Annales Univ. Sci. Budapest., Sect. Comp. 11 (1991) 175-191

ON MATRIX METHODS FOR OPTIMIZATION OF GENERALIZED AUTOMATA

M.K. TCHIRKOV

1. Introduction. Automata optimization is one of the most important (as in theory so in practice) problem of mathematical automata theory.In this paper the classical problem of an automaton states minimization is investigated for finite generalized automata over arbitrary field. The optimization methods proposed are founded on construction of two basis matrices of an automaton and on using of them for special trunsformation of the automaton. Since a generalized automaton is natural generalization of some special classes of finite automata thus these methods may be used (sometimes with some remarks) for optimization of special classes automata. Such methods, for example, are very important for optimization of some systems and processes which may be described in terms of generalized automata over any field. The conceptions of this paper are the further development of the ideas and methods, stated in [1-5].

2. Basic definitions. By an alphabet X we mean a finite non-empty ordered set of elements. $\mid X \mid$ is the cardinal number of the set X. A finite sequence $w = x_1 x_2 ... x_t$ ($x_i \in X, t \geq 0$) is called a word over X, and $t = \mid w \mid$ is the length of w. We use the notations X^* and X^t for the set of all words over X and for the set of all words of length t over X, respectively.

If F is an arbitrary field (for example, the field of real numbers), the following notations are used for the sets of all m-dimensional row-vectors, m-dimensional column-vectors and $(m \times n)$ - matrices over $F : F^{1,m}, F^{m,1}, F^{m,n}$.

Let X, A, Y be the alphabets of inputs, states and outputs, respectively, and $\mid X \mid = n, \mid A \mid = m, \mid Y \mid = k$. Then finite generalized automaton $\mathcal{A}$ over the field F is a system

$$\mathcal{A} = < X, A, Y, r, R, q >, \tag{1}$$

where $r (\in F^{1,m})$ is the initial vector, $q (\in F^{m,1})$ is the final vector and $R : X \times Y \rightarrow F^{m,m}$ is the transition-output function which presents a mapping of $X \times Y$ into $F^{m,m}$. The mapping R is usually represented by a set of square matrices

$$\{R(x,y)\} = \{R(x,y) \mid R(x,y) \in F^{m,m}, x \in X, y \in Y\}.$$

The domain of this mapping is extended from $X \times Y$ to $X^t \times Y^t$ in the following way

$$R(w,v) = \begin{cases} \prod_{i=1}^{t} R(x_i, y_i) & \text{if } t \neq 0, \\ I(m) & \text{if } t = 0, \end{cases} \tag{2}$$

where $w = x_1 x_2 ... x_t \in X^t$, $v = y_1 y_2 ... y_t \in Y^t$ and $I(m)$ is the unit matrix of type $(m \times m)$.

The generalized mapping $\Phi_{\mathcal{A}}$ induced by a generalized automaton $\mathcal{A}$ is the mapping of $X^* \times Y^*$ into F defined by

$$\Phi_{\mathcal{A}}(w,v) = \begin{cases} rR(w,v)q & \text{if } \mid w \mid = \mid v \mid, \\ 0 & \text{if } \mid w \mid \neq \mid v \mid. \end{cases} \tag{3}$$

Hereafter we use the term automaton to mean a finite generalized automaton.

Let

$$\mathcal{A} < X, A, Y, r, \{R(x,y)\}, q >,$$

$$\mathcal{A}' =< X, A', Y, r', \{R'(x,y)\}, q' >$$

be the automata over the field $F, \mid A \mid = m, \mid A' \mid = m'$.

The initial vector r in the automaton $\mathcal{A}$ (with the final vector q) and the initial vector r' in the automaton $\mathcal{A}'$ (with the final vector q') are said to be equivalent (in notation: $r\mathcal{A}(q) \sim r'\mathcal{A}'(q')$) if for these initial vectors

$$\Phi_{\mathcal{A}}(w,v) = \Phi_{\mathcal{A}'}(w,v) \tag{4}$$

for each $(w,v) \in X^* \times Y^*$.

We say that an automaton $\mathcal{A}$ is initially reduced if for every $r \in F^{1,m}$ and every $r' \in F^{1,m}$

$$r\mathcal{A}(q) \sim r'\mathcal{A}(q) \Longleftrightarrow r = r'. \tag{5}$$

The automata $\mathcal{A}$ and $\mathcal{A}'$ are called initially equivalent (in notation $\mathcal{A}(q) \sim \mathcal{A}'(q')$) if for each $r \in F^{1,m}$ there exists $r' \in F^{1,m'}$ such that for r, r' (4) holds and vice versa.

Every initially reduced automaton $\mathcal{A}'$ such that $\mathcal{A}'(q') \sim \mathcal{A}(q)$ is said to be the initially reduced form of the automaton $\mathcal{A}$ (in notation: $\mathcal{A}' = Red\mathcal{A}(q)$).

Accordingly, the final vector q in the automaton $\mathcal{A}$ (with the initial vector r) and the final vector q' in the automaton $\mathcal{A}'$ (with the initial vector r') are called equivalent (in notation: $(r)\mathcal{A}q \sim (r')\mathcal{A}'q'$) if for q, q' (4) holds.

