Reports of the Academy of Sciences of the USSR

1962

Abstract

Full Text

Reports of the Academy of Sciences of the USSR
1962. Volume 147, No. 6

CYBERNETICS AND CONTROL THEORY

V. I. LEVENSHTEIN

ON THE INVERSION OF FINITE AUTOMATA

(Presented by Academician M. V. Keldysh, 4 VII 1962)

Recently, in papers ([6–8]), the common character of certain problems in coding theory and in the theory of automata has been noted. Thus, questions concerning the mutual unambiguity of coding and the possibility of constructing a decoding device with finite memory have essentially been reduced to questions concerning the mutual unambiguity of finite automata and the possibility of their inversion. True, in doing so it proved necessary to broaden somewhat the concept of an automaton, abandoning the requirement of synchronism, i.e. equality of the lengths of input words and of the corresponding output words. In a paper by Yu. V. Glebskii ([6]), an effective criterion was found for recognizing the mutual unambiguity of a completely defined finite automaton on a finitely enumerable set. The present paper is devoted to the description of effectively verifiable necessary and sufficient conditions for the inversion of (asynchronous) partial finite automata, as well as to methods for constructing inverse automata. In addition, algorithms are constructed in the paper for recognizing certain other properties of partial finite automata.

1. Let $\mathfrak{A} = \{A, B, S, s_{i_0}, f, \varphi\}$ be an arbitrary partial automaton, where $A = \{a_1, \dots, a_m\}$ is the input alphabet; $B = \{b_1, \dots, b_r\}$ is the output alphabet; $S = \{s_1, \dots, s_N\}$ is the set of states; s_{i_0} is the initial state; f and φ are, respectively, the transition function and the output function of the automaton, defined on some subset $\mathfrak{N} \subseteq S \times A$, and

$$f(s_i, a_j) = s_j^i, \qquad \varphi(s_i, a_j) = v_j^i,$$

where $s_j^i \in S$, while v_j^i is some (possibly empty) word in the alphabet B. The automaton $\mathfrak{A}$ is called **synchronous** if every word v_j^i is a letter of the alphabet B. The automaton $\mathfrak{A}$ induces a function $F_{\mathfrak{A}}$, defined on every word

$$\alpha = a_{j_1} a_{j_2} \dots a_{j_l},$$

for which

$$(s_{i_{p-1}}, a_{j_p}) \in \mathfrak{N}, \qquad p = 1, 2, \ldots, l,$$

where

$$s_{i_p} = f(s_{i_{p-1}}, a_{j_p}),$$

and taking on it the value

$$F_{\mathfrak{A}}(\alpha) = v_{j_1}^{i_0} v_{j_2}^{i_1} \ldots v_{j_l}^{i_{l-1}}.$$

In addition, the automaton $\mathfrak{A}$ induces a function $F_{\mathfrak{A}}^{\infty}$, defined on every infinite sequence

$$\bar{\alpha} = a_{j_1} a_{j_2} \ldots,$$

for which

$$(s_{i_{p-1}}, a_{j_p}) \in \mathfrak{N}, \qquad p = 1, 2, \ldots,$$

where

$$s_{i_p} = f(s_{i_{p-1}}, a_{j_p}),$$

and taking on it the value

$$F_{\mathfrak{A}}^{\infty}(\bar{\alpha}) = v_{j_1}^{i_0} v_{j_2}^{i_1} \ldots.$$

We shall denote the domains of definition of the functions $F_{\mathfrak{A}}$ and $F_{\mathfrak{A}}^{\infty}$ by $I_{\mathfrak{A}}$ and $I_{\mathfrak{A}}^{\infty}$, respectively. The state into which the automaton passes from the state s_i under the action of an input word α will be denoted by $f(s_i, \alpha)$. Two automata $\mathfrak{A}$ and $\widetilde{\mathfrak{A}}$ will be considered equivalent if, for every word α in the alphabet A, one has*

$$F_{\mathfrak{A}}(\alpha) = F_{\widetilde{\mathfrak{A}}}(\alpha).$$

Since we shall be interested in automata only up to equivalence, it may be assumed without loss of generality that all states of the automaton $\mathfrak{A}$ are distinguishable and that for any state $s_i \in S$ there exists a word $\alpha \in I_{\mathfrak{A}}$ such that

