# CYBERNETICS AND CONTROL THEORY

1970

SovietRxiv

**Abstract**

**Full Text**

# CYBERNETICS AND CONTROL THEORY

**Ya. M. Barzdin'**

# ON DECODING AUTOMATA IN THE ABSENCE OF AN UPPER BOUND ON THE NUMBER OF STATES

*(Presented by Academician V. M. Glushkov, 22 V 1969)*

By **automata** we shall mean finite initial Mealy automata with numbered states $q_1, q_2, q_3, \ldots$. Unless otherwise specified, the state $q_1$ will be considered the initial state. In what follows we shall assume that all automata have one and the same input alphabet $X = \{x_1, \ldots, x_a\}$, where $a = \text{const} \geq 2$, and one and the same output alphabet $Y = \{y_1, \ldots, y_b\}$, where $b = \text{const} \geq 2$. The operator realized by an automaton $\mathfrak{M}$ in the initial state $q_i$ will be denoted by $T(\mathfrak{M}, q_i)$.

Let $\mathfrak{M}$ be an automaton about whose internal structure (state diagram) nothing is known, including no upper bound on the number of states (or such a bound is known, but in the present formulation one is not allowed to use it). Such an automaton will be called a **black box** (b.b.). It is assumed that experiments can be carried out with a b.b. As is well known, there is no experiment by means of which an arbitrary b.b. could be decoded. However, as shown in (1), there exists a multiple experiment (of course, a branching one) which makes it possible to decode the "majority" of b.b.' s. The purpose of the present note is to obtain an analogous result in the case of **simple** experiments and to estimate the length of the corresponding simple experiments.

The precise formulation of the decoding problem is connected with a refinement of the concept of an algorithm over a b.b. (a decoding algorithm). Substantively, by an **algorithm** $\Omega$ over a b.b. we shall mean an effective prescription describing a simple branching experiment and indicating how, from the result of this experiment, to construct the corresponding automaton (which presumably works in the same way as the given b.b.). More precisely, the algorithm $\Omega$, applied to a b.b. $\mathfrak{M}$, works in steps. At each step the algorithm $\Omega$ performs with the b.b. $\mathfrak{M}$, which is in the state in which it remained after the preceding step, a certain simple homogeneous experiment, and, depending on the result of this experiment as well as on the result of the experiments performed at the preceding steps, does one of two things: a) either it gives the result $\Omega(\mathfrak{M})$—some automaton (for example, in the form of a diagram); b) or it constructs an input word which determines the simple experiment to be carried out at the next step. The state of the automaton $\mathfrak{M}$ into which it passes as a result of applying the

algorithm $\Omega$ will be denoted by $q_\Omega$. We shall say that the algorithm $\Omega$ **decodes** the b.b. $\mathfrak{M}$ if
$T(\Omega(\mathfrak{M}), q_1) = T(\mathfrak{M}, q_\Omega)$, i.e., if the algorithm $\Omega$ uniquely determines the further behavior of the b.b. $\mathfrak{M}$. The length of the simple experiment performed by the algorithm $\Omega$ applied to $\mathfrak{M}$ will be denoted by $\Omega'(\mathfrak{M})$. Put $\Omega^*(k) = \max \Omega'(\mathfrak{M})$, where the maximum is taken over all automata $\mathfrak{M}$ having $k$ states.

Let $\mathcal{L}$ be the class of all pairwise distinct* automata (b.b.' s), $\{\mathcal{L}_\lambda\}$ some partition of the class $\mathcal{L}$ into finite subclasses, and $\mathcal{L}_\lambda^\Omega -$

* Two automata $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are considered identical if and only if any two states with identical numbers, both in the diagram of $\mathfrak{M}_1$ and in the diagram of $\mathfrak{M}_2$, are connected by edges that are oriented in the same way and have identical labels.

