Acta Informatica 17, 245 – 265 (1982

Deleting the Root of a Heap*

Ernst E. Doberkat[1]

Department of Mathematics and Computer Science, Clarkson College of Technology, Potsdam, N.Y. 13676, USA

Summary. The average behavior of the familiar algorithm for root deletion is considered, when every heap has the same probability to occur. The analysis centers around the notion of a viable path in the tree representation, i.e. such a path the label which replaces the label of the root may be allowed to travel when the heap is reconstructed. In case the size of the heap is a power of 2 it is shown that both the expected number of comparisons and of interchanges are asymptotically equal to the respective numbers in the worst case.

1. Introduction

In this paper some aspects of root deletion of a heap are investigated from an expected performance point of view. This algorithm is well known and works as follows: suppose you have a heap of size N such that the label of the root is the smallest element in this heap. If you want to extract this smallest element, then you will copy it elsewhere, and the label of the rightmost node at the bottom will be placed at the root. Since the heap property is violated in doing this, this label will percolate the tree down, until it has reached a position in which the heap property is reconstructed. This process implies that the second smallest element is now on top of the heap, and if the process of deleting the root and reconstructing the heap property is repeated until all elements in the heap are processed, the elements of the heap are arranged in ascending order. Thus root deletion is at the heart of the selection phase of heapsort and of the use of heaps as priority queues. Let us pick two applications in which root deletion occurs in a natural way.

---

* Some results were presented at the 19th Allerton Conference on Communications, Control, and Computing, University of Illinois, 1981, and the Second World Conference on Mathematics at the Service of Man, Las Palmas, Canary Islands, Spain, 1982

1 Some of this work was done at the Fernuniversität, Hagen, West Germany

0001-5903/82/0017/0245/$04.20

Following a suggestion due to Aho, Hopcroft, and Ullman ([1], p. 174), heaps underly the classical algorithms to construct minimal spanning trees, and shortest paths, respectively, due to Dijkstra. Both algorithms require the knowledge of that edge in the graph which has the lowest cost, and hence which has to be processed next. Thus deleting the root in this situation means that the next edge will be investigated, and joined the spanning tree, or the path constructed so far.

In his discussion on the implementation of priority queues Habermann suggests to realize these queues as heaps ([7], Sect. 6.5). Hence if we are given a set of jobs awaiting service from some processor, all of them endowed with a priority, the root of the heap corresponds to that job which will be served next. In this context root deletion means that the job in question receives its service, and that the remaining jobs are arranged in such a manner that the job next to be processed is readily available.

In this paper it is assumed that the heap that serves as an input to the algorithm is chosen at random in the sense that every heap has the same probability to occur. In the combinatorial approach (see [9]) this would be realized by considering all permutations of $\{1, \ldots, N\}$ that have the heap property and assigning equal probability to each of them. The analysis then would proceed by combinatorial means. However, it turns out in this case that a continuous model has some technical advantages. Here one considers a subset of all real vectors with the heap property and works with e.g. the uniform distribution. The algorithm is represented as a map, and the effect of the algorithm on the originally given distribution is represented then as an image measure with respect to this representing map. The evaluation of this image measure is done by the Change of Variables Formula of classical Real Analysis. This approach works well, too, in case of heap construction, and in inserting a new element into a heap (see [4, 5]).

The main idea in the present analysis is to characterize those paths in the heap, the element replacing the originally given root of the heap may travel in order to reconstruct the heap property; call such a path a viable one. After having done this characterization, it turns out to be necessary to compute the probabilities with which the element in question does indeed follow such a viable path. These probabilities are computed explicitly in terms of the binary expansion of N and of some intrinsic properties of the viable path, viz., the height of the subtree which is rooted at the first node on the path that is not on the some particular path, and the number of nodes that are on the path in question, but not on the path leading from node N to the root of the tree. Using these probabilities and the length of the paths, the expected numbers of interchanges, and of comparisons necessary for the algorithm may be computed. This is done in case N is a power of 2, say $N=2^n$, and in this case the following is shown: the expected number of interchanges equals $n-1+o(1)$, and the expected number of comparisons equals $2n-1+o(1)$. Note that the worst case is n, resp. $2n$.

The plan of the paper is as follows: in Sect. 2 some preparations for the analysis are made, in particular some anatomical details of heaps are quoted from the corresponding chapter in [9], the algorithm is given more formally and the probabilistic assumptions are made explicit. Section 3 deals with distributional aspects, it is shown how root deletion affects the uniform distribution, and how the resulting distribution looks like. In Sect. 4 we calculate the probabilities with which the replacement of the root follows a given viable path, and in Sect. 5 the leading terms of the asymptotic expansions for the expected numbers of interchanges and of comparisons are computed in case N is a power of 2. The concluding remarks in Sect. 6 discuss some other distributions than the uniform one, and give some suggestions for further work.

2. Preparations

A N-dimensional real array $b[1 \dots N]$ is said to be a *heap* iff $b[\lfloor i/2 \rfloor] \leqq b[i]$ holds for any i such that $2 \leqq i \leqq N$. If $\{1, \dots, N\}$ is represented as a tree such that 1 is the root, and $\lfloor i/2 \rfloor$ is the father of node i, and if furthermore $b[i]$ serves as a label for node i, then b is a heap iff no node has a smaller label, than its father. The tree representation will be important in the sequel, and some notation will be used with respect to it, cp. [9].

To begin with, if $1 \leqq i \leqq N$, let $g(N, i)$ denote the size of the subtree rooted at i, i.e., the number of descendants of i in the tree corresponding to $\{1, \dots, N\}$ including i;

$$G(N, i) := \prod \{g(N, j);\, j \text{ is a node in the subtree rooted at } i\}$$

denotes the product of the respective subtree sizes. A *special node* in the tree is a node which is on the path from N to the root 1; define $t(N, 0) := N$, and $t(N, i+1) := \lfloor t(N, i)/2 \rfloor$ for $0 \leqq i < \lfloor \log_2 N \rfloor$, then $\{t(N, i);\ 0 \leqq i \leqq \lfloor \log_2 N \rfloor\}$ constitutes the special path. Assume that N equals 43, then the following figure displays the representing tree. The special path is shown with heavier lines. Let

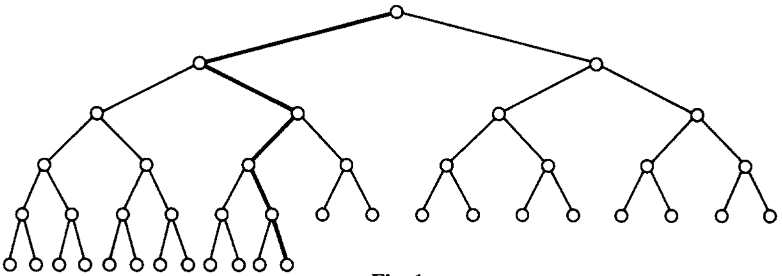

Fig. 1

N have the binary representation $(1 b_{n-1} \dots b_0)_2$, then it is easily seen that $t(N, i) = (1 b_{n-1} \dots b_i)_2$ holds, and from [9], Exercise 5.2.3.20 it is derived that we have

$$g(N, t(N, i)) = (1 b_{i-1} \dots b_0)_2. \tag{1}$$

The following algorithm will be investigated in the sequel:

Algorithm Rootdeletion

Input: a heap $b[1 \ldots N]$

Output: x, the smallest element of the heap, and the heap $b[1 \ldots N-1]$ which is the originally given heap with its root deleted and the heap property reconstructed

Method:
1. [Copy the root]
   $x := b[1]$; $b[1] := b[N]$;
2. [Reconstruct]
   $i := 1$;
   loop: *if* i is no leaf and a son of i contains a smaller label than i does
   then
   $k :=$ son of i with the smallest label;
   interchange $b[i]$ and $b[k]$;
   $i := k$;
   goto loop
   endif.

