INFORMATION AND CONTROL 24, 201–211 (1974)

The Dimension of Stability of Stochastic Automata

WOJCIECH RYTTER

Warsaw University, Institute of Mathematical Machines, Warsaw, Palac Kultury i Nauki VIII p., Poland

The concept of the dimension of stability of stochastic automata is introduced and examined. We are concerned with the minimal-state form of the automaton and the dimension of stability. Natural automata are defined, for which the dimension of stability is not less than the difference between the number of states and the number of states of its minimal-state equivalent. For minimal state automata, the upper bound on the dimension of stability is given, and we demonstrate minimal-state automata for which the dimension of stability reaches this upper bound. An estimate of the dimension of stability of automata that are not necessarily minimal state is based on this result for minimal-state automata.

1. INTRODUCTION

The stochastic automaton which is considered in this paper is the generalization of the deterministic automaton of Rabin's type. The internal structure of an n-state stochastic automaton is given by a set of stochastic square matrices or, equivalently, by the set of linear transformations of the set T^n into T^n, where T^n is the set of all possible probability distributions of states of the n-state automaton. The output behavior of the stochastic automaton is given by the output function.

By stability we denote the ability of stochastic automaton changing its internal structure without a change of the output behavior. We introduce the idea of the set of stable perturbations which characterizes stability of automata. Since stable perturbations are vectors in ordinary are Euclidean space, we can introduce the concept of the dimension of stability as a dimension of linear subspace generated by stable perturbations.

2. NOTIONS AND NOTATIONS

Let $\Sigma = \{\sigma_1, \sigma_2, ..., \sigma_n\}$ and let Σ^* be the set of all finite sequences of elements in Σ, the empty sequence included. The elements of Σ^* are said to be the words under the alphabet Σ and denoted by u, v, w. If ξ is a vector then ξ_i denotes its ith entry.

DEFINITION 1. The set $T^n = \{\xi \in E^n: \sum_{i=1}^{n} \xi_i = 1; \xi_i \geqslant 0 \text{ for } i = 1,..., n$ is said to be the stochastic simplex and its elements are stochastic vectors. The matrix is called stochastic whenever all its rows are stochastic vectors.

DEFINITION 2. The n-state stochastic automaton A (abbreviated by s.a. A) is the pair $\langle M_A, \eta_A \rangle$, where M_A is a set of $n \times m$ stochastic matrices indexed by the elements of Σ, such that to every $\sigma \in \Sigma$ there corresponds exactly one matrix in M_A indexed by σ. By $A(\sigma)$ we shall mean the matrix belonging to M_A, indexed by σ.

η_A is a n-dimensional column-vector whose entries equal 1 or 0.

The n-dimensional stochastic vector whose ith entry equals 1 and other entries 0 is denoted by s_i^n and is called the ith state of the s.a. A. T^n is called the set of distributions of states of the s.a. A.

For $u = \sigma_1 \cdots \sigma_m$ where $\sigma_i \in \Sigma$ for $i = 1,..., m$, we define

$$A(u) = A(\sigma_1) \cdots A(\sigma_m).$$

$A(\Lambda)$ is the unit matrix. δ_A is the transition function and λ_A is the output function for an s.a. A where

$$\delta_A: T^n \times \Sigma^* \to T^n; \qquad \delta_A(\xi, u) = \xi A(u);$$
$$\lambda_A: T^n \times \Sigma^* \to E^1; \qquad \lambda_A(\xi, u) = \xi A(u)\eta_A = \delta_A(\xi, u)\eta_A;$$

for $\xi \in T^n, u \in \Sigma^*$.

DEFINITION 3. Two s.a. A, B are strongly undistinguishable (denoted by $A \equiv B$) if the following conditions simultaneously hold. (1) A, B have the same number of states; (2) $\eta_A = \eta_B$; (3) for every $u \in \Sigma^*$, $\xi \in T^n$, $\lambda_A(\xi, u) = \lambda_B(\xi, u)$.

