The Influence of Memory on Creating Order

Ulrich Tamm
Dept. of Mathematics
University of Bielefeld
P. O. Box 100131
33501 Bielefeld, Germany

Abstract. A simple strategy for creating order in a special multi-user model is analyzed. In each time unit $s \geq 2$ bits arrive at an organizer who in the same time unit has to choose one bit as his output, preferably he will choose a 1, which may be one of the arriving bits or a 1 stored in some memory device. Contrasting to the original model for creating order introduced in [1] the number of bits in the memory device now may vary in time. We shall study the influence of the maximum size of the memory device on the expected occurrence of the first 0 in the sequence produced by the organizer.

I. Introduction

Ahlswede, Ye, and Zhang [1] introduced the following model for creating order in sequence spaces. We are given a box containing (a fixed number) β balls labelled by letters from an alphabet of size α. In each time unit a person $\mathcal{O}$ – denoted as organizer – takes out one ball of the box which is replaced by a new ball thrown into the box by a second person $\mathcal{I}$. The aim of the organizer is to reduce the space of possible output sequences. As a measure for the efficiency of the ordering process the number of possible output sequences and the entropy of the output space have been studied (cf. also [2], [3]). A related model in which the output sequence is regarded as a message from person $\mathcal{I}$ to a decoder $\mathcal{D}$ was considered by several authors studying the permuting channel, e. g. [4], [5], and [6].

In this paper a multi–user version of the original model for creating order is discussed. Now there are $s \geq 2$ persons $\mathcal{I}_1, \ldots, \mathcal{I}_s$, say, throwing balls labelled either 0 or 1 into the box. We shall present the model using a slightly different terminology. In each time unit s sources $\mathcal{I}_1, \ldots, \mathcal{I}_s$ produce one bit each. These s bits arrive at an organizer who in the same time unit has to choose one bit for output. This bit may be one of the s arriving bits or a bit stored in some memory device (the box), in which the bits not used so far may be stored. The organizer follows a simple strategy: if it is possible the output must be a 1. So if one of the arriving bits is a 1, the organizer will put out a 1 for sure. The bits not used for output he may store in the memory device. If all the s sources produce a 0, then the organizer will take a look at the memory. If there is still a 1 contained he will put out a 1, otherwise he must put out a 0.

We assume that in the beginning the memory device is empty. At some point it may occur that no further 1 can be stored, since the device is full of 1's (a 0 may be replaced by a 1). In this case there is a maximum size or capacity of M bits which cannot be superceded. Of theoretical interest is also the not very realistic model, in which the memory device can store infinitely many bits.

Observe that contrasting to the original model for creating order the size of the memory device (or box) now may vary in time. A natural question is: how much influence does the maximum size of the memory have on the behaviour of the sequence of bits arranged by the organizer? Of course, in the strategy considered the organizers aim is to produce the all–one sequence and we shall study how well he can manage to achieve this goal. As a new measure for the influence of the memory we consider the expected value of the first occurence of a 0 in this sequence. We shall denote this expectation as

$$E_0 = \sum_{t=1}^{\infty} t \cdot \text{Prob(first 0 at time t)}. \tag{1}$$

Further, we shall denote by

$$\underline{x}^i = (x_1^i, x_2^i, \ldots), \tag{2}$$

the sequences of bits produced by the sources $\mathcal{I}_i$, respectively and by

$$\underline{z} = (z_1, z_2, \dots) \tag{3}$$

the sequence arranged by the organizer according to the above strategy. The results below are mainly derived from properties of the sequence

$$\underline{u} = (u_1, \dots u_s, u_{s+1}, \dots, u_{2s}, u_{2s+1}, \dots) = (x_1^1, \dots, x_1^s, x_2^1, \dots, x_2^s, x_3^1, \dots) \tag{4}$$

which is obtained by merging the sequences $\underline{x}^i, i = 1, \dots, s$ into one sequence $\underline{u}$, where the bits in the positions $\equiv i \bmod s$ are those produced by $\mathcal{I}_i$. (Observe that the order of the incoming bits at each time unit does not matter, since the organizer waits until all the s bits have arrived. So $\underline{u}$ may also be defined by any other merging procedure which assigns the postions $(s-1)t+1, \dots, st$ to the bits arriving in time unit t).
Our first result is obtained for two memoryless, symmetrical sources, i. e., $\mathcal{I}_1$ and $\mathcal{I}_2$ produce 1 and 0 with equal probability $\frac{1}{2}$.In this case it is

