ANNALES ACADEMIAE SCIENTIARUM FENNICAE

Series A

I. MATHEMATICA

429

ON PROBABILISTIC AUTOMATA AND THEIR GENERALIZATIONS

BY

PAAVO TURAKAINEN

HELSINKI 1968
SUOMALAINEN TIEDEAKATEMIA

doi:10.5186/aasfm.1969.429

Communicated 10 May 1968 by P. J. Myrberg and K. Inkeri

KESKUSKIRJAPAINO
HELSINKI 1968

Preface

I express my deepest gratitude to Professor A. Salomaa for the invaluable guidance and support he has given me at all stages of my work. My gratitude is also due to Professors K. Inkeri and O. Hellman as well as my colleagues for their interest in my investigation.

For the revision of the English manuscript, I am grateful to Lector M. S. Hardy.

The financial support from the Osk. Huttunen Foundation and Suomen Kulttuurirahaston Varsinais-Suomen Rahasto is gratefully acknowledged. I am also indebted to the Finnish Academy of Sciences for accepting this publication for inclusion in the Annals of the Academy.

Turku, May, 1968

Paavo Turakainen

Contents

INTRODUCTION

There has been a rapidly increasing amount of research concerning finite automata, most of which has been carried out by electrical engineers, logicians and mathematicians during the last ten years. One reason for the widespread interest in this subject is the fact that finite automata form a part of the theoretical background of digital computers.

Finite deterministic automata are characterized by the fact that the state transitions have a deterministic behavior. Papers concerning these automata are very numerous. One of the most important results, due to KLEENE [8], is that a language can be represented in a finite deterministic automaton if and only if it is regular.

The notion of a probabilistic automaton as a generalization of a finite deterministic automaton was introduced by BUKHARAEV [1]—[3], CARLYLE [4] and RABIN [13]. In such an automaton the state transitions have a stochastic behavior. This generalization is essential with respect to the family of representable languages; RABIN [13] showed that also non-regular languages can be represented in finite probabilistic automata.

In the present paper, languages representable in finite probabilistic automata are investigated. Following SALOMAA [17], we shall call them stochastic languages. The notion of a finite probabilistic automaton as well as the language represented in it is defined as in [12]. Thus the automata investigated do not necessarily have a fixed initial state, as in [13], but an initial probability distribution over the set of all internal states.

In the first chapter, finite probabilistic automata over a one-letter alphabet are investigated. We are mainly interested in conditions under which non-regular languages can be represented in these automata.

PAZ [12] gave an example of a three-state probabilistic automaton over a one-letter alphabet with a fixed initial state in which a non-regular language can be represented. We generalize his result by giving in Theorem 1 a necessary and sufficient condition for a three-state probabilistic automaton to represent a non-regular language.

In Theorem 3 we give a finite upper bound for the number of cut-points representing a non-regular language in an n-state probabilistic automaton. (Another approach to the same question is found in [15].) This bound can be determined by means of the so-called normal form of the transition matrix.

Furthermore, the corresponding exceptional cut-points can be calculated.

The exceptional cut-points are considered in § 5 and § 6. More precisely, we investigate conditions under which (a) no one of them represents a non-regular language (Theorems 4, 5 and 6), (b) each of them represents a non-regular language (Theorem 7), (c) at least one of them represents a non-regular language (Theorems 8 and 9). These conditions are closely associated with the eigenvalues of the transition matrix. The problem (a) has also been considered in [12], where it is assumed that the automaton has a fixed initial state and a single final state.

It should be noted that Theorems 1 and 3 have been presented in [19]. The proof of Theorem 1 in this work has been essentially simplified by our general theory.

In the second chapter, finite probabilistic automata and so-called generalized probabilistic automata and generalized automata are considered.

In § 8 we present some theorems on stochastic languages. It is first established that every stochastic language can be represented in a probabilistic automaton with any cut-point η such that $0 < \eta < 1$. The restriction $\eta > 0$ is essential, since only regular languages can be represented with the cut-point 0.

It is not known whether or not the family of stochastic languages is closed under any of the Boolean operations. In § 8 we establish some partial results on the closure under these operations. It is proved that the union and the intersection of a stochastic language and a regular language are both stochastic languages representable in the same automaton. As a consequence of this result we get a sufficient condition for the complement of a stochastic language to be stochastic.

By the definition, a probabilistic automaton has a fixed set F of final states, to which there corresponds a column vector π_F consisting of 0's and 1's only. In § 9 we replace π_F by a column vector with arbitrary real components. This generalization is due to Page [11]. However, it is not essential as far as the family of representable languages is concerned; in Theorem 16 we prove that a language can be represented in a generalized probabilistic automaton if and only if it can be represented in a probabilistic automaton.

By means of generalized probabilistic automata we introduce in § 10 a certain subfamily $\mathcal{L}(\neq)$ of stochastic languages which contains all regular languages as a proper subfamily. It is established that this family is closed under union and intersection. In addition, it is verified that the intersection of a stochastic language and a language belonging to $\mathcal{L}(\neq)$ is a stochastic language. We do not know whether or not $\mathcal{L}(\neq)$ is a proper subfamily of stochastic languages.

By the definition, the initial distribution vector and the transition matrices of a generalized probabilistic automaton are stochastic. In § 11 we replace them by a row vector and matrices with arbitrary real elements. The automaton thus obtained is called a generalized automaton. Our main result (Theorem 21) is that a language can be represented in a generalized automaton if and only if it can be represented in a probabilistic automaton, i.e., if and only if it is a stochastic language. This result is then used for two applications concerning the mirror image of a stochastic language (Theorem 22) and the right derivatives of languages (Theorem 23).

Finally, in § 12 criteria for the realizability of mappings by probabilistic automata and by generalized automata are considered.

DEFINITIONS AND NOTATIONS

By an *alphabet* I we mean a finite non-empty set. The set of words, including the empty word Λ, over the alphabet I is denoted by $W(I)$. Subsets of $W(I)$ are called *languages* over I. We often identify words with their unit sets.

The length of a word $P \in W(I)$ is denoted by $l(P)$.

The sum or union of two languages L_1 and L_2 is denoted by $L_1 + L_2$, their intersection by $L_1 \cap L_2$, and their product or catenation by $L_1 L_2$. The complement of a language L with respect to $W(I)$ is denoted by $\bar{L}$. In addition, we use the notation $L_1 - L_2 = L_1 \cap \bar{L}_2$. The *iteration* of a language L is defined by

$$L^* = \sum_{i=0}^{\infty} L^i .$$

Here L^0 denotes the language $\{\Lambda\}$.

A language L over I is called *regular* if it is obtained from the empty language and the elements of I by finitely many applications of the operations sum, product and iteration. Otherwise, L is *non-regular*.

Definition 1. A *finite probabilistic automaton* over the alphabet I is an ordered quadruple $\mathfrak{PA} = (S, M, \pi_0, F)$, where $S = \{s_1, \ldots, s_n\}$ is a finite non-empty set (the set of *internal states*), M is a mapping of I into the set of stochastic $n \times n$ matrices, $\pi_0 = (p_1, \ldots, p_n)$ is an n-dimensional stochastic row vector (the *initial distribution*) and F is a non-empty subset of S (the set of *final states*).

Matrices $M(x)$ $(x \in I)$ are called *transition matrices*. The domain of M is extended from I to $W(I)$ by defining

$$M(\Lambda) = E_n \quad (n \times n \text{ identity matrix}) ,$$
$$M(x_1 x_2 \cdots x_k) = M(x_1) M(x_2) \cdots M(x_k) ,$$

where $k \geqq 2$ and $x_i \in I$.

Let π_F be the n-dimensional column vector whose ith component equals 1 if $s_i \in F$ and 0 otherwise. The language *represented* in $\mathfrak{PA}$ with the *cut-point* η, $0 \leqq \eta < 1$, is defined by

$$L(\mathfrak{PA}, \eta) = \{P \in W(I) \mid \pi_0 M(P) \pi_F > \eta\} .$$

For $0 \leqq \eta < 1$, a language L is *η-stochastic* if and only if there exists a $\mathfrak{PA}$ such that $L = L(\mathfrak{PA}, \eta)$. A language L is *stochastic* if and only if, for some η, it is η-stochastic. For a given $\mathfrak{PA}$, a cut-point η is called *non-regular* if and only if $L(\mathfrak{PA}, \eta)$ is a non-regular language.

For any matrix A, we shall use the notation A^T to mean the transpose of A.

