COMPARISONS IN HOARE'S FIND ALGORITHM

PETER KIRSCHENHOFER AND HELMUT PRODINGER

Abstract. We study the number of comparisons in Hoare's Find algorithm. Using trivariate generating functions, we get an explicit expression for the variance of the number of comparisons, if we search for the jth element in a random permutation of n elements. The variance is also asymptotically evaluated under the assumption that j is proportional to n. Similar results for the number of passes (recursive calls) are given, too.

1. Introduction

The traditional sorting method Quicksort [11, 6] uses a pivot element and a partitioning strategy in order to bring the pivot element into its correct position and having only smaller elements to the left and larger elements to the right, so that the procedure can be applied recursively to the smaller subfiles. There is a relation between Quicksort and binary search trees, as can be seen in the example in Figure 1. Subfiles are indicated by rectangular boxes, and the leftmost element will be the pivot element.

The corresponding binary search tree is depicted in Figure 2.

If one only wants the find the jth ranked element (which is j, since we assume that the elements in question are $\{1, 2, \dots, n\}$), one uses the same partitioning strategy, but follows only the path which contains the sought element. This is the same as to say that one goes down recursively in only *one* subfile. This procedure is called Hoare's Find algorithm [5].

There is some renewed interest in Hoare's find algorithm, as can be seen from the following references, [12, 14, 8, 15]. For some general background on binary search trees we cite [13]; see also [7].

Knuth [10] has already computed the average number of comparisons $C_{n,j}$. For this, it is assumed that every permutation of $\{1, 2, \dots, n\}$ is equally likely, and that the partitioning phase needs $n-1$ comparisons. Then there is the recursion

$$C_{n,j} = n - 1 + \frac{1}{n} \sum_{1 \le k < j} C_{n-k,j-k} + \frac{1}{n} \sum_{j < k \le n} C_{k-1,j} \ . \tag{1}$$

The solution is

$$C_{n,j} = 2\Big(n + 3 + (n+1)H_n - (j+2)H_j - (n+3-j)H_{n+1-j}\Big) \ . \tag{2}$$

Here, $H_n = \sum_{1 \le k \le n} \frac{1}{k}$ is the nth harmonic number; we will also need harmonic numbers of order two, $H_n^{(2)} = \sum_{1 \le k \le n} \frac{1}{k^2}$.

More generally, one can also set up an analogous recursion for the *probability generating functions*, viz.

$$\mathcal{C}_{n,j}(v) = \frac{v^{n-1}}{n} \left[\sum_{1 \le k < j} \mathcal{C}_{n-k,j-k}(v) + 1 + \sum_{j < k \le n} \mathcal{C}_{k-1,j}(v) \right] \ . \tag{3}$$

Date: April 28, 1996.

