Ann. Inst. Statist. Math.
39 (1987), Part A, 69-84

RANDOM SEQUENTIAL BISECTION AND ITS ASSOCIATED BINARY TREE

MASAAKI SIBUYA AND YOSHIAKI ITOH

(Received Mar. 24, 1986; revised July 24, 1986)

Summary

Random sequential bisection is a process to divide a given interval into two, four, eight, $\cdots$ parts at random. Each division point is uniformly distributed on the interval and conditionally independent of the others. To study the asymptotic behavior of the lengths of subintervals in random sequential bisection, the associated binary tree is introduced.

The number of internal or external nodes of the tree is asymptotically normal. The levels of the lowest and the highest external nodes are bounded with probability one or with probability increasing to one as the number of nodes increases to infinity.

The associated binary tree is closely related to random binary tree which arises in computer algorithms, such as binary search tree and quicksort, and one-dimensional packing or the parking problem.

1. Introduction

Divide the interval $(0, x)$ into two subintervals at random so that the division point is uniformly distributed on the interval. Similarly divide each of the subintervals. New division points are independent of each other under the condition that the subintervals are given. Repeat the procedure endlessly, all the division points being conditionally independent. This process is called "random sequential bisection".

Formally, let $(U_{d\,2j}, 0\leqq j<2^{d-1})_{d=1}^{\infty}$ be a doubling sequence of independent random variables following the $(0, 1)$ uniform distribution. Let X_{dj}, $0\leqq j<2^d$, $d=0, 1, 2, \cdots$, denote lengths of the subintervals after d-th step, defined by $X_{00}=x$,

$$X_{d\,2j}=X_{d-1\,j}U_{d\,2j}\,, \qquad \text{and}$$

Key words and phrases: Random binary tree, random division, random packing, parking problem, random spacings.

$$X_{d\,2j+1}=X_{d-1\,j}U_{d\,2j+1}\,, \qquad \text{where } U_{d\,2j}+U_{d\,2j+1}=1\,.$$

We are interested in the asymptotic behavior of X_{dj}'s.

Our bisection process is described by an infinite complete binary tree. The root is labeled as $X_{00}=x$, and the roots of its left and right subtrees are labeled as X_{10} and X_{11}, respectively. In general, a node at the $(d-1)$-th level labeled as $X_{d-1\,j}$ has two son nodes which are labeled as $X_{d\,2j}$ and $X_{d\,2j+1}$.

Now suppose that we stop to divide subintervals shorter than one, and continue bisection as long as the subintervals are not shorter than one. Thus the bisection process terminates in finite steps with probability one, leaving a finite number of subintervals, all shorter than one. Correspondingly we have now an associated finite 'binary tree' which is defined recursively as follows.

A binary tree in general is a finite set of nodes, which is partitioned, if not empty, into a triple; 'root', 'left subtree' and 'right subtree', where the root is a subset of a single node and the subtrees are binary trees. A node may have only a left or a right tree. This feature is different from a usual tree which appears, for example, in cluster analysis. The notion has emerged from the computer techniques such as binary search tree and quicksort. See Knuth ([4], Vol. I, Chapter 2, Section 3) and Section 5 of this paper.

In the infinite complete binary tree whose nodes are labeled by subinterval lengths, keep the nodes with a label of at least one, calling them "internal nodes", which form a binary tree. The internal nodes may have, as their son nodes, internal nodes and non-internal nodes. Also keep these non-internal nodes, calling them "external nodes", and erase all the other nodes. See Figure 1. The external nodes correspond to remaining subintervals in the random sequential bisection with the stopping rule, while the internal nodes correspond to intermediate subintervals to be eventually divided.

All X_{dj}'s are proportional to the starting length x, and we could start from the unit interval obtaining subinterval lengths $Z_{dj}=X_{dj}/x$, and these standardized quantities appear in later discussions. Rather, if x is increased in the above construction of finite binary tree, then one external node, with the largest label among the external ones, becomes an internal node, and its two son nodes external nodes. When length x is increased, the binary tree grows randomly. We call this evergrowing random binary tree "the binary tree associated with random sequential bisection" or simply "the associated tree". See Section 5 for other models of the randomly growing binary tree.

Our objective is now better defined as the study of the asymptotic

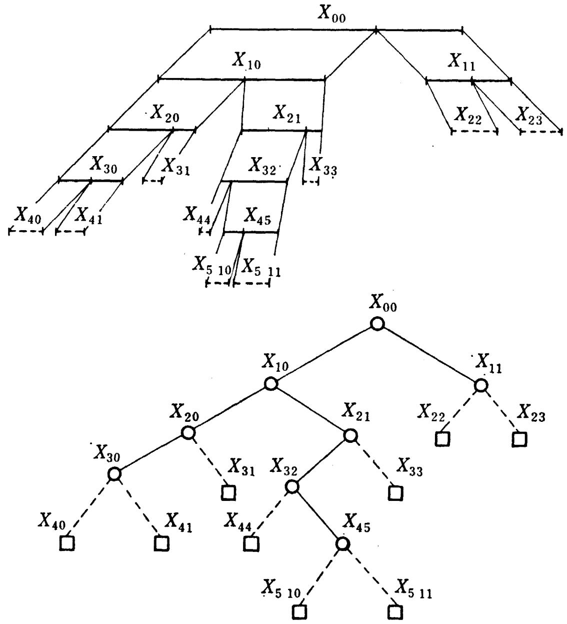

Fig. 1. Random sequential bisection and the associated binary tree.
X_{dj}: subinterval lengths/node labels
○: internal nodes ($X_{dj} \geqq 1$)
□: external nodes (sons of internal nodes and $X_{dj} < 1$)