$$\text{Prob(first 0 at time } t-1) = \frac{1}{4^t} \cdot a(0, t-1), \tag{5}$$

where $a(0, t-1)$ denotes the number of sequences $\underline{u}$ produced by the two sources leading to the all-one sequence as output with actual memory size (= number of 1's in the device) 0 at time $t-1$. Clearly an empty memory at time $t-1$ is a necessary condition for the occurrence of the first 0 at time unit t, which then happens with probability $\frac{1}{4}$ (4^{t-1} is the number of all possible sequences $\underline{u}$ until time $t-1$).
Analogously, the numbers $a(m, t)$ are defined for actual memory size $m = 1, \dots, M, t = 1, 2, \dots$. Let us dnote by $\vec{u}_t = (u_{2t-1}, u_{2t})$ the two bits arriving at time t. Now observe that in each time unit the source inputs $\vec{u}_t = 10$ and $\vec{u}_t = 01$ do not change the actual size of the memory, $\vec{u}_t = 00$ decreases the memory size by one bit (and is forbidden for $m = 0$, since in this case the organizer must put out a 0) and $\vec{u}_t = 11$ increases the memory size by one bit if $m < M$ (and does not change the memory size if the capacity $m = M$ is reached. So we have a random walk in the memory size expressed by the recursion formulae for the numbers $a(m, t)$

$$\begin{pmatrix} a(0,t) \\ \vdots \\ a(M,t) \end{pmatrix} = A_{M+1} \cdot \begin{pmatrix} a(0,t-1) \\ \vdots \\ a(M,t-1) \end{pmatrix} \tag{6}$$

where

$$A_n = \begin{pmatrix} 2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 2 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 3 \end{pmatrix}, \tag{7}$$

with initial matrices $A_1 = (3)$, $A_2 = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$.

Theorem 1: If $\mathcal{I}_1$ and $\mathcal{I}_2$ are memoryless, symmetrical sources, then
i)

$$E_0 = \frac{1}{4} \sum_{i=1}^{M+1} c_i \left(1 - \frac{\lambda_i^{(M+1)}}{4}\right)^{-2} \tag{8}$$

where $\lambda_i^{(M+1)}, i = 1, \ldots, M+1$ are the eigenvalues

$$\lambda_i^{(M+1)} = 4 \cdot \sin \frac{i\pi}{2M+1} \tag{9}$$

of the matrix A_{M+1} and $c_i, i = 1, \ldots, M+1$ are appropriate constants.

ii) E_0 does not exist if the size of the memory $M(t)$ at time t is bounded by a function $f(t)$ which exceeds every positve integer M from some time $t_0(M)$ on.

We shall prove Theorem 1 in the next section where we shall also take a closer look at the transition matrices defined in (7) and further matrices of a similar structure. The matrices A_n have been studied before in geometry and they can also be obtained from the squares of the transition matrices of the random walk which occurs in the study of the Brownian motion. Their characteristic polynomials are known to be combinations of Chebyshev polynomials. Further in Section 2 for small capacities $M = 0, 1$ the numbers $a(0, t)$ will be determined exactly for small capacities $M = 0, 1$ and also the case in which two identical nonsymmetric sources give bits to the organizer will be considered.

Part (ii) of Theorem 1 states that E_0 does not exist if the number of bits $M(t)$ that the memory device is able to store at time t is an arbitrarily slowly growing function (in fact, the condition is even weaker). If every incoming bit can be stored, i. e., $M(t) = t$ is linear in time, then there is a second proof based on some property of the merged sequence $\underline{u}$, which also extends to the situation of s incoming bits in each time unit.

Theorem 2: Let there be s identical sources $\mathcal{I}_1$, ..., $\mathcal{I}_s$ producing one bit each per time unit with $\text{Prob}(X = 1) = p$, $\text{Prob}(X = 0) = 1 - p$. If the memory device can store every incoming bit, then the expected value for the occurence of the first 0 in the sequence arranged by an organizer (if he puts out a 1 if possible) is

$$E_0 \begin{cases} = \infty, & p = \frac{1}{s} \\ \leq \infty, & p < \frac{1}{s} \end{cases} \tag{10}$$

Theorem 2 will be proved in Section 3. It turns out that the numbers $a(0, t)$ (defined as before for the finite memory case) are generalized Catalan numbers

$$C_t^{(s)} = \frac{1}{(s-1) \cdot t + 1} \cdot \binom{st}{t} \tag{11}$$

Finally, in Section 4 some generalizations and open problems will be stated.