This algorithm is a variation of Floyd's method to construct heaps ([1], p. 90), and if repeated N times, it returns the array sorted in ascending order. The question which will be investigated below is the expected number of interchanges and comparisons that are necessary in the execution of this algorithm. In order to calculate these expectations, some preparations are needed.

It will be assumed that every heap has the same probability to occur, but rather than doing so for permutations, we take the heaps to be chosen at random from

$$H(N) := \{(x_1, \ldots, x_N); 0 \leqq x_i \leqq 1, \text{ the vector is a heap}\}.$$

Then to say that every heap has the same probability to be considered as an input to the algorithm means that for any Borel set $A \subseteq H(N)$

$$\mathrm{Prob}(A) := \chi_N \cdot \lambda^N(A) \tag{2}$$

defines the probability in question. Here χ_N denotes the product of all subtree sizes, i.e.

$$\chi_N := \prod\{\mathrm{g(N,i)}; 1 \leqq i \leqq N\} \qquad (= G(N, 1)),$$

and λ^N is N-dimensional volume or, more technical, Lebesgue measure restricted to (the Borel sets of) $H(N)$. It is shown in [4], Theorem 1 that if heaps are constructed with Floyd's algorithm, and the inputs are uniformly distributed, the resulting heaps are, thus have a distribution according to Eq. (2). Note that with probability 1, every heap has mutually different components since

$$\lambda^N(\{x \in H(N); x_i = x_j \text{ for some } i \neq j\}) = 0.$$

Thus we may and do neglect heaps in which some components are equal. In Sect. 6 we will discuss how this model is related to the usual assumption of randomness, in which the inputs are permutations of $\{1, \ldots, N\}$ having the heap property, every of these occurring with probability χ_N^{-1}.

A helpful tool for the subsequent investigations will be the *Change of Variables Formula* from Integral Calculus as a generalization of the familiar situation in case $N=1$, in which variables of integration are substituted. It is stated here for the reader's convenience and reads as follows: let U and V be open subsets of $\mathbb{R}^N$, such that $U=g[V]$ for some continuously differentiable homeomorphism g, then we have for any integrable map ϕ: $U \to \mathbb{R}$ the equality

$$\int_U \phi \, d\lambda^N = \int_V \phi \circ g \cdot |\det g'| \, d\lambda^N, \tag{3}$$

see [10], 8.27. This formula will be applied below in the following way: g will represent the action of the algorithm, transforming some part V of the input domain in a unique way to an output range U. Then the formula describes how the underlying distribution will be transformed by the algorithm.

Finally it will be convenient to describe the way $b[N]$ percolates the tree by a word $v \in \{0,1\}^*$, where 0 means that it is interchanged with a left son, and 1 denotes an interchange with a right son. Let for such a v be v_j the prefix of length j, i.e. if $v=c_1 \ldots c_k$, $1 \leqq j \leqq k$, then $v_j=c_1 \ldots c_j$; for the sake of completeness v_0 is defined as the empty word e. Given such a word v, let $Z(v)$ be the endnode of the path represented by v, i.e., if $s(v)$ is the value of the binary number corresponding to v (e.g. $s(000101)=(101)_2=5$), then $Z(v) := 2^{|v|}+s(v)$.

3. Distributional Aspects

Let for $v \in \{0,1\}^*$ as the path to be followed $H(N,v)$ be the set of all $x \in H(N)$ such that the N-th component x_N does indeed follow v. Hence $x \in H(N,v)$ holds iff

$$x_N > \min\{x_{Z(v_{j-1}0)}, x_{Z(v_{j-1}1)}\} = x_{Z(v_j)}$$

holds for $1 \leqq j \leqq |v|$. Thus if $x \in H(N,v)$, then x_N will be interchanged with $x_{Z(v_j)}$ for $1 \leqq j \leqq |v|$, but it will *not* be interchanged with any son of $x_{Z(v)}$, provided there is one. Now call v a *viable path* for x_N iff $H(N,v) \neq \emptyset$, and denote the set of all viable paths by $\mathscr{V}_N$. Note that not every path in the tree corresponding to $\{1, \ldots, N\}$ starting at 1 can be a viable path for N: any part $b_{n-1} \ldots b_j$ of the special path excluding $\lfloor N/2 \rfloor$ *must not* be a viable path. This is so because the heap property implies that

$$x_N > x_{(1 b_{n-1} \ldots b_j)_2}$$

is true for $1 \leqq j \leqq n-1$, hence if x_N would stop at $(1 b_{n-1} \ldots b_j)_2$, this would imply

$$x_N < x_{(1 b_{n-1} \ldots b_j b_{j-1})_2},$$

which is clearly a contradiction. On the other hand, $b_{n-1} \ldots b_1$ constitutes a viable path, and it is easily seen that $v \in \mathscr{V}_N$ holds iff both $Z(v) \leqq N-1$, and $v \notin \{b_{n-1} \ldots b_i;\ 2 \leqq i \leqq n-1\}$ are true. Note that $H(N)$ equals the disjoint union of $H(N,v)$ for $v \in \mathscr{V}_N$.

The action of the algorithm on $H(N,v)$ is described above; denote this action as T_N, hence T_N maps $H(N)$ into $R(N-1)$, where

$$R(N-1):=\{(x_1,\dots,x_{N-1},y);(x_1,\dots,x_{N-1})\in H(N-1),0\leqq y\leqq x_1\}.$$

In order to describe the distributional effects of T_N, some subsets of $R(N-1)$ will be needed; these subsets will depend on viable paths for N and will shown to be the bijective images of the sets of the form $H(N,v)$. To begin with, let

$$R(N-1,b_{n-1}\dots b_1):=R(N-1),$$

and if v is a viable path which does not stop in $\lfloor N/2\rfloor$, then

$$R(N-1,v):=\{(x_1,\dots,x_{N-1},y)\in R(N-1);x_{Z(v)}>x_{\lfloor N/2\rfloor}\}$$

is defined. Consider for example the case $N=6$, and $X:=(1/5,\ 1/3,\ 1/4,\ 1/1,\ 1/2,\ 1/6)$, then $X\in R(5,v)$ for $v=0$, 00, 01, 1 (note that $R(5,1)=R(5)$). X is displayed in the following figure with the mentioned paths indicated by heavier lines. It

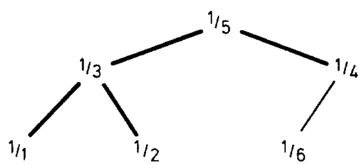

Fig. 2

will be seen in a moment that X lies in the image of exactly the classes $H(N,0)$, $H(N,00)$, $H(N,01)$, and $H(N,1)$ under T_N.