the set of those automata from $\mathcal{L}_\lambda$ that the algorithm $\Omega$ deciphers. Consider the ratio $P_\lambda = |\mathcal{L}_\lambda^\Omega|/|\mathcal{L}_\lambda|$. We shall say that the algorithm $\Omega$ **deciphers automata** (a.e.) **with frequency** $1-\varepsilon$ **under the given partition** $\{\mathcal{L}_\lambda\}$, if $P_\lambda \geq 1-\varepsilon$ for all $\lambda$. In the present note we shall consider the partition under which two automata belong to one and the same class if they differ only in their output functions (we shall call such a partition the partition by graph). This partition is sufficiently fine. Therefore, if we show that a.e. one can decipher with frequency $1-\varepsilon$ under the given partition, then all the more this will be true also for coarser partitions, for example such as the partition by the number of states, when any two automata belong to one subclass if they have the same number of states. Henceforth, when saying that an algorithm $\Omega$ deciphers a.e. uniformly with frequency $1-\varepsilon$, we shall mean precisely the partition by graph. By an automaton graph we shall mean the graph obtained from the diagram of an automaton if the output labels are erased (and the input labels are left). Obviously, from an automaton graph $G$ with $k$ vertices, by placing output labels in all possible ways, one can obtain $b^{ak}$ pairwise distinct automata. We denote the class of such automata by $G$. Thus, if an algorithm $\Omega$ deciphers a.e. uniformly with frequency $1-\varepsilon$, this means that for every automaton graph $G$ one has $|G^\Omega|/|G| \geq 1-\varepsilon$. The main result of the present note is

**Theorem 1*.** *For every $\varepsilon > 0$ there exists an algorithm $\Omega$ which deciphers a.e. uniformly with frequency $1-\varepsilon$ and has $\Omega^*(k) \leq C_\varepsilon k^C$, where $C$ is a constant independent of $k$ and $\varepsilon$, and $C_\varepsilon$ is a constant independent of $k$, but dependent on $\varepsilon$.*

We shall say that an input word $x$ **sufficiently distinguishes** automata $\mathfrak{M}_1$ and $\mathfrak{M}_2$ if the following is true: either $\mathfrak{M}_1$ and $\mathfrak{M}_2$, when the word $x$ is applied, produce different output words, or $\mathfrak{M}_1$ and $\mathfrak{M}_2$, after the application of the words $x$, implement identical operators (i.e. $T(\mathfrak{M}_1, q_{1x}) = T(\mathfrak{M}_2, q_{1x})$). By the **total degree of distinguishability** of automata $\mathfrak{M}_1$ and $\mathfrak{M}_2$ we shall understand the distinguishability of the automaton $\mathfrak{M}_1 + \mathfrak{M}_2$, which is obtained from $\mathfrak{M}_1$ and $\mathfrak{M}_2$ if they are regarded as one automaton.

**Lemma 1**.** *For every set $U$ of automata and every natural number $s$ there exists an input word of length at most $[4a^{s+1}\ln|U|]$ that sufficiently distinguishes any two automata from $U$ having total degree of distinguishability at most $s$.*

Let us prove the lemma. Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be automata having total degree of distinguishability at most $s$. First we estimate from below the number of input words of length $2s$ that sufficiently distinguish $\mathfrak{M}_1$ and $\mathfrak{M}_2$. Obviously, for any input word $v$ of length $l(v)$ there exists an input word of length $l(v)+s$, beginning with $v$ and sufficiently distinguishing $\mathfrak{M}_1$ and $\mathfrak{M}_2$; denote it by $(v)\langle s\rangle$ (if there are several such, then one of them). Note that if $l(v)\leq s$, then the number of possible continuations of the word $(v)\langle s\rangle$ to length $2s$ is equal to $a^{s-l(v)}$, and all these continuations will sufficiently distinguish $\mathfrak{M}_1$ and $\mathfrak{M}_2$. We now describe one procedure for selecting words sufficiently distinguishing $\mathfrak{M}_1$ and $\mathfrak{M}_2$.

* An analogous assertion for a coarser partition (by the number of states) was proved by the author of the present note jointly with M. P. Vasilevskii.

** Theorems 1, 2, and 3 of paper ([2]) follow directly from this lemma. For this it is enough to take into account only the following: a) if $U_k$ is the class of all pairwise nonidentical automata with $k$ states, then $|U_k|=(b^k)^{ak}$; b) the total degree of distinguishability of two automata from $U_k$ is at most $2k-1$, i.e. not greater than the degree of restoration (for Theorems 1 and 2 of ([2])); c) the total degree of distinguishability of two automata from $U_k$ that differ only by the choice of the initial state is at most $k-1$, i.e. not greater than the ordinary degree of distinguishability (for Theorem 3 of ([2])). Also from this lemma and from the estimate of the degree of restoration for almost all automata ([3]) there follows the following fact: almost all automata with $k$ states can be deciphered (in the sense mentioned above) by a simple unbranched experiment of length $k^C$, where $C$ is a constant.