We say that an automaton $\mathcal{A}$ is finally reduced if for every $q \in F^{m,1}$ and every $q' \in F^{m,1}$

$$(r)\mathcal{A}q \sim (r)\mathcal{A}q' \Longleftrightarrow q = q'. \tag{6}$$

The automata $\mathcal{A}$ and $\mathcal{A}'$ are called finally equivalent (in notation: $(r)\mathcal{A} \sim (r')\mathcal{A}'$) if for each $q \in F^{m,1}$ there exists $q' \in F^{m'1}$ such that for q, q' (4) holds and vice versa.

Every finally reduced automaton $\mathcal{A}'$ such that $(r')\mathcal{A}' \sim (r)\mathcal{A}$ is called the finally reduced form of the automaton $\mathcal{A}$ (in notation: $\mathcal{A}' = Red(r)\mathcal{A}$).

The pair of the vectors (r, q) in the automaton $\mathcal{A}$ and the pair of the vectors $(r'q')$ in the automaton $\mathcal{A}'$ are called equivalent (in notation:$r\mathcal{A}q \sim r'\mathcal{A}q'$) if for these pairs (4) holds.

We say that an automaton $\mathcal{A}$ is in a minimal form if there is not any automaton $\mathcal{A}'$ such that (4) holds and $| A' | < | A |$.

By a minimal form of an automaton $\mathcal{A}$ we mean every automaton $\mathcal{A}'$ which is in a minimal form and such that (4) holds (in notation: $\mathcal{A}' = Min\mathcal{A}$).

In accordance with above definitions some problems of generalized automata optimization may be formulated. Let $\mathcal{A}$ be the automaton (1). It is necessary to construct:

a) an initially reduced form of the automaton $\mathcal{A}$;

b) a finally reduced form of the automaton $\mathcal{A}$;

c) a minimal form of the automaton $\mathcal{A}$.

3. Basis matrices of an automaton. For stochastic automata the concept of a $\mathcal{L}$-basis matrix is known [1-3,5]. The furthermost generalization of this concept for a generalized automaton (1) over the field F may be made in the following way [4].

Let $\mathcal{L}(\mathcal{A}q)$ be the m-dimensional vector space generated by the set of column-vectors

$$\mathcal{L}(\mathcal{A}q) = \{h_q(w, v) \mid h_q(w, v) =$$
$$= R(w, v)q, w \in X^*, v \in Y^*, \mid w \mid = \mid v \mid\}. \tag{7}$$

Then a matrix $H(\mathcal{A}q)$ of type $(m \times dim\mathcal{L}(\mathcal{A}q))$, which consists of a system of linearly independent column-vectors of the space $\mathcal{L}(\mathcal{A}q)$, is called the $\mathcal{L}(q)$-basis matrix of the automaton $\mathcal{A}$.

Now, let $\mathcal{L}(r\mathcal{A})$ be the m-dimensional vector space generated by the set of row-vectors

$$\mathcal{L}(r\mathcal{A}) = \{h_r(w,v) \mid h_r(w,v) = \\ = rR(w,v), w \in X^*, v \in Y^*, \mid w \mid = \mid v \mid\}. \tag{8}$$

Then a matrix $H(r\mathcal{A})$ of type $(dim\mathcal{L}(r\mathcal{A}) \times m)$, which consists of a system of linearly independent row-vectors of the space $\mathcal{L}(r\mathcal{A})$, is called the $\mathcal{L}(r)$-basis matrix of the automaton $\mathcal{A}$.

For construction of the matrices $H(\mathcal{A}q)$ and $H(r\mathcal{A})$ may be used, for example, the simplified procedure which was proposed by the author in the books [1,2].

Now we are going to study a few properties of the matrices $H(\mathcal{A}q)$ and $H(r\mathcal{A})$. Let $H^+(\mathcal{A}q)$ and $H^+(r\mathcal{A})$ are the pseudo inverse matrices [6] of $H(\mathcal{A}q)$ and $H(r\mathcal{A})$, respectively. Since the columns of the matrix $H(\mathcal{A}q)$ are linearly independent and the rows of the matrix $H(r\mathcal{A})$ are linearly independent thus the matrices $H^+(\mathcal{A}q), H^+(r\mathcal{A})$ may be given by the expressions

$$H^+(\mathcal{A}q) = (H^*(\mathcal{A}q)H(\mathcal{A}q))^{-1}H^*(\mathcal{A}q), \\ H^+(r\mathcal{A}) = H^*(r\mathcal{A})(H(r\mathcal{A})H^*(r\mathcal{A}))^{-1}, \tag{9}$$

where H^* is the conjugate matrix of the matrix H and H^{-1} is the inverse matrix of the matrix H [6].

It follows from (9) that

$$H^+(\mathcal{A}q)H(\mathcal{A}q) = I(rankH(\mathcal{A}q)),$$

$$H(r\mathcal{A})H^+(r\mathcal{A}) = I(rankH(r\mathcal{A})). \tag{10}$$

In accordance with the definition of a pseudo-inverse matrix [6]

$$H(r\mathcal{A})H^+(r\mathcal{A})H(r\mathcal{A}) = H(r\mathcal{A}), \\ H(\mathcal{A}q)H^+(\mathcal{A}q)H(\mathcal{A}q) = H(\mathcal{A}q)$$

and therefore

$$h_r(w,v)H^+(r\mathcal{A})H(r\mathcal{A}) = h_r(w,v), \\ H(\mathcal{A}q)H^+(\mathcal{A}q)h_q(w,v) = h_q(w,v) \tag{11}$$

for every $h_r(w,v) \in \mathcal{L}(r\mathcal{A})$ and every $h_q(w,v) \in \mathcal{L}(\mathcal{A}q)$, respectively.

