Z. Wahrscheinlichkeitstheorie 3, 279—285 (1965)

Finite Abstract Random Automata

By

OCTAV ONICESCU and SILVIU GUIAȘU

§ 1. Introduction

The abstract theory of automata considered as a chapter of abstract algebra has recently known a rather important development. The principal results obtained are brought up in the synthesizing work [3]. Seen from the viewpoint of abstract algebra an automaton in its general form is an object

$$\{A, X, Y, \delta, \lambda\}$$

composed by three abstract, nonempty sets: A, X, Y and two functions $\delta : A \times X \to A$, $\lambda : A \times X \to Y$.

A is the set of states, X the set of inputsignals, Y the set of outputsignals, δ the transitionfunction and λ the outputfunction of the automaton. Knowing the order of inputsignals and the function δ, it means the knowledge of a function $\delta^* : A \times X \to A \times X$ effecting that to a state and an inputsignal of the automaton, there corresponds a certain state and a certain inputsignal of the next moment.

But automata don't always act in accordance with the well defined law. The human brain as well as the calculating automata are far from acting in a univocally determined manner. Disturbances (noises) of most differing nature alter univocally determined functioning of the automaton. At a first approximation one may leave aside the existance of such disturbances. But going deeper into the matter one cannot bar the existance these disturbances, hence the necessity for introducing probabilities in automata and exploring adequate ways as to reduce as much as possible the influence of disturbances. As disturbances are of a statistical character, the introducing of quantities specific for the probability theory is indispensable for the study of abstract automata; and first of all consider the results of the information theory.

Let be

$$\Delta = \{\delta^* \mid \delta^* : A \times X \to A \times X\}$$

and

$$\Lambda = \{\lambda \mid \lambda : A \times X \to Y\}$$

In the case of a nonrandom automaton, a certain $\delta^* \in \Delta$ and a certain $\lambda \in \Lambda$ are corresponding to the automaton. In the case of a random automaton these mappings are substituted by two random processes in which at any moment t and for any pair (a, x) correspond, with a determined probability, a mapping from Δ and Λ respectively.

The disturbances being of complex nature we practically can get them by experimentally determining the probabilities of the above mentioned mappings.

Our paper will deal but with finite abstract automata because, on one hand, they are forming the most widespread class of automata and on the other hand, as they may represent a step towards the more general automata.

In § 2 we shall define the finite abstract random automaton investigating the quantities that are determining the functioning of the automaton and in § 3 we shall show a way to the reducing of disturbances i. e. the reducing of disturbances by passing from the current functioning schema to the schema of extended functioning.

§ 2. Finite abstract random automata. Definition and functioning

Definition 1. A finite abstract random automaton is an object

$$\{A, X, Y, p^{(0)}_{A\times X}, p_{A\times X}(\cdot\mid(a,x)), p_Y(\cdot\mid(a,x))\} \tag{1}$$

composed by three nonempty sets A, X, Y and three probabilities $p^{(0)}_{A\times X}$, $p_{A\times X}(\cdot\mid(a,x))$, $p_Y(\cdot\mid(a,x))$. The set A will be denoted as the set of states of the automaton, the set X is the set of input signals, and Y the set of output signals; $p^{(0)}_{A\times X}$ is the initial probability of states and input signals; $p_{A\times X}(\cdot\mid(a,x))$ the probability of transition and $p_Y(\cdot\mid(a,x))$ the probability of output.

$p^{(0)}_{A\times X}$ is defined on the cartesian product $A\times X$; $p_{A\times X}(\cdot\mid(a,x))$ is likewise a probability on $A\times X$ conditioned by a system $(a,x)\in A\times X$. This probability must be more completely written $p_{A\times X}((a',x')\mid(a,x))$ and represents the transition probability of the automaton from its state a with the input signal x, in the state a' with the input signal x', in the next moment; $p_Y(y\mid(a,x))$ is defined on the elements of set Y, conditioned by the system (a,x) and represents the probability of the outputsignal y, if in the previous moment the automaton was in the state a with the inputsignal x.

An automaton has two distinctly different communication channels. The first one is $A\times X\to A\times X$ and the second leads from $A\times X\to Y$, the first being the transition channel, the latter the output channel. The probabilities $p_{A\times X}(\cdot\mid(a,x))$ characterize the disturbances on the transition channel and $p_Y(\cdot\mid(a,x))$ the disturbances in the output channel. If $p_{A\times X}(\cdot\mid(a,x))$ takes the values 0 or 1 only, for each $(a,x)\in A\times X$ and $(a',x')\in A\times X$, on the transitionel channel, there are no disturbances at all. Analogically for $p_Y(\cdot\mid(a,x))$.