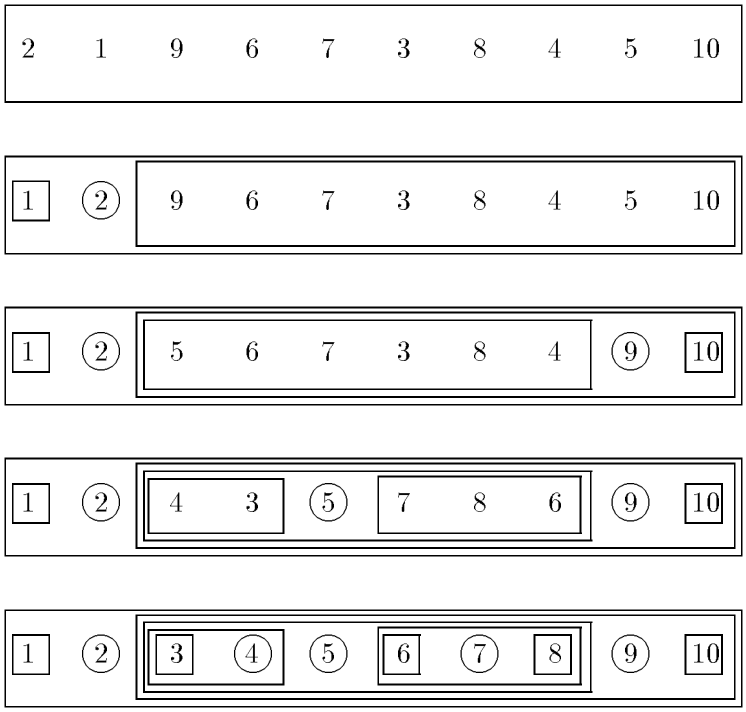

FIGURE 1. The partition process

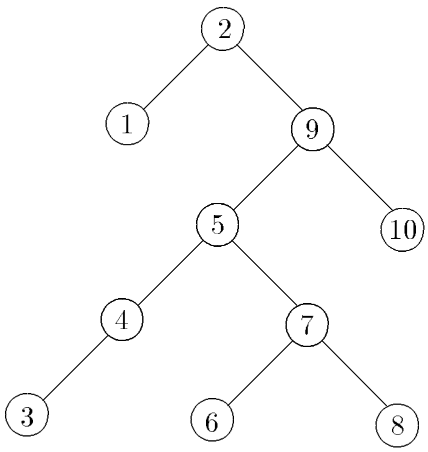

FIGURE 2. The corresponding binary search tree with 10 nodes

In this paper we want to compute the *variance* of the number of comparisons.

This requires an incredible amount of computations (by MAPLE, with careful human guidance). We cannot present all the details, but rather present the cornerstones of the whole procedure. We have to deal a lot with harmonic numbers, and use many identities about them which can be found e.g. in [9, 3, 4].

2. Analysis

Following a suggestion of Philippe Flajolet, we rather think about binary search trees (as explained above) and use a trivariate generating function.

The coefficient of $z^n u^j v^m$ in $F(z;u,v)$ is the probability that it takes m comparisons to find node j, under the usual assumption that all permutations of n elements are equally likely. For example, the tree in Figure 2 (obtained by the permutation in Figure 1) contributes

$$\frac{z^{10}}{10!}\Big(u^1v^9+u^2v^9+u^3v^{22}+u^4v^{22}+u^5v^{21}+u^6v^{23}+u^7v^{23}+u^8v^{23}+u^9v^{16}+u^{10}v^{16}\Big)\ . \tag{4}$$

Recursion (3) translates into the fundamental equation

$$\frac{\partial}{\partial z}F(z;u,v)=\frac{F(zv;u,v)}{1-zv}+\frac{u}{(1-zv)(1-zuv)}+\frac{u\,F(zv;u,v)}{1-zuv}\ . \tag{5}$$

We can see this by multiplying the recursion (3) by $nz^{n-1}u^j$ and summing over n and j. However, with a little experience, one can deduce it directly from the underlying binary search trees.

When speaking of differential equations, we will assume that z is the variable, whereas u and v are parameters.

Note the special value

$$F(z;u,1)=\frac{zu}{(1-z)(1-zu)}=\sum_{1\le j\le n}u^jz^n\ . \tag{6}$$

If we derive the fundamental equation (5) w.r.t. v, and set $v=1$, we get a differential equation for the generating function $G(z,u)$ of the average values $C_{n,j}$, which Maple is able to solve[1], subject to the initial condition $G(0,u)=0$. It is

$$\begin{aligned}G(z,u)=&\frac{2}{(1-u)(1-z)(1-zu)}\left[(u-2)\log\frac{1}{1-zu}+u(2u-1)\log\frac{1}{1-z}\right]\\&+\frac{2zu(z^2u-2zu+3-2z)}{(1-z)^2(1-zu)^2}\ ,\end{aligned} \tag{7}$$

which gives the coefficients $C_{n,j}$. (Knuth derived them using an elementary method.)

