ANNALES ACADEMIAE SCIENTIARUM FENNICAE

Series A

I. MATHEMATICA

506

ON THE MINIMIZATION OF LINEAR SPACE AUTOMATA

BY

PAAVO TURAKAINEN

HELSINKI 1972
SUOMALAINEN TIEDEAKATEMIA

doi:10.5186/aasfm.1972.506

ISBN 951-41-0033-6

Communicated 10 December 1971 by ARTO SALOMAA.

KESKUSKIRJAPAINO
HELSINKI 1972

On the minimization of linear space automata

For an n-state generalized automaton (cf. [3]), the state space is the n-dimensional real vector space R^n of n-dimensional row vectors, and the transition matrices can be considered as linear mappings from R^n into itself. Instead of generalized automata, it is natural to consider linear space automata, where R^n is replaced by an arbitrary finite or infinite dimensional linear space V over the field R of real numbers, and the final vector is replaced by a linear function from V into R. Finite dimensional linear space automata accept exactly the same languages as finite probabilistic automata.

For any given linear space automaton with a fixed initial vector, we consider the problem, how a minimal linear space automaton generating the same word function can be effectively constructed. Another solution to the same problem has been given by Carlyle and Paz [1], who showed that a word function is generated by a generalized automaton if and only if it is of finite rank; the rank (if finite) is the number of states in the minimal generalized automaton generating the word function. In our method, we define an equivalence relation over the state space and construct a new automaton by using the linear space of equivalence classes. A connection between two minimal linear space automata generating the same word function is established. The minimization problem is considered also in a more general form, where linear space automata do not have any fixed initial vector. For a given automaton A, the task is to construct an automaton A_1 of minimal dimension such that, for any initial vector of A, there exists an equivalent initial vector of A_1, and conversely. The case where A and A_1 have to be probabilistic automata, has been considered by Starke [2].

1. Preliminaries. Let $I = \{x_1, \ldots, x_k\}$ be a finite alphabet and $W(I)$ the set of all words over I, including the empty word λ. The length of a word P is denoted by $\lg(P)$. Let R be the field of real numbers and R^n the real linear space of n-tuples $(a_1, \ldots, a_n)$ where $a_i \in R$. For any linear space V, we use the notations $\dim V$ and O_V for the dimension of V and the zero vector of V, respectively.

Definition. *A linear space automaton* (l.s.a) over the alphabet I is an ordered quadruple

$$A = (V, \{M(x_1), \ldots, M(x_k)\}, \pi, f)$$

where V is a linear space (*state space*) over R, $M(x_i)$ is a linear mapping from V into V $(i = 1, \ldots, k)$, $\pi \in V$ (*initial vector*), and f is a linear function from V into R (*final function*).

If $\dim V$ is finite, then A is called a *finite dimensional* linear space automaton. Let $M(\lambda)$ be the identity mapping in V. For each word $P = y_1 \cdots y_n$ $(y_i \in I)$, we define

$$M(P) = M(y_1) \cdots M(y_n) .$$

Thus, $M(P)$ is a linear mapping in V. The image of any vector $X \in V$ under the mapping $M(P)$ is denoted by $XM(P)$.

Let η be a real number. We say that a linear space automaton A *accepts* the language

$$L(A, \eta) = \{P \in W(I) \mid f(\pi M(P)) > \eta\}$$

with the cut-point η.

An n-state generalized automaton can be rewritten as a linear space automaton when the state set is replaced by R^n and the final vector is considered as the corresponding linear function.

