GLASNIK MATEMATIČKI Vol. 22 (42) (1987), 269 – 278.

A CONTRIBUTION TO THE ANALYSIS OF IN SITU PERMUTATION

P. Kirschenhofer, H. Prodinger and R. F. Tichy, Vienna, Austria

Abstract. There is a simple algorithm to replace (x_1, \ldots, x_n) by $(x_{p(1)}, \ldots, x_{p(n)})$, where $\pi = (p(1), \ldots, p(n))$ is a permutation of $\{1, 2, \ldots, n\}$, essentially without further storage requirements. This paper continues some research work by D. E. Knuth about a characteristic parameter of this algorithm. Using generating function techniques alternative derivations for several results of Knuth as well as a number of new theorems are obtained.

1. Introduction

Let $\pi = \begin{pmatrix} 1 & \dots & n \\ p(1), \dots, p(n) \end{pmatrix}$ be a permutation of the numbers 1, 2, ..., n and let us consider the following part of a program:

for
$$j:=1$$
 to n do
begin $k:=p(j)$;
while $k > j$ do
 $k:=p(k)$
end;

These instructions can be used to check whether j is a cycle leader, i. e. the smallest number in its cycle. For this, one has to ask ,k=j? after passing the while-loop.

The detection of the cycle leader is useful if one wants to permute an array $x[1], \ldots, x[n]$ along the permutation π essentially without further storage requirements (in situ permutation). For each cycle (i_1, \ldots, i_k) the elements $x[i_1], \ldots, x[i_k]$ should be replaced by $x[p(i_1)], \ldots, x[p(i_k)]$. If we do that iff i_1 is the cycle leader, this will be done exactly once for each cycle. The complete algorithm was developed by Mac Leod [5] and analyzed by Knuth [4]:

Mathematic subject classification (1980): 05 A 15, 68 E 05.

Key words and phrases: In situ permutation, generating functions, functional differential equations, asymptotic expansions, combinatorial algorithms.

One of the three interesting parameters of this analysis is denoted by $a(\pi)$ and equals the number of times the instruction ,k:=p(k)" is executed. Knuth [4] has shown that

$$0 \leq a(\pi) \leq \binom{n}{2}; \tag{1.2}$$

the average of $a(\pi)$ is

$$(n+1)H_n - 2n;$$
 (1.3)

the variance of $a(\pi)$ is

$$2n^{2} - (n+1)^{2} H_{n}^{(2)} - (n+1) H_{n} + 4n \qquad (1.4)$$

(where $H_n^{(s)} = \sum_{1 \le k \le n} k^{-s}$ denotes the *n*-th harmonic number of degree s, $H_n^{(1)} = H_n$).

In this paper we exploit a method which allows us to get these quantities by less computation. Furthermore, we are able to determine the s-th factorial moment of $a(\pi)$ asymptotically, viz.

$$n^{s}\log^{s}n + (\gamma - 2)sn^{s}\log^{s-1}n + O(n^{s}\log^{s-2}n), \quad n \to \infty$$
 (1.5)
where $\gamma = .57721...$ is Euler's constant.

Since the s-th moment is just a linear combination of the j-th factorial moments for $j \leq s$, we obtain the same asymptotic expansion for the s-th moment.

To stress the method of our treatment in a few words, we introduce certain generating functions $G_n(z)$, obtain a recursion for them, which does not allow getting a simple explicit expression for $G_n(z)$; from this recursion we obtain differential equations for the generating functions of the s-th factorial moments, from which we can derive the above asymptotic expansion.

2. Generating functions

Assume that $\pi = q(n)$ is the canonical representation of the permutation π as a product of cycles in the way described in Knuth [3, p. 176]. In the following we always represent a permutation in this way; it is known that

$$a(\pi) = \operatorname{card} \{(i, j): 1 \le i < j \le n, q(i) < q(k) \text{ for all } k \text{ with } i < k \le j\}.$$
 (2.1)

By a_{nk} we denote the number of permutations π of *n* elements such that $a(\pi) = k$ and by

$$G_n(z) = \sum_{k \ge 0} a_{nk} z^k / n!$$

the corresponding probability generating function.