If we derive the fundamental equation (5) two times w.r.t. v, and set $v=1$, and use the already computed function $G(z,u)$, we get a differential equation for the generating function $D(z,u)$ of the second factorial moments $D_{n,j}$. Its solution is

$$\begin{aligned}D(z,u)=&\frac{4(2-u)^2}{(1-z)(1-zu)(1-u)^2}\log^2\frac{1}{1-zu}\\&+\frac{4u(1-2u)^2}{(1-z)(1-zu)(1-u)^2}\log^2\frac{1}{1-z}\\&-\frac{z^2u(-10z^3u^2-5z^2u^2+6zu^2-15u-5z^2u+38zu-15+6z)}{(1-z)^3(1-zu)^3}+\end{aligned} \tag{8}$$

→ next page

[1] We cannot refrain from mentioning a peculiarity of Maple. It frequently replaces $\log(1-z)$ by $\log(z-1)+i\pi$ &c. To avoid this unpleasant feature, we had to replace z by $-z$ and u by $-u$.

$$
\begin{aligned}
&+\frac{8u(2-u)(1-2u)}{(1-z)(1-zu)(1-u)^2}\log\frac{1}{1-zu}\log\frac{1}{1-u}\\
&+\frac{8(2-u)(1-2u)}{(1-z)(1-zu)(1-u)^2}\log\frac{1}{1-z}\log\frac{1}{1-u}\\
&+\frac{8u(2-u)(1-2u)}{(1-z)(1-zu)(1-u)^2}\log\frac{1}{1-zu}\log(-u)\\
&-\frac{2u(-5z^2u+19z^2u^2-11zu^2-10zu+z+3u+3)}{(1-u)(1-z)^2(1-zu)^2}\log\frac{1}{1-z}\\
&-\frac{10z^2u^2-38z^2u-2zu^2+20zu+22z-6u-6}{(1-u)(1-z)^2(1-zu)^2}\log\frac{1}{1-zu}\\
&-\frac{8u(2-u)(1-2u)}{(1-z)(1-zu)(1-u)^2}\Big(\operatorname{dilog}\Big(-\frac{(1-z)u}{1-u}\Big)-\operatorname{dilog}\Big(-\frac{u}{1-u}\Big)\Big)\\
&-\frac{8(2-u)(1-2u)}{(1-z)(1-zu)(1-u)^2}\Big(\operatorname{dilog}\Big(\frac{1-zu}{1-u}\Big)-\operatorname{dilog}\Big(\frac{1}{1-u}\Big)\Big)\ .
\end{aligned}
$$

Let us recall the definition of the dilogarithm [1],

$$\operatorname{dilog} z=\int_1^z\frac{\log t}{1-t}\,dt\ . \tag{9}$$

The appearence of a term $\log(-u)$ looks strange at first sight, but it compensates with the dilog–terms.

In order to get the coefficients, we rewrite $D(z,u)$ as follows (with the heavy use of Maple),

$$
\begin{aligned}
D(z,u)=&\frac{9}{(1-z)^3}\sum_{j\ge1}u^j\\
&+\log^2\frac{1}{1-z}\sum_{j\ge1}u^j\Big\{\frac{4(j-4)}{(1-z)^2}+\frac{4(z^{j-3}-z)}{(1-z)^3}+\frac{4z^{j-3}(-1-2z+z^2)}{1-z}\Big\}\\
&+\frac{1}{(1-z)^2}\sum_{j\ge1}u^j\Big\{-8(j+2)H_j+21j-25\Big\}\\
&+\frac{1}{(1-z)^2}\log\frac{1}{1-z}\sum_{j\ge2}u^j\Big\{(8j-16H_{j-1}-26)z^j\\
&\qquad+\Big(-16j+70+24H_{j-1}-\frac{8}{j-1}\Big)z^{j-1}\\
&\qquad+\sum_{i=2}^{j-1}\Big(8H_{j-i}-\frac{8i}{j-i}+16-\frac{16}{i+1}+\frac{40}{i}-\frac{16}{i-1}\Big)z^{j-i}\\
&\qquad+\Big(-8H_j+\frac{24}{j}-12-\frac{16}{j-1}\Big)\Big\}\\
&+\frac{1}{(1-z)^2}\log\frac{1}{1-z}\cdot u(10-18z)+
\end{aligned}
\tag{10}
$$