behavior of the associated tree as x increases to infinity.

In Section 2 we show that the expected numbers of internal and external nodes (with label larger than w) at the d-th level are written using the Poisson distribution (Theorem 1). In Section 3 the variances of total numbers of internal and external nodes (with label larger than w) are calculated and their asymptotic normality is shown (Theorem 2). In Section 4, as $d=c\log x$ increases the external nodes are shown to be confined in a range of levels with probability approaching one (Theorem 3). The highest and the lowest level of external nodes are similarly confined (Theorem 4). In Section 5 other views of the associated tree are explained, and the associated tree is shown to be an upper approximation of the discrete random binary tree. Finally, in Section 6, random sequential bisection and random packing are regarded as special cases of a more general bisection process.

2. Expected number of nodes at the d-th level

In the binary tree $T(x)$ associated with the random sequential bisection starting from $(0, x)$, let $N_i(x, d)$ and $N_e(x, d)$ denote the numbers of the internal and the external nodes at the d-th level respectively, and let $m_i(x, d)$ and $m_e(x, d)$ denote their expected values respectively.

If the first division point of the interval $(0, x)$ is Y, $0<Y<x$, then $N_i(x, d)=N_i(Y, d-1)+N_i(x-Y, d-1)$, and the expectation of this equality shows that for $1\leqq x<\infty$ and $d=1, 2,\cdots$

$$m_i(x, d)=\frac{1}{x}\int_0^x\{m_i(y, d-1)+m_i(x-y, d-1)\}dy \tag{2.1}$$
$$=\frac{2}{x}\int_0^x m_i(y, d-1)dy ,$$

with

$$m_i(x, 0)=\begin{cases} 0, & \text{if } 0\leqq x<1, \\ 1, & \text{if } 1\leqq x<\infty. \end{cases} \tag{2.2}$$

The recursion equation (2.1) starting from (2.2) gives

$$m_i(x, d)=\begin{cases} 0, & \text{if } 0\leqq x<1, \\ 2^d\dfrac{1}{x}\displaystyle\sum_{k=d}^{\infty}\frac{(\log x)^k}{k!}, & \text{if } 1\leqq x<\infty, \end{cases} \tag{2.3}$$

for $d=0, 1, 2,\cdots$.

In any binary tree, $N_i(x, d-1)$ internal nodes have $2N_i(x, d-1)$ son nodes, among which $N_i(x, d)$ are internal, therefore for $d=1, 2,\cdots$

$$N_e(x, d)=2N_i(x, d-1)-N_i(x, d) . \tag{2.4}$$

The expectation of this equality shows that for $d=1, 2,\cdots$,

$$m_e(x, d)=2m_i(x, d-1)-m_i(x, d)=2^d\frac{1}{x}\frac{(\log x)^{d-1}}{(d-1)!} . \tag{2.5}$$

The value of $m_e(x, 0)$ is undefined at present.

Summarizing (2.3) and (2.5),

THEOREM 1. *Among the possible 2^d nodes at the d-th level, $1\leqq d$, of the associated tree $T(x)$, the proportions of the expected number of the internal and the external nodes are the Poisson probabilities*

$$\frac{1}{x}\sum_{k=d}^{\infty}\frac{(\log x)^k}{k!} \quad \textit{and} \quad \frac{1}{x}\frac{(\log x)^{d-1}}{(d-1)!}$$

respectively.

The implication of Theorem 1 is discussed in Section 3, and an extension of (2.5) is given in the remaining part of this section. The expected number $m_e(x, d)$ satisfies the recursion equation

$$m_e(x, d)=\begin{cases} 0, & \text{if } 0\leqq x<1, \\ \dfrac{2}{x}\displaystyle\int_0^x m_e(t, d-1)dt, & \text{if } 1\leqq x<\infty, \end{cases} \tag{2.6}$$

for $d=1, 2, \cdots$ with the initial condition

$$m_e(x, 0)=\begin{cases} 1, & \text{if } 0\leqq x<1, \\ 0, & \text{if } 1\leqq x<\infty. \end{cases} \tag{2.7a}$$

Because of the difference of the initial condition $m_e(x, d)$ vanishes in $(0, 1)$ and the integration in (2.6) can be limited to the interval $(1, x)$, but not in (2.1). The initial condition (2.7a) looks just like a conventional rule since an external node cannot appear at the root. However, the condition (2.7a) reflects the fact that external nodes can appear at the roots of the subtrees.

