Opgave A NAW 5/5 nr. 2 jun 2004

Jaap Spies

June 2004

The problem

Introduction

The sequence 333111333131333111333... is identical to the sequence of its block lengths. Compute the frequency of the number 3 in this sequence.

Solution

This sequence is known as the Kolakoski-(3,1) sequence. See N.J.A. Sloane's On-Line Encyclopedia of Integer Sequences, sequence number $A064353$, which is in fact the Kolakoski-(1,3) sequence, different only in the first position. See [1].

Michael Baake and Bernd Sing wrote: Unlike the (classical) Kolakoski sequence on the alphabet {1,2}, its analogue on {1,3} can be related to a primitive substitution rule. See [2] and [3]. We base our calculations on section 2 of this paper.

Let $A = 33$, $B = 31$ and $C = 11$. In the case of $\mathrm{Kol}(3, 1)$ the substitution σ and the matrix M of the substitution are given by

$$\sigma : \begin{array}{lcl} A & \mapsto & ABC \\ B & \mapsto & AB \\ C & \mapsto & B \end{array} \quad \text{and} \quad M = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \tag{1}$$

where $m_{ij} = 1$ if and only if there is corresponding mapping in σ, for instance $A \mapsto ABC$ corresponds to the fist column of M, etcetera.

An infinite fixed point can be obtained as follows:

$$A \mapsto ABC \mapsto ABCABB \mapsto \dots \tag{2}$$

This corresponds to

$$333111333131\dots \tag{3}$$

which is the unique infinite Kol$(3,1)$. The matrix M is primitive because M^3 has only positive entries. The characteristic polynomial $P(\lambda)$ of M is

$$P(\lambda) = \lambda^3 - 2\,\lambda^2 - 1, \tag{4}$$

and has one real root λ_1 and two complex roots $\lambda_{2,3}$. We have

$$2.205569 \approx \lambda_1 > 1 > |\lambda_2| = |\lambda_3| \approx 0.67 \tag{5}$$

According to the Perron-Frobenius Theorem there is a positive eigenvector to λ_1. We easily verify that $\mathbf{x_1} = (\lambda_1, \lambda_1^2 - \lambda_1, 1)^T$ is such an eigenvector. Starting with $\mathbf{x}(0) = (1,0,0)^T$ we define

$$\mathbf{x}(k+1) = M\mathbf{x}(k) \tag{6}$$

The asymptotical behavior of this system will be of the form $\mathbf{x}(n) = c{\cdot}(\lambda_1)^n\mathbf{x_1}$ for some value of c.

From $\mathbf{x}(n)$ we can calculate the number of A's, B's and C's. In $A = 33$ there are two 3's, etcetera, so we can easily calculate the relative frequencies of the letters of the alphabet. The frequency of the '3':

$$\rho_3 = \frac{2\cdot\lambda_1 + 1\cdot(\lambda_1^2 - \lambda_1) + 0\cdot 1}{2\cdot(\lambda_1^2 + 1)} \approx 0.6027847150 \tag{7}$$

References

[1] http://www.research.att.com/projects/OEIS?Anum=A064353
[2] Baake, Sing: Kolakoski-(3,1) is a (deformed) Model Set, Canad. Math. Bull. 47, No. 2, 168–190 (2004)
[3] See also http://arxiv.org/abs/math.MG/0206098