The following statement is true.

Theorem 1. Let $\mathcal{A}$ be an automaton (1) and $H(r\mathcal{A})$ be its $\mathcal{L}(r)$-basis matrix. Then for arbitrary $q, q' \in F^{m,1}$

$$(r)\mathcal{A}q \sim (r)\mathcal{A}q' \Leftrightarrow H(r\mathcal{A})q = H(r\mathcal{A})q'. \tag{12}$$

Proof. It follows from (4), (8) that

$$(r)\mathcal{A}q \sim (r)\mathcal{A}q' \Leftrightarrow h_r(w,v)q = h_r(w,v)q'$$

for every $(w,v) \in X^* \times Y^*$.Then

$$(r)\mathcal{A}q \sim (r)\mathcal{A}q' \Leftrightarrow hq = hq'$$

for each $h \in \mathcal{L}(r\mathcal{A})$. Since the rows of the matrix $H(r\mathcal{A})$ form a basis of the vector space $\mathcal{L}(r\mathcal{A})$ thus (12) holds. Q.e.d.

If we take (4), (7), the next statement may be proved by the analogous method.

Theorem 2. Let $\mathcal{A}$ be an automaton (1) and $H(\mathcal{A}q)$ be its $\mathcal{L}(q)$-basis matrix. Then for arbitrary $r, r' \in F^{1,m}$

$$r\mathcal{A}(q) \sim r'\mathcal{A}(q) \Leftrightarrow rH(\mathcal{A}q) = r'H(\mathcal{A}q). \tag{13}$$

4. Some theorems on reduced forms. Let us now prove some statements on reduced forms of a generalized automaton.

Theorem 3. An automaton $\mathcal{A}$ is finally reduced if and only if *rank* $H(r\mathcal{A}) = \mid A \mid$.

Proof. It follows from (12) that

$$(r)\mathcal{A}q \sim (r)\mathcal{A}q' \Leftrightarrow H(r\mathcal{A})(q - q') = 0.$$

Then, accordingly with the definition, the automaton $\mathcal{A}$ is finally reduced if and only if the system

$$H(r\mathcal{A})(q - q') = 0$$

has no non-trivial solution $(q - q') \neq 0$ and therefore if and only if $rank H(r\mathcal{A}) = \mid A \mid$. Q.e.d.

The analogous statement may be proved for initially reduced automata similarly.

Theorem 4. An automaton $\mathcal{A}$ is initially reduced if and only if *rank* $H(\mathcal{A}q) = \mid A \mid$.

The following property of a finally reduced automaton is important.

Theorem 5. If an automaton

$$\mathcal{A} < X, A, Y, r_A, \{R_A(x,y)\}, q_A >$$

is finally reduced, then there does not exist an automaton

$$\mathcal{B} =< X, B, Y, r_B\{R_B(x,y)\}, q_B >$$

such that $(r_A)\mathcal{A} \sim (r_B)\mathcal{B}$ and $\mid B \mid < \mid A \mid$.

Proof. Let us assume that the inverse proposition is true and there exists an automaton $\mathcal{B}$ such that $(r_B)\mathcal{B} \sim (r_A)\mathcal{A}$ and $\mid B \mid = m_B < \mid A \mid = m_A$. Since the automaton $\mathcal{A}$ is finally reduced thus, by Theorem 3, $rank H(r_A\mathcal{A}) = m_A$.

Let

$$q_A^{(i)} = (0, ..., 0, 1, 0, ..., 0)^T, \quad i = \overline{1, m_A},$$

be degenerate final column-vectors such that the element 1 is in the position number i. Assume now that

$$(r_A)\mathcal{A}q_A^{(i)} \sim (r_B)\mathcal{B}q_B^{(i)}, \quad \overline{1, m_A},$$

where

$$q_B^{(i)} = (q_1^{(i)}, q_2^{(i)}, ..., q_{m_B}^{(i)})^T, \quad \overline{1, m_A}.$$

Then in accordance with (3), (4), (8)

$$h_r^{(A)}(w,v)q_A^{(i)} = h_r^{(B)}(w,v)q_B^{(i)}, \quad \overline{1, m_A},$$

or

$$h_r^{(A)}(w,v)I = h_r^{(B)}(w,v)Q_B, \tag{14}$$

where $Q_B = (q_B^{(1)} q_B^{(2)} ... q_B^{(m_A)})$ is the matrix of type $(m_B \times m_A)$.

Now we take the pairs of the words $(w_\nu, v_\nu), \nu = \overline{1, m_A}$, in such a manner that the row-vectors $h_r^{(A)}(w_\nu, v_\nu), \nu = \overline{1, m_A}$, are linearly independent and therefore these row-vectors are the rows of $\mathcal{L}(r_A)$-basis matrix $H(r_A\mathcal{A})$ of the automaton $\mathcal{A}$ and, by Theorem 3,

$$rank H(r_A\mathcal{A}) = m_A. \tag{15}$$

From (14) we have

$$H(r\mathcal{A}) = H_B Q_B, \tag{16}$$

where the rows of the matrix H_B are the row-vectors $h_r^{(B)}(w_\nu, v_\nu)$, $\nu \doteq \overline{1, m_A}$.