The equation (2.6) can be extended to the expected number $m_e(x, d, w)$ of external nodes which is at the d-th level and whose label is at least w $(0\leqq w<1)$. That is, $m_e(x, d, w)$ satisfies the same recursion equation (2.6) (actually $m_e(x, d)=m_e(x, d, 0)$) with the initial condition

$$m_e(x, 0, w)=\begin{cases} 0, & \text{if } 0\leqq x<w, \\ 1, & \text{if } w\leqq x<1, \\ 0, & \text{if } 1\leqq x<\infty. \end{cases} \tag{2.7b}$$

The solution to (2.6), with $m_e(x, d)$ replaced by $m_e(x, d, w)$, and (2.7b) is

$$m_e(x, d, w)=2^d(1-w)\frac{1}{x}\frac{(\log x)^{d-1}}{(d-1)!}, \tag{2.8}$$

if $0\leqq w<1\leqq x<\infty$ and $d=1, 2, \cdots$.

The above expression (2.8) is just $1-w$ times of (2.5). This is clear from the fact that the marginal distribution of a label X_{dj} of an external node, length of a remaining subinterval of random sequential bisection with the stopping rule, is the $(0, 1)$ uniform distribution. To prove this, in the formal definition of X_{dj} in Section 1, let $j=(j_1 j_2\cdots j_d)$ be the binary expression of a nonnegative integer index j, possibly having leading zeroes. Namely, $X_{dj}=x\prod_{k=1}^{d} U_{k(j_1\cdots j_k)}$, where d is the smallest integer such that $x\prod_{k=1}^{d} U_{k(j_1\cdots j_k)}$ is less than one, or

$$-\sum_{k=1}^{d}\log U_{k(j_1\cdots j_k)}>\log x .$$

Since the left-hand side forms a Poisson process with intensity one, the difference $-\sum \log U_{k(j_1\cdots j_k)}-\log x=-\log X_{dj}$ is a standard exponential variable under the above inequality condition, and X_{dj} is conditionally a (0, 1) uniform random variable.

3. Asymptotic size of the associated tree

We study further total numbers $N_i(x)=\sum_{d=0}^{\infty} N_i(x, d)$ of the internal nodes and $N_e(x, w)=\sum_{d=1}^{\infty} N_e(x, d, w)$ of the external ones, whose labels are at least w, of the associated tree $T(x)$. By summing up (2.3) and (2.5) respectively, their expected values $m_i(x)$ and $m_e(x, w)$ are, if $1\leqq x<\infty$,

$$m_i(x)=2x-1 , \qquad \text{and} \qquad m_e(x, w)=2(1-w)x . \tag{3.1}$$

The relationship between these values comes also from the fact that $N_e(x, 0)=N_i(x)+1$ in any binary tree because of (2.4), and from the discussion at the end of the last section. Note that $m_i(x)$ is the solution to the integral equation

$$m_i(x)=\frac{2}{x}\int_0^x m_i(y)dy+1 , \qquad 1\leqq x<\infty \tag{3.2}$$

with

$$m_i(x)=0 , \qquad 0\leqq x<1 ,$$

and $m_e(x, w)$ ($m_e(x)=m_e(x, 0)$) is the solution to

$$m_e(x, w)=\frac{2}{x}\int_0^x m_e(y, w)dy , \tag{3.3}$$

with the convention

$$m_e(x, w)=\begin{cases} 0 , & 0\leqq x<w , \\ 1 , & w\leqq x<1 . \end{cases}$$

The equation (3.3) is obtained from (3.2) by putting $m_e(x, 0)=m_i(x)+1$. See Section 6 where the equations are extended.

Let $v_e(x, w)$ denote the variance of $N_e(x, w)$. If $1\leqq x$ and the first division point of $(0, x)$ is Y. Then $v_e(x, w)$ is the solution to

$$\begin{aligned} v_e(x, w)=&\mathrm{E}^Y [\mathrm{Var} [N_e(Y, w)+N_e(x-Y, w) \mid Y]] \\ &+\mathrm{Var}^Y [\mathrm{E} [N_e(Y, w)+N_e(x-Y, w) \mid Y]] . \end{aligned}$$

Because of the additivity of the conditional variance, this is equal to

$$\frac{1}{x}\int_0^x (v_e(y, w)+v_e(x-y, w))dy+\mathrm{Var}^Y [m_e(Y, h)+m_e(x-Y, h)] \tag{3.4}$$

$$=\frac{2}{x}\int_0^x v_e(y, w)dy+2\{\mathrm{Var}^Y[m_e(Y, h)]$$
$$+\mathrm{Cov}^Y[m_e(Y, h), m_e(x-Y, h)]\}.$$

Some results of this simple but tedious integration are shown in Appendix 1.