Hereafter, we use the term *probabilistic automaton* to mean a finite probabilistic automaton.

CHAPTER I

PROBABILISTIC AUTOMATA OVER A ONE-LETTER ALPHABET

§ 1. Preliminary remarks on stochastic matrices

1.1. In this chapter we investigate probabilistic automata whose alphabet consists of a single letter x. Our considerations are closely associated with certain properties of the transition matrix $M(x)$ and its powers $M(x)^m$. Therefore, we need some results concerning stochastic matrices. The terminology we use below is the same as in [6].

1.2. A *permutation* of a square matrix M is a permutation of the rows of M combined with the same permutation of the columns. M is called *reducible* if there is a permutation which transforms it into the form

$$\widetilde{M} = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix}$$

where B and D are square matrices. Otherwise, M is called *irreducible.*

Let M be a stochastic $n \times n$ matrix. Its eigenvalues satisfy the condition $|\lambda| \leqq 1$. Furthermore, $\lambda_1 = 1$ is an eigenvalue. For the eigenvalues of modulus 1, we have the following lemma, which is an immediate consequence of a theorem of Frobenius (cf. [6], Vol. 2, p. 53).

Lemma 1. *If M is an irreducible stochastic matrix, then $\lambda = 1$ is a simple root of the characteristic equation of M. Moreover, if M has h eigenvalues $\lambda_1, \ldots, \lambda_h$ of modulus 1, then these values are all distinct and are roots of the equation $\lambda^h - 1 = 0$, i.e.,*

$$\lambda_{\nu+1} = \exp(2\pi\nu i/h) \quad (\nu = 0, \ldots, h-1).$$

The number h is called the *index of imprimitivity* of M. The arguments of $\lambda_1, \ldots, \lambda_h$, denoted by $\arg\lambda_1, \ldots, \arg\lambda_h$, are of the form $2\pi r$ where r is a rational number. We say that $\arg\lambda$ $(\lambda \neq 0)$ is *rational in degrees* if it can be expressed in this form. Otherwise, $\arg\lambda$ is *irrational in degrees*.

With suitable permutations a stochastic matrix M can be transformed into its *normal form* (cf. [6], Vol. 2, p. 75)

$$(1.1) \qquad \widetilde{M} = \begin{bmatrix} M_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & M_2 & \ldots & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \ldots & \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & \ldots & M_g & 0 & \ldots & 0 \\ M_{g+1,1} & M_{g+1,2} & \ldots & M_{g+1,g} & M_{g+1} & \ldots & 0 \\ \cdot & \cdot & \ldots & \cdot & \cdot & \ldots & \cdot \\ M_{s1} & M_{s2} & \ldots & M_{sg} & M_{s,g+1} & \ldots & M_s \end{bmatrix}$$

where $M_1, \ldots, M_s$ are irreducible matrices and for each j, $g+1 \leqq j \leqq s$,

$$M_{j1} + M_{j2} + \ldots + M_{j,j-1} \neq 0.$$

The normal form is uniquely determined up to a permutation of the blocks and permutations within the diagonal blocks. Note that $M_1, \ldots, M_g$ are stochastic matrices but $M_{g+1}, \ldots, M_s$ are not. The eigenvalues of $M_{g+1}, \ldots, M_s$ are of modulus less than 1, since in each of these matrices at least one row sum is less than 1.

Denote by $h_1, \ldots, h_g$ the indices of imprimitivity of the matrices $M_1, \ldots, M_g$. Let h be the least common multiple of them, in symbols, $h = \text{l.c.m.}(h_1, \ldots, h_g)$. Then $\lambda = 1$ is the only eigenvalue of $(\widetilde{M})^h$ with modulus 1. It is obvious that, for each natural number m, the eigenvalues of M^m and $(\widetilde{M})^m$ are exactly the same. Consequently, $\lambda = 1$ is the only eigenvalue of M^h with modulus 1. This implies that the limit $\lim_{m\to\infty} (M^h)^m$ exists (cf. [6], Vol. 2, p. 93). We shall use this result in the following section.

1.3. Denote by $\lambda_1, \ldots, \lambda_r$ the distinct eigenvalues of a stochastic $n \times n$ matrix M. Then the following formula holds for the powers of M (cf. [6], Vol. 1, p. 107):

$$M^m = \sum_{k=1}^{r} \frac{1}{(m_k - 1)!} \left[\frac{d^{m_k - 1}}{d\lambda^{m_k - 1}} \left(D_k(\lambda) \lambda^m \right) \right]_{\lambda = \lambda_k} . \tag{1.2}$$

Here $D_k(\lambda)$ is an $n \times n$ matrix depending on λ, and m_k is the exponent of $\lambda - \lambda_k$ in the minimal polynomial of M. Denote by $p_{ij}^{(m)}$ the $(i\,,j)$th element of M^m. Calculating the required derivatives in (1.2), we obtain

$$p_{ij}^{(m)} = \sum_{k=1}^{r} \omega_{ijk}(m) \lambda_k^m , \tag{1.3}$$

where $\omega_{ijk}(m)$ is a polynomial in m of degree smaller than m_k. Let $\lambda_1, \ldots, \lambda_t$ be the distinct eigenvalues of M with modulus 1. Then $m_1 = \ldots = m_t = 1$ (cf. [6], Vol. 2, p. 86). Formula (1.3) can now be written in the form

$$p_{ij}^{(m)} = \sum_{k=1}^{t} f_{ijk}(\lambda_k) \lambda_k^m + \varepsilon_{ij}(m) ,$$

where $f_{ijk}(\lambda_k)$ is the $(i\,,j)$th element of the matrix $D_k(\lambda_k)$ and $\lim_{m \to \infty} \varepsilon_{ij}(m) = 0$. As we have already mentioned, the eigenvalues of M and $\widetilde{M}$ are exactly the same. This implies, by Lemma 1, that $\arg \lambda_1, \ldots, \arg \lambda_t$ are rational in degrees, because $\lambda_1, \ldots, \lambda_t$ are eigenvalues of the irreducible stochastic matrices $M_1, \ldots, M_g$. Hence

$$w_{ij}(m) = \sum_{k=1}^{t} f_{ijk}(\lambda_k) \lambda_k^m$$

is a periodic function of m having only a finite number of distinct values. Let h be as in section 1.2. Then the limit $\lim_{m \to \infty} M^{mh + \nu}$ exists for each ν, $0 \leqq \nu \leqq h - 1$. This implies that h is the period of $w_{ij}(m)$ $(i\,, j = 1, \ldots, n)$. Thus, we have obtained the formula

$$p_{ij}^{(m)} = w_{ij}(m) + \varepsilon_{ij}(m) \tag{1.4}$$

where $\lim_{m \to \infty} \varepsilon_{ij}(m) = 0$ and $w_{ij}(m)$ is a periodic function of m, the period being $h = \text{l.c.m.}(h_1, \ldots, h_g)$.

If all of the eigenvalues of M are simple, i.e., $r = n$, then we have the formula (cf. [5], p. 431)

$$p_{ij}^{(m)} = \sum_{k=1}^{n} \Big(\sum_{\nu=1}^{n} x_\nu^{(k)} y_\nu^{(k)} \Big)^{-1} x_i^{(k)} y_j^{(k)} \lambda_k^m , \tag{1.5}$$

where $x_i^{(k)}$ is the ith component of the column eigenvector corresponding to λ_k, and $y_j^{(k)}$ is the jth component of the row eigenvector corresponding to λ_k.

Formulas (1.4) and (1.5) are often used below.

§ 2. Lemmas

2.1. In what follows, we present some lemmas which are often needed in our considerations.

Lemma 2. *Every regular language L over the alphabet $\{x\}$ can be expressed in the form*

$$L = L_1 + (x^{m_1} + \ldots + x^{m_k})(x^u)^*$$

where L_1 is a finite language, $k \geqq 0$ and $u > 0$.

This lemma is an immediate consequence of Theorem 1 in [14]. If L is an infinite language, then necessarily $k > 0$.