The matrix $H_B$is of type $(m_B \times m_A)$ and the matrix Q_B is of type $(m_A \times m_B)$, where $m_B < m_A$. Therefore, by Sylvester law of nullity [6],

$$rank(H_B Q_B) \leq m_B < m_A. \tag{17}$$

But (17) contradicts with (15), (16). Therefore our assumption is untrue and such an automaton $\mathcal{B}$ does not exist. Q.e.d.

The analogous theorem may be proved for initially reduced, automata similarly.

Theorem 6. If an automaton

$$\mathcal{A} < X, A, Y, r_A, \{R_A(x,y)\}, q_A >$$

is initially reduced, then there does not exist an automaton

$$\mathcal{B} =< X, B, Y, r_B, \{R_B(x,y)\}, q_B >$$

such that $\mathcal{A}(q_A) \sim \mathcal{B}(q_B)$ and $|B| < |A|$.

5. Methods for construction of reduced forms. Let us prove the following statement.

Theorem 7. Let $\mathcal{A}$ be an automaton

$$\mathcal{A} =< X, A, Y, r_A, \{R_A(x,y)\}, q_A >$$

and $H(\mathcal{A}q_A)$ be its $\mathcal{L}(q_A)$-basis matrix. If an automaton

$$\mathcal{B} =< X, B, Y, r_B, \{R_B(x,y)\}, q_B >$$

is such that for $x \in X$, $y \in Y$

$$R_B(x,y) = H^+(\mathcal{A}q_A) R_A(x,y) H(\mathcal{A}q_A), \tag{18}$$

$$r_B = r_A H(\mathcal{A}q_A), q_B = H^+(\mathcal{A}q_A) q_A,$$

then

(a) $\mathcal{B} = Red\mathcal{A}(q_A)$, $|B| = m_B = rank H(\mathcal{A}q_A)$,

$H(\mathcal{B}q_B) = I(m_B)$;

(b) $H(r_B\mathcal{B})$ is the matrix which may be formed of a system of linearly independent rows of the matrix $H(r_A\mathcal{A})H(\mathcal{A}q_A)$ and $rank H(r_B\mathcal{B}) = rank(H(r_A\mathcal{A})H(\mathcal{A}q_A))$;

(c) for every pair $(r,q), r \in F^{1,m_A}, q \in \mathcal{L}(\mathcal{A}q_A)$ there is the pair (r',q') such that

$$r' = rH(\mathcal{A}q_A), \quad q' = H^+(\mathcal{A}q_A)q, \quad r\mathcal{A}q \sim r'\mathcal{B}q'$$

and

$$r \in \mathcal{L}(r_A\mathcal{A}) \Rightarrow r' \in \mathcal{L}(r_B\mathcal{B});$$

(d) for every pair $(r',q'), r' \in F^{1,m_B}, q' \in F^{m_B,1}$ there is the pair (r,q) such that

$$r = r'H^+(\mathcal{A}q_A), \quad q = H(\mathcal{A}q_A)q' \in \mathcal{L}(\mathcal{A}q_A), \tag{19}$$

$$r\mathcal{A}q \sim r'\mathcal{B}q'$$

and

$$r' \in \mathcal{L}(r_B\mathcal{B}) \Rightarrow r \in \mathcal{L}(r_A\mathcal{A}).$$

Proof. First of all let us prove that

$$\mathcal{A}(q_A) \sim \mathcal{B}(q_B).$$

Let r be an arbitrary $r \in F^{1,m_A}$. Then

$$\Phi_{\mathcal{A}}(w,v) = r\prod_{i=1}^{t} R_A(x_i,y_i)q_A.$$

Since $q_A \in \mathcal{L}(\mathcal{A}q_A)$ and

$$h_q^{(A)}(w_\nu, v_\nu) = \prod_{i=\nu}^{t} R_A(x_i,y_i)q_A \in \mathcal{L}(\mathcal{A}q_A) \tag{20}$$

for $x_i \in X, y_i \in Y,\ \nu = \overline{1,t}$, thus in accordance with (11),(18)

$$\Phi_{\mathcal{A}}(w,v) = r\prod_{i=1}^{t} R_A(x_i,y_i)q_A =$$

$$= r\prod_{i=1}^{t}[H(\mathcal{A}q_A)H^+(\mathcal{A}q_A)R_A(x_i,y_i)](H(\mathcal{A}q_A)H^+(\mathcal{A}q_A))q_A =$$

$$= rH(\mathcal{A}q_A)\prod_{i=1}^{t}[H^+(\mathcal{A}q_A)R_A(x_i,y_i)H(\mathcal{A}q_A)]H^+(\mathcal{A}q_A)q_A =$$

$$= r'\prod_{i=1}^{t} R_B(x_i,y_i)q_B = \Phi_{\mathcal{B}}(w,v), \tag{21}$$

where

$$r' = rH(\mathcal{A}q_A). \tag{22}$$

Conversely, let r' be an arbitrary $r' \in F^{1,m_B}$, then in accordance with (18)