270

THEOREM 1. For
$$n \ge 1$$

 $G_n(z) = n^{-1} \cdot \sum_{k=0}^{n-1} z^k G_k(z) G_{n-1-k}(z);$
 $G_0(z) = 1.$

Proof. In the following we write a permutation π of $\{1, \ldots, n\}$ in the form $\pi = \rho | \sigma$, where ρ is a permutation of n-1-k elements and σ a permutation of k elements. It is immediate that

$$a(\pi) = a(\rho) + a(\sigma) + k.$$

Summing up over all permutations π with $a(\pi) = s$ we obtain

$$a_{ns} = \sum_{k=0}^{n-1} \binom{n-1}{k} \sum_{i+j+k=s} a_{n-1-k,i} a_{k,j}.$$

Dividing by n! and multiplying by z^s it follows that

$$a_{ns} z^{s}/n! = n^{-1} \cdot \sum_{k=0}^{n-1} z^{k} \sum_{i+j+k=s} a_{k,j} z^{j} \cdot a_{n-1-k,i} z^{i}/(k!(n-1-k)!).$$

Summing up over $s \ge 0$, Theorem 1 results immediately.

Let us now consider the double generating function H(z, u) defined by

$$H(z, u) = \sum_{n \ge 0} G_n(z) u^n.$$
(2.3)
COROLLARY 1. $\frac{\partial}{\partial u} H(z, u) = H(z, u) \cdot H(z, zu);$
 $H(1, u) = (1 - u)^{-1}.$

Proof. We multiply the recursion in Theorem 1 by nu^{n-1} and sum up over all $n \ge 0$ to get the result. Since $G_n(1) = 1$, the identity for H(1, u) follows.

In the following we consider the s-th factorial moments $\beta_s(n)$ of the random variable given by the probability generating function $G_n(z)$:

$$\beta_s(n) = \frac{d^s}{dz^s} G_n(z) \bigg|_{z=1}.$$
(2.4)

Introducing the generating functions $f_s(u)$ of the s-th factorial moments by

$$f_s(u) = \sum_{n \ge 0} \beta_s(n) u^n, \qquad (2.5)$$

we obtain by Taylor's formula and (2.4)

$$H(z, u) = \sum_{s \ge 0} f_s(u) (z-1)^s / s!.$$
(2.6)

THEOREM 2. For $s \ge 1$

$$f_{s}^{r}(u) - 2(1-u)^{-1} f_{s}(u) = h_{s}(u), \text{ with}$$

$$h_{s}(u) = \sum_{i=1}^{s-1} {\binom{s}{i}} f_{i}(u) \sum_{r=0}^{s-i} {\binom{s-i}{r}} u^{r} f_{s-r-i}^{(r)}(u) + (1-u)^{-1} \sum_{r=1}^{s} {\binom{s}{r}} u^{r} f_{s-r}^{(r)}(u),$$

where $f^{(i)}(u)$ denotes the *i*-th derivative of the function f(u);

$$f_0(u) = (1-u)^{-1}, h_0(u) = -(1-u)^{-2} \text{ and } f_s(0) = 0 \text{ for } s \ge 1.$$

Proof. First note that

$$f_j(zu) = \sum_{k \ge 0} f_j^{(k)}(u) (z-1)^k u^k / k!$$

by Taylor's formula. Inserting (2.6) into the equation of Corollary 1 we get

$$\sum_{s \ge 0} f'_{s}(u) (z-1)^{s} / s! = \left[\sum_{i \ge 0} f_{i}(u) (z-1)^{i} / i! \right] \cdot \left[\sum_{j \ge 0} f_{j}(zu) (z-1)^{j} / j! \right] = \\ = \left[\sum_{i \ge 0} f_{i}(u) (z-1)^{i} / i! \right] \cdot \left[\sum_{j \ge 0} (z-1)^{j} / j! \left(\sum_{k \ge 0} f_{j}^{(k)}(u) (z-1)^{k} u^{k} / k! \right) \right] = \\ = \sum_{m \ge 0} \sum_{i+j+k=m} u^{k} f_{i}(u) f_{j}^{(k)}(u) (z-1)^{m} / (i!j!k!).$$

