ON THE ANALYSIS OF AN ALGORITHM TO GENERATE A RANDOM CYCLIC PERMUTATION

HELMUT PRODINGER

Abstract. Sattolo has presented an algorithm to generate cyclic permutations at random. In this note the two parameters "number of moves" and "distance" are analyzed.

1. Introduction

Sattolo [1] generates a random cyclic permutation as follows: She starts with $12\dots n$, then a random integer between 1 and $n-1$ is chosen, say i, and the numbers in positions i and n are interchanged. Then a random integer between 1 and $n-2$ is chosen, say j, and the numbers in positions j and $n-1$ are interchanged, and so on. After $n-1$ iterations, a random cyclic permutation has been obtained.

Here is an example, with $n = 5$ and the random numbers $4, 1, 2, 1$:

$$
\begin{array}{c}
1\ 2\ 3\ \underline{4}\ \mathbf{5}\\
\underline{1}\ 2\ 3\ \mathbf{5}\ 4\\
5\ \underline{2}\ \mathbf{3}\ 1\ 4\\
\underline{5}\ \mathbf{3}\ 2\ 1\ 4\\
3\ 5\ 2\ 1\ 4
\end{array}
$$

The result is the cyclic permutation $1 \to 3 \to 2 \to 5 \to 4 \to 1$.

In this note we are interested in the (average) number of times digit k moves; in the example we obtain the numbers 1, 1, 2, 1, 3 for $k = 1, \dots, 5$. Altogether that means $1+1+2+1+3 = 8$ moves, which is clear, since each iteration moves exactly 2 digits.

We are also considering the (average) distance digit k travels. In the example we obtain the numbers 3, 1, 2, 1, 5 for $k = 1, \dots, 5$.

2. The number of moves

The probability generating functions $\varphi_{n,k}(u)$ are defined as follows:

$$\varphi_{n,k}(u) = \sum_l a_{n,k;l} u^l = \sum_l \mathbb{P}\{\text{digit } k \text{ makes } l \text{ moves}\} u^l.$$

Date: January 16, 2000.

The probability that position k is never chosen by the random number generator, is given by

$$\frac{n-2}{n-1}\cdot\frac{n-3}{n-2}\cdots\frac{k-1}{k}=\frac{k-1}{n-1}.$$

If position k is chosen by the random generator, digit k is moved once (to the right) and never moves again. If position k is never chosen by the random generator, it starts moving (to the left) when the positions $k+1,\dots,n$ have already been dealt with. This leads to the recursion

$$\varphi_{n,k}(u)=\frac{n-k}{n-1}\,u+\frac{k-1}{n-1}\,\varphi_{k,k}(u),\qquad 1\le k<n.$$

It remains to discuss the instance $k=n$. The recursion is

$$\varphi_{n,n}(u)=\frac{u}{n-1}\Big[\varphi_{n-1,1}(u)+\varphi_{n-1,2}(u)+\cdots+\varphi_{n-1,n-1}(u)\Big],\quad n\ge 2,\ \varphi_{1,1}(u)=1.$$

To understand it, note that one move is made, and, according to the random generator, each position $1,2,\dots,n-1$ might be chosen with probability $\frac{1}{n-1}$. Then digit n is in position k, say, and the further moves are described by $\varphi_{n-1,k}(u)$.

Let us start with the numbers $a_{n,k;1}$. It is easy to see that

$$a_{n,1;1}=1,\quad a_{n,2;1}=1,\quad a_{n,k;1}=\frac{n-k}{n-1},\quad k\ge 3,\quad n\ge 2,\quad a_{1,1;1}=0.$$

The recursions tell us furthermore that

$$a_{n,k;2}=\frac{1}{n-1}\Big[a_{k-1,1;1}+a_{k-1,2;1}+\cdots+a_{k-1,k-1;1}\Big].$$

Hence

$$a_{n,k;2}=\frac{1}{n-1}\Big[\frac{k-1}{2}+\frac{1}{k-2}\Big],\quad n\ge k\ge 3,\quad a_{n,k;2}=0\text{ otherwise.}$$

Similarly

$$a_{n,k;3}=\frac{1}{n-1}\Big[a_{k-1,1;2}+a_{k-1,2;2}+\cdots+a_{k-1,k-1;2}\Big]$$

and thus

$$a_{n,k;3}=\frac{1}{(n-1)(k-2)}\Big[-1+\frac{(k-1)(k-2)}{4}+H_{k-3}\Big],\quad n\ge k\ge 3.$$

We can also compute the next ones as

$$a_{n,k;4}=\frac{1}{(n-1)(k-2)}\Big[-H_{k-3}-\frac14+\frac{(k-1)(k-2)}{8}+\frac12\Big(H_{k-3}^2-H_{k-3}^{(2)}\Big)\Big],\ n\ge k\ge 3,$$

but it is clear the expressions become very messy.