**Step 1.** Consider input words $x=x_i$ of length 1; the number of such words is $a$. Corresponding to each of them, we select the word $(x_i)\langle s\rangle$. Obviously, the number of all possible continuations of all selected words of the form $(x_i)\langle s\rangle$ to length $2s$ is $aa^{s-1}$.

**Step** $k$ $(k\leq s)$. Consider input words $x=x_{i_1}\ldots x_{i_k}$ of length $k$, distinct from the initial segments of length $k$ of the previously selected words; the number of such words is $a^k-a^{k-1}$, since the number of previously selected words is $a^{k-1}$ and they all have different initial segments of length $k$ (even of length $k-1$). Corresponding to each of them, we select the word $(x_{i_1}\ldots x_{i_k})\langle s\rangle$ (thus, the number of selected words of the form $(x_{i_1}\ldots x_{i_k})\langle s\rangle$ will be equal to $a^k-a^{k-1}$). Obviously, the number of all possible continuations of all selected words of the form $(x_{i_1}\ldots x_{i_k})\langle s\rangle$ to length $2s$ is $(a^k-a^{k-1})a^{s-k}$.

As a result we obtain that the total number of input words of length $2s$ that are continuations of words selected during the first $s$ steps is equal to

$$aa^{s-1} + ... + (a^k - a^{k-1})a^{s-k} + ... + (a^s - a^{s-1})a^0 \geq sa^{s-1}$$

(recall that $a \geq 2$). According to what was said above, all these words will distinguish $\mathfrak{M}_1$ and $\mathfrak{M}_2$ residually. Thus, the total number of input words of length $2s$ that do not residually distinguish $\mathfrak{M}_1$ and $\mathfrak{M}_2$ does not exceed $a^{2s} - sa^{s-1}$. Obviously, this result remains valid for any other choice of the initial states of the automata $\mathfrak{M}_1$ and $\mathfrak{M}_2$. Therefore, by induction on $p$ we obtain that the number of input words of length $p \cdot 2s$ that do not residually distinguish $\mathfrak{M}_1$ and $\mathfrak{M}_2$ does not exceed $(a^{2s} - sa^{s-1})^p$.

Now let $U$ be an arbitrary set of automata. Then the number of input words of length $2ps$ that do not residually distinguish at least two automata from $U$ having total degree of distinguishability not greater than $s$ does not exceed

$$C^2_{|U|}(a^{2s} - sa^{s-1})^p < \frac{1}{2}|U|^2(a^{2s} - sa^{s-1})^p.$$

On the other hand, the total number of input words of length $2ps$ is $a^{2ps}$. Therefore, if for some $p_0$

$$\frac{1}{2}|U|^2(a^{2s} - sa^{s-1})^{p_0} \leq a^{2p_0 s},$$

then among the input words of length $2p_0 s$ there will necessarily be a word that residually distinguishes any two automata from $U$ having total degree of distinguishability not greater than $s$. Expressing $p_0$ from the last inequality and taking into account that $|\ln(1 - sa^{-(s+1)})| > sa^{-(s+1)}$, we obtain that as $p_0$ one may take, for example, $[2s^{-1}a^{s+1} \ln |U|]$. Hence Lemma 1 follows.

Let $D_G(x)$ be the set of vertices of the graph $G$ (of the automaton $G$) that are reachable from the vertex $q_1$ by the word $x$, and let $A_G(x)$ be the set of vertices of the graph $G$ (of the automaton $G$) that are reachable from the vertex $q_1$ by the word $x$ or from the vertex $q_1 x$ by arbitrary words. Using the same considerations as above, one can prove the following lemma.

**Lemma 2.** *For every natural number $k$ there exists an input word $b(k)$ of length not greater than $[2ka^{k+1} \ln 2k]$, possessing the following property: whatever automaton graph $G$ we take, if $|A_G(b(k))| \geq k$, then $|D_G(b(k))| \geq k$.*

Let $q_i, q_j$ be vertices of the graph $G$, and let $r$ be a natural number. Denote by $\widetilde{G}(q_i, q_j, r)$ the set of all those automata from $\widetilde{G}$ for which the states $q_i$ and $q_j$ are indistinguishable by input words of length $r$, but are distinguishable by input words of greater length.