Comparing the coefficients of $(z-1)^{s}/s!$ we obtain

$$\begin{aligned} f_{s}^{r}(u) &= \sum_{i+j+k=s} s! \cdot u^{k} f_{i}(u) f_{j}^{(k)}(u) / (i!j!k!) \\ &= \sum_{i=0}^{s} {\binom{s}{i}} f_{i}(u) \sum_{r=0}^{s-i} {\binom{s-i}{r}} u^{r} f_{s-i-r}^{(r)}(u) \\ &= 2 (1-u)^{-1} f_{s}(u) + \sum_{i=1}^{s-1} {\binom{s}{i}} f_{i}(u) \sum_{r=0}^{s-i} {\binom{s-i}{r}} u^{r} f_{s-i-r}^{(r)}(u) + \\ &+ (1-u)^{-1} \cdot \sum_{r=1}^{s} {\binom{s}{r}} u^{r} f_{s-r}^{(r)}(u), \end{aligned}$$

because $\beta_0(n) = 1$ for all n and therefore $f_0(u) = (1-u)^{-1}$.

Since $G_0(z)=1$ we have $\beta_s(0)=0$ for $s \ge 1$ and therefore $f_s(0)=0$ for $s \ge 1$, and the proof of Theorem 2 is complete.

Solving the first order linear differential equation of Theorem 2 we obtain

COROLLARY 2. For $s \ge 1$

$$f_s(u) = (1-u)^{-2} \int_0^u h_s(t) (1-t)^2 dt,$$

where f_s and h_s are as in Theorem 2.

3. The first and second order factorial moments

In principle Corollary 2 allows to compute $f_s(u)$ (and thus $\beta_s(n)$) step by step for any s. To illustrate, we determine the first two moments.

THEOREM 3. With
$$L(u) := -\log(1-u)$$
 we have
 $f_1(u) = L(u) \cdot (1-u)^{-2} - (1-u)^{-2} + (1-u)^{-1}$,
 $f_2(u) = 2L^2(u) \cdot (1-u)^{-3} - 2L(u) \cdot (1-u)^{-3} + 2(1-u)^{-3} - L^2(u) \cdot (1-u)^{-2} - 2(1-u)^{-2}$;
 $\beta_1(n) = (n+1)H_n - 2n$,
 $\beta_2(n) = (n+1)^2(H_n^2 - H_n^{(2)}) - (4n+2)(n+1)H_n + 6n(n+1)$.

Proof. Observing $h_1(u) = u(1-u)^{-3}$ the formula for $f_1(u)$ is immediate; a short computation yields

$$h_2(u) = 2L^2(u)(1-u)^{-4} + 2L(u)(1-u)^{-4} - 2L(u)(1-u)^{-3}$$

from which $f_2(u)$ follows by the formula indicated in Corollary 2.

Expanding $f_1(u)$ resp: $f_2(u)$ we use the following results (compare Greene/Knuth [2, p. 14]):

$$L(u) \cdot (1-u)^{-m-1} = \sum_{n \ge 0} (H_{n+m} - H_m) \binom{n+m}{m} u^n,$$

$$L^2(u) \cdot (1-u)^{-m-1} = \sum_{n \ge 0} ((H_{n+m} - H_m)^2 - (H_{n+m}^{(2)} - H_m^{(2)})) \binom{n+m}{m} u^n.$$

The following special instances are needed for our computations:

$$L(u) \cdot (1-u)^{-2} = \sum_{n \ge 0} \left[(n+1) H_n - n \right] u^n,$$

$$L^2(u) \cdot (1-u)^{-2} = \sum_{n \ge 0} \left[(n+1) (H_n^2 - H_n^{(2)}) - 2nH_n + 2n \right] u^n,$$

$$L(u) \cdot (1-u)^{-3} = \sum_{n \ge 0} \left[\binom{n+2}{2} H_n - (3/4) n^2 - (5/4) n \right] u^n,$$

$$L^2(u) \cdot (1-u)^{-3} = \sum_{n \ge 0} \left[\binom{n+2}{2} (H_n^2 - H_n^{(2)}) - (n/2) (5+3n) H_n + (7/4) n^2 + (9/4) n \right] u^n.$$

Inserting into the formulas for $f_1(u)$ and $f_2(u)$ and simplifying we get the announced results for $\beta_1(n)$ and $\beta_2(n)$.