DEFINITION 4. Let the s.a. A be n-state, $A(\sigma, i)$ denoting the ith row of the matrix $A(\sigma)$, $E_0^n = \{\xi \in E^n: \sum_{i=1}^{n} \xi_i = 0\}$. We say that the vector $\xi \in E_0^n$ is a perturbation of the s.a. A if there exist $\sigma \in \Sigma$ and $1 \leqslant i \leqslant n$ such that $A(\sigma, i) + \xi$ is stochastic. In this case, we construct the automaton $[\sigma, i, \xi]A$ by replacing the ith row in the matrix $A(\sigma)$ by $A(\sigma, i) + \xi$. If A and this constructed automaton are strongly undistinguishable, we say that ξ is a stable perturbation of s.a. A for the ith row of the matrix $A(\sigma)$. In the next part, we shall see that the stable perturbation for ith row of $A(\sigma)$ is also the stable perturbation for all jth rows of $A(\sigma')$ such that $A(\sigma', j) + \xi$ is stochastic, so we define the stable perturbation of the s.a. A as that vector $\xi \in E_0^n$ for which there exist arbitrary $\sigma \in \Sigma$ and $1 \leqslant i \leqslant n$ such that $A(\sigma, i) + \xi$ is stochastic and $[\sigma, i, \xi]A \equiv A$.

Let P_A be the linear subspace generated by all perturbations of s.a. A, Z_A be the linear subspace generated by all stable perturbations of s.a. A.

DEFINITION 5. The dimension of stability of s.a. A (abbreviated by dim st A) equals the dimension of Z_A.

3. RELATIONS BETWEEN STABLE PERTURBATIONS AND VANISHING DIRECTIONS

DEFINITION 6. Vector $\xi \in E_0^n$ is a vanishing direction of an n-state s.a. A iff $\xi \cdot A(u) \cdot \eta_A = 0$ for all $u \in \Sigma^*$. Note that $\xi \cdot A(\Lambda) \cdot \eta_A = 0$ if and only if $\xi \cdot \eta_A = 0$. F_A denotes the set of all vanish direction of s.a. A, F_A is linear subspace of E_0^n.

Let C be an $n \times m$ matrix. It is obvious that $C \cdot A(u) \cdot \eta_A = \bar{0}$ for all $u \in \Sigma^*$ iff all rows of C are vanishing directions of s.a. A.

LEMMA 1. *Let $\eta_A = \eta_B$ and s.a. A, B be n-state. $A \equiv B$ iff for every $\sigma \in \Sigma$ and $1 \leqslant i \leqslant n$, $B(\sigma, i) - A(\sigma, i)$ is a vanishing direction of s.a. A.*

Proof of sufficiency. Let $(B(\sigma, i) - A(\sigma, i)) \in F_A$ for all $1 \leqslant i \leqslant n$. Denote $B(\sigma) - A(\sigma)$ by $C(\sigma)$. Let $u \in \Sigma^*$, $u = \sigma_1, \ldots, \sigma_m$ where $\sigma_i \in \Sigma$ for $i = 1, \ldots, m$. $C(\sigma_i)\, A(w) \cdot \eta_A = 0$ for all $w \in \Sigma^*$. Hence

$$\begin{aligned} B(u) \cdot \eta_A - A(u) \cdot \eta_A = {} & A(\sigma_1) \cdots A(\sigma_{m-1})\, C(\sigma_m) \cdot \eta_A \\ & + A(\sigma_1) \cdot \cdots A(\sigma_{m-2})\, C(\sigma_{m-1})\, A(\sigma_m) \cdot \eta_A \\ & + \cdots C(\sigma_1) \cdots C(\sigma_m) \cdot \eta_A = 0. \end{aligned}$$

Hence $\lambda_A \equiv \lambda_B$ and $A \equiv B$.

Proof of necessity. $A \equiv B$ implies $(B(w) - A(w)) \cdot \eta_A = 0$ for all $w \in \Sigma^*$. Hence $(B(\sigma, i) - A(\sigma, i)) \cdot A(\Lambda) \cdot \eta_A = 0$ for $\sigma \in \Sigma$, $1 \leqslant i \leqslant n$. By induction on the length of u, we can simply obtain that $(B(\sigma, i) - A(\sigma, i))\, A(u) \cdot \eta_A = 0$ for $u \in \Sigma^*$. Hence $B(\sigma, i) - A(\sigma, i)$ is a vanishing direction for $\sigma \in \Sigma$, $1 \leqslant i \leqslant n$. This completes the proof.