3.1 Lemma. *Given $v\in\mathscr{V}_N$, T_N: $H(N,v)\to R(N-1,v)$ is a bijection.*

Proof. 1. It is easily seen that T_N is one to one, when restricted to $H(N,v)$ since T_N operates on all elements in $H(N,v)$ as the same permutation of coordinates. The crucial step is to show that T_N is onto, and this is done by constructing a preimage for an arbitrarily chosen $y\in R(N-1,v)$ and showing that $x\in H(N,v)$. Define x in the following manner:

$$\begin{aligned} x_1&:=y_N,\\ x_{Z(v_{j+1})}&:=y_{Z(v_j)}; \quad 0\leqq j\leqq|v|-1,\\ x_N&:=y_{Z(v)}\\ x_i&:=y_i, \quad \text{otherwise.}\end{aligned}$$

Then $T_N(x)=y$, and x is a member of $H(N)$. For if there exists an index i such that $x_{\lfloor i/2\rfloor}>x_i$, at least one of the indices must be in $A:=\{Z(v_j);\ 0\leqq j\leqq|v|\}\cup\{N\}$, since outside of this set nothing is changed. Now working through all possible combinations (viz., $\{\lfloor i/2\rfloor,i\}\subset A$, $\lfloor i/2\rfloor\in A\not\ni i$, $i\in A\not\ni\lfloor i/2\rfloor$) demonstrates that the assumption leads to a contradiction, hence shows that $x\in H(N)$. Moreover x is easily seen to be a member of $H(N,v)$, since $y_{Z(v)}>y_{\lfloor N/2\rfloor}$ iff $x_N>x_{\lfloor N/2\rfloor}$, provided $v\neq b_{n-1}\dots b_1$. $\square$

An immediate consequence of 3.1 is that we may say for every $z \in R(N-1)$ how many inverse images it has under T_N. If $M(z)$ counts the number of viable paths $v \in \mathscr{V}_N$ such that $z \in R(N-1, v)$, thus if

$$M(z) := \operatorname{card}(\{v \in \mathscr{V}_N; z \in R(N-1, v)\}),$$

then z has exactly $M(z)$ inverse images. From this observation the distribution of T_N is derived. Note that $M(z)$ may be written as

$$M(z) = \sum \{1(R(N-1, v); z); v \in \mathscr{V}_N\},$$

where $1(A; \cdot)$ is the indicator function for the set A.

3.2 Proposition. *Given a Borel set $A \subset R(N-1)$, let*

$$\mu(A) := \operatorname{Prob}(T_N \in A)$$

be the probability that the output of the algorithm is an element of A. Then

$$\mu(A) = \chi_N \cdot \int_A M(z)\, dz \tag{4}$$

holds.

Proof. 1. We demonstrate a seemingly sharper result, viz., that

$$\int_{R(N-1)} f\, d\mu = \chi_N \cdot \int_{R(N-1)} f(z) M(z)\, dz \tag{5}$$

holds for any f: $R(N-1) \to \mathbb{R}$ which is measurable and bounded. Since the indicator function $1(A; \cdot)$ of the Borel set A has these properties, and since plainly

$$\int 1(A; \cdot)\, d\mu = \mu(A),$$

the desired equality (4) is obtained from (5).
2. We have postulated that

$$\operatorname{Prob}(B) = \chi_N \cdot \lambda^N(B)$$

holds for every Borel set $B \subset H(N)$, and from this we have

$$\begin{aligned}\int_{R(N-1)} f\, d\mu &= \chi_N \cdot \int_{H(N)} f(T_N(x))\, dx \\ &= \sum \{\chi_N \cdot \int_{H(N-1, v)} f(T_N(x))\, dx; v \in \mathscr{V}_N\} \\ &= \sum \{\chi_N \cdot \int_{R(N-1, v)} f(z)\, dz; v \in \mathscr{V}_N\},\end{aligned}$$

since by Lemma 3.1 $T_N[H(N, v)] = R(N-1, v)$. Hence the Change of Variable Formula implies

$$\int_{H(N,v)} f(T_N(x))\, |\det T_N'(x)|\, dx = \int_{R(N-1,v)} f(z)\, dz.$$

Now $|\det T_N'(x)|$ is always equal to 1, because T_N only permutes coordinates, thus has the determinant of a permutation matrix as its Jacobian. Observing

$$\int_{R(N-1,v)} f(z)\,dz = \int_{R(N-1)} f(z)\,1(R(N-1,v);z)\,dz$$

and taking the definition of M as well as the additivity of the integral into account, the equality (5) is proved. □

Denote by $\mathscr{T}(x)$ the number of interchanges which are necessary to reconstruct the heap property when the root of $x \in H(N)$ is deleted. Thus $\mathscr{T}(x)$ equals $|v|$ whenever $x \in H(N,v)$, and we have for the expectation $\mathbb{E}(\mathscr{T})$

$$\mathbb{E}(\mathscr{T}) = \chi_N \cdot \sum \{|v|\,\lambda^N(H(N,v)); v \in \mathscr{V}_N\}. \tag{$6_{\mathscr{T}}$}$$

If $\mathscr{S}(x)$ is the number of comparisons which are done in reconstructing the heap property for $x \in H(N)$, and if $x \in H(N,v)$ for $v \in \mathscr{V}_N$, such that $Z(v)$ has i sons $(0 \leqq i \leqq 2)$ in the tree corresponding to $\{1, \ldots, N-1\}$, then $\mathscr{S}(x) = 2|v| + i$ holds. Hence Eq. $(6_{\mathscr{T}})$ implies

$$\begin{aligned} \mathbb{E}(\mathscr{S}) = 2 \cdot \mathbb{E}(\mathscr{T}) &+ b_0 \cdot \chi_N \cdot \lambda^N(H(N, b_{n-1} \ldots b_1)) \\ &+ 2 \cdot \chi_n \cdot \sum \{\lambda^N(H(N,v));\ Z(v) \text{ has two sons w.r.t. } \{1, \ldots, N-1\}\}. \end{aligned} \tag{$6_{\mathscr{S}}$}$$

The second summand comes from the observation that $b_{n-1} \ldots b_1$ is the only viable path which ends in a node with exactly one son, provided $b_0 = 1$; in case $b_0 = 0$, i.e. if N is even, this path ends in a leaf with respect to $\{1, \ldots, N-1\}$.

From Eqs. (6) it is seen that it will be necessary to compute $\lambda^N(H(N,v))$ for $v \in \mathscr{V}_N$. This will be done in the next section.