Theorem 1. *The functioning of a finite random automaton is determined by the probabilities* $p^{(0)}_{A\times X}$, $p_{A\times X}(\cdot\mid(a,x))$, $p_Y(\cdot\mid(a,x))$.

Proof. The functioning schema of the automaton has the following structure

$$\begin{array}{l}
[A\times X, (a,x), p^{(0)}_{A\times X}] \xrightarrow{p_Y(\cdot\mid(a,x))} [Y, y, p^{(1)}_Y]\\
\quad\Big\downarrow p_{A\times X}(\cdot\mid(a,x))\\
[A\times X, (a,x), p^{(1)}_{A\times X}] \xrightarrow{p_Y(\cdot\mid(a,x))} [Y, y, p^{(2)}_Y]\\
\quad\vdots\\
\quad\downarrow\\
[A\times X, (a,x), p^{(n)}_{A\times X}] \xrightarrow{p_Y(\cdot\mid(a,x))} [Y, y, p^{(n+1)}_Y]\\
\quad\Big\downarrow p_{A\times X}(\cdot\mid(a,x))\\
[A\times X, (a,x), p^{(n+1)}_{A\times X}] \xrightarrow{p_Y(\cdot\mid(a,x))} [Y, y, p^{(n+2)}_Y]\,.\\
\quad\vdots
\end{array}$$

The transmission over the transition channel corresponds to a Markov chain, completely determined by the initial probabilities $p^{(0)}_{A\times X}$ and respectively by the transition probabilities: $p_{A\times X}(\cdot\,|\,(a, x))$.

The probability fields determined by the set of states and by the set of outputsignals are varying in time.

The recurrence relations between these probabilities are, as it may be seen in our schema

$$p^{(n+1)}_{A\times X}(a^{n+1}, x^{n+1}) = \sum_{(a^n, x^n)\in A\times X} p^{(n)}_{A\times X}(a^n, x^n)\, p_{A\times X}((a^{n+1}, x^{n+1})\,|\,(a^n, x^n)) \tag{2}$$

$$(n = 0, 1, 2, \ldots)$$

$$p^{(n+1)}_{Y}(y^{n+1}) = \sum_{(a^n, x^n)\in A\times X} p^{(n)}_{A\times X}(a^n, x^n)\, p_Y(y^n\,|\,(a^n, x^n)) \tag{3}$$

where a^n stand for a random state x^n for a random inputsignal and y^n for a random outputsignal at the moment n.

This recurrence relations yield that $p^{(0)}_{A\times X}$, $p_{A\times X}(\cdot\,|\,(a, x))$, $p_Y(\cdot\,|\,(a, x))$ determined all other probabilities of the schema.

Remark. a. Let us assume card. $(A\times X) = m$. If (a', x') is the event q and (a, x) the event r, we can write

$$p_{A\times X}((a', x')\,|\,(a, x)) = \pi_{rq}\,. \tag{4}$$

One may easily calculate the probability that, the state of the automaton should be a' and the inputsignal x', if n steps before the state of the automaton was a and its inputsignal x, with other words one may calculate

$$p^{(n)}_{A\times X}((a', x')\,|\,(a, x)) = \pi^{(n)}_{rq}\,. \tag{5}$$

Indeed, let $\lambda_1, \lambda_2, \ldots, \lambda_s$ be the s distinct roots of the characteristic equation $\Delta(\lambda) = 0$ of the matrix $T = (\pi_{rq})_{\substack{1\leq r\leq m\\ 1\leq q\leq m}}$.

Thus we shall have

$$\Delta(\lambda) = \prod_{k=1}^{s} (\lambda - \lambda_k)^{n_k}$$

where n_k is the multiplicity degree of the root λ_k. Putting

$$\pi^{(n)}_{rq} = \sum_{p=1}^{m} \pi_{rp}\,\pi^{(n-1)}_{pq} = \sum_{p=1}^{m} \pi^{(n-1)}_{rp}\,\pi_{pq}\,, \quad \pi^{(1)}_{rq} = \pi_{rq}\,; \tag{6}$$

denoting by $G_{rq}(\lambda)$ the minor of the element π_{rq} in the determinant $|T - \lambda I|$, D^p_λ being the derivation operator of order p with respect to λ and setting

$$\Psi_j = \frac{|T - \lambda I|}{(\lambda - \lambda_j)^{n_j}}$$

we have according to Perron's formula

$$\pi^{(n)}_{rq} = \sum_{j=1}^{s} \frac{1}{(n_j - 1)!}\, D^{n_j - 1}_\lambda \left[\frac{\lambda^n\, G_{rq}(\lambda)}{\Psi_j(\lambda)}\right]_{\lambda=\lambda_j}. \tag{7}$$

b. These probabilities can be restricted within the states of the automaton by a simple summation only:

$$p_A(a' \mid (a, x)) = \sum_{x' \in X} p_{A\times X}((a', x') \mid (a, x)), \tag{8}$$

i.e. the probability to find the automaton in the state a', if at the previous moment its state was a, and the inputsignal x.