The main result is, if $2\leqq x<\infty$,

$$\text{(3.5)}\qquad v_e(x, w)=\begin{cases}\Big\{12w(1-w)-10(1-w)-1+8(1-w) \\ \qquad\cdot\log\dfrac{2}{1+w}+\dfrac{6}{1+w}(1-w)\Big\}x, & \text{if } 0\leqq w<1/2, \\ (1-w)\Big\{2(-7+6w)+\dfrac{1}{w}+\dfrac{6}{1+w}+8\log\dfrac{2}{1+w}\Big\}x, \\ & \text{if } 1/2\leqq w<1,\end{cases}$$

The variance of $N_i(x)=N_e(x, 0)-1$ is equal to

$$v_e(x, 0)=(8\log 2-5)x\fallingdotseq 0.54517744x, \qquad \text{if } 2\leqq x<\infty.$$

The point is that $v_e(x, w)$ is proportional to x for $2\leqq x$, and $N_i(x)$ is under-dispersion. It is shown in Section 5 that $N_i(x)\geqq[x]$, where $[x]$ denotes the integer part of x, if $0<x$.

Based on the linearity of $m_e(x, w)$ and $v_e(x, w)$ in x, we have the following:

THEOREM 2. *The standardized random variable*

$$(N_e(x, w)-m_e(x, w))/\sqrt{v_e(x, w)}$$

is asymptotically normally distributed as $x\to\infty$. *The same statement holds for* $N_i(x)=N_e(x, 0)-1$.

PROOF. The discussion by Dvoretzky and Robbins ([2], Section 5) can be directly applied. The proof sketched in Appendix 2 makes this paper self-contained.

4. Asymptotic shape of the associated tree

In Sections 2 and 3 some facts on the size of the associated tree $T(x)$ are shown. In this section its shape is discussed. Firstly we note that Theorem 1 actually shows the following fact. If the number of external nodes at a level of $T(x)$ is small, it means either that almost all nodes of the level are internal or that there are only a few internal nodes at the level.

To see this, let $p(x; \mu)$ be the Poisson probability function,

$$p(x;\mu)=e^{-\mu}\mu^x/x!\ .$$

It is shown that its upper and lower tail probabilities are evaluated by the probabilities at the end of the tails;

$$p(x;\mu)<\sum_{y=x}^{\infty}p(y;\mu)<\frac{x+1}{x+1-\mu}p(x;\mu)\,,\qquad \text{if } \mu-1<x\,,$$

and

$$p(x;\mu)<\sum_{y=0}^{x}p(y;\mu)<\frac{x}{\mu-x}p(x;\mu)\,,\qquad \text{if } x<\mu\,.$$

Theorem 1 shows that the expected number of internal and external nodes at the d-th level satisfy

$$m_e(x,d)<2^d-m_i(x,d)<\frac{d}{\log x-d}m_e(x,d)\,,\qquad \text{if } d<\log x\,,$$

and

$$m_e(x,d+1)<m_i(x,d)<\frac{d+1}{d+1-\log x}m_e(x,d+1)\,,\qquad \text{if } \log x<d+1\,.$$

Thus the asymptotic behavior of $m_e(x,d)$ determines that of $m_i(x,d)$.

THEOREM 3. *As x and d increase to infinity satisfying $d=c\log x$,*

$$\left.\begin{array}{r} m_e(x,d) \\ m_i(x,d) \\ (\mathit{if}\ c>1) \\ 2^d-m_i(x,y) \\ (\mathit{if}\ c<1) \end{array}\right\}=\frac{1}{\sqrt{2\pi d}}e^{-d\gamma(c)}(1+O(1/d))\,,$$

where $\gamma(c)=1/c+\log(c/2)-1$. This implies that

$$\lim m_e(x,d)=\lim m_i(x,d)=\begin{cases} 0\,, & \mathit{if}\ \bar{c}\leqq c<\infty\,,\\ \infty\,, & \mathit{if}\ 1<c<\bar{c}\,,\end{cases}$$

and

$$\lim m_e(x,d)=\lim\{2^d-m_i(x,d)\}=\begin{cases} 0\,, & \mathit{if}\ 0<c\leqq\underline{c}\,,\\ \infty\,, & \mathit{if}\ \underline{c}<c<1\,,\end{cases}$$

where the limit means $d=c\log x\to\infty$, and $\bar{c}\doteqdot 4.31107041$ and $\underline{c}\doteqdot 0.373364616$ are solutions to $\gamma(c)=0$ or $c\exp(c^{-1}-1)=2$.

PROOF. Using Stirling's formula we obtain

$$\sqrt{\frac{d-1}{2\pi}}\,\frac{c}{d}\exp\left\{-\gamma(c)d-\frac{1}{12(d-1)}\right\}$$
$$<m_e\left(d, \exp\left(\frac{d}{c}\right)\right)<\frac{1}{\sqrt{2\pi(d-1)}}\exp\left\{-\gamma(c)d-\frac{1}{12d}\right\}.$$

Therefore $m_e(d, \exp(d/c))$ tends to 0 or ∞, if $\gamma(c)\geqq 0$ or $\gamma(c)<0$, respectively. The function $\gamma(c)$ decreases in monotone from $\gamma(+0)=\infty$ to $\gamma(1)=-\log 2$ and increases in monotone to $\gamma(\infty)=\infty$. The above inequalities on $m_i(x, d)$ give the limits.