II. The transition matrices A_n

The transition matrices A_n and further matrices of a similar structure we shall consider in this section are closely related to Chebyshev polynomials $(t_n(x))_{n=1,2,\ldots}$ and Chebychev polynomials of the second kind $(u_n(x))_{n=1,2,\ldots}$. The reason is that these polynomials occur as determinants of special matrices, namely (cf. [7], p. 228)

$$t_n(x) = \frac{n}{2} \cdot \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^i}{n-i} \binom{n-i}{i} (2x)^{n-2i} = \det T_n(x) \tag{12}$$

with

$$T_n(x) = \begin{pmatrix} x & 1 & 0 & \ldots & 0 & 0 & 0 \\ 1 & 2x & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 2x & 1 \\ 0 & 0 & 0 & \ldots & 0 & 1 & 2x \end{pmatrix}, \tag{13}$$

$$u_n(x) = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^i \binom{n-i}{i} (2x)^{n-2i} = \det U_n(x) \tag{14}$$

with

$$U_n(x) = \begin{pmatrix} 2x & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 2x & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 2x & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 2x \end{pmatrix}, \tag{15}$$

Further, in [8] it is shown that

$$u_n(x) + u_{n-1}(x) = \det V_n(x), \tag{16}$$

where

$$V_n(x) = \begin{pmatrix} 2x & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 2x & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 2x & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 2x+1 \end{pmatrix}, \tag{17}$$

Observe that the transition matrix $A_M = V_M(1)$ occurs for the special value $x = 1$.
The following well–known properties of the Chebyshev polynomials (cf. [8] and [9]) will be used later on.

$$t_n(x) = 2x \cdot t_{n-1}(x) - t_{n-2}(x), u_n(x) = 2x \cdot u_{n-1}(x) - u_{n-2}(x) \tag{18}$$

$$t_n(x) = u_n(x) - u_{n-2}(x) \tag{19}$$

$$4 \cdot \sin \frac{i\pi}{2n+1}, i = 1, \dots, , n \text{ are the eigenvalues of } V_n(x) \tag{20}$$

Proof of Theorem 1:

i) With (6)

$$\begin{pmatrix} a(0,t) \\ \vdots \\ a(M,t) \end{pmatrix} = A_{M+1} \begin{pmatrix} a(0,t-1) \\ \vdots \\ a(m,t) \end{pmatrix} = A_{M+1}^t \begin{pmatrix} a(0,0) \\ \vdots \\ a(M,0) \end{pmatrix} = A_{M+1}^t \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{21}$$

since in the beginning the memory device does not contain any 1. A_{M+1} is symmetric, hence there exists an orthogonal matrix $P = (p_{ij})_{i,j=1,\dots,M+1}$ with $P^T = P^{-1}$ such that

$$P \cdot A_{M+1} \cdot P^{-1} = \mathrm{diag}(\lambda_1, \dots \lambda_{M+1}), \tag{22}$$

where $\mathrm{diag}(\lambda_1, \dots, \lambda_{M+1})$ is the diagonal matrix consisting of the eigenvalues of A_{M+1}. Now $A_{M+1}^t = P^{-1} \cdot (\mathrm{diag}(\lambda_1^t, \dots, \lambda_{M+1})^t) \cdot P$. With (21) and $c_i = p_{1i}^2$ then

$$a(0,t) = \sum_{i=1}^{M+1} c_i \cdot \lambda_i^t \tag{23}$$

Now using (23) and $\sum_{t=1}^{\infty} t \cdot x^{t-1} = \frac{1}{(1-x)^2}$

$$E_0 = \sum_{t=1}^{\infty} t \frac{a(0,t-1)}{4^t} = \frac{1}{4} \sum_{i=1}^{M+1} c_i \sum_{t=1}^{\infty} t \cdot (\frac{\lambda_i}{4})^{t-1} = \frac{1}{4} \sum_{i=1}^{M+1} c_i \cdot (1 - \frac{\lambda_i}{4})^{-2} \tag{24}$$