→ next page

$$+\frac{1}{1-z}\sum_{j\ge1}u^j\Big\{(2j^2+18j+16)H_j^2+\Big(-6j^2-14j+14-\frac{16}{j}\Big)H_j \\ -(2j^2+10j+16)H_j^{(2)}-2j^2+25j+10\Big\}\ .$$

The most complicated term in this expansion is the $\frac{1}{1-z}$ term

$$8\frac{(2-u)(1-2u)}{(1-u)^2}\operatorname{dilog}\frac{1}{1-u}+4\frac{(2-u)^2}{(1-u)^3}\log^2\frac{1}{1-u} \\ +2\frac{11-u}{(1-u)^2}\log\frac{1}{1-u}+\frac{u(13-41u+10u^2)}{(1-u)^3}\ . \tag{11}$$

Reading of the coefficient of u^j is pretty much standard except for the following

$$[u^j]8\frac{(2-u)(1-2u)}{(1-u)^2}\operatorname{dilog}\frac{1}{1-u}=4jH_j^2-8\frac{j^2+j+2}{j}H_j+4jH_j^{(2)}+8j\ . \tag{12}$$

For this one we argue as follows. Since

$$\frac{d}{du}\operatorname{dilog}\frac{1}{1-u}=\Big(-\frac{1}{1-u}-\frac{1}{u}\Big)\log\frac{1}{1-u}\ , \tag{13}$$

we find

$$[u^j]\operatorname{dilog}\frac{1}{1-u}=-\frac{1}{j}\Big(H_{j-1}+\frac{1}{j}\Big) \tag{14}$$

The convolution, caused by the factor $\frac{1}{(1-u)^2}$, is then handled by formulæ in [9].

Now we read of the coefficient of u^jz^n in $D(z,u)$, as shown in (10). Using all kind of identities for harmonic numbers (see [9]) we get as an intermediary result

$$\begin{aligned} D_{n,j} &= -2(n+1)(n+8-2j)H_n^2-8(n+1)\,H_n\,H_j \\ &+\frac{(4-32n-12n^2)j^2+(44n^2+12n^3+28n-4)j+16n+16}{j(n+1-j)}H_n \\ &+2(j+8)(j+1)\,H_j^2 \\ &+\Big(-4j^2+(12+8n)j-4n^2+8-12n\Big)\,H_jH_{n+1-j} \\ &-2\frac{(-6n-14)j^3+(14n+17+6n^2)j^2+(-3+6n^2+3n)j+8n+8}{j(n+1-j)}H_j \\ &+\Big(2j^2+(-4n-18)j+2n^2+18n+32\Big)\,H_{n+1-j}^2 \\ &-2\frac{(6n+14)j^3+(-12n^2-25-46n)j^2+(37n+11+32n^2+6n^3)j+8n+8}{j(n+1-j)}H_{n+1-j} \\ &+2(n+1)(n+8-2j)\,H_n^{(2)}-(2j^2+16+10j)\,H_j^{(2)} \\ &+\Big(-2n^2+4nj-18n-2j^2-32+18j\Big)\,H_{n+1-j}^{(2)}+ \end{aligned} \tag{15}$$

→ next page

$$+\frac{10j^4+(-20n-20)j^3+(n^2-66-13n)j^2+(76+9n^3+42n^2+101n)j+32}{2j(n+1-j)}$$
$$+4(n+2-2j)(n+1)\sum_{k=1}^{j}\frac{H_{n-k}}{k}+8(n+1-j)\sum_{k=1}^{j}\frac{H_{n+k-j}}{k}\ .$$

In this representation, there appear two sums, which (most likely) cannot be simplified any further. Also, from this representation, the symmetry $D_{n,j}=D_{n,n+1-j}$ is not at all obvious. Therefore we derived the following lemmata.