For any irrational number γ, the set of numbers $m\gamma$ (mod 1), $m = 1, 2, \ldots$, is everywhere dense in the whole interval [0, 1] (cf. [9], p. 75). If φ is irrational in degrees and q is a natural number, the number $q\varphi$ is irrational in degrees. This implies the following

Lemma 3. *If φ is irrational in degrees, then for any natural number q and any real number ψ, the set of numbers $mq\varphi + \psi$ (mod 2π), $m = 1, 2, \ldots$, is everywhere dense in the whole interval $[0, 2\pi]$.*

2.2. In this section we consider a three-state probabilistic automaton $\mathfrak{PA} = (S, M, \pi_0, F)$ over $\{x\}$ such that $M(x) = [p_{ij}]$ has an imaginary eigenvalue λ_2. The row eigenvector of the eigenvalue $\lambda_1 = 1$ and the column eigenvector of the eigenvalue λ_2 will be denoted by $Y_1 = (y_1, y_2, y_3)$ and $X_2 = (x_1^{(2)}, x_2^{(2)}, x_3^{(2)})^T$, respectively. Since λ_2 is imaginary, it follows that $p_{ii} - 1 < 0$ $(i = 1, 2, 3)$. Using this and the equation $Y_1(M(x) - E_3) = 0$, it is verified that we can choose $y_i \geqq 0$ $(i = 1, 2, 3)$. Thus $y_1 + y_2 + y_3 > 0$, and we may define

$$\begin{aligned} \eta_i &= y_i/(y_1 + y_2 + y_3)\,, \\ \eta_{ij} &= \eta_i + \eta_j\,, \end{aligned} \tag{2.1}$$

where $i, j = 1, 2, 3$ and $i \neq j$.

Lemma 4. *The components of the row and column eigenvectors corresponding to the imaginary eigenvalue* λ_2 *are different from zero.*

Proof. By the equation $(M(x) - \lambda_2 E_3)X_2 = 0$, $x_1^{(2)} = 0$ if and only if

$$p_{12}p_{23} - p_{13}(p_{22} - \lambda_2) = 0 \,. \tag{2.2}$$

If (2.2) holds, then $p_{13} = 0$, since λ_2 is not real. This implies that either $p_{12} = 0$ or $p_{23} = 0$. In both cases, a simple calculation shows that the eigenvalues of $M(x)$ are real, which contradicts our assumption. Thus (2.2) does not hold and, therefore, $x_1^{(2)} \neq 0$. The same argument applies to $x_2^{(2)}$ and $x_3^{(2)}$. The proof is similar for the row eigenvector of λ_2.

Lemma 5. *The equation*

$$p_1 x_1^{(2)} + p_2 x_2^{(2)} + p_3 x_3^{(2)} = 0 \tag{2.3}$$

holds for an initial distribution (p_1 , p_2 , p_3) *if and only if* $(p_1 , p_2 , p_3) = (\eta_1 , \eta_2 , \eta_3)$.

Proof. As we have already remarked, $y_i \geqq 0$ $(i = 1 , 2 , 3)$. Thus $\eta_i \geqq 0$ $(i = 1 , 2 , 3)$. Since, in addition, $\eta_1 + \eta_2 + \eta_3 = 1$, we find that $(\eta_1 , \eta_2 , \eta_3)$ is a stochastic vector. Therefore, it can be chosen as an initial distribution. Furthermore, it is a row eigenvector of λ_1. This implies that the equation (2.3) holds if we choose $(p_1 , p_2 , p_3) = (\eta_1 , \eta_2 , \eta_3)$, because the row and column eigenvectors corresponding to distinct eigenvalues are orthogonal.

Conversely, assume that (2.3) holds for an initial distribution (p_1 , p_2 , p_3). Consider the equation

$$\begin{bmatrix} p_{11} - \lambda_2 & p_{12} & p_{13} \\ p_{21} & p_{22} - \lambda_2 & p_{23} \\ p_{31} & p_{32} & p_{33} - \lambda_2 \\ p_1 & p_2 & p_3 \end{bmatrix} X = 0 \,. \tag{2.4}$$

Since it has a non-trivial solution $X = (x_1^{(2)} , x_2^{(2)} , x_3^{(2)})^T$, every determinant of order 3, formed from the matrix in (2.4), must vanish. Calculating these determinants, we obtain the equations

$$\begin{cases} p_1\lambda_2^2 - (p_1(p_{22} + p_{33}) - p_2p_{21} - p_3p_{31})\lambda_2 + c_1 = 0 \,, \\ p_2\lambda_2^2 - (p_2(p_{11} + p_{33}) - p_1p_{12} - p_3p_{32})\lambda_2 + c_2 = 0 \,, \\ p_3\lambda_2^2 - (p_3(p_{11} + p_{22}) - p_1p_{13} - p_2p_{23})\lambda_2 + c_3 = 0 \,, \end{cases} \tag{2.5}$$

where c_1, c_2 and c_3 are real numbers. According to Lemma 4, $x_i^{(2)} \neq 0$

for each $i = 1, 2, 3$. This and the equation (2.3) imply that at least two of the numbers p_1, p_2, p_3 are different from zero, because $p_1 + p_2 + p_3 = 1 > 0$. Consequently, if $p_1 = 0$ then $p_2 > 0$, $p_3 > 0$ and, by the first of the equations (2.5), $p_2 p_{21} + p_3 p_{31} = 0$, since λ_2 is not real. Thus $p_{21} = 0$ and $p_{31} = 0$. But if $M(x)$ satisfies these conditions, then its eigenvalues are real. This contradicts our assumption. Hence $p_1 \neq 0$. On the other hand, λ_2 satisfies the characteristic equation of $M(x)$. This implies that

$$\lambda_2^2 - (p_{11} + p_{22} + p_{33} - 1)\lambda_2 + \det M(x) = 0 . \tag{2.6}$$

By remembering that $p_1 \neq 0$ and by comparing the equation (2.6) with the first of the equations (2.5), we obtain

$$p_1(p_{22} + p_{33}) - p_2 p_{21} - p_3 p_{31} = p_1(p_{11} + p_{22} + p_{33} - 1) ,$$

since λ_2 is not real. In other words,

$$(p_{11} - 1)p_1 + p_{21}p_2 + p_{31}p_3 = 0 . \tag{2.7}$$

In the same way, by using the second and the third of the equations (2.5), we obtain

$$\begin{aligned} p_{12}p_1 + (p_{22} - 1)p_2 + p_{32}p_3 &= 0 , \\ p_{13}p_1 + p_{23}p_2 + (p_{33} - 1)p_3 &= 0 . \end{aligned} \tag{2.8}$$

From (2.7) and (2.8) it now follows that (p_1, p_2, p_3) is a row eigenvector of $\lambda_1 = 1$. Consequently, $(p_1, p_2, p_3) = (\eta_1, \eta_2, \eta_3)$, since (p_1, p_2, p_3) is a stochastic vector. This completes the proof of Lemma 5.

§ 3. Three-state probabilistic automata

3.1. A three-state probabilistic automaton over a one-letter alphabet is the most simple probabilistic automaton where a non-regular language can be represented. Using the notations (2.1), we give in the following theorem a necessary and sufficient condition for such an automaton to represent a non-regular language.

Theorem 1. *Let* $\mathfrak{PA} = (\{s_1, s_2, s_3\}, M, (p_1, p_2, p_3), F)$ *be a three-state probabilistic automaton over the alphabet* $\{x\}$, *where* F *is a proper subset of the set* $\{s_1, s_2, s_3\}$. *A non-regular language can be represented in* $\mathfrak{PA}$ *if and only if* $M(x)$ *has an imaginary eigenvalue* λ_2 *such that* $\arg \lambda_2$ *is irrational in degrees and* $(p_1, p_2, p_3) \neq (\eta_1, \eta_2, \eta_3)$. *Moreover, if this* *condition is satisfied, then there is exactly one cut-point* η *such that the language* $L(\mathfrak{P}\mathfrak{A}, \eta)$ *is non-regular. This cut-point is* η_i *if* $F = \{s_i\}$ *and* η_{ij} *if* $F = \{s_i, s_j\}$ $(i \neq j)$.

