - \mathbf{E}V_n(q) = \mathbf{E}\{V_n(p) - V_n(q)\} \le \mathbf{E}\left\{ \max_{P\in\mathcal{P}_n} \sum_{e\in P} I_{[p< U_e < q]} \right\}$$

$$\le \mathbf{E}\left\{ \max_{P\in\mathcal{P}_n} \sum_{e\in P} I_{[U_e < \epsilon]} \right\}.$$

Thus,

$$\sup_{p\,q\,|p-q|\le\epsilon} \left| \frac{\mathbf{E}V_n(p)}{n} - \frac{\mathbf{E}V_n(q)}{n} \right| \le \frac{1}{n}\mathbf{E}\left\{ \max_{P\in\mathcal{P}_n} \sum_{e\in P} I_{[U_e \le \epsilon]} \right\} \le \varphi(\epsilon) + o(1)$$

by Lemma 11. Recall that $\varphi(\epsilon) \to 0$ as $\epsilon \to 0$. This implies directly the uniform continuity of $\mathbf{E}V_n(p)/n$ in p. This also implies the uniform continuity of $\mathcal{V}(p)$. □

7. A law of large numbers for $p \in (\alpha, 1 - \alpha)$.

LEMMA 13. *For $p \in (\alpha, 1 - \alpha)$, we have*

$$0 < \mathcal{V}(p) < 1$$

where

$$\mathcal{V}(p) = \lim_{n \to \infty} \frac{\mathbf{E} V_n}{n},$$

and $V_n / \mathbf{E} V_n \to 1$ almost surely as n tends to infinity.

PROOF. Let u be a node of T an n-level b-ary SUM tree with parameter p. We associate with u the value $V'(u)$ related to $V(u)$ in the following manner: $V'(u) \le V(u)$. 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 the results about Pearl trees. The exact definition is given below. We denote by V_n' the random variable defined as the value $V'(u)$ when u is the root of the n-level model. We will show that $\liminf_{n \to \infty} \mathbf{E} V_n'/n > 0$.

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

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

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

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

$$V'(u) = \begin{cases} \max_{v \in A_u} \{V'(v)\} & \text{if } u \text{ at even level in } T \\ \min_{v \in A_u} \{V'(v)\} & \text{if } u \text{ at odd level in } T \end{cases}. \tag{20}$$

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

$$\mathbf{P}\{V_{lN}' = -\infty\} \le \min\left(\varepsilon, \frac{1}{2} b^{-b}\right). \tag{21}$$

For $l = 1$, this is true since

$$\mathbf{P}\{V_N' = -\infty\} = \mathbf{P}\{W_N = 0\} = F_N(0)$$

where F_N is the distribution function of the value of the root of an N-level SUM tree with parameter p, and $W_N = W(u)$ is the value of the root of an N-level tree defined above (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 (21) 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 $q \le (1/2)b^{-b}$. Thus by Lemma 1 we have,

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

For all $m \ge 0$,

$$\mathbf{P}\{V_{2lN+m} < l\} \le \min\left(\varepsilon, \frac{1}{2}b^{-b}\right). \tag{22}$$

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 $b^N b2^{-b^{\lfloor N/2 \rfloor}}$. If the b^N nodes at level $(2l-2)N$ have the value $l-1$, then $\mathbf{P}\{V'(u) = -\infty\} = \mathbf{P}\{W_N = 0\} = F_N(0)$ Thus if we choose N such that $b^{N+1}2^{-b^{\lfloor N/2 \rfloor}} \le \frac{1}{2}\min(\varepsilon, \frac{1}{2}b^{-b})$,

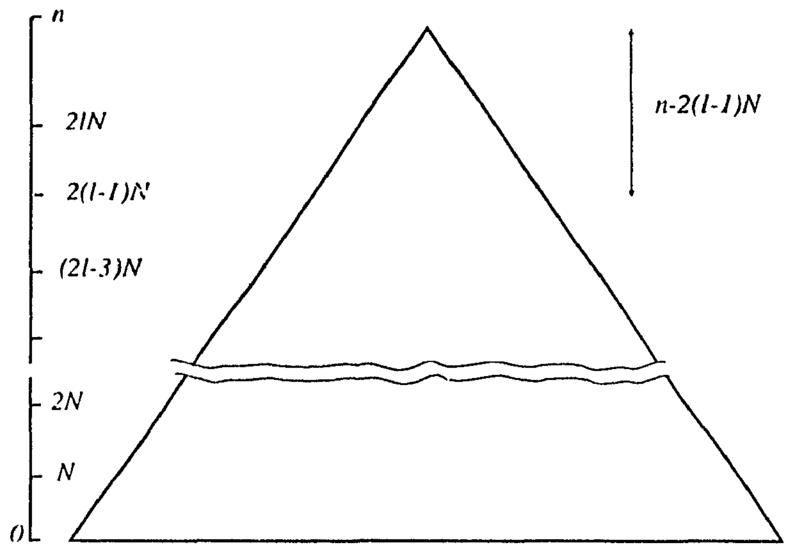
$$\begin{aligned}
\mathbf{P}\{V'(u) = -\infty\} &\le b^{N+1}2^{-b^{\lfloor N/2 \rfloor}} + F_N(0) \\
&< \frac{1}{2}\min\left(\varepsilon, \frac{1}{2}b^{-b}\right) + \frac{1}{2}\min\left(\varepsilon, \frac{1}{2}b^{-b}\right) \\
&= \min\left(\varepsilon, \frac{1}{2}b^{-b}\right).
\end{aligned}$$

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

$$\mathbf{P}\{V'_{lN} = -\infty\} \le \varepsilon.$$

Thus for all $l > 0$,

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



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

We now generalize the result for SUM tree with a number of levels that is not a multiple of N. Let u be the root of T, an n-level b-ary SUM tree with parameter $p > \alpha$. Let $l = \lfloor n/2N \rfloor$. Using (22), with $m = n - 2lN$, we have

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

Thus,

$$\mathbf{E}V_n \geq l(1 - \varepsilon).$$

As a consequence,

$$\begin{aligned}\frac{\mathbf{E}V_n}{n} &\geq (1-\varepsilon)\left(\frac{n}{2N} - 1\right)\frac{1}{n} \\ &= \frac{1-\varepsilon}{2N} - \frac{1-\varepsilon}{n} \\ &= \frac{1-\varepsilon}{2N} - o(1)\end{aligned}$$

so that

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

and finally

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

We can similarly prove that for $p < 1 - \alpha$,

$$\lim_{n\to\infty} \frac{\mathbf{E}V_n}{n} < 1.$$

Thus for all $p > \alpha$ there exist $c > 0$ and $n_0 > 1$ such that for all $n > n_0$, $\mathbf{E}V_n > cn$. According to Lemma 9, for any p, and any $\varepsilon > 0$,

$$\mathbf{P}\left\{|V_n - \mathbf{E}V_n| \geq \varepsilon\right\} \leq 2e^{-2\varepsilon^2/n}.$$

Thus for $n > n_0$,

$$\begin{aligned}\mathbf{P}\left\{\left|\frac{V_n}{\mathbf{E}V_n} - 1\right| \geq \varepsilon\right\} &= \mathbf{P}\left\{|V_n - \mathbf{E}V_n| \geq \varepsilon\mathbf{E}V_n\right\} \\ &\leq \mathbf{P}\left\{|V_n - \mathbf{E}V_n| \geq \varepsilon cn\right\} \\ &\leq 2^{-2\varepsilon^2c^2n}.\end{aligned}$$

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

8. Symmetry.

So far we proved that the $\mathbf{E}V_n/n$ converges to a continuous function of p. Now we study some characteristics of this function.