Lemma 1. [Reciprocity law]

$$\sum_{k=1}^{j}\frac{H_{n-k}}{k}+\sum_{k=1}^{n+1-j}\frac{H_{n-k}}{k}=H_jH_{n+1-j}+H_n^2-H_n^{(2)}-\frac{1}{j(n+1-j)} \tag{16}$$

Lemma 2. [Reciprocity law]

$$\begin{aligned}\sum_{k=1}^{j}\frac{H_{n+k-j}}{k}+\sum_{k=1}^{n+1-j}\frac{H_{j+k-1}}{k}&=\frac{1}{2}\Big(H_j^2+H_j^{(2)}\Big)+\frac{1}{2}\Big(H_{n+1-j}^2+H_{n+1-j}^{(2)}\Big)\\&\quad+H_jH_{n+1-j}+\frac{1}{j(n+1-j)}\\&\quad+\frac{n+1}{j(n+1-j)}\Big(H_n-H_j-H_{n+1-j}\Big)\ .\end{aligned} \tag{17}$$

And here is the formula relating the two types of sums.

Lemma 3.

$$\begin{aligned}\sum_{k=1}^{j}\frac{H_{n-k}}{k}+\sum_{k=1}^{j}\frac{H_{n+k-j}}{k}&=\frac{1}{n+1-j}\Big(H_n-H_j-H_{n+1-j}\Big)+H_jH_n+H_jH_{n+1-j}\\&\quad-H_nH_{n+1-j}+\frac{1}{2}\Big(H_n^2-H_n^{(2)}\Big)+\frac{1}{2}\Big(H_{n+1-j}^2+H_{n+1-j}^{(2)}\Big)\end{aligned} \tag{18}$$

All three lemmata can be easily proved by setting up bivariate generating functions on both sides. The formulæ are 'nice' and easier to *prove* than to *find*. As an example, we give the proof of Lemma 1. In terms of generating functions we have to prove that

$$\begin{aligned}&\frac{1}{(1-u)(1-z)}\log\frac{1}{1-uz}\log\frac{1}{1-z}-\frac{u}{(1-u)(1-uz)}\log^2\frac{1}{1-uz}\\&+\frac{u}{(1-1/u)(1-uz)}\log\frac{1}{1-z}\log\frac{1}{1-uz}-\frac{1}{(1-1/u)(1-z)}\log^2\frac{1}{1-z}\\&=\frac{1}{z(1-uz)(1-z)}\log\frac{1}{1-uz}\log\frac{1}{1-z}+\frac{u}{(1-u)(1-z)}\log^2\frac{1}{1-z}\\&-\frac{u}{(1-u)(1-uz)}\log^2\frac{1}{1-uz}-\frac{1}{z}\log\frac{1}{1-z}\log\frac{1}{1-uz}\ ,\end{aligned} \tag{19}$$

which is easily seen to be true.

Using the lemmata, we can do the following,

$$
\begin{aligned}
(n+1-2j)\sum_{k=1}^{j}\frac{H_{n-k}}{k} &= (n+1-j)\sum_{k=1}^{j}\frac{H_{n-k}}{k} + j\sum_{k=1}^{n+1-j}\frac{H_{n-k}}{k}\\
&\quad - j\Big(H_j H_{n+1-j} + H_n^2 - H_n^{(2)} - \frac{1}{j(n+1-j)}\Big)\,.
\end{aligned}
\tag{20}
$$

This gives us the final result on the second factorial moments.

Proposition 1.