3.2. Before being able to prove this theorem, we have to derive suitable formulas for the probability $\pi_0 M(x^m)\pi_F$ where $\pi_0 = (p_1, p_2, p_3)$. For doing this, we first assume that $F = \{s_i\}$ and $M(x)$ has imaginary eigenvalues λ_2 and λ_3. Then $\lambda_3 = \bar{\lambda}_2$, where the bar denotes complex conjugate. The column eigenvector of the eigenvalue $\lambda_1 = 1$ is $(1, 1, 1)^T$. This together with formula (1.5) implies

$$p_{ki}^{(m)} = \eta_i + R_2 x_k^{(2)} y_i^{(2)} \lambda_2^m + R_3 x_k^{(3)} y_i^{(3)} \lambda_3^m \quad (k = 1, 2, 3), \tag{3.1}$$

where we have used the notation

$$R_j = \Big(\sum_{\nu=1}^{3} x_\nu^{(j)} y_\nu^{(j)} \Big)^{-1} \quad (j = 2, 3).$$

Hence

$$\pi_0 M(x^m)\pi_F = \eta_i + R_2 u^{(2)} y_i^{(2)} \lambda_2^m + R_3 u^{(3)} y_i^{(3)} \lambda_3^m,$$

where we have denoted

$$u^{(j)} = p_1 x_1^{(j)} + p_2 x_2^{(j)} + p_3 x_3^{(j)} \quad (j = 2, 3).$$

From the equations $(M(x) - \lambda_j E_3)X_j = 0$ $(j = 2, 3)$ it follows that the components of the column eigenvector X_3 corresponding to λ_3 can be chosen as the conjugates of the corresponding components of X_2. The same holds for the row eigenvectors Y_2 and Y_3, too. Consequently, the coefficients of λ_2^m and λ_3^m in formula (3.1) are conjugate complex numbers. Thus, this formula can be written in the form

$$\pi_0 M(x^m)\pi_F = \eta_i + u_i \lambda_2^m + \bar{u}_i \lambda_3^m \tag{3.2}$$

where $u_i = R_2 u^{(2)} y_i^{(2)}$.

Lemma 5 is now used. If $(p_1, p_2, p_3) = (\eta_1, \eta_2, \eta_3)$, then $u^{(2)} = 0$ and, therefore, $u_i = 0$. This leads, by (3.2), to the result

$$\pi_0 M(x^m)\pi_F = \eta_i \quad (\pi_0 = (\eta_1, \eta_2, \eta_3)). \tag{3.3}$$

Assume that $(p_1, p_2, p_3) \neq (\eta_1, \eta_2, \eta_3)$. Then, by Lemmas 4 and 5, $u_i \neq 0$, which implies that $\arg u_i$ is defined. We use the notations

$$\arg \lambda_2 = \varphi, \quad \arg u_i = \psi_i.$$

Since $\arg \lambda_3 = -\varphi$ and $\arg \bar{u}_i = -\psi_i$, formula (3.2) gets the form

$$\pi_0 M(x^m)\pi_F = \eta_i + 2|u_i||\lambda_2|^m \cos(m\varphi + \psi_i) \quad (\pi_0 \neq (\eta_1, \eta_2, \eta_3)). \tag{3.4}$$

3.3. Theorem 1 will now be proved. We first assume that $F = \{s_i\}$ ($i = 1$, 2 or 3).

To establish the »if»-part and the last sentence of Theorem 1, suppose that $\varphi = \arg \lambda_2$ is irrational in degrees for an imaginary eigenvalue λ_2 of $M(x)$, and $(p_1, p_2, p_3) \neq (\eta_1, \eta_2, \eta_3)$. We shall show that the language $L(\mathfrak{P}\mathfrak{A}, \eta_i)$ is non-regular. Suppose, on the contrary, that it is a regular language. By our assumptions, formula (3.4) is valid. Since $|u_i||\lambda_2|^m > 0$, it follows that

$$(3.5) \quad L(\mathfrak{P}\mathfrak{A}, \eta_i) - \{\Lambda\} = \{x^m | -\pi/2 < m\varphi + \psi_i \pmod{2\pi} < \pi/2\} .$$

Applying Lemma 3 to the set of numbers $m\varphi + \psi_i \pmod{2\pi}$, $m = 1, 2, \ldots$, we conclude from (3.5) that $L(\mathfrak{P}\mathfrak{A}, \eta_i)$ is an infinite language. Now, Lemma 2 is used. Since $L(\mathfrak{P}\mathfrak{A}, \eta_i)$ is infinite, it follows that $k \geqq 1$ and, for some natural numbers m_1 and u,

$$(3.6) \qquad x^{m_1 + \nu u} \in L(\mathfrak{P}\mathfrak{A}, \eta_i) \quad (\nu = 1, 2, \ldots) .$$

Applying Lemma 3 to the set of numbers $(m_1 + \nu u)\varphi + \psi_i \pmod{2\pi}$, $\nu = 1, 2, \ldots$, we find that, for some natural number ν_1,

$$\pi/2 < (m_1 + \nu_1 u)\varphi + \psi_i \pmod{2\pi} < 3\pi/2 .$$

From formula (3.5) it now follows that the word $x^{m_1 + \nu_1 u}$ does not belong to the language $L(\mathfrak{P}\mathfrak{A}, \eta_i)$. This contradicts (3.6). Consequently, $L(\mathfrak{P}\mathfrak{A}, \eta_i)$ is a non-regular language, whence the »if»-part of Theorem 1 follows.

The last sentence of Theorem 1 will now be proved. By the above considerations, it suffices to show that $L(\mathfrak{P}\mathfrak{A}, \eta)$ is a regular language for any η such that $\eta \neq \eta_i$. Since λ_2 is imaginary, it follows that $M(x)$ is irreducible. If $|\lambda_2| = 1$, then, by Lemma 1, $\varphi = \arg \lambda_2$ is rational in degrees. This contradicts our assumption. Hence $|\lambda_2| < 1$ and, therefore,

$$\lim_{m \to \infty} 2|u_i||\lambda_2|^m \cos (m\varphi + \psi_i) = 0 .$$

This implies that, for any $\eta_i' > \eta_i$ and $\eta_i'' < \eta_i$, $L(\mathfrak{P}\mathfrak{A}, \eta_i')$ is a finite language and $L(\mathfrak{P}\mathfrak{A}, \eta_i'')$ is the complement of a finite language. In both cases, the languages are regular, whence the last sentence of Theorem 1 follows.

For the »only if»-part of Theorem 1, assume that $M(x)$ has an imaginary eigenvalue λ_2 and that $(p_1, p_2, p_3) = (\eta_1, \eta_2, \eta_3)$. Now, formula (3.3) is valid, and clearly $L(\mathfrak{P}\mathfrak{A}, \eta)$ is regular for any cut-point η. The rest of the »only if»-part follows from Theorem 6 which we prove in § 5.

3.4. Finally, we consider the case $F = \{s_i , s_j\}$ $(i \neq j)$. Assume first that $M(x)$ has an imaginary eigenvalue λ_2. We have

$$\pi_0 M(x^m)\pi_F = \sum_{k=1}^{3} p_k(p_{ki}^{(m)} + p_{kj}^{(m)}) .$$

If $(p_1 , p_2 , p_3) = (\eta_1 , \eta_2 , \eta_3)$, then, by (3.3),

$$(\eta_1 , \eta_2 , \eta_3)M(x^m)\pi_F = \eta_i + \eta_j = \eta_{ij} . \tag{3.7}$$

If $(p_1 , p_2 , p_3) \neq (\eta_1 , \eta_2 , \eta_3)$, then $u_k \neq 0$ $(k = 1 , 2 , 3)$ and

$$\sum_{k=1}^{3} p_k(p_{ki}^{(m)} + p_{kj}^{(m)}) = \sum_{k=1}^{3} p_k(1 - p_{ks}^{(m)}) = \eta_{ij} + 2|u_s||\lambda_2|^m \cos(m\varphi + \psi_s') \tag{3.8}$$

where $s \neq i , j$, $\varphi = \arg \lambda_2$ and $\psi_s' = \arg(-u_s)$. By formulas (3.7) and (3.8), the proof is reduced to the corresponding proof for $F = \{s_i\}$. Instead of η_i we have η_{ij}. As in the case $F = \{s_i\}$, the rest of the proof follows from Theorem 6.

3.5. As a consequence of Theorem 1 we establish the following

Theorem 2. *Assume that at least one component of the initial distribution* $\pi_0 = (p_1 , p_2 , p_3)$ *equals zero. A non-regular language can be represented in* $\mathfrak{P}\mathfrak{A} = (\{s_1 , s_2 , s_3\} , M , \pi_0 , F)$, *where* F *is a proper subset of the set* $\{s_1 , s_2 , s_3\}$, *if and only if* $M(x)$ *has an imaginary eigenvalue* λ_2 *such that* $\arg \lambda_2$ *is irrational in degrees.*

Proof. Assume that, for an imaginary eigenvalue λ_2, $\arg \lambda_2$ is irrational in degrees. We have to show that π_0 satisfies the condition of Theorem 1. As in section 3.3, we conclude that $|\lambda_2| = |\lambda_3| < 1$. This implies that the elements of the matrix $\lim_{m \to \infty} M(x)^m$ are positive (cf. [6], Vol. 2, p. 93). On the other hand, by formula (3.1), every row in this matrix equals $(\eta_1 , \eta_2 , \eta_3)$. Thus $\eta_i > 0$ $(i = 1 , 2 , 3)$. Consequently, the condition $(p_1 , p_2 , p_3) \neq (\eta_1 , \eta_2 , \eta_3)$ is satisfied if at least one of the numbers p_1 , p_2 , p_3 equals zero, whence the theorem follows, by Theorem 1.