Lemma 1 and Definition 4 imply:

THEOREM 1. *Vector* $\xi \in E_0^n$ *is a stable perturbation of an n-state s.a. A iff* $\xi \in F_A$ *and* ξ *is a perturbation of an s.a. A.* $A \equiv B$ *iff for every* $\sigma \in \Sigma$, $1 \leqslant i \leqslant n$, $B(\sigma, i) - A(\sigma, i)$ *is a stable perturbation of the s.a. A.*

$$Z_A = F_A \cap P_A .$$

COROLLARY 1. *A stable perturbation* ξ *for ith row of the matrix* $A(\sigma)$ *is also a stable pertubation for all jth rows of* $A(\sigma')$ *such that* $A(\sigma', j) + \xi$ *is stochastic, where* $\sigma, \sigma' \in \Sigma$. *Let an s.a. A, B be n-state and* $\eta_A = \eta_B$. *For number* $a \in [0, 1]$ *we construct the n-state automaton* $aA + (1 - a)B = C$, *where* $\eta_C = \eta_A \cdot C(\sigma) = aA(\sigma) + (1 - a) B(\sigma)$ *for all* $\sigma \in \Sigma$.

THEOREM 2. $A \equiv B \Rightarrow B \equiv aA + (1 - a)B \equiv A$.

Proof. Let $C = aA + (1 - a)B$. $C(\sigma, i) - A(\sigma, i) = (1 - a)(B(\sigma, i) - A(\sigma, i))$ for $1 \leqslant i \leqslant n$, $\sigma \in \Sigma$, is a vanishing direction.
Hence, Lemma 1 implies $C \equiv A$. It completes the proof.

4. RELATIONS BETWEEN STABILITY AND MINIMIZATION

DEFINITION 7. Let an s.a. A, B be n-state and m-state, respectively: $\xi \in T^n$, $\gamma \in T^m$. We say that ξ and γ are stochastic equivalent (denoted by $\xi \equiv^{A,B} \gamma$) iff $\lambda_A(\xi, u) = \lambda_B(\gamma, u)$ for every $u \in \Sigma^*$. For $A = B$ we write $\xi \equiv \gamma$ instead of $\xi \equiv^{A,A} \gamma$ s.a. A, B are stochastic equivalent (denoted by $A \approx B$) iff for every $\xi \in T^n$ there exists $\gamma \in T^m$ such that $\gamma \equiv^{B,A} \xi$ and for every $\gamma \in T^m$ there exists $\xi \in T^n$ such that $\xi \equiv^{A,B} \gamma$. The n-state s.a. A is minimal state iff for every m-state s.a. B $B \approx A$ implies $m \geqslant n$. Relations $\approx$ and $\equiv$ are equivalence relations.

DEFINITION 8.

$$I_d = \{a \cdot x : a \in R, x \in M, a \neq 0\}$$

where $M = \{\xi \in E_0^n$: for some $i \in \{1, 2, ..., n\}$, $\xi_i = -1$ and $\xi_k \geqslant 0$ for all $k \neq i\}$.

The elements of I_d are called internal directions.

It is shown in Paz (1968) that s.a. A is minimal state iff none of states is stochastically equivalent to a convex combination of another states.

Let ξ_1, ξ_2 be distributions of states of s.a. A. $\xi_1 \equiv \xi_2$ iff $\xi_1 - \xi_2$ is a vanishing direction of s.a. A.

If γ is a convex combination of states $s_1{}^m,..., s_{i-1}^m, s_{i+1}^m,..., s_m{}^m$, then $\gamma - s_i{}^m$ is an internal direction. It implies the following.

THEOREM 3. *s.a. A is minimal-state iff $F_A \cap I_d = \varnothing$.*

LEMMA 2. *If n-state s.a. A and m-state s.a. B are stochastic equivalent and s.a. B is minimal state, then there exist $(n - m)$ linearly independent internal directions which are vanishing directions of s.a. A.*

Proof. Let B be the set of all distributions of states of s.a. B stochastically equivalent to any $s_1{}^n,..., s_n{}^n$. We shall prove that $\{s_1{}^m,..., s_m{}^m\} \subset G$. Assume that $s_i{}^m \notin G$ and $1 \leqslant i \leqslant m$. There exists $\xi \in T^n$ such that $\xi \equiv^{A,B} s_i{}^m$ ξ is a convex combination of $s_1{}^n,..., s_n{}^n$. $\xi = \sum_{j=1}^n x_j \cdot s_j{}^n$. For every $1 \leqslant j \leqslant n$ there exists $f(s_j{}^n) \in T^m$ such that $f(s_j{}^n) \equiv^{B,A} s_j{}^n$. Denote $\gamma = \sum_{j=1}^n x_j \cdot f(s_j{}^n)$. For $1 \leqslant j \leqslant n$ $f(s_j{}^n) \neq s_i{}^m$ so $\gamma \neq s_i{}^m$ because $s_i{}^m$ is the vertex of T^m. $\xi \equiv^{A,B} \gamma$ implies $s_i{}^m \equiv \gamma$ and $(s_i{}^m - \gamma) \in F_B$, $s_i{}^m - \gamma$ is an internal direction, hence B is not minimal state.