$$
\begin{aligned}
D_{n,j} &= -2(n+1)(n+6)\,H_n^2\\
&\quad - 8H_n\Big(jH_j + (n+1-j)\,H_{n+1-j}\Big)\\
&\quad + 4\frac{-(3n^2+8n+1)j^2+(n+1)(3n^2+8n+1)j+4(n+1)}{j(n+1-j)}\,H_n\\
&\quad + 2(j+8)(j+1)H_j^2 + 2(n+9-j)(n+2-j)\,H_{n+1-j}^2\\
&\quad + 4\Big(-j^2+(n+1)j-n^2-n+4\Big)H_j\,H_{n+1-j}\\
&\quad - \frac{2}{j(n+1-j)}\Big(-2(3n+7)j^3+(6n^2+14n+13)j^2+(n+1)(6n+1)j+8(n+1)\Big)\,H_j\\
&\quad - \frac{2}{j(n+1-j)}\Big(2(3n+7)j^3-(12n^2+46n+29)j^2\\
&\qquad\qquad + (n+1)(6n^2+26n+15)j+8(n+1)\Big)\,H_{n+1-j}\\
&\quad + 2(n+1)(n+6)\}\,H_n^{(2)}\\
&\quad - 2(j^2+5j+8)\,H_j^{(2)} - 2(j^2-(2n+7)j+n^2+7n+14)\,H_{n+1-j}^{(2)}\\
&\quad + \frac{10j^4-20(n+1)j^3+(n^2-13n-66)j^2+(9n^2+33n+76)(n+1)j+32}{2j(n+1-j)}\\
&\quad + 4n\Big((n+1-j)\sum_{k=1}^{j}\frac{H_{n-k}}{k} + j\sum_{k=1}^{n+1-j}\frac{H_{n-k}}{k}\Big)\,. \qquad \Box
\end{aligned}
\tag{21}
$$

Computing $\mathrm{Var}_{n,j} = D_{n,j} + C_{n,j} - (C_{n,j})^2$, we find the variance.

Theorem 1. *The variance of the number of comparisons needed to find the jth element out of n elements is given by the exact formula*

$$
\begin{aligned}
\mathrm{Var}_{n,j} &= -2(n+1)(3n+8)\,H_n^2\\
&\quad + 8\,H_n\,H_j\Big(n(j+2)+2\Big) + 8\,H_n\,H_{n+1-j}\Big(n(n+3-j)+2\Big)\\
&\quad + 2\frac{-(2n^2+n-9)j^2+(n+1)(2n^2+n-9)j+8(n+1)}{j(n+1-j)}\,H_n\\
&\quad - 2j(j-1)\,H_j^2 - 2(n+1-j)(n-j)\,H_{n+1-j}^2\\
&\quad + 4\Big(j^2-(n+1)j-n^2-5n-8\Big)\,H_j\,H_{n+1-j} +
\end{aligned}
\tag{22}
$$

→ next page

$$
\begin{aligned}
&-\frac{2}{j(n+1-j)}\Big(-(2n+3)j^3+(2n^2+7n+24)j^2-(n+1)(2n+21)j+8(n+1)\Big)\,H_j\\
&-\frac{2}{j(n+1-j)}\Big((2n+3)j^3-(4n^2+8n-15)j^2\\
&\qquad\qquad\qquad+(n+1)(2n^2+3n-18)j+8(n+1)\Big)\,H_{n+1-j}\\
&+2(n+1)(n+6)\,H_n^{(2)}\\
&-2(j^2+5j+8)\,H_j^{(2)}-2(j^2-(2n+7)j+n^2+7n+14)\,H_{n+1-j}^{(2)}\\
&+\frac{10j^4-20(n+1)j^3+(9n^2+31n-6)j^2+(n^2-11n+16)(n+1)j+32}{2j(n+1-j)}\\
&+4n\bigg((n+1-j)\sum_{k=1}^{j}\frac{H_{n-k}}{k}+j\sum_{k=1}^{n+1-j}\frac{H_{n-k}}{k}\bigg)\,. \qquad \square
\end{aligned}
$$

In the instance $n=2N+1$, $j=N+1$ (searching for the *median*), we can use the first reciprocity law to simplify the variance. We get

$$
\begin{aligned}
\mathrm{Var}_{2N+1,N+1}=&-8(2N+5)(N+1)H_{2N+1}^2+16(2N+5)(N+1)H_{2N+1}H_{N+1}\\
&+4\frac{4N^3+9N^2+2N+5}{N+1}H_{2N+1}-8(2N^2+7N+7)H_{N+1}^2\\
&-4\frac{4N^3+9N^2-13N-2}{N+1}H_{N+1}+24(N+1)H_{2N+1}^{(2)}\\
&-4(N^2+7N+14)H_{N+1}^{(2)}+\frac{7N^4+15N^3+9N^2+5N+20}{(N+1)^2}\,.
\end{aligned}
\tag{23}
$$

Asymptotically we get from this

$$
\mathrm{Var}_{2N+1,N+1}\sim N^2\Big(16\log 2-16\log^2 2-\tfrac{2}{3}\pi^2+7\Big)\,. \tag{24}
$$

The numerical constant is 3.823370396 .