Theorem 1. *If* $A = (V, \{M(x_1), \ldots, M(x_k)\}, \pi, f)$ *is a finite dimensional linear space automaton, then the language* $L(A, \eta)$ *is accepted by a finite probabilistic automaton.*

Proof. Let $\dim V = n$, and let $U = \{U_1, \ldots, U_n\}$ be a basis of V. Let $\hat{M}(x_i)$ be the matrix of $M(x_i)$ with respect to the basis U $(i = 1, \ldots, k)$. Let $\pi = a_1 U_1 + \cdots + a_n U_n$, and define $\hat{\pi} = (a_1, \ldots, a_n)$. Finally, let $\hat{f} = (f(U_1), \ldots, f(U_n))^T$. Then $\hat{A} = (S_n, \{\hat{M}(x_1), \ldots, \hat{M}(x_k)\}, \hat{\pi}, \hat{f})$ is an n-state generalized automaton with the state set S_n. Clearly, for any word P,

$$f(\pi M(P)) = \hat{\pi}\, \hat{M}(P)\, \hat{f} . \tag{1}$$

Hence, $L(A, \eta) = L(\hat{A}, \eta)$ for any η, which implies the theorem (cf. [3]).

2. Theorems concerning minimization. Two linear space automata $A_i = (V_i, \{M_i(x_1), \ldots, M_i(x_k)\}, \pi_i, f_i)$, $i = 1, 2$, are called *equivalent* if, for all $P \in W(I)$,

$$f_1(\pi_1 M_1(P)) = f_2(\pi_2 M_2(P)) .$$

If $\dim V_1$ is finite and if $\dim V_2 \geqq \dim V_1$ for every l.s.a A_2 equivalent to A_1, then we say that A_1 is in *minimal form*. From the equation (1) we see that A and $\hat{A}$ are equivalent.

In what follows, we consider an arbitrary l.s.a $A = (V, \{M(x_1), \ldots, M(x_k)\}, \pi, f)$ and construct from it another l.s.a A_1 equivalent to A.

Let V_1 be the set of all finite linear combinations of vectors $\pi M(P)$, $P \in W(I)$. Then V_1 is a linear space. If $X \in V_1$, then clearly, $XM(P) \in V_1$ for any word $P \in W(I)$. Consequently, $(V_1, \{M(x_1), \ldots, M(x_k)\}, \pi, f)$ is a linear space automaton equivalent to A. But, in general, it is not in minimal form. Now, we define an equivalence relation over V_1 as follows. Let X and Y be arbitrary vectors in V_1. Then, define

$X \equiv Y$ if and only if $f(XM(P)) = f(YM(P))$ for all $P \in W(I)$. Clearly, $\equiv$ is an equivalence relation. Furthermore, it is *right invariant*, i.e., if $X \equiv Y$, then $XM(P) \equiv YM(P)$ for any word $P \in W(I)$. For each $X \in V_1$, we denote by $[X]$ the equivalence class to which X belongs. The sum of two classes and the product of a real number and a class are defined as follows:

$$[X] + [Y] = [X+Y] ,$$

$$a[X] = [aX] .$$

These operations are well-defined, since if $X \equiv X_1$ and $Y \equiv Y_1$, then $X + Y \equiv X_1 + Y_1$ and $aX \equiv aX_1$. Here, we have used the fact that f is a linear function and every $M(x_i)$ is a linear mapping in V_1.

It is easy to check that the equivalence classes form a linear space over R with respect to the above operations. The zero vector is the class $[O_V]$ ($O_{V_1} = O_V$). For example,

$$(ab)[X] = [(ab)X] = [a(bX)] = a[bX] = a(b[X]) ,$$

$$[X] + (-1)[X] = [X + (-1)X] = [O_V] .$$

We denote this linear space by $V_1/\equiv$. For each letter $x \in I$, we now define an operation $M_1(x)$ among the equivalence classes:

$$[X]M_1(x) = [XM(x)] .$$

$M_1(x)$ is well-defined, since if $X \equiv Y$, then $XM(x) \equiv YM(x)$, because $\equiv$ is right invariant. It is easy to verify that

$$(a[X] + b[Y])\, M_1(x) = a([X]M_1(x)) + b([Y]M_1(x)) .$$

Consequently, $M_1(x)$ is a linear mapping from $V_1/\equiv$ into itself. For each class $[X]$, we define

$$f_1([X]) = f(X) .$$

Then, f_1 is well-defined and it is a linear function from $V_1/\equiv$ into R.