This contradiction implies $\{s_1{}^m,..., s_m{}^m\} \subset G$. Hence there exist integers $1 \leqslant i_1,..., i_m \leqslant n$ such that $s_{i1}^n,..., s_{im}^n$ are stochastically equivalent to $s_1{}^m,..., s_m{}^m$, respectively. Choose $n - m$ different states $s_{j_1}^n,..., s_{j_{n-m}}^n$ not belonging to $\{s_{i_1}^n,..., s_{i_m}^n\}$. For every $s_{j_k}^n$ there exists $\xi_k \in T^m$ such that $\xi_k \equiv^{B,A} s_{j_k}^n$. ξ_k are convex combinations of states $s_1{}^m,..., s_m{}^m$;

$$\xi_k = \sum_{l=1}^{m} x_{k,l} \cdot s_l{}^m.$$

Hence $\sum_{l=1}^m x_{k,1} \cdot s_{i_l}^n \equiv s_{j_k}^n$ for $k = 1,..., (n - m)$. The vectors

$$\gamma_k = s_{j_k}^n - \sum_{l=1}^{m} x_{k,l} \cdot s_{i_l}^n$$

are linear independent internal directions and vanish directions of s.a. A. The proof is complete.

DEFINITION 9. n-state s.a. A is natural iff there is no $1 \leqslant i \leqslant n$ such that for every $\sigma \in \Sigma$ ith column of the matrix $A(\sigma)$ is zero.

LEMMA 3. *If $\xi \in I_d$ then $\xi \in P_A$ for n-state natural s.a. A.*

Proof. $\xi = (a_1,..., a_{i-1}, -1, a_{i+1},..., a_n)$, where $a_j \geqslant 0$ for $j = 1,..., n$; $j \neq i$. There exist $\sigma \in \Sigma$ and $1 \leqslant k \leqslant n$ such that ith entry of $A(\sigma, k)$ is nonzero. Let this entry equal b_i, then $b_i \cdot \xi$ is such that $A(\sigma, k) + b_i \cdot \xi$ is stochastic, so $b_i \xi \in P_A$. Hence $\xi \in P_A$. This completes the proof.

Lemmas 2 and 3 and Theorem 1 imply the following.

THEOREM 4. *Let s.a. A be natural and n-state. If s.a. A is stochastic and equivalent to m-state minimal-state s.a. B, then* dim st. $A \geqslant n - m$. *If A is not minimal-state then* dim st $A > 0$.

DEFINITION 10. s.a. A is deterministic iff entries of matrices $A(\sigma)$ equal 0 or 1.

THEOREM 5. *If s.a. A is natural and deterministic then A is stable iff A is not minimal state.*

Proof of sufficiency follows from Theorem 4.

Proof of necessity. Let ξ be a stable perturbation of s.a. A $\xi \in F_A$. There exist $\sigma \in \Sigma$ and $1 \leqslant i \leqslant n$ such that $A(\sigma, i) + \xi$ is stochastic. It implies that ξ is of the form $(1_1, ..., a_{i-1}, -a_i, ..., a_n)$ where all a_j are nonnegative and $a_i \neq 0$. Hence $1/a_i \cdot \xi$ is an internal direction belonging to F_A and it implies that A is not minimal state.

THEOREM 6. *If n-state s.a. A is natural and $n \leqslant 3$, then A is stable iff A is not minimal state.*

Proof. Note that if nonzero $\xi \in E_0{}^n$ and $n \leqslant 3$, then ξ is an internal direction. The further proof is analogous to the proof of Theorem 5.