The eigenvalues of A_{M+1} have been listed under (10) and thus (i) is proved.

ii) Observe that the largest eigenvalues $\lambda_{max}^{(M+1)}$ of the matrices A_{M+1} form a sequence converging towards 4, hence the series of expected values for the occurrence of the first 0 $E_0^{(M)}$ is divergent. If the maximum size $M(t)$ of the memory device varies in time such that for every positive integer M it is $M(t) > M$ for some $t_0(M)$ on, then $E_0 > E_0^{(M)}$ for every fixed capacity M (with appropriate rescaling, set $t_0(M) = 0$).

Remark: It is clear from the proof that the results of Theorem 1 also hold if we start with any number $m > 0$ of 1's in the memory device, only the constants c_i will change then. Implicitly this fact has already been used in the proof of part (ii).

The matrices A_M can further be obtained as a submatrix of the squares of matrices

$$B_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 2 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}, \tag{25}$$

with $B_1 = (0)$ and $B_2 = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$.

To see this, observe that the recursion formulae in (6) are obtained by regarding the single time units t of the ordering process, which are represented by the two bits u_{2t-1}, u_{2t} of the sequence $\underline{u}$. It is also possible to obtain a recurrence from the single components of $\underline{u}$. In order to do so, consider the transitions (for $i = 1, 2, \dots; k = 0, 1, \dots, n$)

$$b(i+1,k) = \begin{cases} b(i,k+1) & k = 0, \\ b(i,k-1) + b(i,k+1) & k = 1, \dots n-2 \\ b(i,k-1) + 2 \cdot b(i,k+1) & k = n-1 \\ b(i,k-1), & k = n \end{cases} \tag{26}$$

which just define the matrices B_n. This is just a random walk on the line $\{0, \dots, n\}$ with absorption at the point 0 and reflection at point n, which occurs in the study of the Brownian motion (cf. [10]). Now observe that the recurrence formulae (6) are obtained from (26) by combining two consecutive instances $i, i+1$, i. e., $b(i+2,k)$ is obtained from the $b(i,.)$ via the same recursion as $a(t,m)$ is obtained from the $a(t-1,.)$. Since for even i $b(i,k) = 0$ if k is odd, we have

Proposition 1: Let $B_{2n+1}^2 = (b_{ij})_{i,j=1,\dots 2n+1}$, where B_{2n+1} and A_n are the matrices defined in (25) and (7). Then

$$A_n = (b_{ij})_{i,j \text{ even}}. \tag{27}$$

Moreover, by an appropriate enumeration of lines and columns

$$B_{2n+1}^2 = \begin{pmatrix} A_n & 0 \\ 0 & C_{n+1} \end{pmatrix}. \tag{28}$$

where

$$C_{n+1} = (b_{ij})_{i,j \text{ odd}} = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 2 & 2 \\ 0 & 0 & 0 & \dots & 0 & 1 & 2 \end{pmatrix}, \tag{29}$$

with $C_2 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ and $C_3 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 1 & 2 \end{pmatrix}$.

Analogously, we can consider the matrices

$$\overline{A}_n = \begin{pmatrix} 2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 2 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 2 \end{pmatrix}, \tag{30}$$

$$\overline{B}_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}, \tag{31}$$

$$\overline{C}_n = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 2 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 1 \end{pmatrix}, \tag{32}$$

where following the same lines it can be derived (with $\overline{B}_n^2 = (\overline{b}_{ij})_{i,j=1,\dots,n}$)

$$\overline{B}_{2n+1}^2 = \begin{pmatrix} \overline{A}_n & 0 \\ 0 & \overline{C}_{n+1} \end{pmatrix}. \tag{33}$$

where

$$\overline{A}_n = (\overline{b}_{ij})_{i,j \text{ even}} \tag{34}$$

and

$$\overline{C}_{n+1} = (\overline{b}_{ij})_{i,j \text{ odd}} \tag{35}$$

The matrices B_n and $\overline{B}_n$ are well known as transition matrices for certain random walks (cf. [10] and [11], pp. 238–240) and the matrices A_n and $\overline{A}_n$ have an application in geometry (cf. [8]). Think of the vertices $P_1, \dots, P_N$ of a regular N–gon drawn on a unit circle. Then the eigenvalues of $A_{\frac{N-1}{2}}$ (if N is odd) or $\overline{A}_{\frac{N-2}{2}}$ (if N is even) give the squares of the different distances $\overline{P_i, P_j}, i, j = 1, \dots, N$.