THEOREM 4. *For all $p \in [0, 1]$ we have*

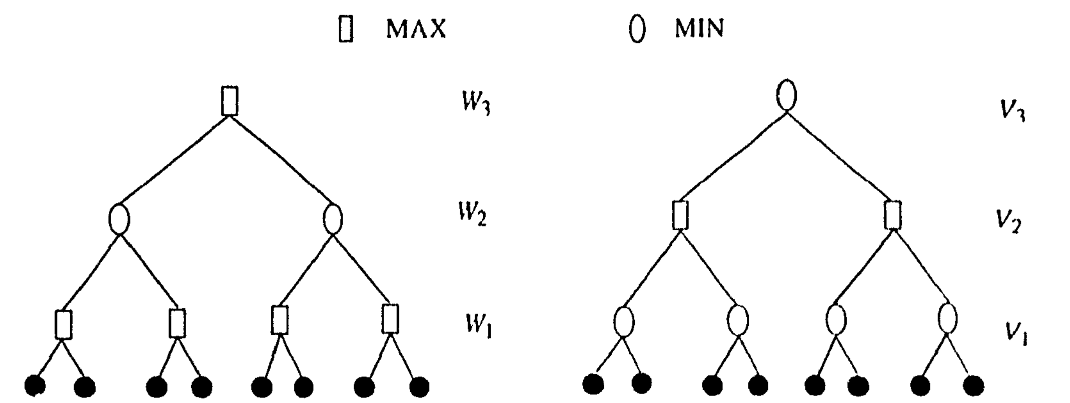
$$\mathcal{V}(1 - p) = 1 - \mathcal{V}(p),$$

where

$$\mathcal{V}(p) = \lim_{n\to\infty} \frac{\mathbf{E}V_n}{n}.$$

This implies that

$$\mathcal{V}(1/2) = 1/2$$



The random variables V_n and W_n

PROOF. Consider the random variables W_n defined recursively as follows:

$$W_0 = 0;$$
$$W_n \stackrel{\mathcal{L}}{=} \begin{cases} \min_{1\le j\le b}\{W_{n-1,j} + X_j\} & \text{if n is even} \\ \max_{1\le j\le b}\{W_{n-1,j} + X_j\} & \text{if n is odd} \end{cases},$$

where all $W_{n-1,j}$ are independent and identically distributed as W_{n-1}. Let V'_n and W'_n be defined as V_n and W_n respectively but with all edge values flipped (0 to 1 and 1 to 0). We use the definition of V'' used at the end of chapter III, $V''_n \stackrel{\mathcal{L}}{=} n - V_n$ and $Y_j = 1 - X_j$. Then we have

$$V''_n \stackrel{\mathcal{L}}{=} \begin{cases} \min_{1\le j\le b}\left\{V''_{n-1,j} + Y_j\right\} & \text{if } n \text{ is even} \\ \max_{1\le j\le b}\left\{V''_{n-1,j} + Y_j\right\} & \text{if } n \text{ is odd} \end{cases}.$$

Hence, $V''_n \stackrel{\mathcal{L}}{=} W'_n$ and $n - V_n \stackrel{\mathcal{L}}{=} W'_n$. By definition of the W'_n's it is also easy to see that

$$W'_{n+1} \stackrel{\mathcal{L}}{=} V'_n + \theta_n$$

for some $\{0, 1\}$-valued random variable θ_n. In fact the tree of W'_{n+1} can be created from V'_n by adding a MAX level at the bottom level, so this will influence the root's value by 0 or 1. Therefore,

$$\mathrm{E}V'_n \le \mathrm{E}W'_{n+1} \le \mathrm{E}V'_n + 1 .$$

Recall also that $\mathrm{E}W'_n = n - \mathrm{E}V_n$. Thus,

$$\frac{\mathrm{E}V'_n}{n} \le 1 + \frac{1}{n} + \frac{\mathrm{E}V_{n+1}}{n} \le \frac{\mathrm{E}V'_n}{n} + \frac{1}{n}.$$

Since

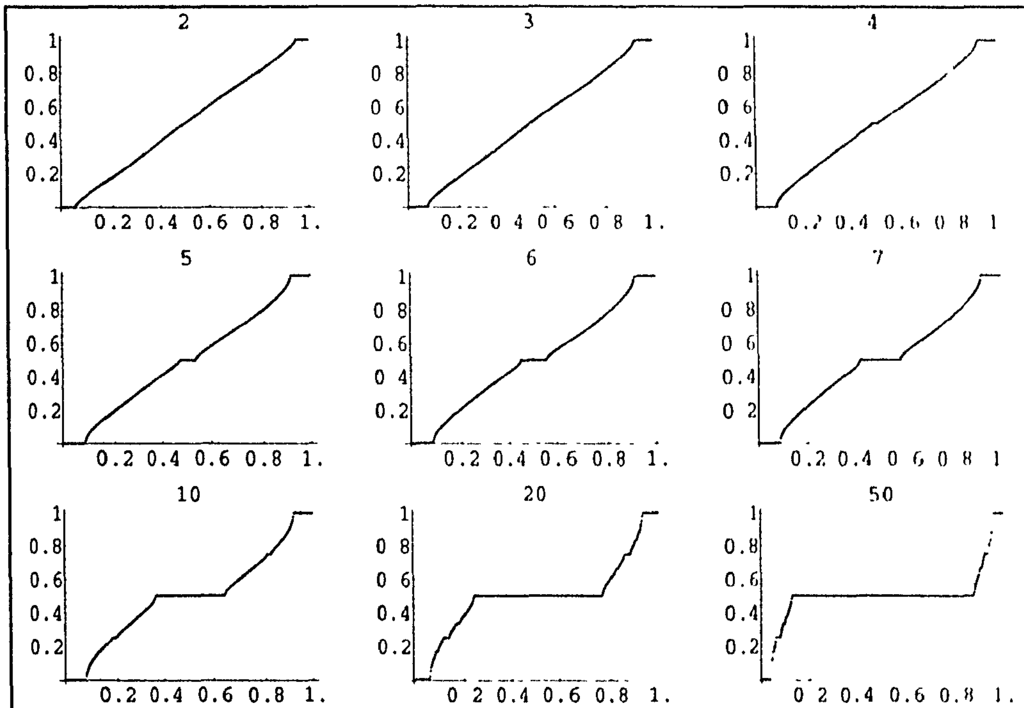
$$\lim_{n\to\infty} \frac{\mathrm{E}V_n}{n} = \mathcal{V}(p), \text{ and } \lim_{n\to\infty} \frac{\mathrm{E}V'_n}{n} = \mathcal{V}(1-p),$$

we have

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

9. Behavior of the expected value around $p=1/2$.

The figure below shows the behavior of $EV_{2n}/2n$ as a function of p for different values of b. We notice the flat part around $p = 1/2$ which grows with b and tends to fill the entire range of p when b tends to infinity. This means that in this range of p the expected value almost does not depend on p. This section includes a proof of this behavior. The idea is to prove that when the number of children becomes big the MAX nodes will almost always add a one to the value and the MIN nodes almost always a zero. Thus, for a $2n$ depth tree the expected value is close to n. The computations show that this behavior appears for $b > 2$.