In the following example, the automaton is not natural, not minimal-state and not stable. Let

$$\Sigma = \{\sigma\}, \qquad \eta_A = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad A(\sigma) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$(0, 1, -1)$ is the internal direction belonging to F_A, hence A is not a minimal state. Assume now that A is stable, so there exists a stochastic perturbation $\xi = (\xi_1, \xi_2, \xi_3) \in E_0{}^3$, $\xi \neq \bar{0}$ such that $\xi \cdot \eta_A = 0$ and $\xi_1 + \xi_2 + \xi_3 = 0$, hence $\xi_2 = -\xi_3 \neq 0$. Hence, the vector $A(\sigma, i) + \xi$ is not stochastic for $1 \leqslant i \leqslant n$ because one of its entries is negative. It follows that A is not stable.

THEOREM 7. *If an s.ami.nimA is al-state and n-state then* dim st $A \leqslant n - 3$.

Proof. Theorem 1 implies $\dim F_A \geqslant \dim \operatorname{st} A$ and Theorem 3 implies that no internal direction belongs to F_A.

We prove that $\dim F_A \leqslant n - 3$. Assume that there exist $(n - 2)$ linearly independent vectors $\xi_1, \ldots, \xi_{n-2}$ belonging to F_A. For $1 \leqslant i \leqslant n - 2$ let ξ_i' denotes $(n - 3)$-dimensional vector obtained from ξ_i by rejecting three last entries. $\xi_1', \ldots, \xi_{n-2}'$ are not linearly independent in E^{n-3} and there exists a nonzero vector $(x_1, \ldots, x_{n-2})$ such that $\sum_{i=1}^{n-2} x_i \cdot \xi_i' = \bar{0}$. Denote

$$\xi = \sum_{i=1}^{n-2} x_i \cdot \xi_i .$$

This vector is nonzero because $\xi_1, \ldots, \xi_{n-2}$ are linearly independent. Only three last entries y_1, y_2, and y_3 of ξ are nonzero, hence the vector $(y_1, y_2, y_3) \in E_0^3$ and it implies that (y_1, y_2, y_3) is an internal direction, hence ξ is an internal direction belonging to F_A. This contradiction completes the proof.

DEFINITION 11. s.a. B is complementary to s.a. A iff s.a. A, B have the same number of states, the same transition matrices and $\eta_B = \bar{1} - \eta_A$ where $\bar{1}$ is the column-vector with all entries equal to 1. Note that every vector $\xi \in E_0^n$ satisfies $\xi \cdot \bar{1} = 0$. Hence, Theorem 1 simply implies the following:

LEMMA 4. *If ξ is a stable perturbation of s.a. A and B is complementary to A then ξ is a stable perturbation of s.a. B.*

THEOREM 8. *If an n-state s.a. A is minimal state and $n \geqslant 5$, then* $\dim \operatorname{st} A \leqslant n - 4$.

Proof. Let η_A have no less than 3 zero entries. If this is not true, we can analogously consider the automaton complementary to A which satisfies this condition, and the theorem will then follow Lemma 6. Let the three last entries be zero. There exist $1 \leqslant i \leqslant n - 3$ such that ith entry of η_A is nonzero, because A is minimal-state. For simplicity, let it be for $i = 1$. It is sufficient to prove that $\dim F_A \leqslant n - 4$. Assume that $n - 3$ linearly independent vectors $\xi_1, \ldots, \xi_{n-3}$ belong to F_A. Analogously to the proof of Theorem 8 we form vectors $\xi_1', \ldots, \xi_{n-3}'$. If $\xi_1', \ldots, \xi_{n-3}'$ are not linearly then analogously to the proof of Theorem 8, we show that there is an internal direction belonging to F_A, hence these vectors are linearly independent and form the base of E^{n-3}. The vector s_1^{n-3} is a linear combination of $\xi_1', \ldots, \xi_{n-3}'$: $s_1^{n-3} = \sum_{i=1}^{n-3} x_i \cdot \xi_i'$. Consider the vector $\gamma = \sum_{i=1}^{n-3} x_i \cdot \xi_i$; $\gamma \in F_A$ but $\gamma \cdot \eta_A = 1$. This contradiction completes the proof.

Let C_n denote the set of all linear subspaces H of E_0^n for which there exists a zero-one column-vector η, $\eta \neq \bar{0}$, $\eta \neq \bar{1}$ such that $H \subset \ker \eta$.