From this property it is immediate that the largest eigenvalues $\lambda_{max}^{(n)}$ of the matrices A_n form a strictly increasing sequence with $\lim_{n\to\infty} \lambda_{max}^{(n)} = 4$. The reason is that the largest distance $\overline{P_1, P_{\frac{N-1}{2}}}$ tends to the diameter 2 of the unit circle. This property was the central argument in the proof of Theorem 1, where it was derived analytically using an explicit formula for the eigenvalues of A_n.

Now let us take a look at the characteristic polynomials.

Proposition 2: For the characteristic polynomials of the matrices we consider holds

$$\chi_{\overline{B}_n}(\lambda) = u_n(-\frac{\lambda}{2}) = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^{n-i} \binom{n-i}{i} \lambda^{n-2i} \tag{36}$$

$$\chi_{B_n}(\lambda)=t_n(-\frac{\lambda}{2})=\sum_{i=0}^{\lfloor\frac{n}{2}\rfloor}(-1)^{n-i}(\binom{n-i}{i}+\binom{n-i-1}{i-1})\lambda^{n-2i} \tag{37}$$

$$\chi_{A_n}(\lambda^2)=\frac{(-1)^{n+1}}{\lambda}\chi_{B_{2n+1}}(\lambda),\ \chi_{\overline{A}_n}(\lambda^2)=\frac{(-1)^{n+1}}{\lambda}\chi_{\overline{B}_{2n+1}}(\lambda), \tag{38}$$

$$\chi_{C_{n+1}}(\lambda)=-\lambda\chi_{A_n}(\lambda),\ \chi_{\overline{C}_{n+1}}(\lambda)=-\lambda\chi_{\overline{A}_n}(\lambda), \tag{39}$$

$$\mid\frac{1}{\lambda}(\chi_{B_{2n+1}}(\lambda))^2\mid=\mid\chi_{B^2_{2n+1}}(\lambda^2)\mid,\ \ \mid\frac{1}{\lambda}(\chi_{\overline{B}_{2n+1}}(\lambda))^2\mid=\mid\chi_{\overline{B}^2_{2n+1}}(\lambda^2)\mid, \tag{40}$$

Proof: By definition, (36) follows from the fact that $B_n = U_n(0)$.
Using the properties (18) and (19) of the Chebyshev polynomials, (37) can be derived via

$$\chi_{B_n}(\lambda)=-\lambda\cdot\chi_{\overline{B}_{n-1}}(\lambda)-2\cdot\chi_{\overline{B}_{n-2}}(\lambda)=-\lambda\cdot u_{n-1}(\frac{\lambda}{2})-2\cdot u_{n-2}(\frac{\lambda}{2})=u_n(\frac{\lambda}{2})-2\cdot u_{n-2}(\frac{\lambda}{2})$$

The characteristic polynomials in (38) can be derived using the recurrence $\chi_{A_n}(\lambda)=(2-\lambda)\cdot\chi_{A_{n-1}}(\lambda)-\chi_{A_{n-2}}(\lambda)$ which is the same as for $\chi_{\overline{A}_n}$, only the initial values differ: $\chi_{A_1}(\lambda)=3-\lambda,\ \chi_{A_2}(\lambda)=\lambda^2-5\lambda+5$, whereas $\chi_{\overline{A}_1}(\lambda)=2-\lambda,\ \chi_{\overline{A}_2}(\lambda)=\lambda^2-4\lambda+3$.
(39) can be derived using $\chi_{C_n}(\lambda)=(2-\lambda)\cdot\chi_{T_{n-1}(1)}(\lambda)-2\cdot\chi_{T_{n-2}(1)}(\lambda)$ and $\chi_{\overline{C}_n}(\lambda)=(1-\lambda)\cdot\chi_{T_{n-1}(1)}(\lambda)-\chi_{T_{n-2}(1)}(\lambda)$.
Finally (40), follows from (28), (33), (38), and (39).