$EV_{2n}/2n$ function of p for different b.

THEOREM 5. *For all b there exists $\beta \in (0, 1/2]$ such that*

$$\lim_{n\to\infty} \frac{\mathbf{E}V_n}{n} \begin{cases} = \frac{1}{2} & \text{if } p \in [\beta, 1-\beta] \\ < \frac{1}{2} & \text{if } p \in [0, \beta) \\ > \frac{1}{2} & \text{if } p \in (1-\beta, 1] \end{cases} .$$

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 $[0, 1]$. Assume that $L(p, b) \neq 0$. Then, for $n \geq \sqrt{2/L^b(p, b)}$,

$$\left\lfloor \frac{n}{2} \right\rfloor - \sqrt{n \log n} \leq \mathbf{E}V_n \leq \left\lfloor \frac{n}{2} \right\rfloor + \sqrt{n \log n},$$

$\beta \geq p$, 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 $0 < \beta < 1/2$ (hence, the flat part exists).

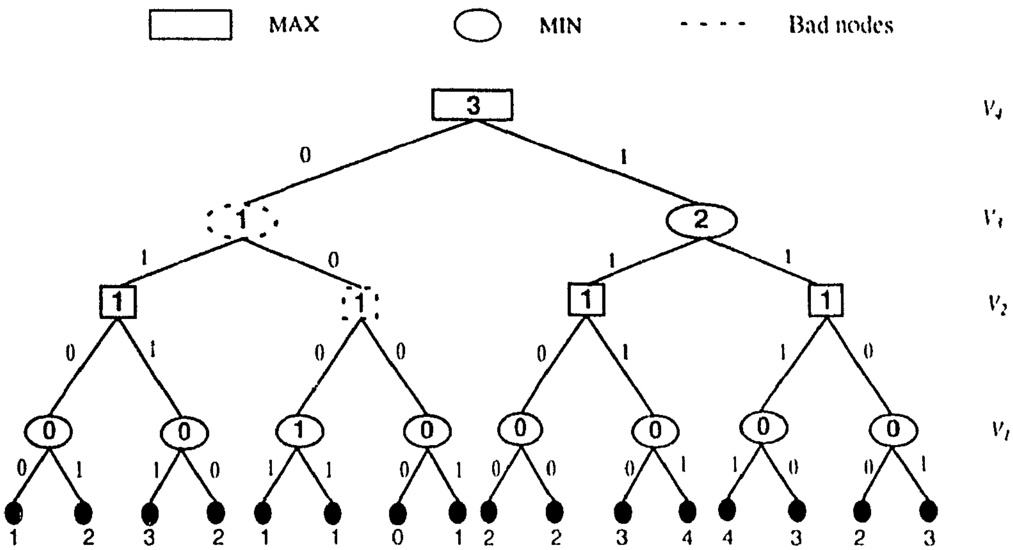
PROOF.

Part 1: The lower bound.

LEMMA 14. *Assume that $L(p, b) \neq 0$ where $L(p, b)$ be the largest root of $1 - (1 - px^b)^b = x$. Then, for $n \geq \sqrt{2/L^b(p, b)}$,*

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

PROOF. We consider a random $2n$-level b-ary SUM 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 bottom level where all the MAX nodes bring a one.



Tree with good and bad nodes.

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

$$\mathrm{P}\{V_{2n} \geq n\} \geq p_n .$$

Furthermore, we have a simple recursion:

$$p_0 = 1 ,$$

and

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

where

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

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:

$$\inf_n \mathrm{P}\{V_{2n} \geq n\} \geq L(p, b),$$
$$\inf_n \mathrm{P}\{V_{2n+1} \geq n\} \geq L^b(p, b).$$

Therefore,

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

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

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

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

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

Part 2: The upper bound.

In the b-ary tree we mark the node good or bad according to the following rules. The leaves are all good. A MAX node is marked good if all the children are good. A MIN node is marked good if there exists at least one good child whose edge is zero. Thus, the root is marked good if there exists a path from the root to the bottom where all the MIN nodes bring a zero. If a node at height $2n$ is good we have $V_{2n} \leq n$. Also $V_{2n+1} \leq n$. Let q_n denote the probability that a node at level n is good. We set $q = 1 - p$. Then

$$\mathbf{P}\{V_{2n} \leq n\} \geq q_n$$

And the recursion is

$$q_0 = q_1 = 1,$$
$$q_{2n+1} = f(q_{2n-1}),$$

with

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

Just as for the lower bound, we get the inequality

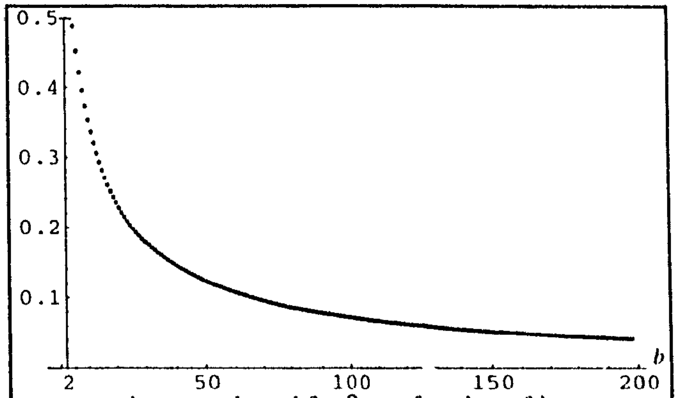
$$\mathbf{E}V_n \leq \left\lfloor \frac{n}{2} \right\rfloor + \sqrt{n \log n}. \square$$

Part 3: Behavior when b tends to infinity.

LEMMA 15. *For $b \geq 8$, $\beta < 1/2$. Also,*

$$\lim_{b \to \infty} \beta = 0.$$

Thus the flat part exists and tends to the full range as $b \to \infty$.



An upper bound for β as a function of b.

PROOF. We show that for all $p \in (0, 1/2]$, there exists B such that $L(p, b) > 0$ for all $b > B$. This implies $\beta \leq p$. Define $g(x) = f(x) - x$. For $x \in (0, 1/2]$, we have

$$g(x) = 1 - x - (1 - px^b)^b.$$

Taking $x = \sqrt[b]{\frac{2}{3}}$, we see that

$$\begin{aligned} g\left(\sqrt[b]{\frac{2}{3}}\right) &= 1 - e^{\frac{1}{b}\log\frac{2}{3}} - \left(1 - \frac{2p}{3}\right)^b \\ &\underset{b \to \infty}{\sim} \frac{\log(3/2)}{b} - \left(1 - \frac{2p}{3}\right)^b \\ &\underset{b \to \infty}{\sim} \frac{\log(3/2)}{b}. \end{aligned}$$