By the above considerations,

$$A_1 = (V_1/\equiv, \{M_1(x_1) , \ldots , M_1(x_k)\}, [\pi], f_1)$$

is a linear space automaton. In addition,

$$[\pi]M_1(P) = [\pi M(P)]$$

for any word $P \in W(I)$. Hence, we have, for all $P \in W(I)$,

$$f_1([\pi]\, M_1(P)\,) = f(\pi\, M(P)\,) .$$

Hence, we have established the following result.

Theorem 2. *A linear space automaton* $A = (V, \{M(x_1) , \ldots , M(x_k)\}, \pi, f)$ *is equivalent to the linear space automaton* $A_1 = (V_1/\equiv, \{M_1(x_1) , \ldots , M_1(x_k)\}, [\pi], f_1)$.

Remark. Let $X \in V_1$ be arbitrary. Then $[X]M_1(P) = [XM(P)]$ for all $P \in W(I)$. Thus, any vector $X \in V_1$ is equivalent to the vector $[X] \in V_1/\equiv$, i.e., $f(XM(P)\,) = f_1([X]M_1(P)\,)$ for any word P. Hence, if $V_1 = V$, then, for any vector X of A, there exists an equivalent vector of A_1, and conversely (cf. Theorem 7).

Theorem 3. *A linear space automaton* $A = (V, \{M(x_1) , \ldots , M(x_k)\}, \pi, f)$ *is equivalent to a finite dimensional linear space automaton if and only if* $\dim V_1/\equiv$ *is finite.*

Proof. By Theorem 2, the condition is sufficient. Assume that A is equivalent to the l.s.a

$$B = (W, \{N(x_1) , \ldots , N(x_k)\}, \beta, \varphi)$$

where $\dim W$ is finite. Let A_1 be as in Theorem 2. Assume that $[X_1] , \ldots , [X_u]$ are linearly independent equivalence classes. Since every vector X_i belongs to V_1, it can be written as a finite linear combination of some vectors $\pi M(P)$. Consequently, there exist words $P_1 , \ldots , P_r$ such that, for each $i = 1 , \ldots , u$,

$$X_i = \sum_{j=1}^{r} a_{ij}\pi M(P_j) . \tag{2}$$

Since A and B are equivalent, we have

$$\varphi(\beta N(P_jP)\,) = f(\pi M(P_jP)\,) \tag{3}$$

for each $j = 1, \ldots, r$ and for each word $P \in W(I)$. Using (2) and (3) we find that

$$\begin{aligned} f(X_i M(P)) &= \sum_{j=1}^{r} a_{ij} f(\pi M(P_j P)) \\ &= \sum_{j=1}^{r} a_{ij} \varphi(\beta N(P_j P)) \\ &= \varphi\left(\left(\sum_{j=1}^{r} a_{ij} \beta N(P_j) \right) N(P) \right). \end{aligned}$$

Consequently, for every word P,

$$\varphi(Y_i N(P)) = f(X_i M(P)), \tag{4}$$

where we have denoted

$$Y_i = \sum_{j=1}^{r} a_{ij} \beta N(P_j).$$

Let us show that $Y_1, \ldots, Y_u$ are linearly independent vectors of W. Assume that

$$c_1 Y_1 + \cdots + c_u Y_u = O_W.$$

This implies that, for every word P,

$$\varphi(c_1 Y_1 N(P) + \cdots + c_u Y_u N(P)) = 0.$$

In other words,

$$c_1 \varphi(Y_1 N(P)) + \cdots + c_u \varphi(Y_u N(P)) = 0.$$

Using the equation (4), this can be written in the form

$$f((c_1 X_1 + \cdots + c_u X_u) M(P)) = 0.$$

But this means that

$$[c_1 X_1 + \cdots + c_u X_u] = [O_V],$$

because P was an arbitrary word. Consequently,

$$c_1 [X_1] + \cdots + c_u [X_u] = [O_V].$$

This is possible only if $c_1 = \cdots = c_u = 0$, because the classes $[X_i]$ were linearly independent. Thus, we get the result that $Y_1, \ldots, Y_u$ are linearly independent. This implies that, if $\dim V_1/\equiv$ is infinite, so is $\dim W$, because u can be chosen as large as we want. But this is a contradiction. Our theorem is thus proved.