4. Computing Path Probabilities

Fix for the moment a viable path $v \in \mathscr{V}_N$. Then from Lemma 3.1 and from the proof of Proposition 3.2 it is clear that

$$\lambda^N(H(N,v)) = \lambda^N(R(N-1,v))$$

holds. Now fix $y \in [0,1]$, and define the set $R(N-1,v)(y)$ by

$$(x_1, \ldots, x_{N-1}) \in R(N-1,v)(y)$$

iff

$$(x_1, \ldots, x_{N-1}, y) \in R(N-1,v).$$

Fubini's Theorem ([10], p. 150f) implies

$$\lambda^N(H(N,v)) = \int_0^1 \lambda^{N-1}(R(N-1,v)(y))\,dy.$$

Let

$$K(N-1,v) = \{(x_1, \ldots, x_{N-1}) \in H(N-1); x_{Z(v)} > x_{\lfloor N/2 \rfloor}\}$$

be the set of all $(N-1)$-heaps in which the label of node $Z(v)$ is greater than the label of $\lfloor N/2 \rfloor$; for the sake of completeness

$$K(N-1, b_{n-1} \dots b_1) := H(N-1)$$

is defined. Then

$$\begin{cases} K(N-1,v) \to R(N-1,v)(y) \\ (x_1, \dots, x_{N-1}) \mapsto (y + x_1(1-y), \dots, y + x_{N-1}(1-y)) \end{cases}$$

constitutes a differentiable homeomorphism with Jacobian $(1-y)^{N-1}$.

Thus we have

$$\begin{aligned} \lambda^N(R(N-1,v)) &= \lambda^{N-1}(K(N-1,v)) \cdot \int_0^1 (1-y)^{N-1}\,dy \\ &= \frac{1}{N} \lambda^{N-1}(K(N-1,v)), \end{aligned}$$

and the task of computing $\lambda^N(R(N-1,v))$ is reduced to that of computing $\lambda^{N-1}(K(N-1,v))$. In case $v = b_{n-1} \dots b_1$ (or $b_{n-1} \dots b_1 0$, provided N is odd) this is done easily, since the respective sets equal $H(N-1)$, which has χ_{N-1}^{-1} as its $(N-1)$-dimensional volume. In the general case, however, some hairy computations will be needed. We will proceed as follows: to begin with, we fix labels on the special path with respect to N, but excluding N, and on the nodes of v. With these fixed labels in mind we have a look at the probability for all heaps which may have these labels at the respective nodes; then we will integrate over all possible values which have been fixed and obtain in this way the $(N-1)$-dimensional volume of $K(N-1,v)$ by means of Fubini's Theorem.

By the latter Theorem, we have

$$\begin{aligned} &\lambda^{N-1}(K(N-1,v)) \\ &= \int_{1 \geqq y_1 \geqq \dots \geqq y_n \geqq 0} \lambda^{N-1-n}(\{x \in K(N-1,v); x_{t(N,i)} = y_i \text{ for } 1 \leqq i \leqq n\}). \end{aligned} \tag{7}$$

Thus for a fixed chain $1 \geqq y_1 \geqq \dots \geqq y_n \geqq 0$ the $(N-1-n)$-dimensional volume

$$\lambda^{N-1-n}(\{x \in K(N-1,v); x_{t(N,i)} = y_i \text{ for } 1 \leqq i \leqq n\})$$

of all those heaps in $K(N-1,v)$ that are labeled on their special path by this chain should be known.

Let us have a look at a possibility to obtain a heap labeling in $K(N-1,v)$, when the special path $\{t(N,i);\ 1 \leqq i \leqq n\}$ (thus excluding N, of course) has already a label. First, we cut the path $\{t(N,i);\ 1 \leqq i \leqq n\}$ out of the tree representing $\{1, \dots, N-1\}$ and are left with a forest of smaller trees. These trees are all rooted at brothers of the special nodes $t(N,n-1)$, ..., $t(N,1)$, and of $t(N,0)$, if N is odd. The following figure might help to clarify things (the nodes $t(N,1)$, ..., $t(N,n)$ are marked by heavier dots, the respective brothers by asterisks, and the path v in question by heavy lines). Now consider a tree in this residual forest with the property that no part of v is in it. This tree may be

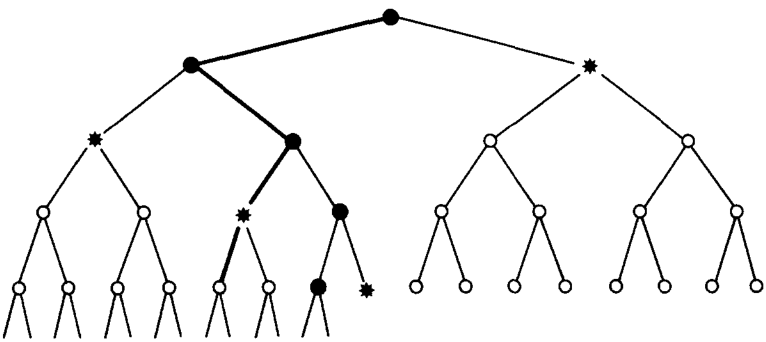

Fig. 3

labeled as a heap under no other constraint, than that the label of its root is not smaller, than the label of its father (in the 'big' tree). All trees but one are labeled in this manner. The remaining one contains that part of the path v which lies not on the special path. In order not to come in conflict with the heap property the root is labeled as are the roots of the other trees, viz., with a number not smaller than the label of its father. But in order not to come into conflict with the requirement that the heap thus produced is a member of $K(N)-1, v)$, the endnode is assigned a number that is not smaller than y_1. Having in mind these restrictions and repeating the considerations from above, this yields a labeling of this subtree, too. Now we must make sure that the described labeling procedure may be traced probabilistically.

It will be helpful to have the following formulae at our disposal

$$\lambda^k(\{(x_1, \ldots, x_k) \in H(k);\ z \leqq x_1\}) = \frac{(1-z)^k}{\chi_k} \tag{8}$$

with $0 \leqq z \leqq 1$. In case $z=0$, Eq. (8) reflects the fact that heap construction with Floyd's algorithm preserves uniform distribution, in case $z=1$ both sides are zero, hence the case $0<z<1$ must be considered. Since the set in question is the image of $H(k)$ under the map

$$(x_1, \ldots, x_k) \mapsto (z+(1-z)\cdot x_1, \ldots, z+(1-z)\cdot x_k),$$

the Jacobian of which is $(1-z)^k$, the equality follows immediately from the Change of Variables Formula.

If $\varepsilon(i) \geqq 0$, $\hat{\varepsilon}(i) := \varepsilon(i)+1$, $0 \leqq \alpha$, $\beta \leqq 1$, we have

$$\int_\beta^1 dx_1 \int_\alpha^{x_1} dx_2 \ldots \int_\alpha^{x_{k-1}} dx_k \prod_{i=1}^{k} (1-x_i)^{\varepsilon(i)} = \sum_{a=1}^{k} \frac{(-1)^{a-1}(1-\beta)^{\hat{\varepsilon}(1)+\ldots+\hat{\varepsilon}(a)}(1-\alpha)^{\hat{\varepsilon}(a+1)+\ldots+\hat{\varepsilon}(k)}}{\left[\prod_{r=1}^{a} \sum_{i=r}^{a} \hat{\varepsilon}(i)\right] \cdot \left[\prod_{t=a+1}^{k} \sum_{i=a+1}^{t} \hat{\varepsilon}(i)\right]}. \tag{9}$$

This formula may be proved by induction on k.

In carrying out the program sketched above, it will be necessary to compute the contribution of certain subtrees to the measure $\lambda^{N-1}(K(N-1, v))$. This is done in the next lemma.