The assumption of Theorem 2 is satisfied for three-state probabilistic automata having a fixed initial state, since then π_0 is a co-ordinate vector.

3.6. In some cases, the question whether or not $\arg \lambda$ is irrational in degrees can be solved by the following result of Olmsted [10].

Lemma 6. *If* φ *is rational in degrees, then the only rational values of* $\cos \varphi$ *are* 0, $\pm 1/2$, ± 1.

The eigenvalues of a stochastic 3×3 matrix $M(x) = [p_{ij}]$ are roots of the equation

$$(\lambda - 1)(\lambda_2 - (p_{11} + p_{22} + p_{33} - 1)\lambda + \det M(x)) = 0 .$$

If an eigenvalue λ_2 of $M(x)$ is imaginary, then

$$\cos(\arg \lambda_2) = (p_{11} + p_{22} + p_{33} - 1)/2\sqrt{\det M(x)} .$$

According to Lemma 6, $\arg \lambda_2$ is irrational in degrees if $\cos(\arg \lambda_2)$ is rational and is different from $0, \pm 1/2, \pm 1$. Theorem 1 now implies that if λ_2 is imaginary, $\cos(\arg \lambda_2)$ is rational and $(p_1, p_2, p_3) \neq (\eta_1, \eta_2, \eta_3)$, then the language $L(\mathfrak{PA}, \eta)$ is regular for every cut-point η if and only if

$$(p_{11} + p_{22} + p_{33} - 1)/2\sqrt{\det M(x)} = 0 , \pm 1/2 , \pm 1.$$

As an example, consider the probabilistic automaton $\mathfrak{PA} = (\{s_1, s_2, s_3\}, M, \pi_0, F)$ over $\{x\}$ where F consists of one or two states and

$$M(x) = \begin{bmatrix} 3/8 & 1/4 & 3/8 \\ 1/8 & 1/4 & 5/8 \\ 1/3 & 0 & 2/3 \end{bmatrix} .$$

A straightforward computation shows that eigenvalues λ_2 and λ_3 are imaginary and $\cos(\arg \lambda_2) = 7/12$. Thus, by Lemma 6, $\arg \lambda_2$ is irrational in degrees. The row eigenvector of $\lambda_1 = 1$ is

$$Y_1 = (12/37 , 4/37 , 21/37) .$$

According to Theorem 1, a non-regular language can be represented in $\mathfrak{PA}$ if and only if $\pi_0 \neq Y_1$. Cut-points η_i and η_{ij} are immediately obtained from Y_1.

§ 4. An upper bound for the number of non-regular cut-points

4.1. For any three-state probabilistic automaton over $\{x\}$, the number of non-regular cut-points is at most one. In what follows, we consider this number for general probabilistic automata over $\{x\}$ and derive a finite upper bound for it.

Let $\mathfrak{PA} = (S, M, \pi_0, F)$ be an n-state probabilistic automaton over the alphabet $\{x\}$, where $\pi_0 = (p_1, \ldots, p_n)$ and $M(x) = [p_{ij}]$. By formula (1.4).

$$\pi_0 M(x^m)\pi_F = \sum_{s_j \in F} \sum_{i=1}^{n} p_i w_{ij}(m) + \sum_{s_j \in F} \sum_{i=1}^{n} p_i \varepsilon_{ij}(m) .$$

Denote the sums on the right by $W(m)$ and $\varepsilon(m)$. Then

$$(4.1) \qquad \pi_0 M(x^m)\pi_F = W(m) + \varepsilon(m) .$$

The values of the periodic function $w_{ij}(m)$ are real and non-negative, because $\lim_{m\to\infty} \varepsilon_{ij}(m) = 0$ and $p_{ij}^{(m)} \geqq 0$. Hence $\varepsilon_{ij}(m)$ is real, too. It follows that $W(m)$ and $\varepsilon(m)$ are real. Furthermore, the values of $W(m)$ are non-negative and $\lim_{m\to\infty} \varepsilon(m) = 0$. Denote by h the period of $w_{ij}(m)$ $(i, j = 1, \ldots, n)$. It is easily verified that also the function $W(m)$ is periodic, the period being h.

Lemma 7. *If* η *is not a value of* $W(m)$, *then* $L(\mathfrak{PA}, \eta)$ *is a regular language.*

Proof. Let $w_1, \ldots, w_h$ be the values of $W(m)$, arranged so that $w_1 \leqq \cdots \leqq w_h$. If $\eta = w_h + \delta\ (\delta > 0)$, then $L(\mathfrak{PA}, \eta)$ is a finite language, since $\lim_{m\to\infty} \varepsilon(m) = 0$. By the same reason, $L(\mathfrak{PA}, \eta)$ is the complement of a finite language if $\eta = w_1 - \delta\ (\delta > 0)$. Finally, assume that, for some i, $w_i < \eta < w_{i+1}$. Let $\delta = \min(\eta - w_i, w_{i+1} - \eta)$. There exists a natural number m_0 such that $|\varepsilon(m)| < \delta$ for all $m > m_0$. Using this, it is immediately verified that $L(\mathfrak{PA}, \eta)$ is of the form

$$L(\mathfrak{PA}, \eta) = L_1 + (x^{m_{i+1}} + \ldots + x^{m_h})(x^h)^* ,$$

where L_1 is a finite language containing only words P with $l(P) \leqq m_0$. Hence $L(\mathfrak{PA}, \eta)$ is a regular language. This completes the proof of Lemma 7.

4.2. Lemma 7 shows that the number of non-regular cut-points does not exceed the period h of $W(m)$. If $\lambda = 1$ is the only eigenvalue of $M(x)$ with modulus 1, then the limit $\lim_{m\to\infty} \pi_0 M(x^m)\pi_F$ exists (cf. section 1.2). This implies, by formula (4.1), that $W(m)$ is constant, i.e., its value does not depend on m. By Lemma 7, we now have

Corollary 1. *If* $\lambda = 1$ *is the only eigenvalue of the transition matrix* $M(x)$ *with modulus* 1, *then there exists at most one non-regular cut-point.*

With suitable permutations $M(x)$ can be transformed into its normal form (1.1). As before, denote the indices of imprimitivity of the irreducible matrices $M_1, \ldots, M_g$ by $h_1, \ldots, h_g$. Then, as noticed in sections 1.3 and 4.1, the period h of $W(m)$ equals the number l.c.m.$(h_1, \ldots, h_g)$. This together with Lemma 7 implies the following general result (for another approach to the same question, see [15]).

Theorem 3. *Let* $\mathfrak{P}\mathfrak{A} = (S, M, \pi_0, F)$ *be a probabilistic automaton over the alphabet* $\{x\}$. *The number of non-regular cut-points does not exceed the number* $h = \text{l.c.m.}(h_1, \ldots, h_g)$ *where* $h_1, \ldots, h_g$ *are the indices of imprimitivity of the matrices* $M_1, \ldots, M_g$ *in the normal form* (1.1) *of* $M(x)$.

This theorem reveals a very special property of probabilistic automata over a one-letter alphabet, not possessed by probabilistic automata over an arbitrary alphabet; it is well-known (cf. [13], [12]) that even in certain two-state probabilistic automata over the alphabet consisting of two letters, an infinite number of non-regular languages can be represented.

§ 5. Probabilistic automata and regular languages

5.1. According to Lemma 7, a necessary condition for a cut-point to be non-regular is that it is a value of the periodic function $W(m)$. As we already saw in Theorem 1, this condition is not sufficient. In this and the following paragraph we investigate conditions under which (a) no value of $W(m)$ is non-regular, (b) every value of $W(m)$ is non-regular, (c) at least one value of $W(m)$ is non-regular. We use the earlier notations and label the distinct eigenvalues of the transition matrix $M(x)$ so that

$$1 = |\lambda_1| = \cdots = |\lambda_{n_0}| > |\lambda_{n_0+1}| \geqq \cdots \geqq |\lambda_r| .$$

5.2. It is obvious that if $\varepsilon(m)$ vanishes identically for sufficiently large values of m, then the language $L(\mathfrak{P}\mathfrak{A}, \eta)$ is regular for any cut-point η. In the case where $\varepsilon(m)$ does not identically vanish, we first establish the following theorem concerning the problem (a) of section 5.1.