4. Asymptotic results

Although, in principle, Corollary 2 allows to determine $f_s(u)$ explicitly for any s, terms get more and more complicated as s gets large. So we confine ourselves for general s to give the two leading terms of the asymptotic expansion of $f_s(u)$ about the singularity u=1. It turns out to be a crucial point in the derivation of the desired result that $f_s(u)$ is a linear combination of functions of the type $L^i(u) \cdot (1-u)^{-j-1}$ (with L from Theorem 3):

In the following we denote by $\Re_{p,q}(u)$ an unspecified linear combination of terms of the form $L^i(u)(1-u)^{-j-1}$ where *i*, *j* are integers with either j < q and *i* arbitrary, or j = q and $i \leq p$. With this notation we have

THEOREM 4. For
$$s \ge 0$$

 $f_s(u) = s! L^s(u) \cdot (1-u)^{-s-1} + \Re_{s-1,s}(u).$

Proof. We proceed by induction and start with s=0: $f_0(u)=(1-u)^{-1}$, and the theorem is valid in this case.

Assuming that the theorem is correct for all j with $0 \le j \le s-1$, we prove that the same holds for s. We will frequently use the fact that for

$$g(u) = cq!L^{p}(u) \cdot (1-u)^{-q-1} + \mathcal{R}_{p-1,q}(u)$$
 (c a constant)

the derivatives $g^{(i)}(u)$ fulfill

$$g^{(i)}(u) = c (q+i)! L^{p}(u) \cdot (1-u)^{-q-i-1} + \mathcal{R}_{p-1,q+i}(u).$$

Especially we have for $j \leq s-1$

$$f_{j}^{(i)}(u) = (j+i)! L^{j}(u) \cdot (1-u)^{-j-i-1} + \mathcal{R}_{j-1,j+i}(u).$$

Inserting into the formula for $h_s(u)$ in Theorem 2 we get

$$h_{s}(u) = \sum_{i=1}^{s-1} {\binom{s}{i}} [i!L^{i}(u) \cdot (1-u)^{-i-1} + \mathcal{R}_{i-1,i}(u)] \sum_{r=0}^{s-i} {\binom{s-i}{r}} u^{r} \times \\ \times [(s-i)!L^{s-i-r}(u) \cdot (1-u)^{-s+i-1} + \mathcal{R}_{s-i-r-1,s-i}(u)] + \\ + (1-u)^{-1} \sum_{r=1}^{s} {\binom{s}{r}} u^{r} [s!L^{s-r}(u) \cdot (1-u)^{-s-1} + \mathcal{R}_{s-r-1,s}(u)].$$

It follows by a short consideration that all remainder terms $\mathscr{R}_{p,q}(u)$ as well as the second sum give a contribution of the form $\mathscr{R}_{s-1,s+1}(u)$. The other terms contribute

$$s!L^{s}(u) \cdot (1-u)^{-s-2} \cdot \sum_{i=1}^{s-1} (1+u/L(u))^{s-i} = s! (s-1) L^{s}(u) \cdot (1-u)^{-s-2} + \mathscr{R}_{s-1,s+1}(u),$$

hence $h_s(u)$ is of the same type.

A contribution to the analysis...

Using Corollary 2 we get

$$f_{s}(u) = (1-u)^{-2} \cdot \int_{0}^{u} s! (s-1) L^{s}(t) \cdot (1-t)^{-s} dt + (1-u)^{-2} \cdot \int_{0}^{u} \Re_{s-1,s-1}(t) dt =$$

= $s! L^{s}(u) \cdot (1-u)^{-s-1} + \Re_{s-1,s}(u)$

by integration by parts.

It should be remarked that from Theorem 4 the leading term of $\beta_{\epsilon}(n)$ for $n \to \infty$ is

$$\beta_s(n) \sim n^s \cdot \log^s n, \tag{4.1}$$

either by observing that $L^{s}(u)$ varies slowly at infinity and applying Hardy-Littlewood-Karamata's Tauberian Theorem (e.g. [1]) or by the explicit knowledge of the coefficients of functions of the following type (compare Zave [6]):

$$L^{p}(u) \cdot (1-u)^{-q-1} = \sum_{n \ge 0} P_{p}(H_{n+q}^{(1)} - H_{q}^{(1)}, \dots, H_{n+q}^{(p)} - H_{q}^{(p)}) \cdot \binom{n+q}{q} u^{n},$$
(4.2)

where $P_p(s_1, \ldots, s_p)$ is defined by $P_0 = 1$ and

 $P_p(s_1,\ldots,s_p) = (-1)^p Y_p(-s_1,-s_2,-2s_3,\ldots,-(p-1)!s_p)$

with Y_p the p-th Bell polynomial.