$$s_i = f(s_{i_0}, \alpha).$$

Fig. 1

Figure 1: Fig. 1

We shall call the automaton $\mathfrak{A}$ **mutually unambiguous** if, for any two distinct words α_1 and α_2 from $I_{\mathfrak{A}}$, one has

$$F_{\mathfrak{A}}(\alpha_1) \neq F_{\mathfrak{A}}(\alpha_2).$$

We shall call the automaton $\mathfrak{A}$ **mutually unambiguous in the weak**

---

* Here, as usual, $F_{\mathfrak{A}}(\alpha) = F_{\widetilde{\mathfrak{A}}}(\alpha)$ means that either both functions are undefined on the word α, or they are defined on this word and take identical values.

sense, or an automaton **without loss of information***, if for any two distinct words α_1 and α_2 from $I_{\mathfrak{A}}$ such that $F_{\mathfrak{A}}(\alpha_1) = F_{\mathfrak{A}}(\alpha_2)$, the states $f(s_{i_0}, \alpha_1)$ and $f(s_{i_0}, \alpha_2)$ are distinct. We shall call the automaton $\mathfrak{A}$ **one-to-one at infinity** if for any two distinct sequences $\overline{\alpha}_1$ and $\overline{\alpha}_2$ from $I_{\mathfrak{A}}^{\infty}$ one has $F_{\mathfrak{A}}^{\infty}(\overline{\alpha}_1) \neq F_{\mathfrak{A}}^{\infty}(\overline{\alpha}_2)$.

For simplicity of the subsequent formulations we impose the following restriction on the domain of definition of the automaton $\mathfrak{A}$: every word from $I_{\mathfrak{A}}$ is the beginning of at least two distinct sequences from $I_{\mathfrak{A}}^{\infty}$. Under this restriction the following is true.

Lemma 1. *If the automaton $\mathfrak{A}$ is one-to-one at infinity, then it is one-to-one in the weak sense.*

It is obvious that a one-to-one automaton is one-to-one in the weak sense. In Fig. 1 a diagram is given of an automaton that is one-to-one in the weak sense, but is neither one-to-one nor one-to-one at infinity.

Fig. 1

If for the automaton $\mathfrak{A}$ there exists an automaton $\mathfrak{B}$ such that $F_{\mathfrak{B}}^{\infty} = (F_{\mathfrak{A}}^{\infty})^{-1}$, then we shall call the automaton $\mathfrak{A}$ **invertible**, and the automaton $\mathfrak{B}$ will be called **inverse** to $\mathfrak{A}$.

It is easy to verify that for a synchronous automaton $\mathfrak{A}$ there exists a synchronous automaton inverse to $\mathfrak{A}$ if and only if the automaton $\mathfrak{A}$ is one-to-one (cf. ([3,4])). In this case the diagram of the inverse automaton is obtained from the diagram of the automaton $\mathfrak{A}$ by interchanging the input and output letters assigned to the edges. We note that in the general case one-to-one-ness of an automaton is not necessary even for inversion of a synchronous automaton (see Figs. 2 and 3).

The difficulties arising in the general case stem mainly from the following circumstances: 1) the class of functions $F_{\mathfrak{A}}^{\infty}$ induced by automata is substantially wider than the class of functions $F_{\mathfrak{A}}^{\infty}$ induced by synchronous automata; 2) for

every automaton $\mathfrak{A}$ there exists a countable number of pairwise nonequivalent automata inducing $F_{\mathfrak{A}}^{\infty}$ (in contrast to the automaton, unique up to equivalence, in the synchronous case).