4.1 Lemma. *Let $P_1, \ldots, P_r$ be a path in a complete binary tree $\mathcal{T}$ of depth s (hence $\mathcal{T}$ has $m := 2^s - 1$ nodes) such that P_1 is the root of $\mathcal{T}$. Then the m-dimensional Lebesgue measure of all those heap labelings for $\mathcal{T}$ such that the label for P_1 is not smaller, than a prescribed $Y_1 \in [0, 1]$ and the label for P_r is not smaller, than a given $Y_2 \in [0, 1]$ equals*

$$\frac{\prod_{i=1}^{r} (2^{s+1-i} - 1)}{\chi_m} \cdot \sum_{a=1}^{r} \frac{(1-Y_2)^{2^{s-r+a}-1} \cdot (1-Y_1)^{2^{s-r}(2^r - 2^a)}}{(2^{s-r+a} - 1)\, 2^{(s-r)(r-1)} \prod_{\substack{t=1 \\ t \neq a}}^{r} (2^t - 2^a)}.$$

Proof. 1. Remember that heaps must not have components outside the closed interval $[0, 1]$. Let y_i be a label to node P_i such that $0 \leqq y_1 \leqq \ldots \leqq y_r \leqq 1$, and let Q_i be the brother of P_i, $2 \leqq i \leqq r$. Then the Lebesgue measure of all possible heap labels which respect $y_1, \ldots, y_r$ equals

$$\prod_{i=1}^{r} g(m, P_i) \cdot \frac{\prod_{i=1}^{r-1} (1-y_i)^{g(m, Q_{i+1})}}{G(m, P_1)} \cdot (1-y_r)^{g(m, P_r) - 1}. \tag{$*$}$$

This is easily deduced from Eq. (8), since only the subtrees of $\mathcal{T}$, which are rooted at the nodes Q_i, and the subtrees of $\mathcal{T}$ which are rooted at the left, or the right son of P_r, have to be taken into consideration. Moreover it is noted that

$$G(m, P_1) = \prod_{i=1}^{r-1} g(m, P_1) \cdot \prod_{i=2}^{r} G(m, Q_i) \cdot G(m, P_r)$$

holds.

2. The equality in question is now proved by an application of Eq. (9) with $\beta = Y_2$, $\alpha = Y_1$, $x_i = y_{r+1-i}$, $\varepsilon(i) = g(m, Q_{r+2-i})$ if $i \geqq 2$ and $\varepsilon(1) = g(m, P_r) - 1$. The looked for measure equals the integral over $(*)$ with respect to all $y_1, \ldots, y_r$, which satisfy

$$\begin{aligned} Y_2 &\leqq y_r \leqq 1, \\ y_{r-1} &\leqq y_r, \\ &\vdots \\ y_1 &\leqq y_2, \\ Y_1 &\leqq y_1, \end{aligned}$$

or, equivalently

$$\begin{aligned} Y_2 &\leqq y_r \leqq 1, \\ Y_1 &\leqq y_{r-1} \leqq y_r, \\ Y_1 &\leqq y_{r-2} \leqq y_{r-1}, \\ &\vdots \\ Y_1 &\leqq y_1 \leqq y_2. \end{aligned}$$

In evaluating sums involving $\hat{\varepsilon}(i)$, it is to be noted that

$$g(m, P_i) = g(m, Q_i) = 2^{s+1-i} - 1. \quad \square$$

Now let us return to our fixed path v, and let us state some characteristics associated with it:

* $P(v)$ is the smallest index j such that $b_{n-1} \dots b_j$ is a prefix of v,
* $\rho(v)$ is the number of nodes which are on v but not on the special path $b_{n-1} \dots b_0$,
* $R(v)$ is the number of nodes which are on the path from the endnode of v to a leaf including that endnode.

Hence we see that

$$R(v) = P(v) - \rho(v) + b_{P(v)-1}.$$

The viable path v may be written as

$$v = b_{n-1} \dots b_{P(v)} d_1 \dots d_{\rho(v)},$$

where $d_1 \neq b_{P(v)-1}$. Note that the case $P(v) = n$ is not excluded. Moreover

$$\sigma(v) := R(v) + \rho(v) - 1$$

is the number of nodes in a path starting from $(1b_{n-1} \dots b_{P(v)} d_1)_2$ to a leaf (including its start node), thus the tree rooted at $(1b_{n-1} \dots b_{P(v)} d_1)_2$ has $2^{\sigma(v)} - 1$ nodes. Consider as an example $N = 52$ and the path $v = 1001$:

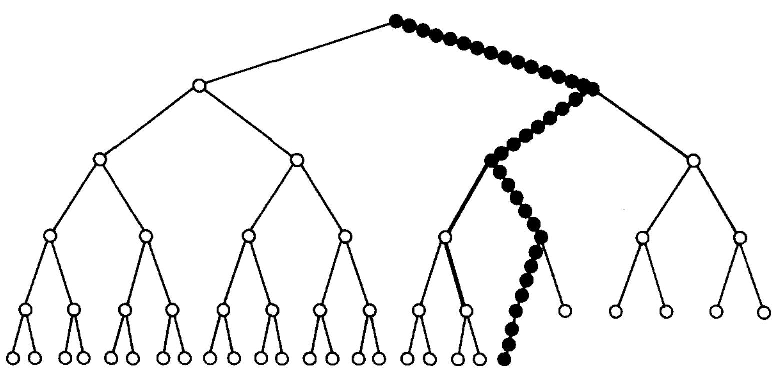

Fig. 4

In the displayed tree, the special path 10100 is characterised by small dots, the path v by heavier lines. We have $n = 5$, $P(v) = 3$ (since 10 is a prefix of v, but 101 is not), $R(v) = 2$, $\sigma(v) = 3$, $\rho(v) = 2$.

Now let T_i be that son of the special node $t(N, i)$ which is different from $t(N, i-1)$, $2 \leqq i \leqq n$; in case N is odd, the node $t(N, 0)$ has a brother, too, hence T_1 is defined. In case N is even, however, care must be observed since T_1 is not defined, but for the sake of convenience

$$g(N, T_1) := 0,$$
$$G(N, T_1) := 1$$

is defined in this case.

Now let y_i be a fixed label of $t(N, i)$ for $1 \leqq i \leqq n$ such that $1 \geqq y_1 \geqq \ldots \geqq y_n \geqq 0$. The contribution of the subtree rooted at the node T_i to the volume of all those heaps which have $y_1, \ldots, y_n$ as the labels is

$$\frac{(1-y_i)^{g(N-1, T_i)}}{G(N-1, T_i)},$$

in case $i \neq P(v)$, and

$$\frac{\prod_{i=1}^{\rho(v)} (2^{\sigma(v)+1-i}-1)}{G(N-1, T_{P(v)})} \sum_{a=1}^{\rho(v)} \frac{(1-y_1)^{2^{\sigma(v)-\rho(v)+a}-1} \cdot (1-y_{P(v)})^{2^{\sigma(v)-\rho(v)}(2^{\rho(v)}-2^a)}}{(2^{\sigma(v)-\rho(v)+a}-1) \cdot 2^{(\sigma(v)-\rho(v))(\rho(v)-1)} \cdot \prod_{\substack{t=1 \\ t \neq a}}^{\rho(v)} (2^t - 2^a)}$$

in case $i = P(v)$. In the latter case the subtree rooted at $T_{P(v)}$ takes over the rôle of $\mathcal{T}$ in Lemma 4.1, and the path mentioned there is that part of v which is not on $b_{n-1} \ldots b_1$. Consequently the heap labelings that are considered must satisfy the condition that the label of the root $T_{P(v)}$ is not smaller, than $y_{P(v)}$ (otherwise the heap condition would be violated), and that the label of the endnote $(1 b_{n-1} \ldots b_{P(v)} d_1 \ldots d_{\rho(v)})_2$ is not smaller, than y_1 (otherwise the defining condition for $K(N-1, v)$ would be violated). The case $i \neq P(v)$ is treated by means of Eq. (8), since any labeling of the subtree rooted at T_i which satisfies the heap condition is constrained only by the property that the label of T_i must not be smaller than y_i is. In passing note that

$$g(N-1, T_i) = g(N, T_i),$$
$$G(N-1, T_i) = G(N, T_i)$$

holds for all i, since no T_i is a special node with respect to $\{1, \ldots, N\}$.