There exists B such that for all $b > B$,

$$g\left(\sqrt[b]{\frac{2}{3}}\right) > 0.$$

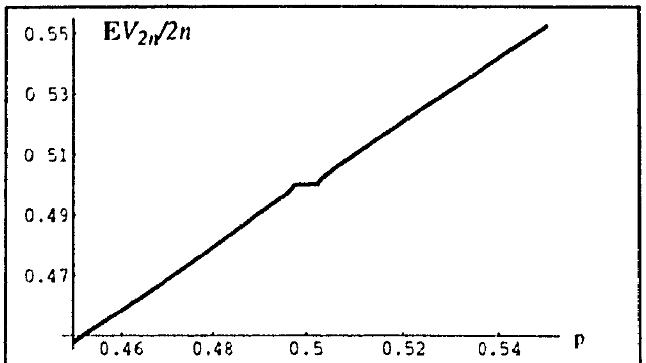
Note also that for any b,

$$g(1) \leq 1, \; g(0) = 0, \; g'(0) = -1.$$

Thus for $b > B$, g has a fixed point in $[\sqrt[b]{2/3}, 1)$, and $\beta \leq p$. Finally for $p = 1/2$ and $b = 8$, we have

$$g\left(\sqrt[b]{\frac{2}{3}}\right) > 0.$$

Thus for $b = 8$, $\beta < 1/2$. Then Lemma 15 is proved. The figure above shows an upper bound of β computed as the smallest p such that $L(p, b) > 0$. The numerical computations show that the flat part appears for $b = 3$. See also the figure below. □



$EV_{2n}/2n$ for b=3 and p between .45 and .55 and n=500.

V

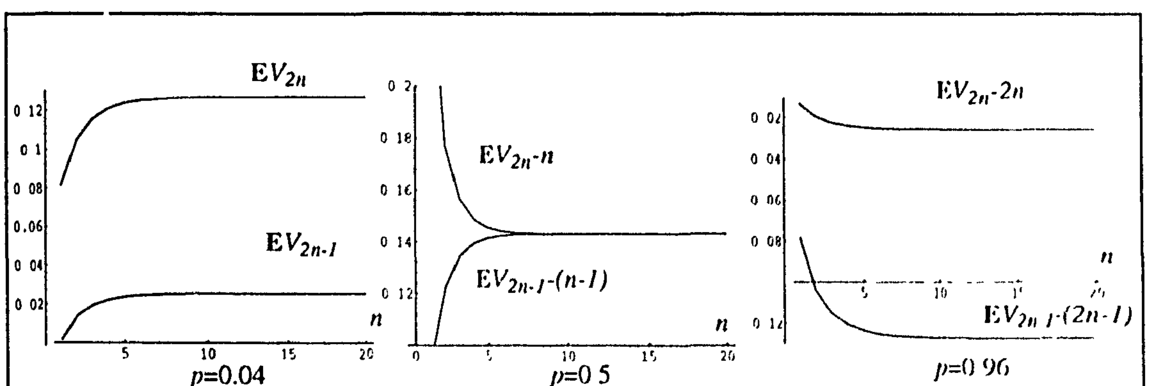
NUMERICAL RESULTS AND DISCUSSION

Simulation of MIN-MAX trees does not provide reliable results, as the number of computation needed for such simulation is exponential in the depth of the tree. In fact, simulations can only be performed for depths up to 20. We are therefore forced to use a numerical method based on the recurrences given in the text We can compute the various distribution functions in polynomial time as a function of the depth. The results are reliable and the convergence is fast These functions have been computed with a Pascal program. We used *Mathematica* to analyze the data. The distribution function and the expected value of V_n are computed for trees up to 2000 levels for 500 different values of p These numerical results give a good illustration of the theoretical properties of the SUM model proved in the preceding parts and allow some conjectures on other properties In this chapter we use the notation introduced before:

- V_n: the root's value of the n-level b-ary SUM model with parameter p,
- F_n: distribution function of V_n and f_n its discrete density,
- $\mathcal{V}(p)$: the limit of $\mathrm{E}V_n/n$ as n tends to infinity,
- α: the smallest p such that $F_n(0)$ has a non-zero limit as n tends to infinity

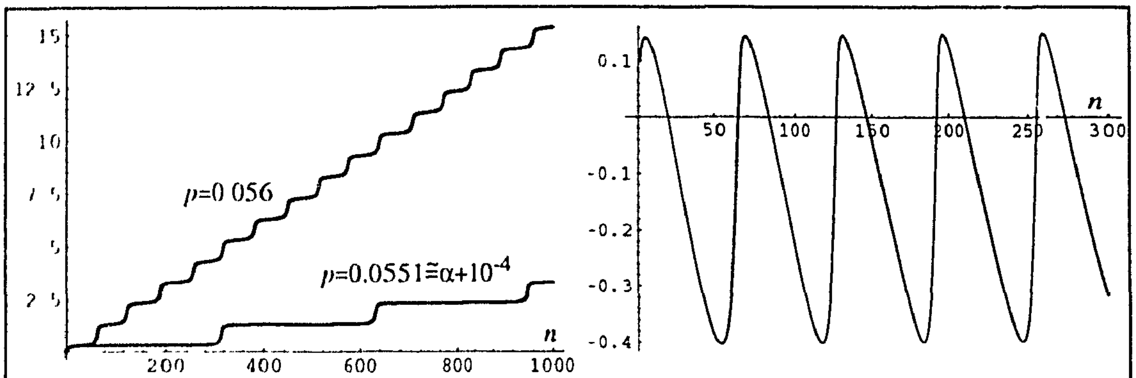
1. Expected value for $b = 2$.

First we present the evolution of the expected value $\mathrm{E}V_n/n$ as a function of the depth n for fixed p. We have three different cases as a function of p, depending upon whether $p \in [0, \alpha]$, $p \in (\alpha, 1-\alpha)$ or $p \in [1-\alpha, 1]$.



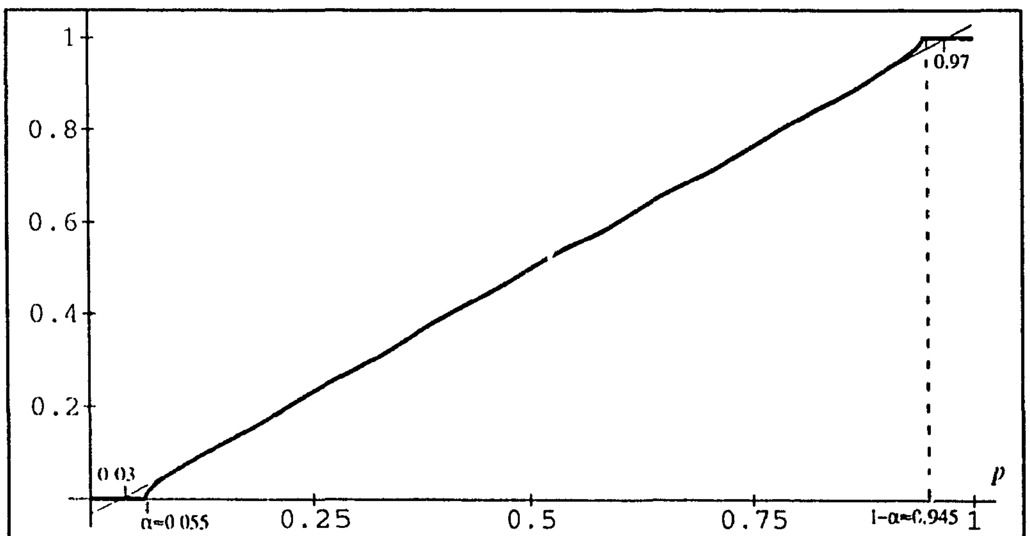
The three different cases.