The numbers $H_n = \sum_{k=1}^{n} \frac{1}{k}$ and $H_n^{(2)} = \sum_{k=1}^{n} \frac{1}{k^2}$ are harmonic numbers of the first and second order.

Now we proceed to the expected numbers of moves $E_{n,k} = \varphi'_{n,k}(1)$. The recursions are

$$E_{n,k} = \frac{n-k}{n-1} + \frac{k-1}{n-1} E_{k,k},$$

and

$$E_{n,n} = 1 + \frac{1}{n-1}\Big[E_{n-1,1} + E_{n-1,2} + \cdots + E_{n-1,n-1}\Big].$$

It is a simple task to find the answers

$$E_{n,k} = \frac{n+2k-5}{n-1}, \qquad k \geq 2, \quad E_{n,1} = 1, \quad n \geq 2, \quad E_{1,1} = 0$$

and prove them by induction. Since this can be done by Maple, it is omitted.

Note that

$$\sum_{k=1}^{n} E_{n,k} = 1 + \sum_{k=2}^{n} \frac{n+2k-5}{n-1} = 2(n-1),$$

as it should.

The second factorial moments can be computed in a similar fashion;

$$\varphi''_{n,k}(1) = \frac{1}{n-1}(8k - 4H_{k-2} - 16)$$

for $k \geq 2$, otherwise 0.

Theorem 1. *The average number of moves of digit k in Sattolo's algorithm and the variance are given by*

$$\begin{aligned} E_{n,k} &= \frac{n+2k-5}{n-1}, \qquad k \geq 2, \quad E_{n,1} = 1, \quad n \geq 2, \quad E_{1,1} = 0, \\ V_{n,k} &= \frac{2(k-2)(3n+1-2k)}{(n-1)^2} - \frac{4H_{k-2}}{n-1}, \qquad k \geq 2, \quad V_{n,1} = 0. \quad \square \end{aligned}$$

3. The distance

It is not too hard to change things appropriately.

The recursions for the probability generating functions are now

$$\begin{aligned} \varphi_{n,k}(u) &= \frac{u^{n-k}}{n-1} + \frac{n-2}{n-1}\varphi_{n-1,k}(u), \qquad 1 \leq k < n, \\ \varphi_{n,n}(u) &= \frac{1}{n-1}\Big[\varphi_{n-1,1}(u)u^{n-1} + \varphi_{n-1,2}(u)u^{n-2} + \cdots + \varphi_{n-1,n-1}(u)u\Big], \quad n \geq 2, \\ \varphi_{1,1}(u) &= 1. \end{aligned}$$

From this we get the expectation

$$E_{n,k} = \frac{n}{2} - k + 1 + \frac{k(3k-5)}{2(n-1)}, \quad k \geq 2, \qquad E_{n,1} = \frac{n}{2}, \quad n \geq 2, \quad E_{1,1} = 0,$$

and the second factorial moment

$$E^{(2)}_{n,k} = \frac{n^2}{3} - n(k - \tfrac{1}{3}) + k(k-1) + \frac{(k-2)(14k^2+k-21)}{18(n-1)}, \quad k \geq 2,$$
$$E^{(2)}_{n,1} = \frac{n(n-2)}{3}, \quad n \geq 2, \qquad E^{(2)}_{1,1} = 0.$$

Theorem 2. *The average distance that digit k travels in Sattolo's algorithm and the variance are given by*

$$E_{n,k} = \frac{n}{2} - k + 1 + \frac{k(3k-5)}{2(n-1)}, \quad k \geq 2, \qquad E_{n,1} = \frac{n}{2}, \quad n \geq 2, \quad E_{1,1} = 0,$$
$$V_{n,k} = \frac{(n-1)^2}{12} - \frac{18k^2-30k+1}{12} + \frac{(4k-7)(17k^2-13k-6)}{18(n-1)} - \frac{k^2(3k-5)^2}{4(n-1)^2}, \quad k \geq 2,$$
$$V_{n,1} = \frac{n(n-2)}{12}. \quad \square$$

Averaging over all digits we find the average distance that a random element travels:

$$\frac{1}{n}\sum_{k=1}^{n} E_{n,k} = \frac{n+1}{2} - \frac{1}{n}.$$

References

[1] S. Sattolo. An algorithm to generate a random cyclic permutation. *Information Processing Letters*, 22:315–317, 1986.

Helmut Prodinger, The John Knopfmacher Centre for Applicable Analysis and Number Theory, Department of Mathematics, University of the Witwatersrand, P. O. Wits, 2050 Johannesburg, South Africa, email: helmut@gauss.cam.wits.ac.za, homepage: http://www.wits.ac.za/helmut/index.htm