Theorem 4. *Let* $\mathfrak{P}\mathfrak{A} = (\{s_1, \ldots, s_n\}, M, (p_1, \ldots, p_n), F)$ *be a probabilistic automaton over the alphabet* $\{x\}$. *Assume that, for some integers* s, q *and* m_0, $|\lambda_{s-1}| > |\lambda_s| = \cdots = |\lambda_{s+q}| > |\lambda_{s+q+1}|$ *and*

$$(5.1) \qquad \sum_{s_j \in F} \sum_{i=1}^{n} \sum_{k=s}^{s+q} p_i\, \omega_{ijk}(m) \lambda_k^m \neq 0 \quad \textit{for all} \quad m > m_0 ,$$

but

$$(5.2) \qquad \sum_{s_j \in F} \sum_{i=1}^{n} \sum_{k=n_0+1}^{s-1} p_i\, \omega_{ijk}(m) \lambda_k^m = 0 \quad \textit{for all} \quad m > m_0 .$$

If $\arg \lambda_s, \ldots, \arg \lambda_{s+q}$ *are rational in degrees, then* $L(\mathfrak{P}\mathfrak{A}, \eta)$ *is a regular language for any cut-point* η.

Proof. By Lemma 7, we may assume that the cut-point η is a value of $W(m)$. Since $\lim_{m \to \infty} \varepsilon(m) = 0$, there exists an integer $m_1 > m_0$ such that

$$\{m | m > m_1 , \ \pi_0 M(x^m) \pi_F > \eta\} = \{m | m > m_1 , \ W(m) > \eta\}$$
$$\cup \{m | m > m_1 , \ W(m) = \eta , \ \varepsilon(m) > 0\} .$$

Denote the sets on the right by M_2 and M_3, respectively. Now we have

$$L(\mathfrak{P}\mathfrak{A} , \eta) = L_1 + \{x^m | m \in M_2\} + \{x^m | m \in M_3\} , \tag{5.3}$$

where L_1 is a finite language containing only words P with $l(P) \leqq m_1$.

It is easily verified that the language

$$L_2 = \{x^m | m \in M_2\}$$

is regular. Now it is sufficient to prove that the language

$$L_3 = \{x^m | m \in M_3\}$$

is regular. We first derive a suitable form for $\varepsilon(m)$.

Since $m_1 > m_0$, it follows from the assumption (5.2) that, for $m > m_1$,

$$\varepsilon(m) = \sum_{s_j \in F} \sum_{i=1}^{n} \sum_{k=s}^{s+q} p_i \, \omega_{ijk}(m) \lambda_k^m + \sum_{s_j \in F} \sum_{i=1}^{n} \sum_{k=s+q+1}^{r} p_i \, \omega_{ijk}(m) \lambda_k^m , \tag{5.4}$$

where $|\lambda_s| = \ldots = |\lambda_{s+q}|$ and $|\lambda_k| < |\lambda_s|$ for each $k = s + q + 1 , \ldots , r$. Let m^{k_1} be the highest power of m in the first sum of (5.4) whose coefficient does not identically vanish. By the assumption (5.1), this power exists. By taking $\lambda_s^m \, m^{k_1}$ as a factor of $\varepsilon(m)$, we have

$$\varepsilon(m) = \lambda_s^m \, m^{k_1} (U_1(m) + u_1(m)) \quad (m > m_1) \tag{5.5}$$

where $\lambda_s^m \, U_1(m)$ is the afore-mentioned coefficient of m^{k_1} and $u_1(m)$ is a sum for which $\lim_{m \to \infty} u_1(m) = 0$. The function $U_1(m)$ has the form

$$U_1(m) = a_1 \exp (i \pi r_1 m) + \ldots + a_\mu \exp (i \pi r_\mu m) \tag{5.6}$$

where $\mu \leqq q + 1$, $a_1 , \ldots , a_\mu$ are complex constants and $r_1 , \ldots , r_\mu$ are rational numbers. The last statement holds, because $\arg \lambda_s , \ldots , \arg \lambda_{s+q}$ are rational in degrees. It implies that $U_1(m)$ is a periodic function of m having only a finite number of distinct values. Let

$$M' = \{m | m > m_1 , \ U_1(m) \neq 0\} .$$

Since $\lim_{m \to \infty} u_1(m) = 0$, we conclude that $U_1(m) + u_1(m) \neq 0$ whenever $m \in M'$ and m is large enough. We may assume the integer m_1 to be so large that $U_1(m) + u_1(m) \neq 0$ for all $m \in M'$. This implies that $\varepsilon(m) \neq 0$ whenever $m \in M'$. Thus, $\arg \varepsilon(m)$ is defined for all $m \in M'$.

Furthermore, since $\varepsilon(m)$ is real, it follows that $\arg \varepsilon(m) = \nu(m)\pi$ where $\nu(m)$ is an integer depending on m. We obtain the result

$$(5.7) \qquad \arg \varepsilon(m) = \arg \lambda_s^m + \arg (U_1(m) + u_1(m)) = \nu(m)\pi \quad (m \in M') .$$

Since $\arg \lambda_s$ is rational in degrees, we have $\arg \lambda_s^m = 2\pi ma/b$, where a and b are integers. This implies, by formula (5.7), that

$$(5.8) \qquad \arg (U_1(m) + u_1(m)) = \nu(m)\pi - 2\pi ma/b \quad (m \in M') .$$

Consequently, the function $\arg (U_1(m) + u_1(m))$ $(m \in M')$ has only a finite number of distinct values (mod 2π). On the other hand,

$$\arg (U_1(m) + u_1(m)) = \arg U_1(m) + \varphi(m) \quad (m \in M') ,$$

where $\lim_{m\to\infty} \varphi(m) = 0$, because $\lim_{m\to\infty} u_1(m) = 0$ and the values of the periodic function $U_1(m)$ are different from zero for all $m \in M'$. Using this result and formula (5.8), it can be proved that if $m_2 > m_1$ is large enough, then $\varphi(m) = 0$ whenever $m > m_2$ and $m \in M'$. Consequently,

$$\arg (U_1(m) + u_1(m)) = \arg U_1(m) \quad (m > m_2 , \quad m \in M') .$$

For $\varepsilon(m)$ we thus obtain from (5.5),

$$(5.9) \qquad \varepsilon(m) = |\lambda_s|^m |U_1(m) + u_1(m)| m^{k_1} \exp (i(\arg \lambda_s^m + \arg U_1(m))) \qquad (m > m_2 , \quad m \in M') .$$

In order to prove that L_3 is a regular language, we first consider the language

$$L_4 = \{x^m | m > m_2 , \quad \varepsilon(m) > 0\} ,$$

which we express in the form

$$L_4 = \{x^m | m > m_2 , \; m \in M' , \; \varepsilon(m) > 0\} + \{x^m | m > m_2 , \; U_1(m) = 0 , \; \varepsilon(m) > 0\}.$$

Our intention is to prove that L_4 is a regular language. Denote the languages on the right by L_5 and L_6, respectively. Thus $L_4 = L_5 + L_6$. By formula (5.9),

$$L_5 = \{x^m | m > m_2 , \quad m \in M' , \quad \arg \lambda_s^m + \arg U_1(m) \pmod{2\pi} = 0\} .$$

From the periodicity of $\arg \lambda_s^m$ (mod 2π) and $\arg U_1(m)$ (mod 2π) $(m \in M')$ it follows that also the function $\arg \lambda_s^m + \arg U_1(m)$ (mod 2π) $(m \in M')$ is periodic. This implies that L_5 is a regular language. Consequently, L_4 is regular if

$$L_6 = \{x^m | m > m_2 , \quad U_1(m) = 0 , \quad \varepsilon(m) > 0\}$$

is regular.