For $p = 1/2$ we can observe that $\mathbf{E}V_{2n} - n$ converges to a finite limit. For general p we observe that $\mathbf{E}V_n - n\mathcal{V}(p)$ stays finite when n tends to infinity. When p slightly bigger than α, $\mathbf{E}V_n - n\mathcal{V}(p)$ oscillates as we can see in the next figure.



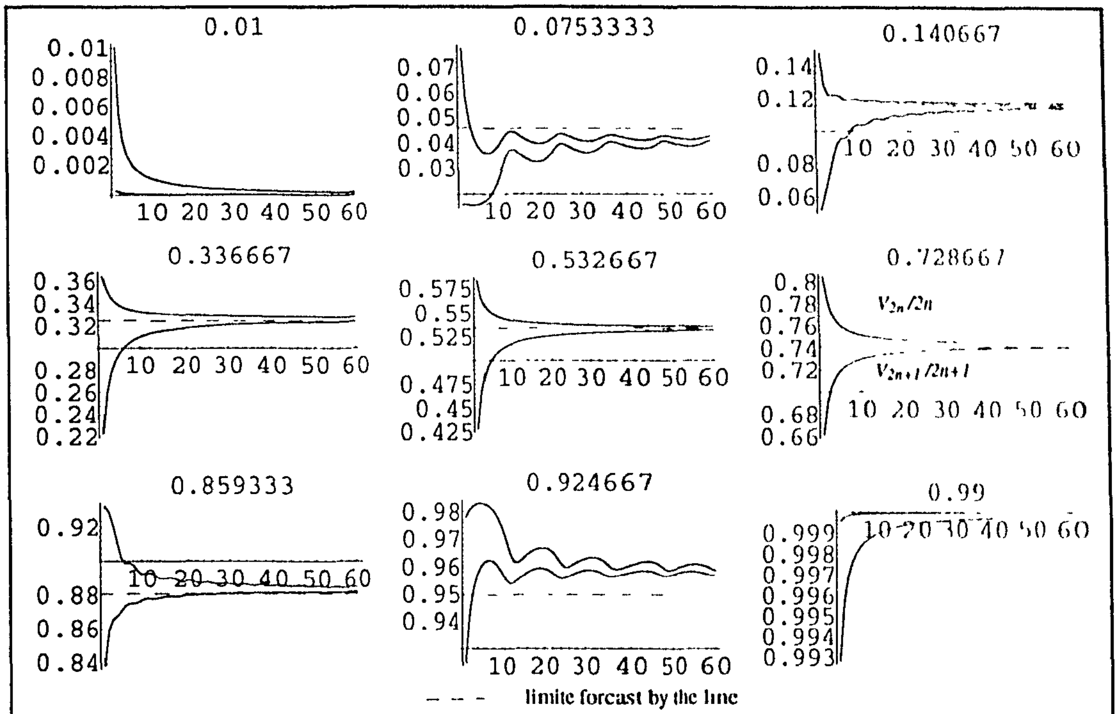
$\mathbf{E}V_{2n}$ for $p>\alpha$ close to α and $b=2$ $\quad$ $\mathbf{E}V_{2n}-nV(p)$ for $p=0\ 056>\alpha$

The conjecture that $\mathbf{E}V_n - n\mathcal{V}(p)$ is finite has not be proved. Instead, we only proved that $\mathbf{E}V_n/n$ has a limit $\mathcal{V}(p)$ when n tends to infinity and that this limit is continuous. The figure below shows a numerical approximation of the function $\mathcal{V}(p)$ computed for $n = 2000$. The function has breakpoints at $p = \alpha$ and $p = 1 - \alpha$ as expected. At these points the function looks non differentiable but it is continuous. For $p \in [0, \alpha]$ the function is equal to zero, and for $p \in [1 - \alpha, 1]$, it is equal to 1. Between the two breakpoints the function is close to a line with a slope of 1.062. The slope of the line joining the points $(0.3, \mathcal{V}(0.3))$ and $(0.7, \mathcal{V}(0.7))$ is a little bit over one. Also, as $\mathcal{V}(1/2) = 1/2$, we see that $\mathbf{E}V_n < pn$ for $p < 1/2$, $\mathbf{E}V_n > pn$ for $p > 1/2$, when n is large enough.



$\mathbf{E}V_{2000}/2000$ as a function of p (500 points), slope of the line: 1.062.

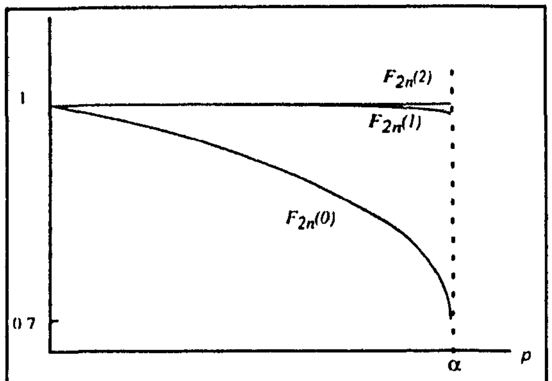
The next figure shows the convergence of EV_n/n for different p. We notice that the convergence is not monotone when p is near α or $1-\alpha$.



$EV_{2n}/2n$ and $EV_{2n+1}/2n+1$ function of n, for 9 values of p between 0.01 and 0.99, b=2.

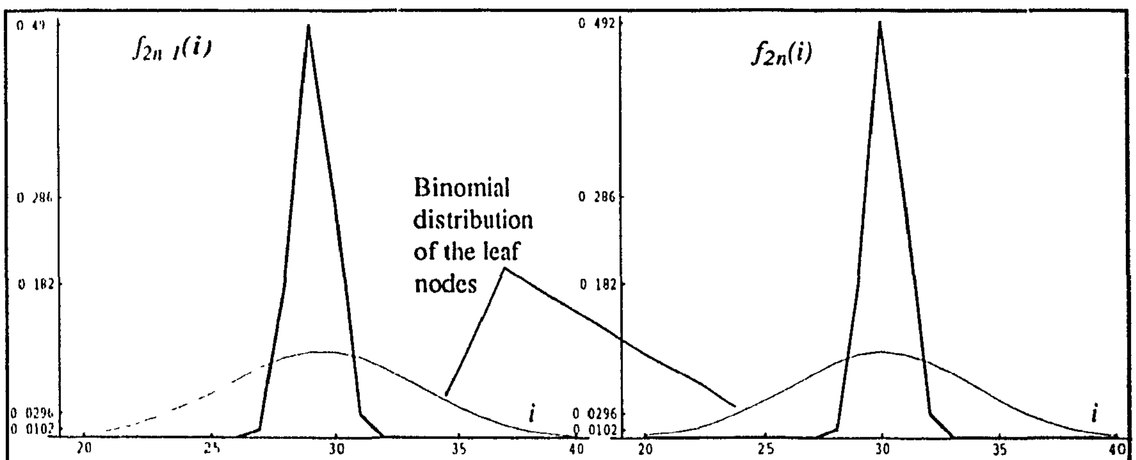
2. Distribution functions for b=2.