Observe that for the coefficients r_i and $\overline{r}_i$ in the polynomials $\chi_{B_n}(\lambda)=\sum_{i=0}^n r_i\lambda^i$ and $\chi_{\overline{B}_n}(\lambda)=\sum_{i=0}^n\overline{r}_i\lambda^i$ it is

$$\sum_{i=0}^n\mid r_i\mid=L_n,\sum_{i=0}^n\mid\overline{r}_i\mid=F_n \tag{41}$$

where F_n and L_n denote the n–th Fibonacci and Lucas number, respectively defined by $F_n=F_{n-1}+F_{n-2}$ and $L_n=L_{n-1}+L_{n-2}$ with initial values $F_1=F_2=1$ and $L_1=1,L_2=3$.
The Fibonacci numbers also occur in the analysis of the expected value E_0 for maximum memory size $M=1$. Obviously, if there is no memory ($M=0$), then the number $a(0,t)=3^t$, where 3 is also the largest eigenvalue of $A_1=(3)$. In this case $E_0=\sum_{t=1}^\infty t\cdot(\frac{3}{4})^{t-1}\cdot\frac{1}{4}=4$ is just the expected value of the geometric distribution with parameter $\frac{1}{4}$, since with probability $\frac{1}{4}$ two 0's arrive at time t.

Proposition 3: For maximum memory size $M=1$ the numbers $a(0,t)$ and $a(1,t)$ are

$$a(0,2t)=5^t\cdot F_t,\ a(1,t)=5^t\cdot F_{t+1},\ a(0,2t+1)=5^t\cdot(2F_t+F_{t+1}),\ a(1,2t+1)=5^t\cdot(F_t+3F_{t+1}) \tag{42}$$

Proof:
We have the transition matrices $A_M=\begin{pmatrix}2&1\\1&3\end{pmatrix}$ Now $A_M^2=\begin{pmatrix}5&5\\5&10\end{pmatrix}=5\cdot\begin{pmatrix}1&1\\1&2\end{pmatrix}$, from which follows that $A_M^{2t}=5^t\cdot\begin{pmatrix}1&1\\1&2\end{pmatrix}^t=5^t\cdot\begin{pmatrix}F_t&F_{t+1}\\F_{t+1}&F_{t+2}\end{pmatrix}^t$ by a well–known property of the Fibonacci numbers.

Finally let us take a look at the transition matrices in the case where two identical nonsymmetric sources give bits to the organizer. If p is the probality for the sources to produce a 1, the transition matrices can be obtained as in (28) from the square of the matrices

$$B_n' = \begin{pmatrix} 0 & 1-p & 0 & \dots & 0 & 0 & 0 \\ p & 0 & 1-p & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & p & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & p & 0 \end{pmatrix}, \tag{43}$$

The matrices B_n' have been studied intensively in [10]. An exact determination of their eigenvalues seems to be quite difficult in the nonsymmetric case, however it is clear from Theorem 2 that for $M \longrightarrow \infty$ the sequence of expected values $E_0(M)$ is not divergent for $p < \frac{1}{2}$. The characteristic polynomial again is a combination of Chebyshev polynomials.

Propositon 4: The characteristic polynomial of B_n' is

$$\chi_{B_n'}(\lambda) = -\lambda(p(1-p))^{\frac{n-1}{2}} \cdot u_{n-1}(\frac{\lambda}{\sqrt{p(1-p)}}) - p(p(1-p))^{\frac{n-2}{2}} u_{n-2}(\frac{\lambda}{\sqrt{p(1-p)}}) \tag{44}$$

where $\overline{B}_n'$ is the matrix obtained from B_n' by changing the element $b_{n-1,n}'$ from 1 to $1-p$.