LEMMA 5. *If $H \in C_n$ then there exists n-state s.a. A such that $F_A = Z_A = H$.*

Proof. Let H be a linear subspace of E_0^n and $H \subset \ker \eta$, where η is zero-one column-vector such that $\eta \neq \bar{0}$, $\eta \neq \bar{1}$ and let $\xi_1, \xi_2, \ldots, \xi_k$ be an ortonormal system of vectors generating H. Let $\xi_1, \xi_2, \ldots, \xi_k, \gamma_1, \gamma_2, \ldots, \gamma_l$, where $k + l = n - 1$, be an ortonormal system of vectors generating E_0^n. We have that E_0^n is not a subspace of $\ker \eta$, because $\eta \neq \bar{1}$, $\eta \neq \bar{0}$, and therefore there exists $1 \leqslant i \leqslant l$ such that $\gamma_i \cdot \eta \neq 0$, for simplicity let it be for $i = 1$, so that $\gamma_1 \cdot \eta \neq 0$. Let hyp T^n be the $(n - 1)$-dimensional hyperplane containing T^n. hyp $T^n = \{x \in E^n : \sum_{i=1}^n x_i = 1\}$. hyp T^n is the result of the translation of hyperplane E_0^n by the vector $\beta = (1/n)\bar{1}$. Hence every point ξ of hyp T^n is of the form

$$\xi = \sum_{i=1}^{k} x_i \cdot \xi_i + \sum_{j=1}^{l} z_j \cdot \gamma_j + \beta,$$

where x_i, z_j are real numbers. Let us consider a position of the simplex T^n on hyp T^n in the topology of hyp T^n. β is an interior point of T^n (we can consider a vector as a point). There exist $(n - 1)$-dimensional spheres K_1 and K_2 in hyp T^n with center at β such that $K_1 \subset T^n$, $K_2 \supset T^n$ because the diameter of T^n is finite. Let r, R denote radii of K_1, K_2 respectively.

$$\left(\sum_{j=1}^{l} z_j\right)^2 \leqslant l^2 \cdot \sum_{j=1}^{l} z_j^2$$

holds for every numbers $z_1, z_2, \ldots, z_l$. Let Q be a linear transformation $Q: E^n \to E^n$ such that $Q(\beta) = \beta$, $Q(\xi_i) = \mu \cdot \xi_i$ for $1 \leqslant i \leqslant k$ and $Q(\gamma_j) = (\mu/l)\gamma_1$ for $1 \leqslant j \leqslant l$, where $\mu = r/2R$.

Such a linear transformation exists and is uniquely determined because $\xi_1, \ldots, \xi_k, \gamma_1, \ldots, \gamma_l, \beta$, form the base of E^n.

We shall prove that for $\xi \in T^n$, $Q(\xi)$ is an interior point of T^n. Let $\xi \in T^n$ then exist $x_1, \ldots, x_k, z_1, \ldots, z_l$ such that

$$\xi = \sum_{i=1}^{k} x_i \cdot \xi_i + \sum_{j=1}^{l} z_j \cdot \gamma_j + \beta$$

and consequently

$$Q(\xi) = \sum_{i=1}^{k} x_i \cdot \mu \cdot \xi_i + \sum_{j=1}^{l} \frac{\mu}{l} \cdot z_j \cdot \gamma_1 + \beta,$$

hence $Q(\xi) \in \text{hyp } T^n$ as well as

$$|Q(\xi) - \beta| = \mu \cdot \sqrt{\sum_{i=1}^{k} x_i^2 + \left(\sum_{j=1}^{l}\left(\frac{z_j}{l}\right)\right)^2} \leqslant \mu \cdot \sqrt{\sum_{i=1}^{k} x_i^2 + \sum_{j=1}^{l} z_j^2}$$

$$= \mu \cdot |\xi - \beta| \leqslant \mu \cdot R = \tfrac{1}{2}r.$$

$|Q(\xi) - \beta| \leqslant (1/2)r$ implies that $Q(\xi)$ is an interior point of T^n in hyp T^n, so that $Q(\xi)$ is a stochastic vector with all entries nonzero.

Let us take the matrix $\bar{Q}$ of the linear transformation Q. ith row of $\bar{Q}$ equals $Q(s_i^n)$ and hence $\bar{Q}$ is stochastic matrix with all entries nonzero.