The limit $m_i(x, d)\to 0$ in the theorem means the probability of the following equivalent events approaches one.

$$A_U(x, d):\quad N_i(x, d)=0 \Longleftrightarrow \max_j X_{dj}<1 \Longleftrightarrow H(x)\leqq d\,,$$

where $H(x)$ denotes the highest level of the external nodes of $T(x)$. Similarly, $2^d-m_i(x, d)\to 0$ means that with probability approaching one,

$$A_L(x, d):\quad N_i(x, d)=2^d \Longleftrightarrow \min_j X_{dj}\geqq 1 \Longleftrightarrow h(x)>d\,,$$

occurs, where $h(x)$ denotes the lowest level of the external nodes of $T(x)$. Thus, the external nodes are located at the levels between $\underline{c}\log x$ and $\bar{c}\log x$. In terms of $Z_{dj}=X_{dj}/x$, lengths of subintervals starting from $(0, 1)$, this means the following:

COROLLARY. *With probability approaching one as* $d\to\infty$,

$$1/\bar{c}<(-\log Z_{dj})/d<1/\underline{c}\,.$$

This corollary gives an upper bound of the distribution of $\max_j Z_{dj}$ and a lower bound of that of $\min_j Z_{dj}$. Rough bounds from the opposite sides are obtained by modifying the random sequential bisection as follows.

Instead of doubling the number of subintervals at each step, keep only one subinterval, always the larger or the smaller one. In the case of smaller subintervals the length V_d of the interval at the d-th step is

$$V_d=x\prod_{j=1}^{d} S_j$$

where (S_j) is a sequence of independent random variables following the uniform distribution on the interval $(0, 1/2)$. The logarithm

$$-\log(2^d V_d/x)=-\sum_{j=1}^{d}\log(2S_j)$$

is the sum of d independent exponential random variables with the mean equal to one, and follows the gamma distribution. Thus

$$(-\log V_d-(1+\log 2)d+\log x)/\sqrt{d}$$

is asymptotically standard normal.

In the case where larger parts are kept, the length W_d at the d-th step is

$$W_d=x\prod_{j=1}^{d} T_j\,,$$

where (T_j) is a sequence of independent random variables following the uniform distribution on $(1/2, 1)$. Since $-\log T_j$ has the mean $\mu_T=1-\log 2$ and variance $\sigma_T^2=1-2(\log 2)^2$,

$$(-\log W_d-\mu_T d+\log x)/\sigma_T\sqrt{d}$$

is asymptotically standard normal.

In a realization of sequential random bisection, V_d is larger than $\min_j X_{dj}$ and W_d is smaller than $\max_j X_{dj}$. That is, V_d is stochastically larger than $\min_j X_{dj}$, and $\max_j X_{dj}$ than W_d. Since $-\log V_d/d\to 1+\log 2$ and $-\log W_d/d\to\mu_T$, we obtain a result:

THEOREM 4.

$$-\log\max_j Z_{dj}/d<1-\log 2 \qquad \textit{and} \qquad 1+\log 2<-\log\min_j Z_{dj}/d$$

almost surely as d increases to infinity.

Remark. The inequalities of Theorem 4 are equivalent to

$$1/(1-\log 2)<H(x)/\log x \qquad \text{and} \qquad h(x)/\log x<1/(1+\log 2)\,,$$

respectively, and they are complements to the corollary of Theorem 3, which means, with probability approaching one as $d\to\infty$,

$$1/\bar{c}<-\log\max_j Z_{dj}/d \qquad \text{and} \qquad -\log\min_j Z_{dj}/d<1/c\,.$$

5. More on the associated tree

The ever-growing binary tree associated with random sequential bisection can be viewed in different ways. Suppose sticks of length L_{nj}, $j=1, 2,\cdots, n$, shorter than one, grow at the same rate, that is $\mathrm{d}L_{nj}/L_{nj}$ is independent of subscripts. If the longest one, say L_{nj}, reaches one the stick breaks into two parts of lengths $L_{n+1\,j}=L_{nj}U_n$ and $L_{n+1\,n+1}=L_{nj}(1-U_n)$, where $\{U_n\}$ is a sequence of independent $(0, 1)$ uniform random variables. The others are renamed as $L_{n+1\,k}=L_{nk}$, $k\neq j$.

Starting from $L_{11}(0)=0$, growing at the rate $\mathrm{d}\,L_{nj}(t)/\mathrm{d}\,t=L/t$, "grow and break stick lengths" $\{L_{nj}(t),\ j=1,2,\cdots,n\}$ with $\sum_{j=1}^{n} L_{nj}(t)=t$ is, at a time instant, is the same set as the remaining subinterval lengths of random sequential bisection with the stopping rule starting from an interval of length t.