2. To construct effective criteria that make it possible to recognize the indicated properties of an automaton $\mathfrak{A}$, we introduce the sets $R_n(\mathfrak{A})$ $(n = 1, 2, ...)$, analogous to the classes of Sardinas and Patterson ([1]). The sets $R_n(\mathfrak{A})$ are defined inductively; their elements are ordered triples of the form (β, i, h), where β is the end of some word v_j^k, and i and h are state numbers. The set $R_1(\mathfrak{A})$ is defined as the totality of all elements (β, i, h) for which there exist words $v_{j_1}^{i_1}$ and $v_{d_1}^{i_1}$ $(j_1 \neq d_1)$ such that $s_i = f(s_{i_1}, a_{j_1})$, $s_h = f(s_{i_1}, a_{d_1})$ and $v_{j_1}^{i_1}\beta = v_{d_1}^{i_1}$. The set $R_{n+1}(\mathfrak{A})$ $(n = 1, 2, ...)$ is defined as the totality of all elements (β', i', h') for which there exist an element $(\beta, i, h) \in R_n(\mathfrak{A})$ and a word v_j^i such that either $s_{i'} = s_h$, $s_{h'} = f(s_i, a_j)$ and $\beta\beta' = v_j^i$, or $s_{i'} = f(s_i, a_j)$, $s_{h'} = s_h$ and $\beta = v_j^i\beta'$.

Lemma 2. *An element (β, i, h) belongs to $R_n(\mathfrak{A})$ if and only if there exist numbers $k \geq 1$, $l \geq 1$, $k + l = n + 1$; $i_1, ..., i_k$,*

---

* The latter name was proposed by Huffman ([3]), who investigated such automata in detail in the synchronous case. This name reflects the fact that a completely defined automaton $\mathfrak{A}$ is one-to-one in the weak sense if and only if there exists an experiment ([2]) of bounded length with the automaton $\mathfrak{A}$, run by an arbitrary (unknown to the experimenter) word α, which makes it possible to determine the word α from the output $F_{\mathfrak{A}}(\alpha)$ and the results of the experiment.

$\tilde{i}_{k+1} = i$, $j_1, ..., j_k$; $h_1, ..., h_l$, $h_{l+1} = h$, $d_1, ..., d_l$ such that $s_{i_{t+1}} = f(s_{i_t}, a_{j_t})$, $t = 1, ..., k$; $s_{h_{p+1}} = f(s_{h_p}, a_{d_p})$, $p = 1, ..., l$; $u\, v_{j_1}^{i_1} v_{j_2}^{i_2} ... v_{j_k}^{i_k}\beta = v_{d_1}^{h_1} v_{d_2}^{h_2} ... v_{d_l}^{h_l}$, where $i_1 = h_1$, $j_1 \neq d_1$, and the word β is the end of the word $v_{d_1}^{h_1}$.

We shall agree to denote the empty word by Λ, the length of a word β in the alphabet B by $\lambda(\beta)$, and the maximum length of the words v_j^i by $\lambda_{\max}$. From Lemma 2 it follows that there exist no more than $N^2(m\lambda_{\max} + 1)$ elements that can constitute the sets $R_n(\mathfrak{A})$. Consequently, these sets begin to repeat periodically starting from some point.

Theorem 1. For the automaton $\mathfrak{A}$ to be one-to-one, it is necessary and sufficient that all words v_j^i be nonempty and that no set $R_n(\mathfrak{A})$, for $n \leq Nm(N\lambda_{\max} - 1) + 2$, contain elements of the form (Λ, i, h).

Theorem 2. For the automaton $\mathfrak{A}$ to be one-to-one in the weak sense, it is necessary and sufficient that no set $R_n(\mathfrak{A})$, for

$$n \leq Nm(N\lambda_{\max} - 1) + N(N-1)/2 + 2,$$

contain elements of the form (Λ, i, i).

Theorem 3. For the automaton $\mathfrak{A}$ to be one-to-one at infinity, it is necessary and sufficient that all sets $R_n(\mathfrak{A})$ be empty starting from some

$$n \leq Nm(N\lambda_{\max} - 1) + N(N-1)/2 + 1.$$

Theorems 1-3 lead to algorithms for recognizing the properties of the automaton $\mathfrak{A}$ under study, consisting in the successive construction and examination of a bounded number of sets $R_n(\mathfrak{A})$. These algorithms are generalizations of the corresponding algorithms from ([1,5]).

Theorem 4. For the automaton $\mathfrak{A}$ to be invertible, it is necessary and sufficient that it be one-to-one at infinity.

The necessity of the condition of Theorem 4 is obvious. Suppose now that the automaton $\mathfrak{A}$ is one-to-one at infinity. Then, by Theorem 3, there exists a minimal number $n(\mathfrak{A})$ such that all sets $R_n(\mathfrak{A})$ are empty for $n \geq n(\mathfrak{A})$. We shall give methods for constructing automata $\mathfrak{A}'$ and $\mathfrak{A}''$, inverse to $\mathfrak{A}$.