We construct the automaton $A = \langle M_A, \eta_A \rangle$ for $\Sigma = \{\sigma\}$ $A(\sigma) = \bar{Q}$, $\eta_A = \eta$. The construction of Q implies that $Q(H) \subset H \subset \ker \eta_A$ and for $\gamma \in (E_0^n - H)$ we have $Q(\gamma) \notin \ker \eta_A$. Hence our constructed automaton satisfies $Z_A = F_A = H$. This completes the proof.

THEOREM 9. *There exists a* 4-*state automaton* A *such that* dim st $A = 1$.

Proof. Let H be the one-dimensional subspace generated by the vector $\xi = (1, 1, -1, -1)$. $H \subset E_0^4$, $H \in C_4$, H contains no internal direction. Hence our thesis follows from Lemma 5.

LEMMA 6. *For every* $n \geqslant 4$ *there exists a linear subspace* $G_n \subset E_0^n$ *containing no internal direction, such that* dim $G_n = n - 3$.

Proof.

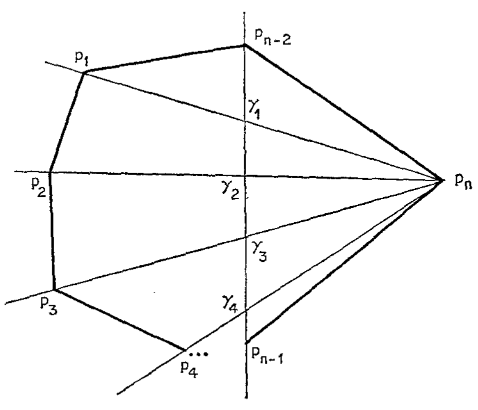

FIG. 1. Construction of the points $\gamma_1, \gamma_2, \ldots, \gamma_{n-3}$.

Denote the straight line connecting the points p_i, p_j by $\overline{p_i p_j}$ and the open segment of the line $\overline{p_i p_j}$ between the points p_i and p_j by $(p_i p_j)$.

Let $p_1, p_2, \ldots, p_n$ be convex independent points in E^2, such that the sets $\{p_1, p_2, \ldots, p_{n-3}\}$ and $\{p_n\}$ are situated on the different sides of the straight line $\overline{p_{n-1} p_{n-2}}$ (see Fig. 1). Such points do exist. For $1 \leqslant i \leqslant n-3$ denote the intersection point of the lines $\overline{p_i p_n}$ and $\overline{p_{n-2} p_{n-1}}$ by γ_i.

From convex independence of the points $p_1, p_2, \ldots, p_n$ it follows that for $1 \leqslant i \leqslant n-3$, $\gamma_i \in (p_{n-2}, p_{n-1})$, so that there exists a number $x_i > 0$ such that $\gamma_i = x_i \cdot p_{n-2} + (1 - x_i) \cdot p_{n-1}$, simultaneously $\gamma_i \in (p_i, p_n)$ so that for some $z_i > 0$ we have $\gamma_i = z_i \cdot p_i + (1 - z_i) \cdot p_n$. For $1 \leqslant i \leqslant n-3$ let us construct the n-dimensional vector ξ^i such that $\xi_i^i = z_i$, $\xi_n^i = (1 - z_i)$, $\xi_{n-1}^i = -(1 - x_i)$, $\xi_{n-2}^i = -x_i$ and $\xi_j^i = 0$ for $j \notin \{i, n-2, n-1, n\}$. The vector ξ^i is of the form $\xi^i = (..0.., z_i, ..0.., -x_i, -(1 - x_i), 1 - z_i)$ where $z_i, x_i > 0$ and z_i is on the ith place. It is obvious that the system of vectors $\xi^1, \xi^2, \ldots, \xi^{n-3}$ is linearly independent and $\xi^i \in E_0^n$ for $1 \leqslant i \leqslant n-3$. Let us take the linear subspace G_n generated by the vectors $\xi^1, \ldots, \xi^{n-3}$ then $\dim G_n = n - 3$ and $G_n \subset E_0^n$.