Corollary. *If* $\dim V_1/\equiv$ *is finite, then the language* $L(A\ ,\eta)$ *is accepted by a finite probabilistic automaton for any cut-point* η.

Theorem 4. *If* $\dim V_1/\equiv$ *is finite, then the above constructed linear space automaton* A_1 *is in minimal form.*

Proof. Let B be as in the proof of Theorem 3. Choose $u = \dim V_1/\equiv$ in this proof. Since $\dim W \geqq u$, we obtain the desired result.

Thus, we have seen that, if A is a finite dimensional l.s.a, then A_1 is equivalent to it and, furthermore, A_1 is in minimal form. A_1 can be converted to a generalized automaton according to the proof of Theorem 1.

3. Construction of a minimal linear space automaton. In this section, we shall show, how the l.s.a A_1 considered above can be constructed effectively from A. We assume that A is rewritten as a generalized automaton

$$A = (V, \{M(x_1), \ldots , M(x_k)\}, \pi, f)$$

where $V = R^n$, each $M(x_i)$ is an $n\times n$ matrix, π is a row vector, and f is written as a linear function. The basis of V_1 can be determined, because it is well-known that it suffices to consider vectors $\pi M(P)$ with $\lg(P) \leqq n-1$, only. Let $Z_1 = \pi M(P_1), \ldots , Z_r = \pi M(P_r)$ be a basis, thus obtained. Hence, $\lg(P_i) \leqq n-1$ $(i = 1, \ldots , r)$. Vectors $[Z_1], \ldots , [Z_r]$ span the space $V_1/\equiv$. We want to form a basis of $V_1/\equiv$ from these vectors. In order to examine linear independence, write the equation

$$a_1[Z_1] + \cdots + a_r[Z_r] = [(0, \ldots , 0)] \tag{5}$$

or, in another form,

$$[a_1Z_1 + \cdots + a_rZ_r] = [(0, \ldots , 0)]\ .$$

This equation holds if and only if, for every word $P \in W(I)$,

$$f(\ (a_1Z_1 + \cdots + a_rZ_r)M(P)\ ) = 0\ . \tag{6}$$

Denote $Z = a_1Z_1 + \cdots + a_rZ_r$. For any word $P \in W(I)$, the vector $ZM(P)$ is a linear combination of vectors $ZM(Q)$ with $\lg(Q) \leqq r-1$. This implies that (6) holds if and only if it holds for every Q with $\lg(Q) \leqq r-1$. Let these Q's be $Q_1, \ldots , Q_s$. Thus, (5) holds if and only if, for each $i = 1, \ldots , s$,

$$a_1f(Z_1M(Q_i)\ ) + \cdots + a_rF(Z_rM(Q_i)\ ) = 0\ . \tag{7}$$

Here, the numbers $a_1, \ldots , a_r$ are unknown, and the coefficients $f(Z_jM(Q_i))$ can be calculated.

Let m be the largest number such that at least one determinant with m rows, formed from the coefficient matrix,

$$\begin{bmatrix} f(Z_1M(Q_1)) & \cdots & f(Z_rM(Q_1)) \\ f(Z_1M(Q_2)) & \cdots & f(Z_rM(Q_2)) \\ \cdot & \cdots & \cdot \\ f(Z_1M(Q_s)) & \cdots & f(Z_rM(Q_s)) \end{bmatrix} \tag{8}$$

is different from zero. (If $m = r$, (7) has only the trivial solution $a_1 = \cdots = a_r = 0$, so that the classes $[Z_i]$ are linearly independent.) Let $Z_{i_1}, \ldots, Z_{i_m}$ be the Z_i-vectors that occur in the columns of this determinant. Then $[Z_{i_1}], \ldots, [Z_{i_m}]$ form a basis of $V_1/\equiv$. Other $[Z_i]$-classes can be determined from the equations (7) in terms of the classes $[Z_{i_j}]$. It may happen that all coefficients of some a_i equal zero in the equations (7). This reveals the fact that $[Z_i]$ equals $[(0, \ldots, 0)]$.