In order to prove that L_6 is a regular language, we first note that if 0 is not a value of $U_1(m)$ then L_6 is the empty language. Hence it is regular. Assume secondly that 0 is a value of $U_1(m)$ and denote

$$N' = \{m \mid m > m_2\,, \quad U_1(m) = 0\}\,.$$

Let m^{k_2} be the highest power of m in the first sum of (5.4) whose coefficient does not vanish for all $m \in N'$. By the assumption (5.1), this power exists. The following formula for $\varepsilon(m)$ $(m \in N')$ is obtained in the same manner as formula (5.5):

$$\varepsilon(m) = \lambda_s^m\, m^{k_2}(U_2(m) + u_2(m)) \quad (m \in N')\,.$$

Here $\lambda_s^m\, U_2(m)$ is the afore-mentioned coefficient of m^{k_2}, and $\lim_{m\to\infty} u_2(m) = 0$. The function $U_2(m)$ is of the form (5.6). Thus it is periodic and has only a finite number of distinct values. Continuing in the same way as we did after formula (5.6), we obtain

$$\begin{aligned} L_6 = L' &+ \{x^m \mid m > m_3\,, \quad m \in N'\,, \quad U_2(m) \neq 0\,, \quad \varepsilon(m) > 0\} \\ &+ \{x^m \mid m > m_3\,, \quad m \in N'\,, \quad U_2(m) = 0\,, \quad \varepsilon(m) > 0\} \end{aligned}$$

for some m_3 and some finite language L' containing only words P with $m_2 < l(P) \leqq m_3$. Denote the last two languages on the right by L_7 and L_8. The language L_7 can be proved regular almost in the same way as L_5. If there are values $m > m_3$ such that $m \in N'$ and $U_2(m) = 0$, then the above procedure is repeated by defining m^{k_3}, $U_3(m)$ and so on. By the assumption (5.1), this procedure must end. Finally, the language L_4 gets the form

$$L_4 = L'' + L_5 + L_7 + \ldots + L_t\,,$$

where L'' is a finite language and $L_5\,, L_7\,, \ldots\,, L_t$ are regular languages.

Now we have

$$L_3 = (L_0 + L_4) \cap \{x^m \mid m > m_1\,, \quad W(m) = \eta\}\,,$$

where L_0 is a finite language. Since $W(m)$ is a periodic function of m, it follows that the right member of the intersection is a regular language. Consequently, L_3 is regular, too. Since, by (5.3), $L(\mathfrak{P}\mathfrak{A}\,, \eta) = L_1 + L_2 + L_3$, we find that $L(\mathfrak{P}\mathfrak{A}\,, \eta)$ is a regular language.

The proof of Theorem 4 is now complete.

5.3. We omit the assumption (5.2) of Theorem 4 and prove the following theorem concerning the problem (a) of section 5.1.

Theorem 5. *Let* $\mathfrak{PA} = (\{s_1, \ldots, s_n\}, M, (p_1, \ldots, p_n), F)$ *be a probabilistic automaton over the alphabet* $\{x\}$. *Assume that, for some integers* s, q *and* m_0, $|\lambda_{s-1}| > |\lambda_s| = \ldots = |\lambda_{s+q}| > |\lambda_{s+q+1}|$ *and*

$$\sum_{s_j \in F} \sum_{i=1}^{n} \sum_{k=s}^{s+q} p_i \omega_{ijk}(m) \lambda_k^m \neq 0 \quad \textit{for all} \quad m > m_0 .$$

If $\arg \lambda_s, \ldots, \arg \lambda_{s+q}$ *and* $\arg \lambda$, *where* λ *runs through all eigenvalues of* $M(x)$ *satisfying the condition* $1 > |\lambda| > |\lambda_s|$, *are rational in degrees, then* $L(\mathfrak{PA}, \eta)$ *is a regular language for any cut-point* η.

Proof. As agreed in section 5.1, the distinct eigenvalues of $M(x)$ are labelled so that

$$1 = |\lambda_1| = \cdots = |\lambda_{n_0}| > |\lambda_{n_0+1}| \geqq \cdots \geqq |\lambda_r| .$$

We write $\pi_0 M(x^m) \pi_F$ in the form

$$\pi_0 M(x^m) \pi_F = W(m) + S_1(m) + \ldots + S_\gamma(m) + S_{\gamma+1}(m) + \varepsilon_1(m) ,$$

where, for each $i = 1, \ldots, \gamma$, $S_i(m)$ corresponds to the eigenvalues $\lambda_{n_{i-1}+1}, \ldots, \lambda_{n_i}$ with the same modulus and $S_{\gamma+1}(m)$ corresponds to $\lambda_s, \ldots, \lambda_{s+q}$.

By Lemma 7, we may assume that the cut-point η is a value of $W(m)$. If $S_1(m) + \ldots + S_\gamma(m)$ vanishes for all $m > m_0$, then the theorem follows from Theorem 4. In the remaining case we may assume that no one of the functions $S_1(m), \ldots, S_\gamma(m)$ vanishes for all $m > m_0$.

Let m^{k_1} be the highest power of m in $\lambda_{n_0+1}^{-m} S_1(m)$ whose coefficient does not vanish for all $m > m_0$. This power exists because of the assumption made above. Denote the coefficient by $U_1(m)$. Then $U_1(m)$ is a periodic function of m, since $\arg \lambda_{n_0+1}, \ldots, \arg \lambda_{n_1}$ are rational in degrees (cf. the proof of Theorem 4). If 0 is not a value of $U_1(m)$, then $L(\mathfrak{PA}, \eta)$ can be proved regular in the same manner as in Theorem 4. If 0 is a value of $U_1(m)$, then we choose the highest power of m in $S_1(m)$ whose coefficient does not vanish for all $m > m_0$ such that $U_1(m) = 0$. If this power does not exist, then we choose the first function $S_\mu(m)$ of the sequence $S_2(m), \ldots, S_{\gamma+1}(m)$ for which such a power exists. This function can be found because of the assumption of the theorem. We apply the same procedure to $S_\mu(m)$ as to $S_1(m)$. If this procedure does not end with $S_\mu(m)$, then we choose a new function from the sequence $S_{\mu+1}(m), \ldots, S_{\gamma+1}(m)$. The procedure ends at the latest with $S_{\gamma+1}(m)$. Finally, the language $L(\mathfrak{PA}, \eta)$ gets the form

$$(5.10) \qquad L(\mathfrak{PA}, \eta) = L_0 + \{x^m | m > m_1, \ W(m) > \eta\} + \{x^m | m > m_1, \ W(m) = \eta\} \cap (L_1 + \ldots + L_t) ,$$

where L_0 is a finite language and $L_1, \ldots, L_t$ are regular languages. Consequently, $L(\mathfrak{PA}, \eta)$ is a regular language, whence Theorem 5 follows.

5.4. Assume for a moment that all the eigenvalues of the transition matrix $M(x)$ have arguments rational in degrees. If the assumption of Theorem 5 is satisfied, then $L(\mathfrak{PA}, \eta)$ is regular for any cut-point η. If it is not satisfied, then we write

$$\pi_0 M(x^m)\pi_F = W(m) + S_1(m) + \ldots + S_\mu(m)$$

where the functions $S_i(m)$ are defined as in the proof of Theorem 5. Let η be a value of $W(m)$. We use the method of the proofs of Theorems 4 and 5. It is verified that the procedure, applied there, ends, because there are only finitely many functions $S_i(m)$. Furthermore, it ends without making use of functions $S_j(m)$ which contain eigenvalues λ such that $\arg \lambda$ is irrational in degrees, because such eigenvalues do not exist for $M(x)$. The language $L(\mathfrak{PA}, \eta)$ is of the form (5.10). Hence it is regular. We have thus established the following

Theorem 6. *Let* $\mathfrak{PA} = (S, M, \pi_0, F)$ *be a probabilistic automaton over the alphabet* $\{x\}$. *If the eigenvalues of the transition matrix* $M(x)$ *have arguments rational in degrees, then* $L(\mathfrak{PA}, \eta)$ *is a regular language for any cut-point* η.

As an immediate consequence we have

Corollary 2. *If the eigenvalues of the transition matrix are real, then* $L(\mathfrak{PA}, \eta)$ *is a regular language for any cut-point* η.

Since the eigenvalues of a stochastic 2×2 matrix are always real, we have

Corollary 3. *Only regular languages can be represented in two-state probabilistic automata over a one-letter alphabet.*

Corollary 3 and the example of section 3.6 show that a three-state probabilistic automaton over a one-letter alphabet is the most simple probabilistic automaton where a non-regular language can be represented.

§ 6. Probabilistic automata and non-regular languages

6.1. In what follows, the representability of non-regular languages in probabilistic automata over a one-letter alphabet is considered.