First we consider the case $p \in [0,\alpha]$ ($\alpha \cong 0.055$). As we proved, for all i, $F_{2n}(i)$ converges to a non-zero limit $F_\infty(i)$. Since $F_\infty(0)$ is the root of a fourth-order polynomial, one could compute (e.g., with *Mathematica*) its symbolic formulation as a function of p. Then via recurrences we may obtain a symbolic formulation of all $F_\infty(i)$. This has been done and the graphs are shown in the figure below. $F_\infty(0)$ is smallest when $p=\alpha$ in which case $F_\infty(0) = \sqrt[3]{256}/9 \cong 0.705512$. This implies that if $p < \alpha$, the probability of having zero as root's value is greater than 70% even when the expected value of the leaves tends to infinity (the expected value of the leaves is np). The probability that the root's value is greater than one is less than 1.4%, and it is virtually impossible that its value is greater than two.



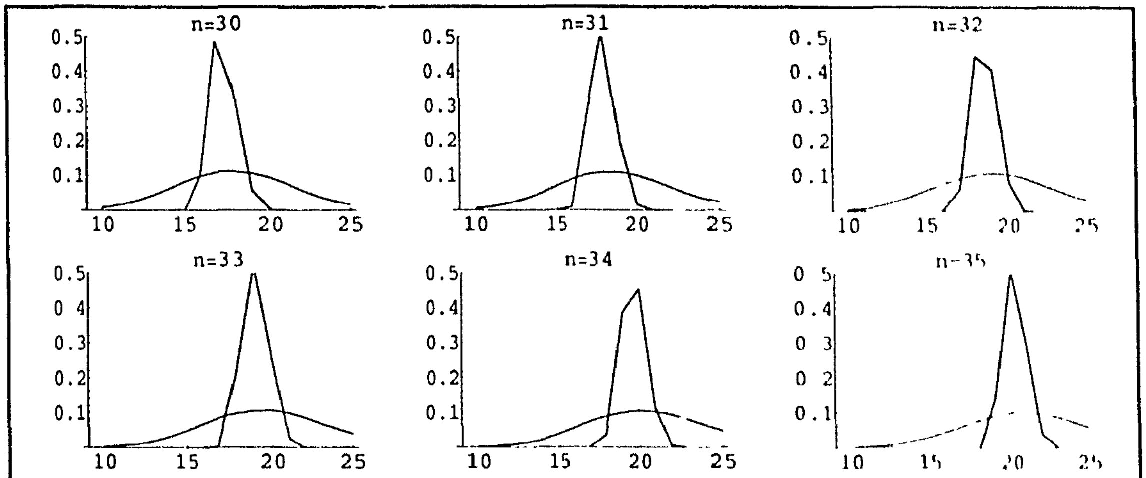
Limit of $F_\infty(0)$, $F_\infty(1)$, $F_\infty(2)$, $b=2$

We now look at the case $p = 1/2$. $F_n(i)$ tends to zero for all i. We observed that $f_{2n}(n+i)$ and $f_{2n-1}(n+i)$ have non-zero limits as n tends to infinity. The figures below show the estimated limits of $f_{2n}(n+i)$ and $f_{2n-1}(n+i)$. The distributions are highly concentrated: V_{2n} and V_{2n+1} are very likely in $[n-2, n+2]$ while $\mathrm{P}\{V_{2n} = n\} > 45\%$. Note also that $f_{2n-1}(n+i)$ looks similar to $f_{2n}(n+i-1)$. With the graphs we can compare these distributions with the distributions of the leaf values, which are much less concentrated. Furthermore we observe that $f_{2n}(n+i)$ has a limit for all i, and that the variance of V_n stays finite when n tends to infinity.



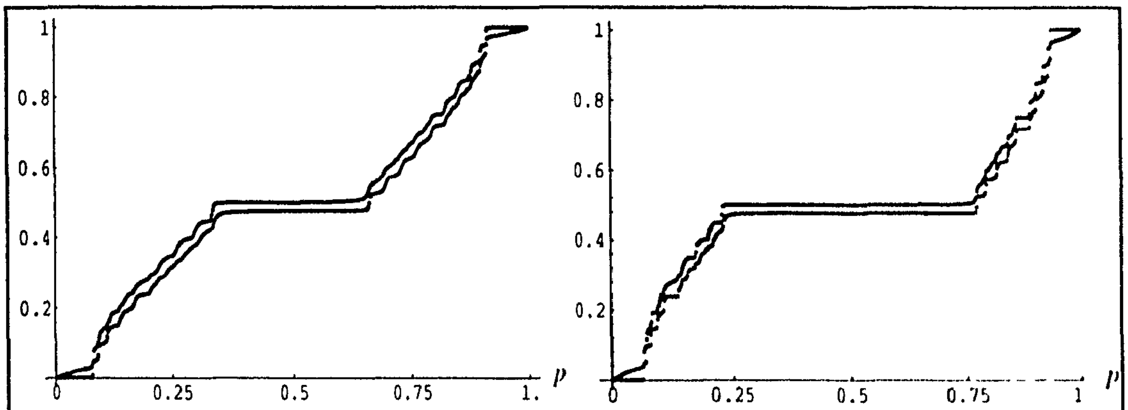
$f_{2n}(i)$ and $f_{2n-1}(i)$, for $n=30$ and $p=0.5$.

For general $p > \alpha$, the distribution of V_{2n} is concentrated around the value $2n\mathcal{V}(p)$. The figure below shows density functions of the leaf values and of the root's value for $p = 0.3$. The variance of V_n seems to stay finite when n tends to infinity. Since $\mathcal{V}(p)$ is close to p, the distributions of the root's value and of the leaf values look concentrated around the same values. Note however that $\mathcal{V}(p) \neq p$ at all p except $p = 1/2$, $p = 0$ and $p = 1$.



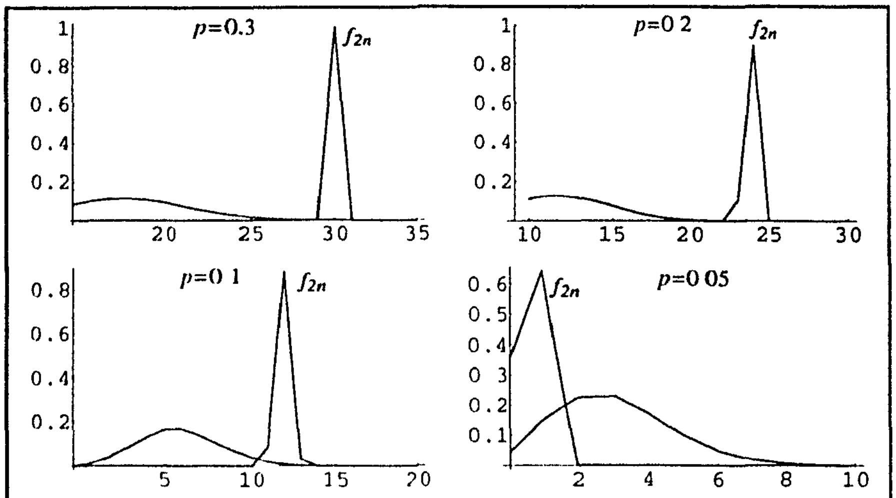
$f_{2n}(i)$, for p=0.3 and n from 30 to 35.