From Fubini's Theorem now it is inferred that

$$\lambda^{N-1}(K(N-1, v)) = \frac{\prod_{i=1}^{\rho(v)} (2^{\sigma(v)+1-i}-1)}{\prod_{i=1}^{n} G(N, T_i)}$$
$$\cdot \sum_{a=1}^{\rho(v)} \left[(2^{\sigma(v)-\rho(v)+a}-1) \cdot 2^{(\sigma(v)-\rho(v))(\rho(v)-1)} \cdot \sum_{\substack{t=1 \\ t \neq a}}^{\rho(v)} (2^t - 2^a) \right]^{-1}$$
$$\cdot \int_0^1 dy_n \int_{y_n}^1 dy_{n-1} \ldots \int_{y_2}^1 dy_1 \prod_{i=1}^{n} (1-y_i)^{\eta(a, i)},$$

where the exponents $\eta(a, i)$ are defined by

$$\eta(a, i) := \begin{cases} g(N, T_1) + 2^{\sigma(v) - \rho(v) + a} - 1, & \text{if } i = 1 \\ 2^{\sigma(v) - \rho(v)}(2^{\rho(v)} - 2^a), & \text{if } i = P(v) \\ g(N, T_i), \quad \text{otherwise} & \end{cases}$$

Now it is not difficult to see that

$$\int_0^1 dx_n \int_{x_n}^1 dx_{n-1} \dots \int_{x_2}^1 dx_1 \prod_{i=1}^n (1 - x_i)^{\gamma(i)} = \prod_{i=1}^n \left[\sum_{j=1}^i \hat{\gamma}(j)\right]^{-1}$$

holds, where $\gamma(i) \geqq 0$, $\hat{\gamma}(i) = \gamma(i) + 1$, as above. In order to evaluate the integral above, note that

$$g(N, T_i) = 2^{i-1+b_{i-1}} - 1$$

and

$$\sum_{i=1}^k 2^{i-1+b_{i-1}} = (1 b_{k-1} \dots b_1 b_0)_2 - 1$$

holds. Since

$$\sum_{i=1}^t \hat{\eta}(a, i) = \begin{cases} 2^{\sigma(v) - \rho(v) + a} + (1 b_{t-1} \dots b_0)_2 - 2, & \text{if } t < P(v) \\ 2^{\sigma(v)} + (1 b_{t-1} \dots b_0)_2 - 1 - 2^{P(v)-1}(1 + b_{P(v)-1}), & \text{if } t \geqq P(v), \end{cases}$$

the Lebesgue measure in question equals

$$\frac{\prod_{i=1}^{\rho(v)} (2^{\sigma(v)+1-i} - 1)}{2^{(\sigma(v) - \rho(v))(\rho(v) - 1)} \cdot \prod_{i=1}^n G(N, T_i)}$$
$$\cdot \prod_{t=P(v)}^n [2^{\sigma(v)} + (1 b_{t-1} \dots b_0)_2 - 1 - 2^{P(v)-1} \cdot (1 + b_{P(v)-1})]^{-1}$$
$$\cdot \sum_{a=1}^{\rho(v)} \left[(2^{\sigma(v) - \rho(v) + a} - 1) \cdot \prod_{\substack{t=1 \\ t \neq a}}^{\rho(v)} (2^t - 2^a)\right.$$
$$\left.\cdot \prod_{t=1}^{P(v)-1} (2^{\sigma(v) - \rho(v) + a} + (1 b_{t-1} \dots b_0)_2 - 2)\right]^{-1}.$$

This expression may be simplified a little bit by noting that both

$$\chi_N = \prod_{i=1}^n G(N, T_i) \cdot \prod_{i=1}^n g(N, t(N, i)),$$

and

$$\sigma(v) = P(v) - 1 + b_{P(v)-1}$$

hold.

Substituting these values, we see that the probability for the path v to be followed does depend only on its characteristics $\sigma(v)$ and $\rho(v)$, and not on, say, the number of times it turns left of right or an another geometrical characterization relative to the tree.

5. Computing the Expectations in Case $N=2^n$

We will assume in this section that $N=2^n$, i.e. that N is the leftmost node on level n. Denote by $p(r,s)$ the probability that x_N follows a path v that has the characteristics $\sigma(v)=s$ and $\rho(v)=r$, hence

$$p(r,s)=N^{-1}\prod_{i=1}^{n}2^i\cdot\prod_{i=1}^{r}(2^{s+1-i}-1)\cdot\prod_{j=s+1}^{n}(2^j-1)^{-1}$$
$$\cdot\,2^{-(s+1)\cdot(r-1)}\cdot\sum_{a=1}^{r}\left[\prod_{j=0}^{s}(2^{s+1-a}+2^j-2)\prod_{\substack{t=1\\t\neq a}}^{r}(2^{-t}-2^{-a})\right]^{-1}$$

holds, since $b_i=0$, $0\leqq i\leqq n-1$. There are 2^{r-1} paths with these characteristics and any of them has length $n-s-1+r$. Thus the expected number $\mathbb{E}(\mathscr{T})$ of interchanges equals

$$\mathbb{E}(\mathscr{T})=\sum_{s=1}^{n-1}\sum_{r=1}^{s}2^{r-1}\cdot(n-s-1+r)\cdot p(r,s)+O\left(\frac{1}{N}\right)$$

since r may take values between 1 and s, and s may range from 1 to $n-1$. The $O(N^{-1})$ term takes that parts of the special path with respect to N that are viable paths into account.