With the information on the structure of the remainder term in Theorem 4 it is possible to determine the second term in the expansion of $f_s(u)$ about u=1 explicitly:

THEOREM 5. For $s \ge 0$

$$f_s(u) = s! L^s(u) \cdot (1-u)^{-s-1} + s! s (H_s - 2) L^{s-1}(u) \cdot (1-u)^{-s-1} + \mathcal{R}_{s-2,s}(u).$$

Proof. From Theorem 4 we know that $f_i(u)$ is of the form

$$f_i(u) = i!L^i(u) \cdot (1-u)^{-i-1} + a_i i!L^{i-1}(u) \cdot (1-u)^{-i-1} + \Re_{i-2,i}(u)$$

with some constant a_i . Observing that

$$f'_{i}(u) = (i+1)!L^{i}(u)(1-u)^{-i-2} + (i+a_{i}(i+1))i!L^{i-1}(u)(1-u)^{-i-2} + \mathscr{R}_{i-2,i+1}(u),$$

$$f^{(j)}_{i}(u) = (i+j)!L^{i}(u)(1-u)^{-i-j-1} + \mathscr{R}_{i-1,i+j}(u). \quad (j \ge 2)$$

and inserting these formulas in the definition of $h_s(u)$ (Theorem 2) we obtain

$$\begin{split} h_{s}(u) &= \sum_{i=1}^{s-1} {s \choose i} [i!L^{i}(u)(1-u)^{-i-1} + a_{i}i!L^{i-1}(u)(1-u)^{-i-1} + \mathcal{R}_{i-2,i}(u)] \times \\ &\times [(s-i)!L^{s-i}(u)(1-u)^{-s+i-1} + a_{s-i}(s-i)!L^{s-i-1}(u)(1-u)^{-s+i-1} + \\ &+ (s-i)(s-i)!L^{s-i-1}(u)(1-u)^{-s+i-1} + \mathcal{R}_{s-i-2,s-i}(u)] + \\ &+ (1-u)^{-1} \sum_{r=1}^{s} {s \choose r} [s!L^{s-r}(u)(1-u)^{-s-1} + \mathcal{R}_{s-r-1,s+1}(u)] = \\ &= s!(s-1)L^{s}(u)(1-u)^{-s-2} + s!L^{s-1}(u)(1-u)^{-s-2} \times \\ &\times \left[s + \sum_{i=1}^{s-1} (a_{i} + a_{s-i} + s - i) \right] + \mathcal{R}_{s-2,s+1}(u). \end{split}$$

On the other hand we have

$$f'_{s}(u) - 2(1-u)^{-1}f_{s}(u) = (s+1)!L^{s}(u)(1-u)^{-s-2} + (s+a_{s}(s+1))s!L^{s-1}(u)(1-u)^{-s-2} - (-2s!L^{s}(u)(1-u)^{-s-2} - 2a_{s}s!L^{s-1}(u)(1-u)^{-s-2} + \Re_{s-2,s+1}(u) = s!(s-1)L^{s}(u)(1-u)^{-s-2} + (s+a_{s}(s-1))s!L^{s-1}(u)(1-u)^{-s-2} + \Re_{s-2,s+1}(u).$$

Comparing the coefficients of $s!L^{s-1}(1-u)^{-s-2}$ we obtain the recurrence relation

$$(s-1)a_s = {\binom{s}{2}} + 2 \cdot \sum_{i=1}^{s-1} a_i.$$

Subtracting this equation from

$$sa_{s+1} = {\binom{s+1}{2}} + 2 \cdot \sum_{i=1}^{s} a_i$$

we derive

$$sa_{s+1} = (s-1)a_s + 2a_s + s,$$

or

$$a_{s+1}/(s+1) = a_s/s + 1/(s+1), \qquad a_1 = -1$$

Summing up we get

$$a_s/s = -1 + \sum_{i=1}^{s-1} (i+1)^{-1} = H_s - 2,$$

276

A contribution to the analysis...

hence

$$a_s = s \left(H_s - 2 \right)$$

and the proof is complete.