Method for constructing the automaton $\mathfrak{A}'$ (see Fig. 3a). Denote by $\mathfrak{N}'$ the set of all pairs (β, i) such that

$$\beta = v_{j_1}^{i_1} v_{j_2}^{i_2} \dots v_{j_k}^{i_k} \gamma_{j_{k+1}}^{i_{k+1}},$$

where $k \geq 0$, $\gamma_{j_{k+1}}^{i_{k+1}}$ is a proper initial segment of the word $v_{j_{k+1}}^{i_{k+1}}$, and the numbers $i_1 = i,\ i_2, \dots, i_k, i_{k+1},\ j_1, j_2, \dots, j_k$ are such that

$$s_{i_{p+1}} = f(s_{i_p}, a_{j_p}), \qquad p = 1, \dots, k.$$

The aggregate of numbers

$$\begin{pmatrix} i_1, i_2, \dots, i_{k+1} \\ j_1, j_2, \dots, j_{k+1} \end{pmatrix}$$

will be called an i-decoding of the word β.

Let

$$\begin{pmatrix} i_1(t), i_2(t), \dots, i_{k(t)+1}(t) \\ j_1(t), j_2(t), \dots, j_{k(t)+1}(t) \end{pmatrix}, \qquad i_1(t) = i,\ t = 1, \dots, M,$$

be all $*$ i-decodings of the word β, and let $k' = k'(\beta, i)$ be the greatest number such that $k(t) \geq k'$, $i_p(t) = i_p$, $j_p(t) = j_p$, $t = 1, \dots, M$; $p = 1, \dots, k'$. By Theorem 3 and Lemma 2 there exists a minimal number

$$T \leq [\, n(\mathfrak{A})/2 \,] \lambda_{\max}$$

(the same for all pairs $(\beta, i) \in \mathfrak{N}'$) such that if, for some t $(t = 1, \dots, M)$, one has $k(t) \geq 1$ and

$$\lambda\Big(v_{j_2(t)}^{i_2(t)} \dots v_{j_{k(t)}(t)}^{i_{k(t)}(t)} \gamma_{j_{k(t)+1}(t)}^{i_{k(t)+1}}\Big) \geq T,$$

then $k(t) \geq 1$ for every t $(t = 1, \dots, M)$ and $k' > 0$. Define on the pairs $(\beta, i) \in \mathfrak{N}'$ the following functions:

$$\psi_1(\beta, i) = i_{k'+1}, \quad \psi_2(\beta, i) = v_{j_{k'+1}}^{i_{k'+1}} \dots v_{j_k}^{i_k} \gamma_{j_{k+1}}^{i_{k+1}}, \quad \psi_3(\beta, i) = a_{j_1} \dots a_{j_{k'}}.$$

Fig. 2. Automaton $\mathfrak{A}$; Fig. 3. Automata inverse to automaton $\mathfrak{A}$ (Fig. 2). a—automaton $\mathfrak{A}'$, b—automaton $\mathfrak{A}''$

Figure 2: Fig. 2. Automaton $\mathfrak{A}$; Fig. 3. Automata inverse to automaton $\mathfrak{A}$ (Fig. 2). a—automaton $\mathfrak{A}'$, b—automaton $\mathfrak{A}''$

In particular, if $k' = 0$, then $\psi_1(\beta, i) = i,\ \psi_2(\beta, i) = \beta,\ \psi_3(\beta, i) = \Lambda$. Denote by $\mathfrak{M}_i$ the set of words β such that $(\beta, i) \in \mathfrak{N}'$ and $\psi_3(\beta, i) = \Lambda$. As the states of the automaton $\mathfrak{A}'$ take the symbols q_β^i, where $\beta \in \mathfrak{M}_i,\ i = 1, \ldots, N$, including the symbol $q_\Lambda^{i_0}$ (even if $\Lambda \notin \mathfrak{M}_{i_0}$), which we shall regard as the initial state. It is obvious that the number of states does not exceed

$$N \cdot \frac{r^{T+\lambda_{\max}} - 1}{r - 1}.$$

The transition and output functions f' and φ' of the automaton $\mathfrak{A}'$ are defined on the pairs (q_β^i, b_j), for which $(\beta b_j, i) \in \mathfrak{N}'$, in the following way:

$$f'(q_\beta^i, b_j) =$$

$$= q_{\psi_2(\beta j, i')}^{\psi_1(\beta j, i')} \varphi'\left(q_\beta^i, b_j\right) = \psi_3(\beta b_j, i).$$

The construction is completed by merging the indistinguishable states of the automaton and by discarding those states that cannot be reached from the initial state $q_\Lambda^{i_0}$.