Proof: The characteristic polynomial of the matrix $\overline{B}_n'$ obeys the three–term recurrence

$$\chi_{\overline{B}_n'}(\lambda) = -\lambda \cdot \chi_{\overline{B}_{n-1}'}(\lambda) - p(1-p) \cdot \chi_{\overline{B}_{n-2}'}(\lambda) \tag{45}$$

which as solution yields the weighted Chebyshev polynomial of the second kind

$$\chi_{\overline{B}_n'} = (p(1-p))^{\frac{n}{2}} \cdot u_n(\frac{\lambda}{\sqrt{p(1-p)}}) \tag{46}$$

Now observe that the characteristic polynomial χ_{B_n}' is

$$\chi_{B_n'}(\lambda) = -\lambda \cdot \chi_{\overline{B}_{n-1}'}(\lambda) - p \cdot \chi_{\overline{B}_{n-2}'}(\lambda) \tag{47}$$

III. Every incoming bit can be stored

When the memory device can store every incoming bit, there is a simple criterion on the sequence $\underline{u}$, which guarantees that the organizer can put out the first 0 at time $t+1$. To see this, recall that there are two necessary conditions for the occurence of the first 0 at time unit t: 1) there are no further 1's in the memory device after the t–th bit has been put out by the organizer, 2) up to time t the all-one sequence has been arranged by the organizer. These conditions can be translated into conditions required from the sequence $\underline{u}$, namely

i)

$$\mathrm{wt}(u_1, \dots, u_{st}) = t, \tag{48}$$

ii)

$$\mathrm{wt}(u_1, \dots, u_{si}) \geq i \text{ for } i = 1, \dots, t-1. \tag{49}$$

As usual, here the weight $\mathrm{wt}(\vec{x})$ of a $\{0,1\}$–vector $\vec{x}$ denotes the number of 1's in $\vec{x}$. By condition (i) no 1's can be left in the device, since t 1's have arrived at the organizer and t 1's have been used for output. The second condition (ii) assures that at all time units before it was possible to put out a 1 (by the same argumentation).

The number $a(0,t)$ of all sequences of length s fulfilling the conditions (i) and (ii) are well known in Combinatorial Theory, these are just the generalized Catalan numbers defined in (11) (cf. [12] or [13], pp. 343–350).

Proof of Theorem 2: Recall that the sources $\mathcal{I}_1, \dots, \mathcal{I}_s$ each produce a 1 with probability p. Hence with probability $(1-p)^s$ only 0's arrive at the organizer at time unit $t+1$. In this case he has to put out a 1 if the memory device is empty, which happens by the preceding discussion with probability $a(0,t) \cdot p^t(1-p)^{(s-1)t}$. Hence the expected value for the occurence of the first 0 now is

$$E_0 = \sum_{t=1}^{\infty} (t+1) \cdot a(0,t) \cdot p^t (1-p)^{st} \tag{50}$$

We shall now use Stirling's formula

$$n! \approx \sqrt{2\pi n} \cdot (\frac{n}{e})^n \tag{51}$$

to find that the binomial coefficient $\binom{st}{t}$ in the Catalan numbers $a(0,t)$ is approximately

$$\binom{st}{t} \approx (\frac{s^s}{(s-1)^{s-1}})^t \cdot \frac{1}{\sqrt{t}} \tag{52}$$

Now let $p = \frac{1}{s'}$. Then with (52)

$$E_0 = \sum_{t=1}^{\infty} (t+1) \cdot a(0,t) \cdot (\frac{1}{s'})^t (1-\frac{1}{s'})^{st} = \sum_{t=1}^{\infty} \frac{t+1}{(s-1)t+1} \binom{st}{t} \cdot (\frac{(s'-1)^s}{(s')^{s+1}})^t \approx \sum_{t=1}^{\infty} S^t \cdot \frac{1}{\sqrt{t}} \tag{53}$$

with

$$S = \frac{s^s \cdot (s'-1)^{s-1}}{(s')^s \cdot (s-1)^{s-1}} \tag{54}$$

Now if $p < \frac{1}{s}$ then $S < 1$ and the series (50) converges and hence the expected value E_0 exists. If $p = \frac{1}{s}$ then $S = 1$ and E_0 does not exist, since in this case $\frac{1}{\sqrt{t}}$ is the decisive term in the single summands which again yields a divergent series for (50).