Remark. The matrix (8) is one of the matrices from which the rank of the word function $f(\pi M(P))$ is determined (cf. [1]).

In order to determine the linear mappings $M_1(x_i)$ $(i = 1, \ldots, k)$, it is sufficient to calculate $[Z_{i_t}]M_1(x_i)$ for each t and i. This is done as follows:

$$[Z_{i_t}]M_1(x) = [Z_{i_t}M(x)] = [\sum_{j=1}^{r} b_{tj}Z_j] = \sum_{j=1}^{r} b_{tj}[Z_j] = \sum_{j=1}^{m} a_{tj}[Z_{i_j}] .$$

The coefficients a_{tj} can be calculated for each $x \in I$. The matrix of $M_1(x)$ with respect to the basis $[Z_{i_1}], \ldots, [Z_{i_m}]$ is $\hat{M}_1(x) = (a_{tj})$.
Next, we determine $[\pi]$:

$$[\pi] = \sum_{j=1}^{r} b_j[Z_j] = \sum_{j=1}^{m} a_j[Z_{i_j}] .$$

The row vector corresponding to $[\pi]$ is $\hat{\pi}_1 = (a_1, \ldots, a_m)$. The column vector corresponding to f_1 is

$$\hat{f}_1 = (f(Z_{i_1}), \ldots, f(Z_{i_m}))^T$$

In this manner, we obtain the minimal generalized automaton $\hat{A}_1 = (S_m, \hat{M}_1(x_1), \ldots, \hat{M}_1(x_k), \hat{\pi}_1, \hat{f}_1)$ equivalent to the given generalized automaton A.

Example. Let A be the three-state probabilistic automaton $A = (S_3, \{M(x)\}, (0, 0, 1), (1, 1, 0)^T)$ over $\{x\}$ where

$$M(x) = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 3/4 & 1/4 & 0 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}$$

Here, $V = R^3$ and the final function is determined by the condition $f(\,(a, b, c)\,) = a + b$ for every $(a, b, c) \in R^3$. The basis of V_1 is

$$Z_1 = \pi = (0, 0, 1),$$

$$Z_2 = \pi M(x) = (1/4, 1/2, 1/4),$$

$$Z_3 = \pi M(x^2) = (9/16, 3/8, 1/16),$$

so that $r = 3$ and $V_1 = V = R^3$. The equations (7) are (the denominators have been eliminated)

$$4a_2 + 5a_3 = 0,$$

$$16a_1 + 20a_2 + 21a_3 = 0,$$

$$80a_1 + 84a_2 + 85a_3 = 0.$$

Here, $m = 2$ and, therefore, $\dim V_1/\!\equiv\, = 2$. A solution of the equations is $a_1 = 1$, $a_2 = -5$, $a_3 = 4$. Thus,

$$[Z_1] - 5[Z_2] + 4[Z_3] = [(0, 0, 0)]\,.$$

As a basis of $V_1/\!\equiv$ we can choose, for instance, $[Z_1]$ and $[Z_2]$. Then

$$[Z_1]M_1(x) = [Z_2],$$

$$[Z_2]M_1(x) = [Z_3] = -\tfrac{1}{4}[Z_1] + \tfrac{5}{4}[Z_2].$$

Hence,

$$\hat{M}_1(x) = \begin{bmatrix} 0 & 1 \\ -1/4 & 5/4 \end{bmatrix},\ \hat{\pi}_1 = (1, 0), \hat{f}_1 = (0, 3/4)^T.$$

Consequently, the minimal generalized automaton equivalent to A is $\hat{A}_1 = (S_2, \{\hat{M}_1(x)\}, \hat{\pi}_1, \hat{f}_1)$.