3. Results for b>2.



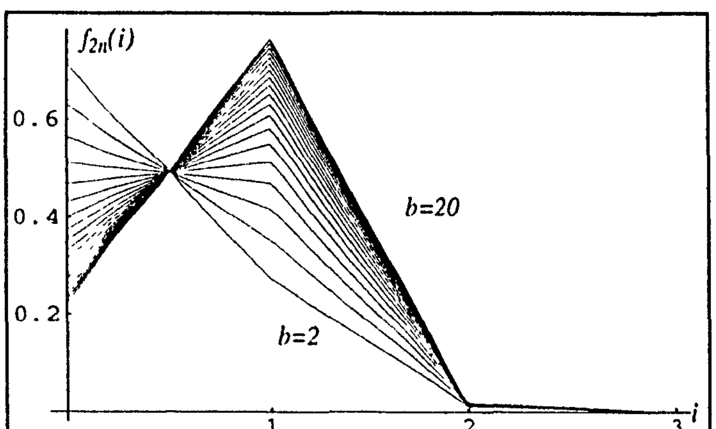
$EV_{20}/20$ and $EV_{21}/21$, for b=10 $EV_{20}/20$ and $EV_{21}/21$, for b=20

Numerical results about the flat part around $p = 1/2$ have already been presented in the preceding chapter (page 44). One could argue that this behavior appears only for huge trees and that it is not reliable for real game trees. This is not the fact. As the convergence is very fast, this behavior appears quickly for small tree as we can see in the above figure.



f_{2n} and the density of the leave value, $2n$=60 and b=20.

The figure above presents density functions of V_{60} for b = 20 and different p. The first observation is that the distribution of the root's value and of the leaf values are concentrated around totally different values, and that the distribution of the root's value is always more concentrated than the leaf's. The first graph shows the density for p = 0.3 (this p is inside the flat part around 1/2). We proved that in this case $EV_{2n}/n \to 1$. In the graph we see that n is almost the only value likely for V_{2n}: $P\{V_{2n} = 30\} > 0.999$, while the probability for a leaf value to be 30 is less than 0.6%. The second and third graphs show the density functions for two values of p outside the flat part These densities are also concentrated around the value $2nV(p)$. The last graph considers a p smaller than α. We can see that the most likely value is 1. In order to have an idea of the distributions for small p we look at the distributions when $p = \alpha$ for different b's. The figure below shows these distribution functions.



The density functions $f_{2n}(i)$ of V_{2n} with n=60 and p=α.

The table below gives the first four values of the density function of V_{2n} when $p = \alpha$ and $n = 60$. It is noteworthy that for b less than five the most likely value is zero and for bigger b it is 1.

	α	f(0)	f(1)	f(2)	f(3)	$F V_{60}$
b=2	0 0551646	0 713362	0 273474	0 0130747	8 89531E 05	0 29989025 9
b=3	0 0755082	0 631249	0 352287	0 0164461	1 80502E 05	0 38523351
b=4	0 083527	0 565241	0 416845	0 0179113	2 57673E 06	0 45267533
b=5	0 0863455	0 511789	0 469898	0 0183134	3 04901E 07	0 50652715
b=6	0 0867173	0 46824	0 513625	0 0181354	3 14547E 08	0 54989589 4
b=7	0 085874	0 431733	0 550518	0 0177491	2 94628E 09	0 58601620 9
b=8	0 0844099	0 40108	0 581711	0 0172083	2 51707E 10	0 61612760 1
b=9	0 0826402	0 374655	0 608684	0 0166615	2 00739E 11	0 642007
b=10	0 0807332	0 351801	0 632101	0 0,6098	1 49547E 12	0 661297
b=11	0 0787843	0 331882	0 652591	0 0155274	1 05027E 13	0 6836458
b=12	0 0768488	0 314287	0 670733	0 0149799	7 99361E 15	0 7006928
b=13	0 0749957	0 298644	0 686901	0 014455	0	0 715811
b=14	0 0731283	0 284391	0 701591	0 0140146	0	0 7296232
b=15	0 0713662	0 271809	0 714659	0 0135318	0	0 7417226
b=16	0 0696786	0 260233	0 726648	0 013119	0	0 752886
b=17	0 0680634	0 249773	0 73752	0 0127071	0	0 7629312
b=18	0 0665206	0 240137	0 747528	0 0123348	0	0 7721976
b=19	0 0650465	0 231383	0 756661	0 0 9566	0	0 7805742
b=20	0 0636392	0 223285	0 76511	0 0116053	0	0 7883206
b=21	0 0622948	0 215829	0 772909	0 0112622	0	0 7951331
b=22	0 0610111	0 208769	0 780256	0 0109752	0	0 8022064
b=23	0 0597852	0 202098	0 787172	0 0107305	0	0 808633
b=24	0 0586105	0 196082	0 79348	0 0104383	0	0 8113566
b=25	0 0574854	0 190524	0 799336	0 0101394	0	0 8196118
b=26	0 0564099	0 185013	0 805046	0 00994096	0	0 8249 2702
b=27	0 0553763	0 180135	0 810201	0 00966346	0	0 8295 2702
b=28	0 0543866	0 17528	0 815245	0 0094751	0	0 8341952
b=29	0 053435	0 170848	0 819901	0 00925083	0	0 83810266
b=30	0 0525216	0 166517	0 824404	0 009078 66	0	0 84256132

Density function $f_{2n}(i)$, for $p=\alpha$, and $n=60$

4. Some remarks.

The existence of the flat parts has some feasible explanation. Let us recall the meaning of the MIN-MAX tree as a game tree, using the bounded look-ahead strategy presented in chapter II. In such a tree with n levels, a leaf node represents the value of a position computed by a static evaluation function. If we assume that the game follows the SUM model, the difference between the efficiencies of the evaluation functions of both players influences the value of p. In fact if your function is better than your opponent's then p is greater than 1/2. According to the previous results, if the number of children is big, the variations of p do not have a lot of influence if it stays inside the flat part around 1/2: the efficiency of the evaluation is dampened. Thus if two perfect players or computers can search down to the same depth and if the efficiency of their evaluation functions are different, they are equally likely to win.

It is interesting to see what happens if we consider the standard probabilistic model with a binomial distribution for the leaf nodes in order to compare the results with the SUM model. Let V be the random variable which defines the leaf values, and assume that,

$$V \stackrel{\text{def}}{=} \frac{Y_{2n}}{2n},$$

where Y_{2n} follows a binomial distribution with $2n$ trials and success probability p. We have [CL91,p.19],

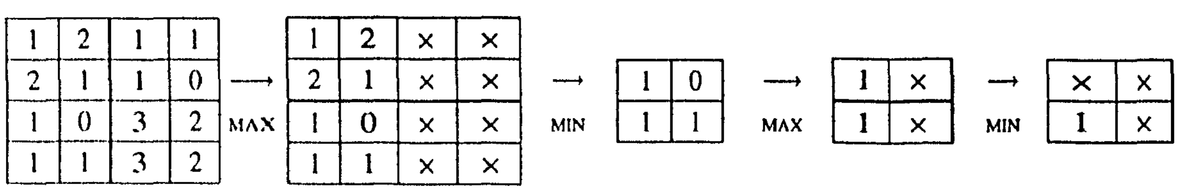
$$P\left\{|V-p| \geq \sqrt{\frac{\log n}{n}}\right\} \leq \frac{1}{\sqrt{n}}.$$

This implies that when n tends to infinity, the distribution function of V tends to a function with a jump at p from 0 to 1. Thus according to chapter 2 the expected value of the root tends to p. It can appear logical that when the leaf values are concentrated around p the expected value of the root is this value. With the SUM model we are far from this behavior. For example, the probability to have zero at a fixed leaf of a n-level SUM tree with parameter $p > 0$ is $(1-p)^n$. This clearly tends to zero when n tends to infinity. In spite of this, the probability of having zero at the root tends to a non-zero limit if $p \leq \alpha$. For example, for a 120-level SUM tree with $b = 2$ and $p = \alpha \approx 0.055$, the probability for a leaf value to be zero is less than 0.2% and the probability for the root's value to be zero is more than 71%.