$$\Phi_{\mathcal{B}}(w,v) = r'\prod_{i=1}^{t} R_B(x_i,y_i)q_B =$$

$$= r'\prod_{i=1}^{t}[H^+(\mathcal{A}q_A)R_A(x_i,y_i)H(\mathcal{A}q_A)]H^+(\mathcal{A}q_A)q_A =$$

$$= r'H^+(\mathcal{A}q_A)R_A(x_1,y_1)\prod_{i=2}^{t}[H(\mathcal{A}q_A)H^+(\mathcal{A}q_A)R(x_i,y_i)]\times$$

$$\times H(\mathcal{A}q_A)H^+(\mathcal{A}q_A)q_A.$$

Since (11) and (20) hold thus, by consecutive application of (11) for q_A and $h_q^{(A)}(w_\nu,y_\nu), \nu = t, t-1, ..., 1$, we have

$$\Phi_{\mathcal{B}}(w,v) = r'H^+(\mathcal{A}q_A)\prod_{i=1}^{t} R_A(x_i,y_i)q_A =$$

$$= r\prod_{i=1}^{t} R_A(x_i,y_i)q_A = \Phi_{\mathcal{A}}(w,v),$$

where $r = r'H^+(\mathcal{A}q_A)$. Therefore $\mathcal{A}(q_A) \sim \mathcal{B}(q_B)$,

$$h_q^{(B)}(w,v) = H^+(\mathcal{A}q_A)h_q^{(A)}(w,v)$$

and

$$H(\mathcal{B}q_B) = H^+(\mathcal{A}q_A)H(\mathcal{A}q_A) = I(rank H(\mathcal{A}q_A)).$$

Then, by Theorem 4, we proved the statement (a).

Now let us find the $rank H(r_B\mathcal{B})$. In accordance with (18)

$$\begin{aligned} h_r^{(B)}(w,v) &= r_B \prod_{i=1}^{t} R_B(x_i,y_i) = \\ &= r_B \prod_{i=1}^{t} [H^+(\mathcal{A}q_A)R_A(x_i,y_i)H(\mathcal{A}q_A)] = \\ &= r_A H(\mathcal{A}q_A)H^+(\mathcal{A}q_A)\times \\ &\times \prod_{i=1}^{t-1} [R_A(x_i,y_i)H(\mathcal{A}q_A)H^+(\mathcal{A}q_A)]R_A(x_t,y_t)H(\mathcal{A}q_A). \end{aligned} \tag{23}$$

Since for every $h \in \mathcal{L}(\mathcal{A}q_A)$ the condition (11) holds thus

$$H(w,v) = \prod_{i=\nu}^{t} R_A(x_i,y_i)H(\mathcal{A}q_A) \Rightarrow$$

$$\Rightarrow H(\mathcal{A}q_A)H^+(\mathcal{A}q_A)H(w,v) = H(w,v).$$

Therefore from (23) we have

$$\begin{aligned} h_r^{(B)}(w,v) &= r_A \prod_{i=1}^{t} R_A(x_i,y_i)H(\mathcal{A}q_A) = \\ &= h_r^{(A)}(w,v)H(\mathcal{A}q_A) \in \mathcal{L}(r_B\mathcal{B}), \end{aligned} \tag{24}$$

and $\mathcal{L}(r_B)$-basis matrix $H(r_B\mathcal{B})$ of the automaton $\mathcal{B}$ may be made of a system of linearly independent rows of the matrix $H(r_A\mathcal{A})H(\mathcal{A}q_A)$. This completes the proof of the statement (b).

Now let us take a pair $(r,q), r \in F^{1,m}, q \in \mathcal{L}(\mathcal{A}q_A)$. Since $q \in \mathcal{L}(\mathcal{A}q_A)$ thus the $\mathcal{L}(q)$-basis matrix of the automaton $\mathcal{A}$ (with q) is just the same as with q_A. Therefore (21), (22) holds for $q_A = q$. From (24) we have

$$r \in \mathcal{L}(r_A\mathcal{A}) \Rightarrow rH(\mathcal{A}q_A) \in \mathcal{L}(r_B\mathcal{B}).$$

This ends the proof of the statement (c).

Conversely, let (r', q') consists of arbitrary $r' \in F^{1,m_B}, q' \in F^{m_B,1}$, then

$$\Phi_{\mathcal{B}}(w,v) = r' \prod_{i=1}^{t} R_B(x_i, y_i) q' =$$

$$= r' \prod_{i=1}^{t} [H^+(\mathcal{A}q_A) R_A(x_i, y_i) H(\mathcal{A}q_A)] q' =$$

$$= r' H^+(\mathcal{A}q_A) \prod_{i=1}^{t-1} [R_A(x_i, y_i) H(\mathcal{A}q_A) H^+(\mathcal{A}q_A)] \times$$

$$\times R_A(x_t, y_t) H(\mathcal{A}q_A) q'.$$

Since $H(\mathcal{A}q_A)q' \in \mathcal{L}(\mathcal{A}q_A)$ thus for $q = H(\mathcal{A}q_A)q'$ and for

$$h_q^{(A)}(w_\nu, v_\nu) = \prod_{\nu=1}^{t} R_A(x_i, y_i) q \in \mathcal{L}(\mathcal{A}q_A), \ \nu = \overline{1,t},$$

(11) holds. Therefore we have

$$\Phi_{\mathcal{B}}(w,v) = r' H^+(\mathcal{A}q_A) \prod_{i=1}^{t} R_A(x_i, y_i) H(\mathcal{A}q_A) q' =$$

$$= r \prod_{i=1}^{t} R_A(x_i, y_i) q = \Phi_{\mathcal{A}}(w,v), \tag{25}$$

where $r = r' H^+(\mathcal{A}q_A)$ and $q = H(\mathcal{A}q_A)q' \in \mathcal{L}(\mathcal{A}q_A)$.