Since $V_1 = R^3$, we know that also the vectors $Z_1 = (1, 0, 0)$, $Z_2 = (0, 1, 0)$, $Z_3 = (0, 0, 1)$ form a basis of V_1. For these vectors, the corresponding calculations are easier, and it is seen that $[Z_2]$ and $[Z_3]$ form a basis of $V_1/\!\equiv$. For the resulting minimal generalized automaton,

$$\hat{M}_1(x) = \begin{bmatrix} 1 & 0 \\ 3/4 & 1/4 \end{bmatrix}.\ \hat{\pi}_1 = (0, 1), \hat{f}_1 = (1, 0)\,.$$

4. On different minimal forms. We prove first the following theorem.

Theorem 5. *For any linear space automaton A over the one-letter alphabet $\{x\}$, there exists an equivalent generalized automaton in minimal form such that the initial vector is $(1, 0, \ldots, 0)$, and the transition matrix is of the form*

$$\begin{bmatrix} 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 1 \\ a_1 & a_2 & a_3 & \cdot & \cdot & \cdot & a_m \end{bmatrix} \tag{9}$$

Proof. We may assume that A is a generalized automaton $(R^m, \{M(x)\}, \pi, f)$ which is already in minimal form. If $[\pi] = [(0, \ldots, 0)]$, we choose the 1×1 matrix (0). Assume that $[\pi] \neq [(0, \ldots, 0)]$. Since A is in minimal form, it follows that $Z_1 = \pi$, $Z_2 = \pi M(x), \ldots, Z_m = \pi M(x^{m-1})$ form a basis of R^m and the classes $[Z_i]$ $(i = 1, \ldots, m)$ form a basis of $R^m/\equiv$. Now we have, for each $i = 1, \ldots, m-1$,

$$\begin{aligned} &[Z_i]M_1(x) = [Z_{i+1}], \\ &[Z_m]M_1(x) = [\pi M(x^m)] = [\sum_{i=1}^{m} a_i Z_i] = \sum_{i=1}^{m} a_i [Z_i]\,. \end{aligned} \tag{10}$$

Note that the numbers a_i are the coefficients of the characteristic polynomial of $M(x) : a_1 + a_2\lambda + \cdots + a_m\lambda^{m-1} - \lambda^m$. From the equations (10) we see that the matrix $\hat{M}_1(x)$ equals the matrix (9). The row vector corresponding to $[\pi]$ is $(1, 0, \ldots, 0)$. Thus, the theorem holds true.

If, for our minimal A, the row sums in $M(x)$ equal 1, then $a_1 + \cdots + a_m = 1$ in the matrix (9), because $M(x)$ satisfies its characteristic equation and the row sums in every $M(x^i)$ equal 1.

For a given linear space automaton, the minimal generalized automaton is not unique. Even in the above method, the result depends on the choice of the basis. In what follows, we establish a connection between different minimal automata equivalent to a given l.s.a.

We say that two generalized automata

$$A = (R^n, \{M(x_1), \ldots, M(x_k)\}, \pi, f)$$
$$B = (R^n, \{N(x_1), \ldots, N(x_k)\}, \beta, \varphi)$$

are *similar*, if there exists a nonsingular $n \times n$ matrix C such that, for each $i = 1, \ldots, k$,

$$\beta = \pi C,\ N(x_i) = C^{-1}M(x_i)C,\ \varphi = C^{-1}f$$

(f and φ are written as column vectors).

It is easy to see that, for different bases, the above method gives minimal forms which are similar. More generally, we have (see also [1])

Theorem 6. *Two generalized automata A and B in minimal form are equivalent if and only if they are similar.*