**Lemma 3*.** *For any automaton graph $G$, any of its vertices $q_i, q_j$, and any natural number $r$, the inequality*

$$|\widetilde{G}(q_i, q_j, r)|/|\widetilde{G}| \leq b^{-r/2}$$

*holds.*

*

From this lemma, in particular, the following fact follows. Say that uniformly almost all automata have degree of distinguishability not exceeding $\varphi(k)$ if

$$\min_{|G|=k} \frac{|\widetilde{G}_{\varphi(k)}|}{|\widetilde{G}|} \to 1 \quad \text{as } k \to \infty,$$

where $|G|$ is the number of vertices of the graph $G$, and $\widetilde{G}_{\varphi(k)}$ is the set of all those automata from $\widetilde{G}$ that have degree of distinguishability not exceeding $\varphi(k)$. Then the following is true: uniformly almost all automata have degree of distinguishability not exceeding $C \log k$. (An estimate of the degree of distinguishability simply for almost all automata is given in (3).)

Now we shall give, in general terms, the idea of the proof of Theorem 1. The corresponding simple experiment is constructed step by step. At each step, proceeding from some number $s$, which substantively denotes the next hypothesis about the number of states of the black box $\mathfrak{M}$, and using Lemmas 1, 2, and 3, an experiment $\zeta_s$ is constructed such that for the "majority" of automata $\mathfrak{M}$ the following holds: a) if $\mathfrak{M}$ has no more than $s$ states (more precisely, $A_{\mathfrak{M}}(\zeta_s) \leq s$), then $\zeta_s$ deciphers $\mathfrak{M}$; b) if $\mathfrak{M}$ has more than $s$ states (more precisely, $A_{\mathfrak{M}}(\zeta_s) > s$), then $\zeta_s$ reaches sufficiently many states of the automaton $\mathfrak{M}$. In the latter case a new hypothesis $s' > s$ is constructed, and the procedure for constructing the experiment continues.

In conclusion, let us consider one more problem. By a **strategy** we shall understand a function of the form $\Phi(x, y, x_i)$, where $x$ is a word in the input alphabet $X$; $y$ is a word in the output alphabet $Y$; $x_i$ is a letter of the alphabet $X$; the value of the function $\Phi(x, y, x_i)$ is a letter of the alphabet $Y$. Substantively, we shall interpret a strategy as a rule for "predicting" the letter that the automaton will output when the letter $x_i$ is supplied, if it is known that before this the input word $x$ was supplied and the automaton transformed it into the output word $y$. Let $\mathfrak{M}$ be an automaton, $\omega = (x(1), x(2), \ldots, x(i), \ldots)$ an infinite input sequence, and $(y(1), y(2), \ldots, y(i), \ldots)$ the corresponding output sequence (which $\mathfrak{M}$ produces when $\omega$ is supplied). We shall say that the strategy $\Phi$ on $\mathfrak{M}$ and $\omega$ makes an error at the $i$-th moment if $\Phi(x(1) \ldots x(i), y(1) \ldots y(i), x(i+1)) \neq y(i+1)$. The number of moments at which the strategy $\Phi$ makes an error will be denoted by $\Phi'_\omega(\mathfrak{M})$. Put $\Phi^*_\omega(k) = \max \Phi'_\omega(\mathfrak{M})$, where the maximum is taken over all automata $\mathfrak{M}$ having $k$ states.

**Theorem 2.** *There exists an effective strategy $\Phi$ such that for any infinite input sequence $\omega$ one has $\Phi^*_\omega(k) \leq Ck \log_2 k$, where $C$ is a constant independent of $k$ and $\omega$.*

This theorem remains valid if, instead of finite automata, one considers Turing machines with input and output channels using, respectively, the alphabets $X$ and $Y$. It is assumed here that the machines begin to operate with a blank tape and that the external alphabet is the same for all machines.

In the case of periodic sequences $\omega$ (with period $p$), as is easy to see, $\Phi^*_\omega(k) \leq C_p k$. However, in the general case, as the following theorem shows, the estimate of Theorem 2 cannot be improved in order of magnitude.

**Theorem 3.** *There exist an infinite input sequence $\omega$ and a constant $C_0$ such that for any strategy $\Phi$ one has $\Phi^*_\omega(k) \geq C_0 k \log_2 k$.*

Computing Center
of P. Stučka Latvian State University
Riga

Received
19 V 1969

## CITED LITERATURE

1. Ya. M. Barzdin, *Problems of Cybernetics*, issue 21, Moscow, 1969.
2. A. A. Muchnik, *Problems of Cybernetics*, issue 20, Moscow, 1968.
3. A. D. Korshunov, *Discrete Analysis*, issue 10, Novosibirsk, 1967, p. 39.

*Note: Figure translations are in progress. See original paper for figures.*