Let us now presume, there are no disturbances in the automaton i.e. $p_A(a' \mid (a,x))$ and $p_Y(y \mid (a, x))$ take the values 0 or 1 only, whatever may be $a' \in A$, $y \in Y$ and $(a, x) \in A \times X$. More precisely in this case only one $a' \in A$ corresponds to the pair (a, x), as well as a single output y. Consequently there are in this case two mappings: A transition mapping: $\delta: A \times X \to A$ and an output mapping $\lambda: A \times X \to Y$. Thus in this case the abstract random automaton reduces to a MEALY automaton:

$$\{A, X, Y, \delta, \lambda\}$$

its functioning schema being

$$\begin{array}{l} A \times X \xrightarrow{\lambda} Y \\ \swarrow \delta \\ A \times X \xrightarrow{\lambda} Y \\ \swarrow \delta \end{array}$$

to which we must add the initial state a_0 for which $p_A^{(0)}(a_0) = \sum_{x \in X} p_{A\times X}^{(0)}(a_0, x) = 1$.

In this case the functioning of the automaton is univocally determined by the mappings δ and λ as it is defined by G. A. MEALY [4] and N. M. GLUSHKOV [3].

§ 3. The entropies and the extension of the automaton

The two entropies. Let us consider the finite abstract random automaton

$$\{A, X, Y, p_{A\times X}^{(0)}, p_{A\times X}(\cdot \mid (a, x)), p_Y(\cdot \mid (a, x))\}.$$

The degree of indetermination — the inputsignals — due to exterior causes — and the outputsignals — due to the disturbances in the channel — are undergoing, is given by the respective entropies.

The transition entropy at the moement m is

$$H^{(m)}(A \times X) = - \sum_{(a^m, x^m) \in A \times X} p_{A\times X}^{(m)}(a^m, x^m) \log p_{A\times X}^{(m)}(a^m, x^m) = T^{(m)} H^{(0)}(A \times X) \tag{9}$$

where $T^{(m)}$ is m-iterate of the transformation T of matrix

$$(p_{A\times X}((a', x') \mid (a, x)))_{\substack{(a, x) \in A \times X \\ (a', x') \in A \times X}}$$

while the output entropy at the same moment is

$$H^{(m)}(Y) = - \sum_{y^m \in Y} p_Y^{(m)}(y^m) \log p_Y^{(m)}(y^m). \tag{10}$$

The recurrence relation (2) and (3) show that these entropies $H^{(m)}(A \times X)$, $H^{(m)}(Y)$ are completely determined at each moment m by the probabilities $p_{A\times X}^{(0)}, p_{A\times X}(\cdot \mid (a, x)) p_Y(\cdot \mid (a, x))$.

The extension. The entropies above defined give us the clue to the following main problem: Given a finite abstract random automaton (1), what could be done in order to reduce the influence of the disturbances to the utmost, i.e. the entropies $H^{(m)}(A\times X)$, $H^{(m)}(Y)$ should be at the moment m reduced as much as possible. We are going to show now, that the reducing of the influence of the disturbances is quite possible, by substituting to the automaton a certain extension of its.

Definition 2. Given a finite abstract random automaton

$$\{A, X, Y, p^{(0)}_{A\times X}, p_{A\times X}(\cdot\,|\,(a, x)), p_Y(\cdot\,|\,(a, x))\}$$

we shall denote the lengthextension n of this automaton — the random automaton —

$$\left\{\prod_{k=1}^{n} A_k, \prod_{k=1}^{n} X_k, \prod_{k=1}^{n} Y_k, P^{(0)}_{\prod\limits_{k=1}^{n}(A_k\times X_k)}, P_{\prod\limits_{k=1}^{n}(A_k\times X_k)}(\cdot\,|\,(a_1, x_1, \ldots, a_n, x_n)), \right.$$
$$\left. P_{\prod\limits_{k=1}^{n} Y_k}(\cdot\,|\,(a_1, x_1, \ldots, a_n, x_n))\right\}$$