Proof. The condition is sufficient, since, clearly, $N(P) = C^{-1}M(P)C$ whenever $P \in W(I)$. We use the symbol V instead of R^n. The space V_1 and the relation $\equiv$ for A and B are denoted, respectively, by $V_1(A)$, $\equiv_A$ and $V_1(B)$, $\equiv_B$. Since A and B are in minimal form, it follows that $V_1(A) = V_1(B) = V$ and $\dim V/{\equiv_A} = \dim V/{\equiv_B} = n$. Let

$$\{Z_i = \pi M(P_i) \mid i = 1, \ldots, n\}$$

be a basis of V. Then the classes $[Z_i]_A$ $(i = 1, \ldots, n)$ form a basis of $V/{\equiv_A}$. It is easy to show that

$$\{[U_i]_B \mid U_i = \beta N(P_i), i = 1, \ldots, n\}$$

is a basis of $V/{\equiv_B}$. Let $x \in I$ be arbitrary. Then

$$[Z_i]_A M_1(x) = [Z_i M(x)]_A = \sum_{j=1}^{n} a_{ij}[Z_j]_A ,$$

$$[U_i]_B N_1(x) = [U_i N(x)]_B = \sum_{j=1}^{n} b_{ij}[U_j]_B ,$$

where $i = 1, \ldots, n$. This implies that $\hat{M}_1(x) = (a_{ij})$ and $\hat{N}_1(x) = (b_{ij})$. Consider the set of equations (7) for the vectors $Z_1, \ldots, Z_n$, $Z_i M(x)$, where i is arbitrarily fixed. The numbers a_{ij} $(j = 1, \ldots, n)$ form a solution of the set, i.e.,

$$a_{i1}f(Z_1M(Q_j)) + \cdots + a_{in}f(Z_nM(Q_j)) - f(Z_iM(x)M(Q_j)) = 0 \tag{11}$$

where $j = 1, \ldots, s$. In the same way, the numbers b_{ij} $(j = 1, \ldots, n)$ satisfy the equations

$$b_{i1}\varphi(U_1N(Q_j)) + \cdots + b_{in}\varphi(U_nN(Q_j)) - \varphi(U_iN(x)N(Q_j)) = 0 \tag{12}$$

where $j = 1, \ldots, s$. On the other hand, the coefficients in (11) and (12) are the same, because A and B are equivalent. Consequently, we have $(a_{ij}) = (b_{ij})$, i.e., $\hat{M}_1(x) = \hat{N}_1(x)$. Since $Z_iM(x) = \sum a_{ij}Z_j$ and $U_iN(x) = \sum b_{ij}U_j$, we conclude that

$$\begin{aligned} \hat{M}_1(x) &= CM(x)C^{-1}, \pi = \hat{\pi}_A C , \\ \hat{N}_1(x) &= DN(x)D^{-1}, \beta = \hat{\pi}_B D , \end{aligned} \tag{13}$$

where the rows of C and D are, respectively, $Z_1, \ldots, Z_n$ and $U_1, \ldots, U_n$.

As in connection with (11) and (12), it is verified that the components of $\hat{\pi}_A$ and $\hat{\pi}_B$ are obtained from the same set of equations of the form (7). Hence $\hat{\pi}_A = \hat{\pi}_B$, which implies $\pi C^{-1} = \beta D^{-1}$. Thus, $\beta = \pi C^{-1} D$. Since $\hat{M}_1(x) = \hat{N}_1(x)$, the equations (13) give $N(x) = (C^{-1}D)^{-1} M(x) C^{-1} D$ whenever $x \in I$. Finally, $Cf = D\varphi$, which gives $\varphi = (C^{-1}D)^{-1} f$. Consequently, A and B are similar.

5. Minimization of linear space automata having no fixed initial vector. In this section, we consider finite dimensional linear space automata A having no fixed initial vector. If an initial vector π is chosen for A, then the resulting l.s.a is denoted by (A, π). Two linear space automata A and B are called *strongly equivalent*, if, for every vector π of A, there exists a vector β of B such that (A, π) and (B, β) are equivalent, and conversely. We say that A is in *minimal form*, if the dimension of its state space does not exceed the dimension of the state space of any l.s.a strongly equivalent to A.

Consider an arbitrary l.s.a $A = (V, \{M(x_1), \ldots, M(x_k)\}, f)$, where f is the final function. The equivalence relation $\equiv$, the mappings $M_1(x_i)$ $(i = 1, \ldots, k)$ and the function f_1 are defined as before. The only difference is that $\equiv$ is defined over the whole space V, instead of V_1. In this manner, we obtain the linear space automaton

$$A_1 = (V/\!\equiv, \{M_1(x_1), \ldots, M_1(x_k)\}, f_1)$$

for which the following theorem holds.