To get a better feeling about the sums $\sum_{k=1}^{j}\frac{H_{n-k}}{k}$, we compute their *average*

$$
\frac{1}{n}\sum_{1\le j\le n}\sum_{k=1}^{j}\frac{H_{n-k}}{k}=\frac{n+1}{n}\left(H_n^2-H_n^{(2)}\right)-H_n+1\,. \tag{25}
$$

It might be even more instructive to compute the average value

$$
\begin{aligned}
&\frac{1}{n}\sum_{1\le j\le n}\left[(n+1-j)\sum_{k=1}^{j}\frac{H_{n-k}}{k}+j\sum_{k=1}^{n+1-j}\frac{H_{n-k}}{k}\right]\\
&\qquad=\frac{(n+2)(n+1)}{n}\left(H_n^2-H_n^{(2)}\right)-\frac{3n+5}{2}H_n+\frac{5n+11}{4}\,.
\end{aligned}
\tag{26}
$$

Now we want to study the asymptotic behaviour of the variance, if $n\to\infty$ and $j=\rho n$, $0<\rho<1$.

Using the well known asymptotics for the harmonic numbers and Euler's summation formula, we find the asymptotic behaviour of the extra sum.

Lemma 4. *As $n \to \infty$ and $j = \rho n$ with $0 < \rho < 1$, we have*

$$\sum_{k=1}^{j} \frac{H_{n-k}}{k} = \log^2 n + (2\gamma + \log\rho)\log n + \left(\gamma^2 + \gamma\log\rho - \operatorname{dilog}(1-\rho)\right) + o(1) \ . \tag{27}$$

Using this intermediary result in our formula for the variance, an elementary calculation gives us the asymptotic evaluation of the variance.

Theorem 2. *The variance of the number of comparisons in Hoare's* FIND *algorithm to find the ρn–th out of n elements $(0 < \rho < 1)$ fulfills for $n \to \infty$*

$$\begin{aligned}\mathrm{Var}_{n,\rho n} \sim n^2 \Big(&- 2\rho^2\log^2\rho + 4(\rho^2-\rho-1)\log\rho\log(1-\rho) - 2(1-\rho)^2\log^2(1-\rho) \\ &- 4\rho\log\rho - 4(1-\rho)\log(1-\rho) - 4\rho\operatorname{dilog}\rho - 4(1-\rho)\operatorname{dilog}(1-\rho) \\ &+ \tfrac{2}{3}\pi^2\rho(1-\rho) + 5\rho(1-\rho) + \tfrac{1}{2}\Big) \ . \qquad \square\end{aligned} \tag{28}$$

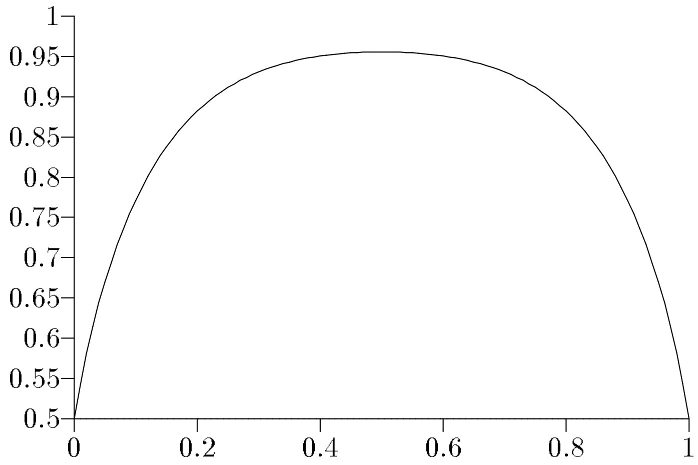

FIGURE 3. The coefficient of n^2 in Theorem 2, as a function of ρ

3. The number of passes

A similar but somewhat easier treatment is possible for the number of passes (recursive calls) instead of the number of comparisons. Therefore we only cite the results.