Combining Theorem 5 with formula (4.2) we reach our final result

THEOREM 6. For $s \ge 0$

$$\beta_s(n) = n^s \cdot \log^s n + s(\gamma - 2) n^s \cdot \log^{s-1} n + \mathcal{O}(n^s \cdot \log^{s-2} n),$$

where $\gamma = .57721 \dots$ denotes Euler's constant.

Proof. From Theorem 5 and (4.2)

$$\beta_{s}(n) = s! P_{s}(H_{n+s}^{(1)} - H_{s}^{(1)}, \dots, H_{n+s}^{(s)} - H_{s}^{(s)}) {\binom{n+s}{s}} + s(H^{s}-2) s! P_{s-1}(H_{n+s}^{(1)} - H_{s}^{(1)}, \dots, H_{n+s}^{(s-1)} - H_{s}^{(s-1)}) {\binom{n+s}{s}} + \mathcal{O}(n^{s} \cdot \log^{s-2} n),$$

since

$$P_{p}(H_{n+s}^{(1)}-H_{s}^{(1)},\ldots,H_{n+s}^{(p)}-H_{s}^{(p)})=\mathcal{O}(n^{s}\cdot H_{n}^{p})=\mathcal{O}(n^{s}\cdot \log^{p} n).$$

Regarding

$$P_{p}(s_{1},\ldots,s_{p})=s_{1}^{p}-\binom{p}{2}s_{1}^{p-2}s_{2}+\ldots$$

we have

$$\beta_{s}(n) = {\binom{n+s}{s}} [s! (H_{n+s} - H_{s})^{s} + s! s (H_{s} - 2) (H_{n+s} - H_{s})^{s-1}] + \\ + \mathcal{O}(n^{s} \cdot \log^{s-2} n) = n^{s} [H_{n+s}^{s} - sH_{n+s}^{s-1} H_{s} + s (H_{s} - 2) H_{n+s}^{s-1}] + \\ + \mathcal{O}(n^{s} \cdot \log^{s-2} n) = n^{2} [(\log (n+s) + \gamma)^{s} - 2s (\log (n+s) + \gamma)^{s-1}] + \\ + \mathcal{O}(n^{s} \cdot \log^{s-2} n) = n^{s} [\log^{s} n + s (\gamma - 2) \log^{s-1} n] + \mathcal{O}(n^{s} \cdot \log^{s-2} n).$$

REFERENCES:

- [1] E. A. Bender, Asymptotic methods in enumeration, SIAM Review 16 (1974), 485-515.
- [2] D. H. Greene and D. E. Knuth, Mathematics for the analysis of algorithms, 2nd edition, Birkhäuser, Boston-Basel-Stuttgart 1982.
- [3] D. E. Knuth, The art of computer programming, Vol. 1: Fundamental algorithms, Addison Wesley, Reading 1972.
- [4] _____, Mathematical analysis of algorithms, in: Information Processing 71, 19-27, North Holland 1972.

1. 1.

- [5] I. D. G. Mac Leod, An algorithm for in-situ permutation, Austral. Comput. J. 2 (1970), 16-19.
- [6] D. A. Zave, A series expansion involving the harmonic numbers, Inform. Process. Lett. 5 (1976), 75-77.

(Received March 1, 1985)

P. Kirschenhofer and H. Prodinger Inst. f. Algebra u. Diskrete Mathematik R. F. Tichy Inst. f. Analysis, Technische Mathematik u. Versicherungsmathematik Technical University of Vienna Wiedner Hauptstr. 8-10 A-1040 Vienna, Austria

PRILOG ANALIZI (IN SITU) PERMUTACIJA

P. Kirschenhofer, H. Prodinger i R. F. Tichy, Beč, Austrija

Sadržaj

Postoji jednostavni algoritam koji zamjenjuje (prevodi) (x_1, \ldots, x_n) sa $(x_{p(1)}, \ldots, x_{p(n)})$ gdje je $\pi = (p(1), \ldots, p(n))$ permutacija od 1, 2, ..., n, koji u biti ne zahtijeva dodatno korištenje memorije.

U ovom redu se nastavljaju istraživanja D. E. Knutha o jednom karakterističnom parametru tog algoritma. Korištenjem tehnika funkcija izvodnica dobiveno je osim alternativnih izvoda nekoliko rezultata Knutha i nekoliko novih rezultata.