If $r' \in \mathcal{L}(r_B\mathcal{B})$, then, by the statement (b), there is such a row-vector $(c_1, c_2, ..., c_\eta), \eta = rank H(r_A\mathcal{A})$, that

$$r' = (c_1, c_2, ..., c_\eta) H(r_A\mathcal{A}) H(\mathcal{A}q_A). \tag{26}$$

In this case we have from (11), (25), (26)

$$\Phi_{\mathcal{B}}(w,v) = (c_1, c_2, ..., c_\eta) H(r_A\mathcal{A}) \prod_{i=1}^{t} R_A(x_i, y_i) q =$$

$$= r\prod_{i=1}^{t} R_A(x_i, y_i)q = \Phi_A(w, v),$$

where $r = (c_1, c_2, ..., c_\eta)H(r_A\mathcal{A}) \in \mathcal{L}(r_A\mathcal{A})$. Q.e.d.

For $\mathcal{L}(r)$-basis matrix of the automaton $\mathcal{A}$ the analogous statements may be proved similarly.

Theorem 8. Let $\mathcal{A}$ be an automaton

$$\mathcal{A} < X, A, Y, r_A, \{R_A(x, y)\}, q_A >$$

and $H(r_A\mathcal{A})$ be its $\mathcal{L}(r_A)$-basis matrix.
If an automaton

$$\mathcal{D} =< X, D, Y, r_D, \{R_D(x, y)\}, q_D >$$

is such that for $x \in X, , y \in Y$

$$R_D(x, y) = H(r_A\mathcal{A})R_A(x, y)H^+(r_A\mathcal{A}),$$

$$r_D = r_A H^+(r_A\mathcal{A}), q_D = H(r_A\mathcal{A})q_A, \tag{27}$$

then

(a) $\mathcal{D} = Red\ (r_A)\mathcal{A}, \mid D \mid= m_D =\ rank\ H(r_A\mathcal{A}),\ H(r_D\mathcal{D}) = I(m_D)$;
(b) $H(\mathcal{D}q_D)$ is the matrix which may be formed of a system of linearly independent columns of the matrix $H(r_A\mathcal{A})H(\mathcal{A}q_A)$ and $rank\ H(\mathcal{D}q_D) =\ rank\ (H(r_A\mathcal{A}) \times H(\mathcal{A}q_A))$;
(c) for every pair $(r, q), r \in \mathcal{L}(r_A\mathcal{A}), q \in F^{m_A,1}$, there is the pair (r', q') such that $r' = rH^+(r_A\mathcal{A}), q' = H(r_A\mathcal{A})q,\ \ r\mathcal{A}q \sim r'\mathcal{D}q'$ and $q \in \mathcal{L}(\mathcal{A}q_A) \Rightarrow q' \in \mathcal{L}(\mathcal{D}q_D)$;
(d) for every pair $(r', q'), r' \in F^{1,m_D}, q' \in F^{m_D,1}$, there is the pair (r, q) such that

$$r = r'H(r_A\mathcal{A}) \in \mathcal{L}(r_A\mathcal{A}),\ q = H^+(r_A\mathcal{A})q',$$

$$r\mathcal{A}q \sim r'\mathcal{D}q'$$

and

$$q' \in \mathcal{L}(\mathcal{D}q_D) \Rightarrow q \in \mathcal{L}(\mathcal{A}q_A).$$

Theorems 7 and 8 and the known methods for computation of matrices $H(r_A\mathcal{A}), H(\mathcal{A}q_A), H^+(r_A\mathcal{A}), H^+(\mathcal{A}q_A)$[1, 2, 6] give us effective methods for construction of the initially or finally reduced forms of generalized automata.

6.Methods for construction of minimal forms. Now let us prove the following statement.