Fix s for the moment, and abbreviate the product

$$\left[\prod_{j=0}^{s}(2^t+2^j-2)\cdot\prod_{j=s+1}^{n}(2^j-1)\right]^{-1}$$

by $f(t)$, then it is not difficult to establish that

$$\sum_{r=1}^{s}2^{r-1}(n-s+r-1)\,p(r,s)$$

equals

$$B_1(s,n)-B_2(s,n),$$

where

$$B_1(s,n):=N^{-1}\cdot\prod_{i=1}^{n}(1-2^{-i})^{-1}\cdot(2^s-1)\cdot\prod_{j=1}^{s}(1-2^{-j})\cdot(n-1)$$

and

$$B_2(s,n):=N^{-1}\cdot\prod_{i=1}^{n}2^i\cdot\left\{2^s\prod_{i=1}^{s}(1-2^{-i})\sum_{a=2}^{s}f(a)\right.$$
$$\left.\cdot\left[\prod_{t=1}^{s-a}(1-2^{-t})\cdot(1-2^{-(a-1)})\right]^{-1}\right\}.$$

In establishing this, we have made use of the equalities

$$\sum_{t=0}^{k-1} 2^t \prod_{j=1}^{t} \frac{1-2^{-k}2^j}{1-2^j} = \delta_{1,k},$$
$$\sum_{t=0}^{k-1} t \cdot 2^t \prod_{j=1}^{t} \frac{1-2^{-k}2^j}{1-2^j} = \left(-\prod_{j=1}^{k-2} (1-2^{-j})\right)(1-\delta_{1,k})$$

(δ denoting Kronecker's symbol), which may be proved by means of [8], (89.18.2). Now

$$\sum_{s=1}^{n-1} B_2(s,n) = 0(1)$$

is not difficult to prove, thus

$$\sum_{s=1}^{n-1} B_1(s,n)$$

remains to be evaluated. For this, the generating function

$$\mathscr{G}(z) := \sum_{n=1}^{\infty} \sum_{s=1}^{n-1} \hat{B}_1(s,n)\, z^n$$

will be explored where

$$\hat{B}_1(s,n) := \sum_{s=1}^{n-1} (2^s-1) \prod_{j=1}^{s} (1-2^{-j}),$$

and in order to investigate $\mathscr{G}$, we consider the generating function

$$\mathscr{D}(z) := \sum_{s=1}^{\infty} \prod_{j=1}^{s} (1-2^{-j})\, z^s$$

for $\left(\prod_{j=1}^{s} (1-2^{-j})\right)_{s \geqq 1}$. $\mathscr{D}$ is a special case of a basic hypergeometric series, in Bailey's notation $\mathscr{D}(z) = {}_1\phi_0(\frac{1}{2}; z)$ with $q=\frac{1}{2}$ (see [2], Chapter VIII), and our considerations will give an analytic continuation of this basic hypergeometric series into the complex plane with $\{2^s; s \geqq 0\}$ omitted. Applying the equations (89.18.4) and then (89.18.2) from [8] and using Euler's partition formula ([9], Exercise 5.1.1.16), we infer that

$$\mathscr{D}(z) = \sum_{t=0}^{\infty} \frac{z^t}{(2^t - z)\, 2^{t(t-1)/2} \prod_{j=1}^{t} (2^j - 1)} \prod_{j=t+1}^{\infty} (1 - z 2^{-j}) - 1$$

holds, provided $|z|<1$. In case $z \neq 2^s$, $s \in \mathbb{N} \cup \{0\}$, the quotient criterion shows that this series converges.

Now

$$\mathscr{G}(z) = \frac{z}{1-z} (\mathscr{D}(2z) - \mathscr{D}(z)),$$

hence $\mathscr{G}$ has simple poles at $z=2^s$, $s\in\mathbb{N}\{-1,0\}$ (what looks like a double pole at $z=1$ turns out to be a simple one); and Darboux's theorem ([3], Theorem 4) implies that

$$\sum_{s=1}^{n-1} \hat{B}_1(s,n)=F^{-1}2^n-F^{-1}+o(1), \qquad \text{as } n\to\infty$$

holds, where

$$F:=\prod_{j=1}^{\infty}(1-2^{-j})^{-1}.$$

By means of the identities [8], (89.18.3) and (89.18.2) the product $\prod_{j=1}^{k}(1-2^{-j})^{-1}$ may be expressed using F, viz.

$$\prod_{j=1}^{k}(1-2^{-j})^{-1}=F\cdot(1-\gamma_k),$$

where

$$\gamma_k:=\sum_{j=1}^{\infty}[2^{kj}\cdot 1\cdot 3\cdot\ldots\cdot(2^j-1)]^{-1}.$$

Thus

$$\sum_{s=1}^{n-1} B_1(s,n)=n-1+o(1),$$

and we have proved

5.1 Theorem. *In case $N=2^n$ the algorithm Rootdeletion requires*

$$n-1+o(1)$$

interchanges on the average, as $n\to\infty$. □

Let us turn our attention to the number $\mathscr{S}$ of necessary comparisons. The Eq. (6) above tells us that we should have a look at the probabilities for those paths which end in a node that has exactly two sons with respect to the tree representing $\{1,\ldots,N-1\}$. No such path v will have the characteristic $\sigma(v)=1$, or $\rho(v)=\sigma(v)$, thus

$$\sum_{s=2}^{n-1}\sum_{r=1}^{s-1} 2^{r-1}p(r,s) \tag{10}$$

should be known. An argumentation very close to that dealing with the divided difference above shows that

$$\begin{aligned}\sum_{r=1}^{s-1} 2^{r-1}p(r,s)=N^{-1}\cdot\prod_{j=1}^{n}2^j\\ \cdot\Bigg\{f(1)\cdot(2^s-1)-\sum_{a=1}^{s}f(s+1-a)\\ \cdot 2^{2s-a}\prod_{i=1}^{s}(1-2^{-i})\cdot\Bigg[\prod_{\substack{t=1\\t\neq a}}^{s}(1-2^{t-a})\Bigg]^{-1}\Bigg\}\end{aligned} \tag{10'}$$

holds, where f is as above. The evaluation of the latter sum requires the knowledge of

$$\sum_{a=1}^{s} 2^{-a}\left[\prod_{j=0}^{s}(2^{s+1-a}+2^j-2)\cdot\prod_{\substack{t=1\\t\neq a}}^{s}(1-2^{t-a})\right]^{-1}$$
$$=\frac{\theta}{2\pi i}\int_{\Gamma} z\cdot\left[(2^{s+1}\cdot z)^s\prod_{t=1}^{s+1}(1-2^t\cdot z)\right]^{-1}dz$$
$$\text{(for some real } \theta,\ |\theta|<1) \tag{11}$$

where the contour Γ of integration is given in the following figure. Here $\varepsilon>0$ is arbitrary, and $0<\eta<2^{-(s+1)}$ holds. By making $M\to\infty$, applying Jordan's Lemma

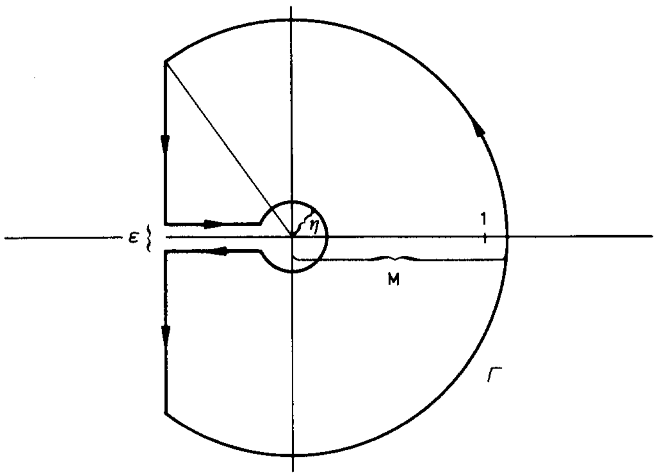

Fig. 5

([11], 6.222, p. 115) together with the additivity of the integral, we see that the latter integral reduces to

$$-\frac{\theta}{2\pi i}\int_{|z|=\eta} 2^{-(s+1)\cdot s}\cdot z^{-(s-1)}\left[\prod_{t=1}^{s+1}(1-2^t\cdot z)\right]^{-1}dz.$$