In order to complete the proof it is sufficient now to prove that G_n contains no internal direction. Let us consider the following set

$$V = \left\{\alpha \in E^n : \sum_{j=1}^{n} \alpha_j \cdot p_j = \bar{0}\right\}.$$

It can be simply verified that V is a linear subspace. For every $1 \leqslant i \leqslant n-3$ we have $\xi^i \in V$ because

$$\sum_{j=1}^{n} \xi_j^i \cdot p_j = z_i \cdot p_i + (1 - z_i) \cdot p_n - x_i \cdot p_{n-2} - (1 - x_i) \cdot p_{n-1}$$

$$= \gamma - \gamma = \bar{0}.$$

Hence $G_n \subset V$. We prove that V contains no internal direction. Assume that $I_d \cap V \neq \varnothing$, so that for some $\epsilon \in M$ and $a \neq 0$, $a \in R$, $a\epsilon \in V$, as well as $\epsilon \in V$. The vector ϵ is of the form $\epsilon = (\epsilon_1, \ldots, \epsilon_{k-1}, -1, \epsilon_{k+1}, \ldots, \epsilon_n)$, $\epsilon_k = -1$ for some $1 \leqslant k \leqslant n$, where $\epsilon_i \geqslant 0$ for $i \neq k$ $1 \leqslant i \leqslant n$, $i \neq k$ and $\sum_{i \neq k} \epsilon_i = 1$.

If $\epsilon \in V$, then $\sum_{i \neq k} \epsilon_i \cdot p_i - p_k = \bar{0}$ and $p_k = \sum_{i \neq k} \epsilon_i \cdot p_i$ so the point p_k is a convex combination of the points $p_1, p_2, \ldots, p_{k-1}, p_{k+1}, \ldots, p_n$. This contradiction proves that $V \cap I_d = \varnothing$.

Hence $G_n \cap I_d = \varnothing$ and G_n satisfies required conditions. This completes the proof.

For $n \geqslant 5$ let H be the set of all vectors of the form $(\xi, 0)$ where $\xi \in G_{n-1}$. $H \in C_n$ and $\dim H = n - 4$. H contains no internal direction. Hence Theorem 9 implies the following.

THEOREM 10. *For every $n \geqslant 5$, there exists an n-state minimal-state s.a. A such that* $\dim \operatorname{st} A = n - 4$.

Using estimations of the dimension of stability of minimal-state automata we can give similar estimations for automata which are not necessarily minimal-state. For an s.a. A let $\min A$ denote the number of states of the minimal-state automaton B which is stochastically equivalent A. This $\min A$ is well defined, because all minimal-state automata stochastically equivalent to A have the same number of states.

THEOREM 11. *If* s.a. *A has n states, then*

(1) *if* $\min A \geqslant 5$ *then* $\dim \operatorname{st} A \leqslant n - 4$

(2) *if* $\min A = 4$ *then* $\dim \operatorname{st} A \leqslant 1$.

Proof. We prove (1). Let m-state minimal-state s.a. B be stochastically equivalent A. Assume that $m \geqslant 5$. It was shown in the proof of Lemma 2 that for every $1 \leqslant k \leqslant m$ there exists $1 \leqslant i_k \leqslant n$ such that $s_k{}^m \equiv^{B,A} s_{i_k}^n$ and $i_k \neq i_{k'}$ for $k \neq k'$. Take the linear transformation $Q: E^n \to E^m$ such that $Q(s_{i_k}^n) = s_k{}^m$ for $1 \leqslant k \leqslant m$, and for other states $s_l{}^n$ of s.a. A, let $Q(s_l{}^n)$ equal arbitrary $\xi_l \in T^m$ such that $s_l{}^n \equiv^{A,B} \xi_l$. This construction is possible because $A \approx B$. B is a minimal-state and the $s_i{}^n$ $\{s_i{}^n: 1 \leqslant i \leqslant n\}$ form the basis of E^n. It can be simply verified that $Q(F_A) \subset F_B$. Hence $Q(Z_A) \subset F_B$. Denote $Q(Z_A) = H$. Assume that $\dim \operatorname{st} A > n - 4$, so $\dim H > (n - 4 - n - m)$ and $\dim H > m - 4$ because the dimension of $\ker Q$ equals $n - m$. Hence $\dim F_B > m - 4$. But in the proof of Theorem 8 it was shown that for an m-state, minimal-state s.a. B $\dim F_B \leqslant n - 4$ if $m \geqslant 5$. This contradiction completes the proof of (1). The proof of (2) is analogous.

RECEIVED: May 15, 1972

REFERENCES

PAGE, C. (1969), Strong stability problems for probabilistic sequential machines, *Information and Control* **15**, 487.

PAZ, A. (1968), Homomorphisms between stochastic sequential machines and related problems, *Math. Systems Theory* **2**, No. 3.