We have the fundamental equation

$$\frac{\partial}{\partial z}F(z;u,v) = \frac{v\,F(z;u,v)}{1-z} + \frac{uv}{(1-z)(1-zu)} + \frac{uv\,F(z;u,v)}{1-zu} \tag{29}$$

The equation can easily be gained from referring to (5): Observing the fact that this time the partitioning phase "costs" 1 (pass) yielding a factor v^1 instead of a "cost" of $n-1$ (comparisons) which corresponded to the substitution $z \to zv$ on the right-hand-side of (5) we find (29).

The expectation is an old result of Arora and Dent [2];

$$E_{n,j} = H_j + H_{n+1-j} - 1 \ . \tag{30}$$

For the variance we get

$$\begin{aligned}\mathrm{Var}_{n,j} &= \frac{2(n+1)}{j(n+1-j)}\,H_n + \left(1 - \frac{2(n+1)}{j(n+1-j)}\right)(H_j + H_{n+1-j}) \\ &\quad - \left(H_j^{(2)} + H_{n+1-j}^{(2)}\right) + \frac{2}{j(n+1-j)} + 2\;. \end{aligned} \tag{31}$$

Acknowledgement. One of us (H. P.) had the opportunity to use the powerful computers at INRIA. Thanks are due to P. Flajolet for the generous invitation.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover, 1973. A reprint of the tenth National Bureau of Standards edition, 1964.
[2] S. Arora and W. Dent. Randomized binary search technique. *Communications of the ACM*, 12:77–80, 1969.
[3] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics.* Addison Wesley, 1989.
[4] D. H. Greene and D. E. Knuth. *Mathematics for the analysis of algorithms.* Birkhauser, Boston, second edition, 1982.
[5] C. A. R. Hoare. Find (Algorithm 65). *Communications of the ACM*, 4:321–322, 1961.
[6] C. A. R. Hoare. Quicksort. *Computer Journal*, 5:10–15, 1962.
[7] H.-K. Hwang. On the average shape of binary search trees. *Unpublished manuscript*, 1990.
[8] P. Kirschenhofer, C. Martínez, and H. Prodinger. Analysis of Hoare's Find algorithm with median–of–three partition. *Submitted*, 1995.
[9] D. E. Knuth. *The Art of Computer Programming*, volume 1: Fundamental Algorithms. Addison-Wesley, 1968. Second edition, 1973.
[10] D. E. Knuth. Mathematical analysis of algorithms. In *Information Processing 71*, pages 19–27. North Holland Publishing Company, 1972. Proceedings of IFIP Congress, Ljubljana, 1971.
[11] D. E. Knuth. *The Art of Computer Programming*, volume 3: Sorting and Searching. Addison-Wesley, 1973.
[12] J. Lent and H. M. Mahmoud. Average–case analysis of multiple quickselect: An algorithm for finding order statistics. *Statistics and Probability Letters (to appear).*
[13] H. M. Mahmoud. *Evolution of Random Search Trees.* John Wiley, New York, 1992.
[14] H. M. Mahmoud, R. M. Modarres, and R. T. Smythe. Analysis of one-sided quicksort: An algorithm for order statistics. *RAIRO, Theoretical Informatics and Applications*, 29:255–276, 1995.
[15] H. Prodinger. Multiple quickselect — Hoare's find algorithm for several elements. *Information Processing Letters*, 56:123–129, 1995.

INSTITUT FÜR ALGEBRA UND DISKRETE MATHEMATIK, TECHNICAL UNIVERSITY OF VIENNA, WIEDNER HAUPTSTRASSE 8–10, A-1040 VIENNA, AUSTRIA.

E-mail address: Peter.Kirschenhofer@tuwien.ac.at and Helmut.Prodinger@tuwien.ac.at