Since 0 is the only singularity of the integrand (in fact, it is a pole of order $s-1$), the integral equals the residuum of it at zero. From [8], (89.18.2), it is seen that

$$\prod_{t=1}^{s+1}(1-2^t\cdot z)^{-1}=\sum_{k\geq 0}2^k z^k\prod_{j=1}^{k}\frac{2^{s+j}-1}{2^j-1},$$

holds and from the Cauchy-d'Alembert test ([11], p. 30) we infer that the series converges in $|z|<2^{-(s+1)}$. Consequently, the integral

$$\frac{1}{2\pi i}\int_{|z|=\eta} z^{-(s-1)}\prod_{t=1}^{s+1}(1-2^t z)^{-1}\,dz$$

equals

$$2^{s-2}\prod_{j=1}^{s-2}\frac{2^{s+j}-1}{2^j-1}=F\cdot\frac{(1-\gamma_s)(1-\gamma_{s-2})}{(1-\gamma_{2s-2})}\cdot 2^{s^2-s-2}.$$

Hence the sum in Eq. (11) is $O(2^{-2s})$. Plugging this into Eq. (10), it is now clear that this sum equals

$$N^{-1}\cdot(1-\gamma_s)^{-1}\cdot(1-\gamma_n)\cdot(2^s-1)+N^{-1}\cdot O(2^{-s}).$$

This implies that we have for the sum in Eq. (10)

$$\begin{aligned}\sum_{s=2}^{n-1}\sum_{r=1}^{s-1}2^{r-1}\cdot p(r,s)&=N^{-1}\cdot(1-\gamma_n)\cdot(N+O(1))\\&=1+O\left(\frac{1}{N}\right).\end{aligned}$$

Since the latter sum equals the contribution of the probabilities for all those paths which have exactly two endnodes in the tree corresponding to $\{1,\ldots,N-1\}$, we may conclude from Eq. $(6_{\mathscr{S}})$ and Theorem 5.1:

5.2 Theorem. *In case $N=2^n$, the algorithm Rootdeletion requires*

$$2n-1+o(1)$$

comparisons on the average. □

An attempt to do a similar asymptotic analysis for other heapsizes than that of powers of two seems to be worthwhile, but shows considerable technical difficulties. In order to get an impression of the size of the numbers involved, the following table lists some expected values for interchanges.

N	$\mathbb{E}(\mathscr{T})$	N	$\mathbb{E}(\mathscr{T})$	N	$\mathbb{E}(\mathscr{T})$
120	5.639309	128	5.701212	255	6.636096
121	5.531967	129	5.597580	256	6.764067
122	5.646428	130	5.729245	257	6.682071
123	5.538227	250	6.717618	258	6.765042
124	5.668911	251	6.631717	259	6.681986
125	5.561248	252	6.735185	260	6.787195
126	5.671459	253	6.648512		
127	5.563767	254	6.722797		

6. Concluding Remarks

It has been assumed so far that the inputs for the algorithm Rootdeletion come from all vectors with components in the closed unit interval that have the heap property, and that these vectors are uniformly distributed. In calculating the path probabilities the corresponding measure has been used extensively. Hence it is surprising to see that a large class of probability distribution will display the same expected behavior, when used as input distributions. In order to make this statement more precise, let us introduce a *symmetric model* (cp. [6]). Such a model consists of a base set A and of a distribution μ with the following properties

(a) $A \subset \mathbb{R}^N$ is a Borel set

(b) if $x \in A$, then $(x_{p(1)}, \dots, x_{p(N)}) \in A$ for every permutation p of $\{1, \dots, N\}$

(c) μ is a probability measure on (the Borel sets of) A such that

$$\mu(B) = \int_B F(x)\, dx,$$

where the density F has the property that $F(x_1, \dots, x_N) = F(x_{p(1)}, \dots, x_{p(N)})$ for every $x \in A$ and every permutation p of $\{1, \dots, N\}$.

Examples for symmetric models include

(i) $A =]a, b[^N$, $F(x_1, \dots, x_N) = \prod_{i=1}^{N} g(x_i)$

for some $-\infty \leqq a < b \leqq +\infty$, $\int_a^b g(x)\, dx = 1$ (this corresponds to the case of independent and identically distributed random variables),

(ii) A is the unit simplex, and μ is a multivariate Dirichlet distribution,

(iii) $A = \mathbb{R}^N$, and μ is a multivariate Cauchy distribution.

On the finite side, we have the usual model of all permutations of $\{1, \dots, N\}$ that enjoy the heap property, assigning every of these permutations the probability χ_N^{-1} (cp. [9], Theorem 5.2.3.H, Eq. (16) on p. 154). It turns out that the following holds.

6.1 Theorem. *Let A, μ be a symmetric model, $H_A(N) := \{x \in A;\ x \text{ is a heap}\}$.*

a) $\mu(H_A(N)) = \chi_N^{-1}$

b) *taking the inputs to Rootdeletion either from $H_A(N)$ with input distribution $\chi_N \cdot \mu$ or from the set of all permutations of $\{1, \dots, N\}$ having the heap property under uniform distribution will result in the same expected number of interchanges and comparisons.*

Proof. See [9], Corollary 4.9, Sect. 5. □

Finally, some further work is to be suggested. Since Rootdeletion affects uniform (or symmetric) distribution, the way to tackle the problem pursued here can be only the first step in analyzing the selection phase of heapsort. Thus a suggestion for further work in this area along the lines sketched here might be to study the probability distributions arising in iterated rootdeletion carefully and to derive from these distributions informations concerning the expected performance of heap selection. Another and more immediate suggestion addresses the higher moments of the random variables involved.

Acknowledgements. The author would like to thank the referees for helping to improve the paper.

References

1. Aho, A.V., Hopcroft, J.E., Ullman, J.D.: The Design and Analysis of Computer Algorithms. Reading: Addison-Wesley, MA 1974
2. Bailey, W.N.: Generalized Hypergeometric Series. Cambridge: University Press 1935
3. Bender, E.A.: Asymptotic Methods in Enumeration. SIAM Review 16 (1974) 485-515
4. Doberkat, E.E.: Some Observations on the Average Performance of Heapsort - Preliminary Report. 21st IEEE FOCS, Syracuse, N.Y. 1980, 229-237
5. Doberkat, E.E.: Inserting a New Element into a Heap. BIT 21 (1981) 255-269
6. Doberkat, E.E.: Continuous Models that are Equivalent to Randomness for the Analysis of Many Sorting Algorithms. Preprint 1981, submitted for publication
7. Habermann, A.N.: Introduction to Operating System Design. Chicago: Science Research Associates 1976
8. Hansen, E.R.: A Table of Series and Products. Prentice-Hall: Englewood Cliffs 1975
9. Knuth, D.E.: The Art of Computer Programming - Vol. 3, Sorting and Searching. Reading: Addison-Wesley, MA 1973
10. Rudin, W.: Real and Complex Analysis. New Delhi: Tata McGraw Hill 1974 (second edition)
11. Whittaker, E.T., Watson, G.N.: A Course of Modern Analysis (Fourth Edition). Cambridge: University Press 1927

Received April 22, 1982