where $A_k = A$, $X_k = X$, $Y_k = Y (k = 1, 2, \ldots, n)$ and the conditional probabilities are given by the equalities

$$\begin{aligned} &P_{\prod\limits_{k=1}^{n}(A_k\times X_k)}((a_1', x_1', \ldots, a_n', x_n')\,|\,(a_1, x_1, \ldots, a_n, x_n)) \\ &= p_{A\times X}((a_1', x_1')\,|\,(a_1, x_1)) \cdots p_{A\times X}((a_n', x_n')\,|\,(a_n, x_n)) \end{aligned} \tag{11}$$

$$P_{\prod\limits_{k=1}^{n} Y_k}((y_1, \ldots, y_n)\,|\,(a_1, x_1, \ldots, a_n, x_n)) = p_Y(y_1\,|\,(a_1, x_1)) \cdots p_Y(y_n\,|\,(a_n, x_n)) \tag{12}$$

because we are presuming that we neither have on the transition-channel nor on the outputchannel any memory; for the initial probability we have

$$P^{(0)}_{\prod\limits_{k=1}^{n}(A_k\times X_k)}(a_1, x_1, \ldots, a_n, x_n) \neq p^{(0)}_{A\times X}(a_1, x_1) \ldots p^{(n-1)}_{A\times X}(a_n, x_n).$$

In other words, in the lengthextension n of an automaton a single state is substituted by a sequence of states having the length n, an inputsignal by a sequence of inputsignals of length n and an outputsignal by a sequence of n outputsignals. The elements of the sets $\prod\limits_{k=1}^{n} A_k$, $\prod\limits_{k=1}^{n} X_k$, $\prod\limits_{k=1}^{n} Y_k$ are named state-words, input-words and output-words respectively, of length n.

The functional element of the schema on the extension of length n of the automaton is the following:

$$\begin{tikzcd}
\left[\prod\limits_{k=1}^{n}(A_k\times X_k), (a_1, x_1, \ldots, a_n, x_n), P^{(m)}_{\prod\limits_{k=1}^{n}(A_k\times X_k)}\right] \arrow[r, "P_{\prod\limits_{k=1}^{n} Y_k}(\cdot\,|\,(a_1, x_1, \ldots, a_n, x_n))"] \arrow[d, "P_{\prod\limits_{k=1}^{n}(A_k\times X_k)}(\cdot\,|\,(a_1, x_1, \ldots, a_n, x_n))"'] & \left[\prod\limits_{k=1}^{n} Y_k, (y_1, \ldots, y_n), P^{(m+1)}_{\prod\limits_{k=1}^{n} Y_k}\right] \\
\left[\prod\limits_{k=1}^{n}(A_k\times X_k), (a_1, x_1, \ldots, a_n, x_n), P^{(m+1)}_{\prod\limits_{k=1}^{n}(A_k\times X_k)}\right]. &
\end{tikzcd}$$

Theorem 2. *Let be*

$$\{A, X, Y, p^{(0)}_{A\times X}, p_{A\times X}(\cdot \mid (a, x)), p_Y(\cdot \mid (a, x))\}$$

a random automaton.

Whatever $\varepsilon > 0$ may be, there is a natural number $n_0(\varepsilon)$ such that in any extension of length $n \geqq n_0(\varepsilon)$ of the random automaton, the indetermination may differ by less than ε from the admissible minimum of indetermination.

Proof. For simplyfying let us denote

$$\bar{H}^{(m)}_{(n)} = H^{(m)}\left(\prod_{k=1}^{n} (A_k \times X_k)\right) \tag{13}$$

$$\bar{\bar{H}}^{(m)}_{(n)} = H^{(m)}\left(\prod_{k=1}^{n} Y_k\right) \tag{14}$$

$$C^*_m = \inf_n \frac{\bar{H}^{(m)}_{(n)}}{n}; \quad C^{**}_m = \inf_n \frac{\bar{\bar{H}}^{(m)}_{(n)}}{n}. \tag{15}$$

The value C^*_m stand for the smallest mean indetermination at the moment m in the first channel, with respect to all extensions of the automaton and C^{**}_m is the smallest mean indetermination at the moment m in the second channel, likewise with respect to all extensions.