If a grow and break sticks process starts from a unit interval and the time is measured by $\log t$, then the process can be regarded as the following diffusion or percolation process on the infinite complete binary tree. That is, liquid run through edges of the tree. Reaching a node at the d-th level sometime, liquid reaches its two son nodes after $-\log U_{nj}$ and $-\log(1-U_{nj})$ time intervals respectively. The pair $(-\log U_{nj},\ -\log(1-U_{nj}))$ is a degenerated bivariate exponential random variable. The time to reach a node corresponds to its counterpart's breaking time. The discussion by Pittel ([6], Section 4) can be applied here.

An associated tree $T(x)$ keeps all the data of the above-mentioned process. For a fixed value of x, $T(x)$ shows the history and state at a certain time instant of the process. However, $T(x)$ can be constructed node by node in a different way. That is, arrange consecutively all subintervals greater than or equal to one to form a line segment, and drop a division point uniformly on it. This corresponds to make and internal node branch, having two son nodes which may be internal or external. Continue the process while there is at least one internal node.

A finite tree $T(x)$ does not retain the order of growth of this type but the order can be reconstructed stochastically as follows. Starting from the root choose left or right son nodes with probabilities X_{10}/X_{00} and X_{11}/X_{00}. If, for example, the left node is chosen, then with probabilities X_{20}/X_{00}, X_{21}/X_{00} and X_{11}/X_{00}, choose one of possible three nodes at random, and so on. The number of possible orders depends on the shape of the tree.

With the random sequential bisection, another type of random tree can be associated as follows. Assume the starting length x be a positive integer. Label the roots of the subtrees as $\tilde{X}_{10}=[X_{10}]$ and $\tilde{X}_{11}=[X_{11}]$, the integer part of X_{10} and X_{11}, respectively. In general, if $\tilde{X}_{d-1\,j}$ is a label of an internal node at the $(d-1)$-th step, its son nodes are labeled as $\tilde{X}_{d\,2j}=[\tilde{X}_{d-1\,j}U_{d\,2j}]$ and $\tilde{X}_{d\,2j+1}=[\tilde{X}_{d-1\,j}U_{d\,2j+1}]$ in the notation of Section 1. Namely, under the condition that $\tilde{X}_{d-1\,j}=k$, $\tilde{X}_{d\,2j}=k-1-\tilde{X}_{d\,2j+1}$ is uniformly distributed on $\{0,1,\cdots,k-1\}$ (with probability one). Unit length being lost at each bisection, the number of internal nodes is n (with probability one).

This "discrete" random binary tree is precisely the concept modeling binary search tree and quicksort (Knuth [4], Vol. III) and it is also associated with random spacings. It is well studied, and its symptotic shape has been shown recently by Robson [8], Mahmoud and Pittel [5], and Pittel [6].

Compare a realization of the discrete tree with that of the associated tree of the same starting length $x=n$, to see $X_{dj}\geqq\tilde{X}_{dj}$. The associated tree is larger than the discrete one in subinterval length, number of internal nodes at a level and in total, and levels of the lowest and the highest external nodes. Among the statements following Theorem 4, for example, $H(x)/\log x<\bar{c}$ and $h(x)/\log x<1+\log 2$ with probability approaching one, are valid for the discrete tree. Robson [8] obtained $\bar{c}$ by a more direct evaluation.

In this associated binary tree, the probability $P_U(x, d)$ of the event $A_U(x, d)$ (see the discussion before Corollary of Theorem 3 in Section 4): $N_i(x, d)=0\Longleftrightarrow\max_j X_{dj}<1\Longleftrightarrow H(x)\leqq d$, satisfies the equation, with P replaced by P_U,

$$P(x, d)=\frac{1}{x}\int_0^x P(y, d-1)P(x-y, d-1)dy \tag{5.1}$$

and the initial condition

$$P_U(x, 0)=\begin{cases}1, & \text{if } 0\leqq x<1,\\ 0, & \text{if } 1\leqq x<\infty.\end{cases} \tag{5.1a}$$

The probability $P_L(x, d)$ of the dual event $A_L(x, d)$: $N_i(x, d)=2^d\Longleftrightarrow\min_j X_{dj}\geqq 1\Longleftrightarrow h(x)>d$, satisfies the same equation (5.1), with P replaced by P_L, and the initial condition

$$P_L(x, 0)=\begin{cases}0, & \text{if } 0\leqq x<1,\\ 1, & \text{if } 1\leqq x<\infty.\end{cases} \tag{5.1b}$$

While in the discrete binary tree the corresponding probabilities $\tilde{P}_U(n, d)$ and $\tilde{P}_L(n, d)$ satisfy a discrete analogue of (5.1).

$$\tilde{P}(n, d)=\frac{1}{n}\sum_{k=0}^{n-1}\tilde{P}(k, d-1)\tilde{P}(n-k-1, d-1) \tag{5.2}$$

and the initial conditions

$$\tilde{P}_U(n, 0)=\begin{cases}1, & \text{if } n=0,\\ 0, & \text{if } n=1, 2,\cdots,\end{cases} \tag{5.2a}$$

and

$$\tilde{P}_L(n, 0)=\begin{cases} 0, & \text{if } n=0, \\ 1, & \text{if } n=1, 2, \cdots. \end{cases} \tag{5.2b}$$

respectively.