Remark: Our model for creating order could be extended to the case in which there are $s > 2$ bits arriving in each time unit at the organizer who then has to put out a number l $(s > l > 1)$ of bits. With the argumentation above, the number $a(0,t)$ in this case is the number of sequences of length st fulfilling the two conditions

i')

$$\mathrm{wt}(u_1, \dots, u_{st}) = l \cdot t, \tag{55}$$

ii')

$$\mathrm{wt}(u_1, \dots, u_{si}) \geq l \cdot i \text{ for } i = 1, \dots, t-1. \tag{56}$$

It is an open problem to determine the number of such sequences for $\gcd(l,s) = 1$, which would in some sense further generalize the Catalan numbers. A rough estimation would be $\binom{st}{lt}$ which is the total number of sequences of length st with weight lt. Following the lines of the proof of Theorem 2, application of Stirlings formula will then demonstrate that E_0 exists for $p < \frac{l}{s}$. In order to analyze the case $p \geq \frac{l}{s}$, however, a sharper estimation for $a(0,t)$ will be needed.

IV. Concluding remarks

We analyzed a model for creating order in sequence spaces in which in each time unit two or more bits arrive at an organizer who then has to choose one bit for output. The case, where two bits are produced by symmetric sources is quite well understood. If the organizer is equipped with a memory device that can store an infinite number of bits, then the expected value for the occurence of the first 0 in the output sequence E_0 does not exist. If no more than M bits can be stored in the memory device, then E_0 exists and can be expressed in terms of the eigenvalues of certain transition matrices.

Further, for the situation in which every incoming bit can be stored there is a critical probability $p = \frac{1}{s}$, which is a threshold for the existence of E_0, when $s \geq 2$ bits arrive at the organizer in each time unit.

Several generalizations are possible. In the s–user model ($s \geq 3$) also the case of a memory device with maximum size M may be considered. It is clear that E_0 exists in this case. If all the s sources produce only zeros in M consecutive time units, then the organizer must put out a 0 for sure. This event has a positive probability and hence a finite expected waiting time, which clearly upperbounds E_0. A further qualitative analysis, however, is much more difficult, since now the transition matrices become more complicated, e. g., a reflection argument in order to obtain the transition matrix from the power of the matrix of an elementary random walk is no longer possible.
For the infinite–memory case an open problem connected to the model in which the organizer has to put out more than one bit per time unit has already been stated at the end of Section 3.

One might also consider nonidentical sources. Even for $s = 2$ sources producing a 1 with probalities p and q, respectively, this seems to be quite hard. For the case that every incoming bit can be stored one might ask for a condition on the probabilities p and q to assure the existence of E_0.

Acknowledgment

The author wishes to thank Miklos Ruszinko for helpful discussions

References

[1] R. Ahlswede, J. P. Ye, and Z. Zhang, "Creating order in sequence spaces with simple machines", *Information and Computation*, vol. 89 (1), pp. 47–94, 1990.

[2] R. Ahlswede and Z. Zhang, "Contributions to a theory of ordering for sequence spaces", *Problems of Control and Information Theory*, vol. 18 (4), pp. 197–221, 1989.

[3] H. D. L. Hollmann and P. Vanroose, "Entropy reduction, ordering in sequence spaces, and semigroups of non–negative matrices", Preprint 95–092, SFB 343, University of Bielefeld, 1995.

[4] R. Ahlswede and A. Kaspi, "Optimal coding strategies for certain permuting channels", *IEEE Trans. Inform. Theory*, vol. IT–33, pp. 310 – 314, 1987.

[5] K. Kobayashi, "Combinatorial structure and capacity of the permuting relay channel", *IEEE Trans. Inform. Theory*, vol. IT–33, pp. 813 – 826, 1987.

[6] P. Piret, "Two results on the permuting mailbox channel", *IEEE Trans. Inform. Theory*, vol. 35, pp. 888–892, 1989.

[7] H. S. M.Coxeter, *Regular polytopes*, Dover, 1973.

[8] D. Y. Savio and E. R. Suryanarayan, "Chebyshev polynomials and regular polygons", *American Mathematical Monthly*, vol. 100, pp. 657 – 661, 1993.

[9] T. J. Rivlin, *The Chebyshev polynomials*, Wiley, 1974.

[10] M. Kac, "Random walk and the theory of Brownian motion", *American Mathematical Monthly*, vol. 54, pp. 369-391, 1947.

[11] F. Spitzer, *Principles of random walks*, Springer 1976.

[12] T. Motzkin, “Relations between hyper–surface cross–ratios, and a combinatorial formula for partitions of a polygon, for permanent preponderance, and for non–associative products”, *Bull. Am. Math. Soc.* vol. 54, pp. 352–360, 1948

[13] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics*, Addison Wesley, 1988.