The method of constructing the automaton $\mathfrak{A}''$ (see Fig. 3b). This method differs from the preceding one only in a different definition of the functions ψ_1, ψ_2, and ψ_3: if $k' = 0$ or $k' \geq 1$ and

$$\lambda\left(v_{j_2}^{i_2} \ldots v_{j_k}^{i_k} v_{j_{k+1}}^{i_{k+1}}\right) < T,$$

then $\psi_1(\beta, i) = i,\ \psi_2(\beta, i) = \beta,\ \psi_3(\beta, i) = \Lambda$; if $k' \geq 1$ and

$$\lambda\left(v_{j_2}^{i_2} \ldots v_{j_k}^{i_k} v_{j_{k+1}}^{i_{k+1}}\right) = T,$$

then

$$\psi_1(\beta, i) = i_{k''+1}, \quad \psi_2(\beta, i) = v_{j_{k''+1}}^{i_{k''+1}} \ldots v_{j_k}^{i_k} v_{j_{k+1}}^{i_{k+1}}, \quad \psi_3(\beta, i) = d_{j_1} \ldots d_{j_{k''}},$$

where k'' is the greatest number such that $1 \leq k'' \leq k'$ and

$$\lambda\left(v_{j_{k''+1}}^{i_{k''+1}} \ldots v_{j_k}^{i_k} v_{j_{k+1}}^{i_{k+1}}\right) = T.$$

Fig. 2. Automaton $\mathfrak{A}$

Fig. 3. Automata inverse to the automaton $\mathfrak{A}$ (Fig. 2). a—automaton $\mathfrak{A}'$, b—automaton $\mathfrak{A}''$

The choice of the automata $\mathfrak{A}'$ and $\mathfrak{A}''$ from the countable set of other automata inverse to $\mathfrak{A}$ is explained by their special significance in coding theory. If one assumes that the automaton $\mathfrak{A}$ is used to encode messages, then the automata $\mathfrak{A}'$ and $\mathfrak{A}''$ may be regarded as decoding automata, the first of which decodes each message with a minimal delay not exceeding T, while the second (in the case of nonempty words v_j^i) decodes each message with a minimal constant delay T. The automaton $\mathfrak{A}'$, and also the automaton $\mathfrak{A}''$ in the indicated case, are determined by these conditions uniquely up to equivalence.

Remark 1. There exist invertible partial automata (see, for example, Fig. 3) for which any automata obtained by completing their definition without extending the output alphabet are not invertible.

Remark 2. It follows from Theorem 4 that if $\mathfrak{A}$ is a finite automaton and the function $(F_{\mathfrak{A}}^{\infty})^{-1}$ is induced by an automaton having an infinite set of states, then it is induced by some finite automaton.

Received
3 VII 1962

References

1. A. A. Sardinas, G. W. Patterson, Conv. Rec. IRE, IT, pt. 8, 104 (1953).
2. E. F. Moore, in: *Automata*, 1956, p. 179.
3. D. A. Huffman, IRE Trans., CT-6, Spec. Suppl., 41 (1959).
4. S. Ginsburg, Trans. Am. Math. Soc., 96, No. 3, 400 (1960).
5. V. I. Levenshtein, DAN, 140, No. 6, 1274 (1961).
6. Yu. V. Glebskii, DAN, 141, No. 5, 1054 (1961).
7. V. I. Levenshtein, DAN, 141, No. 6, 1320 (1961).
8. Al. A. Markov, in: *Problems of Cybernetics*, 8, 1962.

* From Lemma 1 and from the fact that an automaton that is one-to-one in the weak sense cannot output more than $N-1$ empty words in succession, it follows that the number of i-decodings of the word β is finite.

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