Since $\max_j X_{dj}<1$ implies $\max_j \tilde{X}_{dj}<1$ and $\min_j \tilde{X}_{dj}\geqq 1$ implies $\min_j X_{dj}\geqq 1$,

$$P_U(n, d)\leqq \tilde{P}_U(n, d) \quad \text{and} \quad P_L(n, d)\geqq \tilde{P}_L(n, d) \qquad \text{for } 1\leqq d\leqq n.$$

6. Relation to the random packing problem

One dimensional random packing is known as the car parking problem. This problem can be extended to include random sequential bisection as follows. Along a street of length x, cars of length l park at random in vacant intervals longer than v ($v>l$). The usual car parking problem, discussed first by Rényi [7] is the case $v=l=1$, and random sequential bisection with the stopping rule is the case $v=1$ and $l=0$.

In the usual set up of the car parking problem, cars are supposed to arrive one by one. However, since all wide intervals are eventually filled up, it can be viewed as a bisection with deletion of car length interval, and can be associated with a binary tree. A car corresponds to an internal node whose label is the length of the interval into which the car entered, while a gap between cars corresponds to an external node.

The expected number $m_i(x)$ of total internal nodes (cars) satisfies

$$m_i(x+l)=\frac{2}{x}\int_0^x m_i(y)dy+1\ ,$$

with

$$m_i(x)=\begin{cases} 0, & \text{if } 0\leqq x<v, \\ 1, & \text{if } 0\leqq x<v+l, \end{cases}$$

while the expected number $m_e(x)$ of total external nodes (gaps) satisfies

$$m_e(x+l)=\frac{2}{x}\int_0^x m_e(y)dy$$

with

$$m_e(x)=\begin{cases} 1, & \text{if } 0\leqq x<v, \\ 2, & \text{if } 0\leqq x<v+l. \end{cases}$$

The equations are essentially the same since $m_e(x)=m_i(x)+1$. The delay term l, which can be one without loss of generality if positive, in the argument of the left-hand side of the equation characterizes the packing problem.

In the packing problem the distribution of gap lengths has been studied. Bánkövi [1] obtained the expected number of gaps longer than w, and Itoh [3] obtained the distribution of the minimum gap. These are related to Sections 2-4 of this paper, and can be extended further.

DEPARTMENT OF MATHEMATICS, KEIO UNIVERSITY
THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Bánkövi, G. (1962). On gaps generated by a random space filling procedure, *Publ. Math. Inst. Hungar. Acad. Sci.*, **7**, 395-407.

[2] Dvoretzky, A. and Robbins, H. (1964). On the 'parking problem', *Publ. Math. Inst. Hungar. Acad. Sci.*, **9**, 209-225.

[3] Itoh, Y. (1980). On the minimum of gaps generated by one-dimensional random packing, *J. Appl. Prob.*, **17**, 134-144.

[4] Knuth, D. E. (1973). *The Art of Computer Programming*, Vol. I, *Fundamental Algorithms* (2nd edition), Vol. III *Sorting and Searching*, Addison-Wesley.

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

[6] Pittel, B. (1984). On growing binary trees, *J. Math. Analysis and Appl.*, **103**, 461-480.

[7] Rényi, A. (1958). On a one-dimensional problem concerning space-filling, *Publ. Math. Inst. Hungar. Acad. Sci.*, **3**, 109-127.

[8] Robson, J. M. (1979). The height of binary search trees, *The Australian Computer J.*, **11**(4), 151-153.

Note added in proof. After the acceptance of this paper, the authors noticed relationship of our problem with the random version of Kakutani's subdivision problem. Closely related papers are as follows.

Lootgieter, J. C. (1977). *Ann. Inst. Henri Poincaré*, **13**, 385-410.
van Zwet, W. R. (1978). *Ann. Prob.*, **6**, 133-137.
Slud, E. (1978). *Zeit. Wahrscheinlichkeitsth.*, **41**, 341-352.
Pyke, R. (1980). *Ann. Prob.*, **8**, 157-163.

New results in random sequential bisection were obtained in

Devroye, L. (1986). *J. Ass. Comput. Mach.*, **33**, 489-498.