Theorem 9. Let

$$\mathcal{A} < X, A, Y, r_A, \{R_A(x, y)\}, q_A >$$

be an automaton and $H(r_A\mathcal{A}), H(\mathcal{A}q_A)$ be its $\mathcal{L}(r_A)$ - and $\mathcal{L}(q_A)$ - basis matricies, respectively. Then $\mathcal{A}$ is in a minimal form if and only if

$$rank H(r_A\mathcal{A}) = rank H(\mathcal{A}q_A) = \mid A \mid .$$

Proof. The necessity of this theorem conditions is obvious. Let us show the sufficiency. We assume that the inverse proposition is true and therefore an automaton $\mathcal{B} =< X, B, Y, r_B, \{R_B(x, y)\}, q_B >$ exists such that

$$\Phi_{\mathcal{A}}(w, v) = \Phi_{\mathcal{B}}(w, v), \quad w \in X^*, v \in y^*, \tag{28}$$

and $m_B = \mid B \mid < \mid A \mid = m_A$. Let w, v be such that

$$w = w_1 w_2, \ v = v_1 v_2, \ \mid w_1 \mid = \mid v_1 \mid, \ \mid w_2 \mid = \mid v_2 \mid . \tag{29}$$

It follows from (2), (3), (4), (28) that

$$\Phi_{\mathcal{A}}(w, v) = \left[r_A \prod_{i=1}^{|w_1|} R_A(x_i, y_i)\right] \left[\prod_{i=|w_1|+1}^{|w|} R_A(x_i, y_i) q_A\right] =$$

$$= \left[r_B \prod_{i=1}^{|w_1|} R_B(x_i, y_i)\right] \left[\prod_{i=|w_2|+1}^{|w|} R_B(x_i, y_i) q_B\right] = \Phi_{\mathcal{B}}(w, v)$$

and therefore

$$h_r^{(A)}(w_1, v_1) h_q^{(A)}(w_2, v_2) = h_r^{(B)}(w_1, v_1) h_q^{(B)}(w_2, v_2) \tag{30}$$

for arbitrary $w_1, w_2 \in X^*, v_1, v_2 \in Y^*$, which satisfy to the conditions (29).

Now we take $(w_1^{(i)}, v_1^{(i)}), i = \overline{1, m_A}$, in such a way that the row-vectors $h_r^{(A)}(w_1^{(i)}, v_1^{(i)}), i = \overline{1, m_A}$, are linearly independent. Accordingly, we take $(w_2^{(i)}, v_2^{(i)}), i = \overline{1, m_A}$, in such a way that the column-vectors $h_q^{(A)}(w_2^{(i)}, v_2^{(i)}), i = \overline{1, m_A}$, are linearly independent. Then it follows from (30) that

$$H(r_A\mathcal{A})H(\mathcal{A}q_A) = H_B Q_B, \tag{31}$$

where the rows of the matrix H_B are the row-vectors $h_r^{(B)}(w_1^{(i)}, v_1^{(i)}), i = \overline{1, m_A}$, and the columns of the matrix Q_B are the column-vectors $h_q^{(B)}(w_2^{(i)}, v_2^{(i)}), i = \overline{1, m_A}$. The matrix H_B is of type $(m_A \times m_B)$ and the matrix Q_B is of type $(m_B \times m_A)$, where $m_B < m_A$. Therefore, by Sylvester law of nullity [6],

$$rank(H_B Q_B) \leq m_B < m_A. \tag{32}$$

But in accordance with the theorem conditions

$$rank(H(r_A\mathcal{A})H(\mathcal{A}q_A)) = m_A$$

which contradicts with (31), (32). Therefore our assumption is untrue and such an automaton $\mathcal{B}$ does not exist. Q.e.d. The next theorem immediately follows from Theorems 7,8,9.

Theorem 10. Let $\mathcal{A}$ be an automaton and $H(r_A\mathcal{A}), H(\mathcal{A}q_A)$ be its $\mathcal{L}(r_A)$- and $\mathcal{L}(q_A)$-basis matrices, respectively. Let $\mathcal{B} = Red\mathcal{A}(q_A)$ $(\mathcal{B} = Red(r_A)\mathcal{A})$ be the automaton constructed from $\mathcal{A}$ in accordance with Theorem 7 (Theorem 8) and $\mathcal{D}$ be the automaton constructed from $\mathcal{B}$ in accordance with Theorem 8 (Theorem 7). Then

(a) $\mathcal{D} = Min\mathcal{A}$;

(b) $\mid D \mid = m_D = rank(H(r_A\mathcal{A})H(\mathcal{A}q_A))$;

(c) for every pair

$$(r, q), \;\; r \in \mathcal{L}(r_A\mathcal{A}), q \in \mathcal{L}(\mathcal{A}q_A) \tag{33}$$

there is the pair

$$(r', q'), \;\; r' \in F^{1,m_D}, \;\; q' \in F^{m_D,1}, \tag{34}$$

such that

$$r\mathcal{A}q \;\; \sim \;\; r'\mathcal{D}q', \tag{35}$$

and, conversely, for every pair (34) there is the pair (33) such that (35) holds.

This Theorem gives us an effective method for construction of minimal forms of generalized automata.

7. Algorithms of optimization. Let $\mathcal{A}$ be a generalized automaton over the field F.The problems of generalized automata optimization which were formulated in above may be solved now in the following way:

1) To construct $H(r_A\mathcal{A})$ and $H(\mathcal{A}q_A)$ (for example, using the procedure proposed in [1,2]), to find $H(r_A\mathcal{A})H(\mathcal{A}q_A)$ and $rank\ (H(r_A\mathcal{A})\ H(\mathcal{A}q_A))$.

2) If $rank(H(r_A\mathcal{A})H(\mathcal{A}q_A)) <\mid A \mid$, then to calculate $H^+(r_A\mathcal{A})$ and $H^+(\mathcal{A}q_A)$ in accordance with (9).