Theorem 7. *The linear space automaton A_1 is strongly equivalent to the linear space automaton A. Furthermore, A_1 is in minimal form.*

Proof. A and A_1 are strongly equivalent, since, for any $X \in V$, (A, X) and $(A_1, [X])$ are equivalent. In order to show that A_1 is in minimal form, denote first $u = \dim V/\!\equiv$ and let $B = (W, \{N(x_1), \ldots, N(x_k)\}, \varphi)$ be any l.s.a strongly equivalent to A_1. Then B is strongly equivalent to A, too. Let $[X_1], \ldots, [X_u]$ be linearly independent vectors of $V/\!\equiv$, and let $Z_1, \ldots, Z_n$ be a basis of V. Hence, for each $i = 1, \ldots, u$,

$$X_i = \sum_{j=1}^{n} a_{ij} Z_j \,. \tag{14}$$

For each Z_j $(j = 1, \ldots, n)$, there exists a vector $U_j \in W$ such that (A, Z_j) and (B, U_j) are equivalent, i.e.,

$$f(Z_j M(P)) = \varphi(U_j N(P)) \tag{15}$$

for any word $P \in W(I)$.

We show that the vectors

$$(16) \qquad Y_i = \sum_{j=1}^{n} a_{ij} U_j \quad (i = 1, \ldots, u)$$

are linearly independent vectors of W. Hence, assume that

$$c_1 Y_1 + \cdots + c_u Y_u = O_W .$$

Then, for every word $P \in W(I)$,

$$(17) \qquad c_1 \varphi(Y_1 N(P)) + \cdots + c_u \varphi(Y_u N(P)) = 0 .$$

On the other hand, using the equations (14)—(16), it is verified that $\varphi(Y_i N(P)) = f(X_i M(P))$ $(i = 1, \ldots, u)$. Therefore, (17) gets the form

$$c_1 f(X_1 M(P)) + \cdots + c_u f(X_u M(P)) = 0,$$

which gives

$$f((c_1 X_1 + \cdots + c_u X_u) M(P)) = 0$$

for every word $P \in W(I)$. Consequently,

$$[c_1 X_1 + \cdots + c_u X_u] = [O_V]$$

In other words, we have

$$c_1 [X_1] + \cdots + c_u [X_u] = [O_V] ,$$

which is possible only if $c_1 = \cdots = c_u = 0$. This implies that $Y_1, \ldots, Y_u$ are linearly independent. Thus, $\dim W \geqq u = \dim V/\equiv$. The proof is now complete.

If in our previous example the initial vectors are omitted from A and $\hat{A}_1$, then the resulting linear space automata are strongly equivalent, because $V_1 = V$.

Note that $V/\equiv$ does not contain any equivalent vectors. Namely, if $f_1([X] M_1(P)) = f_1([Y] M_1(P))$ for all $P \in W(I)$, then, by the definitions, $f(XM(P)) = f(YM(P))$ which says that $[X] = [Y]$. Using the terminology of Starke [2], this means that A_1 is strongly reduced (stark reduziert). Thus, our theorem says that every l.s.a (probabilistic automaton) A has a strongly reduced l.s.a (generalized automaton) which is equivalent to A. As Starke [2] has shown, this does not hold, if A and A_1 have to be probabilistic automata.

Department of Mathematics
University of Oulu, Finland

References

[1] Carlyle, J. W. and A. Paz: Realizations by stochastic finite automata - J. Comput. System Sci. 5 (1971), 26–40.

[2] Starke, P. H.: Über die Minimalisierung von stochastischen Rabin-Automaten - Elektron. Informationsverarbeit. Kybernetik 5 (1969), 153–170.

[3] Turakainen, P.: Generalized automata and stochastic languages - Proc. Amer. Math. Soc. 21 (1969), 303–309.

Printed April 1972