APPENDIX 1

Computation of $v_e(x, w)$ in (3.4) to obtain the result in (3.5) is sketched in the following. The random variable Y is uniformly distributed on $(0, x)$, then

$$\mathrm{E}\,[m_e(Y, w)]=(1-w)x\,, \qquad 1\leqq x\,,$$

$$\mathrm{E}\,[m_e^2(Y, w)]=\frac{1-w}{x}\left\{1+\frac{4}{3}(1-w)(x^3-1)\right\}\,, \qquad 1\leqq x\,,$$

and

$$
\mathrm{E}\,[m_e(x-Y, w)m_e(Y, w)] = \begin{cases} 0\,, & 1 \leqq x < 2w\,, \\ (x-2w)/x\,, & \max(1, 2w) \leqq x < 1+w\,, \\ \{2w-x+2(1-w)(x-w)^2\}/x\,, & 1+w \leqq x < 2\,, \\ 2(1-w^2)(x^3+1-3w)/3x\,, & 2 \leqq x\,. \end{cases}
$$

Using these, the integration equation (3.4) is solved case by case. When the function $v_e(x, w)$ is determined in an interval of x, the integration in (3.4) on the whole interval is replaced by its definite integral. For $x \leqq 2$, the function $v_e(x, w)$ is as follows.

$$
v_e(x, w) = \begin{cases} 0\,, & 0 \leqq x < 1\,, \\ 2(1-w)x(1-2(1-w)x)\,, & 1 \leqq x < 2w\,, \\ 2(1-(1-w)x)(2(1-w)x-1)\,, & w \leqq 1/2 \text{ and } 1 \leqq x < 1+w\,, \\ -2+\{2(1-w)+1/w\}x-4(1-w)^2x^2\,, & 1 < 2w \leqq x < 1+w\,, \\ 2(1+4w(1-w))+\left\{\dfrac{2}{1+w}(1-4w+w^2)-8(1-w)\log(1+w)\right\}x & \\ \qquad +8(1-w)x\log x-4(1-w)^2x^2\,, & w \leqq 1/2 \text{ and } 1+w \leqq x < 2\,, \\ 2(1+4w(1-w))+\left\{-14+6w+\dfrac{12}{1-w}+\dfrac{1}{w}-8(1-w)\log(1+w)\right\}x & \\ \qquad +8(1-w)x\log x-4(1-w)^2x^2\,, & 1/2 < w \text{ and } 1+w \leqq x < 2\,. \end{cases}
$$

APPENDIX 2

PROOF OF THEOREM 2.

In the third paragraph of Section 5, the second way to construct $T(x)$ node by node was shown. Suppose $n-1$ division points are dropped at random on $(0, x)$ in that way, and the lengths of the subintervals are $\boldsymbol{y}=(y_1, y_2, \cdots, y_n)$. The number $n=n(x)=o(x)$ will be specified later. Construct $T(y_j)$, if $y_j>1$, in any way, without changing the interval if $y_j<1$. The final result is equivalent to unconditional construction of $T(x)$, and

$$
N_e(x, w)=\sum_{i=1}^{n} N_e(y_i, w)\,,
$$

where the terms are conditionally independent.

Put

$$Y_i = N_e(y_i, w)/v_e(x, w)^{1/2}, \qquad i=1, 2, \cdots, n,$$

and

$$Y_0 = -m_e(x, w)/v_e(x, w)^{1/2}.$$

Then

$$Z(x, w) = (N_e(x, w) - m_e(x, w))/v_e(x, w)^{1/2} = \sum_{i=0}^{n} Y_i,$$

and

$$\sum_{i=1}^{n} \mathrm{E}[Y_i] = 0. \tag{A.1}$$

Since $v_e(x, w)$ is equal to $\lambda(w)x$ if $2 \leqq x < \infty$, and is finite if $0 \leqq x < 2$,

$$\sum_{i=0}^{n} v_e(y_i, w) = \lambda(w)x + O(n),$$

for any $\boldsymbol{y}$, and

$$\mathrm{Var}\,[Z(x, w) \mid \boldsymbol{y}] = \sum_{i=0}^{n} \mathrm{Var}\,[Y_i \mid \boldsymbol{y}] = 1 + o(1) \tag{A.2}$$

if $n = o(x)$.

Now put

$$n = [x^{1/2}(\log x)^2] = o(x)$$

and let $B = B(x, n, \eta)$ denote the event

$$\max_i y_i < \eta x^{1/2}.$$

Consider the maximum u of n random spacings of the interval $(0, x)$. For a positive integer k,

$$\Pr\left[u > \frac{2x}{k}\right] \leqq k\left(1 - \frac{1}{k}\right)^n \to 0,$$

if $k = c\sqrt{x}$, where c is a positive constant, as x increases to infinity. Since u is stochastically larger than $\max_i y_i$, which is the longest interval in random division with constraints, $\Pr[B] \to 1$ as x increases. Taking $\eta < \delta\sqrt{\lambda(w)}/2$, the event B means $|\mathrm{E}[Y_i]| < \delta$, and

$$|Y_i - \mathrm{E}[Y_i]| \leqq \delta. \tag{A.3}$$

The conditions (A.1), (A.2) and (A.3) jointly satisfy the conditions of Lemma 2 of Dvoretzky and Robbins [2], a version of the central limit theorem, and proves the asymptotic normality of $Z(x, w)$.