3) If $rankH(\mathcal{A}q_A) <\mid A \mid$ $(rank\ H(r_A\mathcal{A}) <\mid A \mid)$ then to find $\mathcal{B} = Red\mathcal{A}(q_A)(\mathcal{B} = Red(r_A)\mathcal{A})$ using the contruction of Theorem 7 (Theorem 8). If $rank(H(r_A\mathcal{A})H(\mathcal{A}q_A)) =\mid B \mid$, then $\mathcal{B} = Min\mathcal{A}$.

4) If $rank(H(r_A\mathcal{A})H(\mathcal{A}q_A)) <\mid B \mid$, then to find matrix $H(r_B\mathcal{B})$ $(H(\mathcal{B}q_B))$ and to find $\mathcal{D} = Red(r_B)\mathcal{B}(\mathcal{D} = Red\mathcal{B}(q_B))$ using the contruction of Theorem 8 (Theorem 7).In result we have $\mathcal{D} = Min\mathcal{A}$.

Example. Let $\mathcal{A}$ be the generalized automaton over the field of real numbers defined as

$$\mathcal{A} < X, A, Y, r, \{R(x,y)\}, q >,$$

where

$$X = \{x_1\},\quad A = \{a_1, a_2, a_3, a_4\},\quad Y = \{y_1, y_2\},$$
$$r = (0;0;0,5;0,5),\quad q = (1,1,1,1)^T,$$

$$R(x_1,y_1) = \begin{pmatrix} 0,2 & 0,3 & 0,1 & 0,1 \\ 0,1 & 0,4 & 0,1 & 0,1 \\ 0,2 & 0 & 0,1 & 0,1 \\ 0,2 & 0,2 & 0,1 & 0,1 \end{pmatrix},$$

$$R(x_1,y_2) = \begin{pmatrix} 0,1 & 0,2 & 0 & 0 \\ 0,2 & 0,1 & 0 & 0 \\ 0,1 & 0,1 & 0,2 & 0,2 \\ 0,2 & 0 & 0,1 & 0,1 \end{pmatrix}.$$

It is necessary to find $\mathcal{D} = Min\mathcal{A}$.

First of all we construct $\mathcal{L}(r)$-basis matrix and $\mathcal{L}(q)$-basis matrix of $\mathcal{A}$:

$$H(r\mathcal{A}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0,5 & 0,5 \end{pmatrix}, H(\mathcal{A}q) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$H(r\mathcal{A})H(\mathcal{A}q) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0,5 & 0,5 \end{pmatrix}, rank(H(r\mathcal{A})H(\mathcal{A}q)) = 2.$$

Since $| A |> 2$ thus we calculate the pseudo-inverse matrix

$$H^{+}(\mathcal{A}q) = \begin{pmatrix} 0,5 & 0,5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Using the construction of Theorem 7 we have

$$\mathcal{B} =< X, B, Y, r_B, \{R_B(x,y)\}, q_B >,$$

where

$$B = \{b_1, b_2, b_3\},$$

$$r_B = rH(\mathcal{A}q) = (0; 0,5; 0,5),$$

$$q_B = H^{+}(\mathcal{A}q)q = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

$$R_B(x_1, y_1) = H^{+}(\mathcal{A}q)R(x_1, y_1)H(\mathcal{A}q) = \begin{pmatrix} 0,5 & 0,1 & 0,1 \\ 0,2 & 0,1 & 0,1 \\ 0,4 & 0,1 & 0,1 \end{pmatrix},$$

$$R_B(x_1, y_2) = H^{+}(\mathcal{A}q)R(x_1, y_2)H(\mathcal{A}q) = \begin{pmatrix} 0,3 & 0 & 0 \\ 0,2 & 0,2 & 0,2 \\ 0,2 & 0,1 & 0,1 \end{pmatrix},$$

$$H(r_B\mathcal{B}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0,5 & 0,5 \end{pmatrix}, H^{+}(r_B\mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Using now the construction of Theorem 8 we have

$$\mathcal{D} =< X, D, Y, r_D, \{R_D(x,y)\}, q_D >,$$

$$D = \{d_1, d_2\}, \quad r_D = (0,1), \quad q_D = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$R_D(x_1, y_1) = \begin{pmatrix} 0,5 & 0,2 \\ 0,3 & 0,2 \end{pmatrix}, R_D(x_1, y_2) = \begin{pmatrix} 0,3 & 0 \\ 0,2 & 0,3 \end{pmatrix},$$

$$\mathcal{D} = Min\mathcal{A}.$$

REFERENCES

[1] TCHIRKOV, M.K., Elements of general automata theory, Leningrad, 1975 (in Russian).

[2] TCHIRKOV, M.K., Partial automata, Leningrad, 1983 (in Russian).

[3] BUHARAEV, R.G., Elements of probabilistic automata theory, Moskva, 1985 (in Russian).

[4] TCHIRKOV, M.K., SHESTAKOV, A.A., Similarity and minimization of generalized finite automata, *Mathematical problems of informatica*, 158-173, Leningrad, 1987 (in Russian).

[5] PAZ, A., Introduction to probabilistic automata, New-York, 1971.

[6] GANTMAHER, F.R., Matrix theory, Moskva, 1967. (in Russian).

(*Received October* 24, 1988)

M.K. TCHIRKOV
Leningrad State University
USSR