5. An example: The board-splitting game.

The hypotheses of the SUM model can appear quite far from real games as we assume the game to end after exactly $2n$ moves and that for each move there are exactly b choices. Judea Pearl invented a class of games that matches these hypotheses for the standard probabilistic model. This game is called the board-splitting game or P-game. Nau [Nau82] used this kind of game with dependent nodes

The game consists of an $N \times N$ board ($N = b^n$) and each cell contains an integer between zero and $2n$. We call the two players MAX and MIN. A move for MAX consists of cutting the board in b vertical parts and keeping only one of these. MIN does the same but horizontally. The goal of MAX is that the last number after $2n$ turns is the greater, and MIN has the opposite goal. If the value in the last cell is $x \in [0, 2n]$, MAX wins $n - x$ points and MIN wins $x - n$ points. The figure below shows an example.



An illustrative play.

With Pearl's example the cell values are independent and have the same distribution. In our model the cell values are dependent according to the SUM model. Let B_{2n} be a $b^n \times b^n$ board for such a game when MAX plays the next turn. Let B_{2n-1} be a $b^{n-1} \times b^n$ board when MIN plays the next turn.

The recurrence to build such boards is

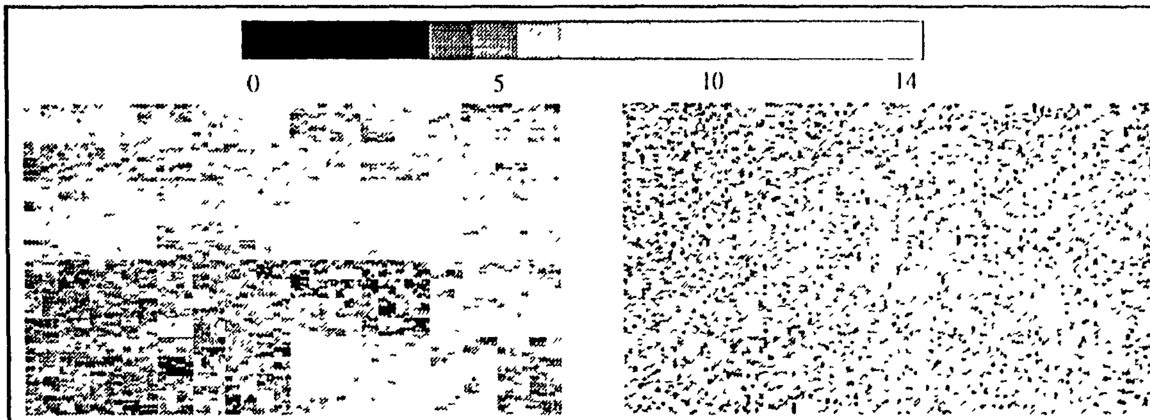
$$\mathcal{B}_0(1,1) = 0\,;$$
$$\mathcal{B}_{2n}(i,j) = \mathcal{B}_{2n-1}(\lceil i/b \rceil, j) + x, \text{ for } 1 \le (i,j) \le b^n\,;$$
$$\mathcal{B}_{2n-1}(i,j) = \mathcal{B}_{2n-2}(i, \lceil j/b \rceil) + x, \text{ for } 1 \le i \le b^{n-1},\ 1 \le j \le b^n,$$

where x is an independent drawing of a Bernoulli variable with mean p. The table below shows one example of such board for $b = 2, p = 1/2$ and $n = 3$.

1	3	3	3	1	2	3	3
2	2	3	4	1	2	2	2
3	3	2	3	3	2	1	2
2	3	2	3	2	1	2	2
3	3	1	1	3	5	4	4
3	2	1	0	3	4	4	3
2	2	2	2	4	4	2	2
2	3	1	2	3	4	2	3

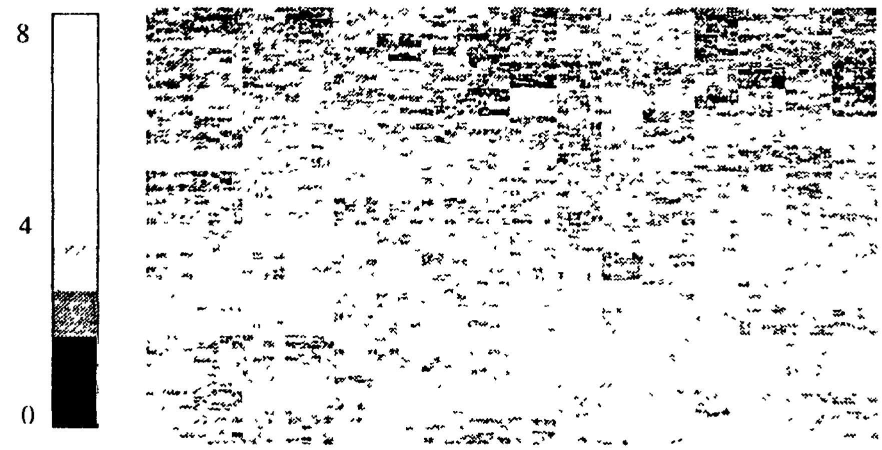
A board for the SUM model, $p = 1/2, b = 2, n = 3$.

We can simulate bigger random boards with *Mathematica*. For $b = 2$ we took $n = 7$, i.e., a 128×128 board. For $b = 4$ we took $n = 4$, i.e., a 256×256 board. To represent these boards we represent the cell values by a gray level, the correspondences between the values and the gray levels being given in the insets. The first figure shows a simulation for $b = 2$. It is interesting to compare this board with an equivalent board simulated with independent values corresponding to the standard model, which independent cell values, each binomial with mean $p = .5$ and fourteen trials. For both models the cell values have the same distribution with mean $2pn$, but in the SUM model they are dependent.



SUM model — Independent model (binomial distribution)

NxN boards for b=2 and N=128, n=7 and p=0.5



Board for the SUM model with b=4, N=256, n=4 and p=.5.

According to our results, if b is big, there is an interval for p around 1/2 where the root's value is highly concentrated around n. For b = 20, p = 0 3 and n = 30, if MAX and MIN play best possible the final cell's value is 30 with a probability greater than 99.9%, even through this value appears only in very few cells (less than 0.6%). However, if MIN and MAX play randomly, the last cell's value is near 10.

REFERENCES.

[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, 1991.

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

[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.

[Nau82] D.S. Nau, "The last player theorem," in: *Artificial Intelligence*, 18: 53-65, 1982.

[Nau82] 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, 1984.

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