We firstly show that

$$C^*_m = \lim_{n\to\infty} \frac{\bar{H}^{(m)}_{(n)}}{n}. \tag{16}$$

Indeed C^*_m being a lower boundary, there exists for any $\varepsilon > 0$ a natural number s, such that

$$\frac{\bar{H}^{(m)}_{(s)}}{s} \leqq C^*_m + \varepsilon.$$

The properties of the entropy yield

$$n \geqq s; (r_n - 1)s \leqq n < r_n s \Rightarrow \bar{H}^{(m)}_{(n)} \leqq \bar{H}^{(m)}_{(r_n s)}$$

hence

$$\bar{H}^{(m)}_{(n)} \leqq r_n \bar{H}^{(m)}_{(s)}$$

implying

$$\frac{\bar{H}^{(m)}_{(n)}}{n} \leqq \frac{r_n \bar{H}^{(m)}_{(s)}}{n} \leqq \frac{r_n}{r_n - 1} \frac{\bar{H}^{(m)}_{(s)}}{s} \leqq \frac{r_n}{r_n - 1} (C^*_m + \varepsilon).$$

Making n to tend to infinity, r_n is likewise tending to infinity and so we have

$$\limsup_n \frac{\bar{H}^{(m)}_{(n)}}{n} \leqq C^*_m + \varepsilon$$

for any ε. Thus

$$\limsup_n \frac{\bar{H}^{(m)}_{(n)}}{n} \leqq C^*_m. \tag{17}$$

On the other hand, in accordance with the definition of C^*_m

$$\frac{\bar{H}^{(m)}_{(n)}}{n} \geqq C^*_m \tag{18}$$

(17) and (18) are yielding

$$\lim_{n\to\infty} \frac{\bar{H}_{(n)}^{(m)}}{n} = C_m^* .$$

Thus, for any $\varepsilon > 0$ there exist $n_m^*(\varepsilon)$ such that for $n \geqq n_m^*(\varepsilon)$

$$0 \leqq \frac{\bar{H}_{(n)}^{(m)}}{n} - C_m^* < \varepsilon . \tag{19}$$

In the same manner one may demonstrate that:

$$C_m^{**} = \lim_{n\to\infty} \frac{\bar{\bar{H}}_{(n)}^{(m)}}{n} \tag{20}$$

i.e. for any $\varepsilon > 0$ there exist $n_m^{**}(\varepsilon)$ such that for $n \geqq n_m^{**}(\varepsilon)$

$$0 \leqq \frac{\bar{\bar{H}}_{(n)}^{(m)}}{n} - C_m^* < \varepsilon .$$

The random automaton has a finite lifetime $[0, M]$.

We are writing

$$n_0(\varepsilon) = \max_{1 \leqq m \leqq M} (n_m^*(\varepsilon), n_m^{**}(\varepsilon)) .$$

Then we have, in any extension of lenth $n \geqq n_0(\varepsilon)$

$$0 \leqq \frac{H^{(m)}\left(\prod\limits_{k=1}^{n} (A_k \times X_k)\right)}{n} - C_m^* < \varepsilon ; 0 \leqq \frac{H^{(m)}\left(\prod\limits_{k=1}^{n} Y_k\right)}{n} - C_m^{**} < \varepsilon$$

whatever $m (1 \leqq m \leqq M)$ may be.

Remark. A random automaton may fulfill its programme when the degree of indetermination does not exceed a certain value H. Should this automaton respond "on the fly" to the disturbances appearing in the automaton, the indeterminations need to be less than H. At any moment an auxiliary device should record the values of the entropies $H^{(m)}(A \times X)$, $H^{(m)}(Y)$. Should one of these values go beyond H, one had to extend the length of the random automaton conveniently.

References

[1] Chinčin, A. J.: Der Begriff der Entropie in der Theorie der Wahrscheinlichkeit. Uspechi mat. Nauk. 8, 3—20 (1953).

[2] Feinstein, A.: Foundations of information theory. New York-Toronto-London 1958.

[3] Gluškov, V. M.: Abstrakte Theorie der Automaten. Uspechi mat. Nauk. **16**, 3—62 (1961).

[4] Mealy, G. H.: A method for synthesizing sequential circuits. Bell System techn. J. **34**, 1045—1079 (1955).

[5] Moisil, Gr. C.: Teoria algebrica a mecanismelor automate. Bucuresti 1959.

[6] Neumann, J. von: The general and logical theory of automata. Massachusetts, Institute of Technology, 1948.

[7] Shannon, C. E.: A mathematical theory of communication. Bell System techn. J. **27**, 379—423, 623—656 (1948).

Institutul de Matematica
al Academiei R. P. R.,
Str. Eminescu 47,
Bucuresti, Rumänien

(Eingegangen am 1. August 1963)