Let us assume that, for a given probabilistic automaton $\mathfrak{PA} = (\{s_1, \ldots, s_n\}, M, (p_1, \ldots, p_n), F)$ over $\{x\}$, the eigenvalues of $M(x)$ are simple. Then, by formula (1.5),

$$p_{ij}^{(m)} = \sum_{k=1}^{n} R_k\, x_i^{(k)}\, y_j^{(k)}\, \lambda_k^m$$

where

$$R_k = \Big(\sum_{\nu=1}^{n} x_\nu^{(k)}\, y_\nu^{(k)}\Big)^{-1} .$$

Consequently,

$$\pi_0 M(x^m)\pi_F = \sum_{k=1}^{n} R_k\, U_k\, V_k\, \lambda_k^m \tag{6.1}$$

where we have denoted

$$U_k = \sum_{i=1}^{n} p_i\, x_i^{(k)} , \quad V_k = \sum_{s_j \in F} y_j^{(k)} .$$

If the coefficients of the eigenvalues λ with $|\lambda| < 1$ do not all vanish in formula (6.1), then by omitting all the vanishing terms this formula can be written in the form

$$\pi_0 M(x^m)\pi_F = W(m) + S(\lambda_s^m, \ldots, \lambda_{s+q}^m) + \varepsilon_1(m) , \tag{6.2}$$

where $|\lambda_s| = \ldots = |\lambda_{s+q}| < 1$ and $\varepsilon_1(m)$ corresponds to eigenvalues λ with $|\lambda| < |\lambda_s|$. If, for some i $(s \leqq i \leqq s + q)$, λ_i is imaginary, then $\bar{\lambda}_i$ is an eigenvalue of $M(x)$. Furthermore, the coefficient of $\bar{\lambda}_i$ in (6.1) is $\bar{R}_i\, \bar{U}_i\, \bar{V}_i$. Since $|\bar{\lambda}_i| = |\lambda_i|$, it follows that $S(\lambda_s^m, \ldots, \lambda_{s+q}^m)$ contains the term $\bar{R}_i\, \bar{U}_i\, \bar{V}_i\, \bar{\lambda}_i^m$. Denoting

$$\varphi_j = \arg \lambda_j , \quad \psi_j = \arg R_j\, U_j\, V_j \quad (j = s, \ldots, s + q) ,$$

we obtain

$$R_i\, U_i\, V_i\, \lambda_i^m + \bar{R}_i\, \bar{U}_i\, \bar{V}_i\, \bar{\lambda}_i^m = 2|R_i\, U_i\, V_i||\lambda_i|^m \cos (m\varphi_i + \psi_i) .$$

In this way, all the conjugate terms in the sum $S(\lambda_s^m, \ldots, \lambda_{s+q}^m)$ are combined. Thus, formula (6.2) gets the form

$$\pi_0 M(x^m)\pi_F = W(m) + |\lambda_s|^m \Big(\sum_{i=s}^{s+p} u_i \cos (m\varphi_i + \psi_i) + \varepsilon_2(m)\Big) \tag{6.3}$$

where $0 \leqq p \leqq q + 1$, $\lim_{m \to \infty} \varepsilon_2(m) = 0$ and

$$u_i = \begin{cases} 2|R_i\, U_i\, V_i| & \text{if } \operatorname{Im} \lambda_i \neq 0 , \\ |R_i\, U_i\, V_i| & \text{if } \operatorname{Im} \lambda_i = 0 . \end{cases}$$

Here $\operatorname{Im} \lambda_i$ denotes the imaginary part of λ_i. Note that, by formula (6.3), $\varepsilon_2(m)$ is real.

6.2. Using formula (6.3) and the notations of the previous section, we now prove the following theorem concerning the problem (b) of section 5.1.

Theorem 7. *Let $\mathfrak{PA} = (S, M, \pi_0, F)$ be a probabilistic automaton over the alphabet $\{x\}$ such that $\pi_0 M(x^m)\pi_F$ is of the form* (6.3). *Assume that φ_s is irrational in degrees and $\varphi_{s+1}, \ldots, \varphi_{s+p}$ are rational in degrees. The language $L(\mathfrak{PA}, \eta)$ is non-regular if η is a value of $W(m)$ and $u_s > z$ where z means the largest modulus of the values of the periodic function*

$$S(m) = \sum_{i=s+1}^{s+p} u_i \cos(m\varphi_i + \psi_i)\,.$$

Remark 1. If $p = 0$, then we define $S(m) \equiv 0$. If the assumption concerning $\varphi_{s+1}, \ldots, \varphi_{s+p}$ is omitted, then the condition concerning u_s can be replaced by the condition $u_s > \limsup_{m\to\infty} |S(m)|$, which is satisfied, for example, if $u_s > u_{s+1} + \ldots + u_{s+p}$.

Proof. Assume that $u_s > z$, i.e., $u_s = z + \delta$ where $\delta > 0$. Denote by h the period of $W(m)$. Let w_r be an arbitrarily fixed value of $W(m)$ and n_r a natural number such that

$$W(n_r + ih) = w_r \quad (i = 0, 1, 2, \ldots)\,.$$

By formula (6.3),

$$\pi_0 M(x^m)\pi_F = W(m) + |\lambda_s|^m (u_s \cos(m\varphi_s + \psi_s) + S(m) + \varepsilon_2(m))\,.$$

Using this, we prove that $L(\mathfrak{PA}, w_r)$ is a non-regular language, which implies the theorem.

We first show that $L(\mathfrak{PA}, w_r)$ is an infinite language. Let m_0 be a natural number such that $|\varepsilon_2(m)| < \delta/2$ for all $m > m_0$. Applying Lemma 3 to the set of numbers $(n_r + ih)\varphi_s + \psi_s \pmod{2\pi}$, $i = 1, 2, \ldots$, we find that there exist infinitely many values of i for which $n_r + ih > m_0$ and

$$\cos((n_r + ih)\varphi_s + \psi_s) > (z + \delta/2)/(z + \delta)\,. \tag{6.4}$$

Since, in addition, $|S(m)| \leq z$, it follows that, for these values of i,

$$\pi_0 M(x^{n_r+ih})\pi_F > w_r + |\lambda_s|^{n_r+ih}((z+\delta)(z+\delta/2)/(z+\delta) - z - \delta/2) = w_r\,. \tag{6.5}$$

This implies that $L(\mathfrak{PA}, w_r)$ is an infinite language.

In order to prove that $L(\mathfrak{P}\mathfrak{A}, w_r)$ is a non-regular language, we assume, on the contrary, that it is regular. Hence, by Lemma 2, $L(\mathfrak{P}\mathfrak{A}, w_r)$ can be expressed in the form

$$L(\mathfrak{P}\mathfrak{A}, w_r) = L_1 + (x^{m_1} + \ldots + x^{m_k})(x^u)^* \tag{6.6}$$

where L_1 is a finite language, $k > 0$ and $u > 0$. Note that k cannot be zero, because $L(\mathfrak{P}\mathfrak{A}, w_r)$ is an infinite language. Let n_0 be a natural number such that $l(P) < n_0$ for every word $P \in L_1$.

There exist natural numbers a and b such that

$$h = au/b\ . \tag{6.7}$$

We apply Lemma 3 to the set of numbers

$$\{(n_r + ibh)\varphi_s + \psi_s \pmod{2\pi} \mid i = 1, 2, \ldots\}\ . \tag{6.8}$$

As earlier in this proof (cf. (6.4) and (6.5)), it is verified that there exists a natural number k_1 for which $n_r + k_1bh > n_0$ and $x^{n_r + k_1bh} \in L(\mathfrak{P}\mathfrak{A}, w_r)$. Hence, by formula (6.6), $n_r + k_1bh = m_\mu + tu$ where $1 \leqq \mu \leqq k$ and $t \geqq 0$. By (6.7), we now have

$$n_r = m_\mu + tu - k_1au\ . \tag{6.9}$$

If we again apply Lemma 3 to the set of numbers (6.8), we find that there exists a natural number $k_2 > k_1$ for which $\varepsilon_2(n_r + k_2bh) < \delta/2$ and

$$\cos((n_r + k_2bh)\varphi_s + \psi_s) < -(z + \delta/2)/(z + \delta)\ .$$

This implies that

$$\pi_0 M(x^{n_r + k_2bh})\pi_F < w_r$$

and, therefore,

$$x^{n_r + k_2au} = x^{n_r + k_2bh} \notin L(\mathfrak{P}\mathfrak{A}, w_r)\ . \tag{6.10}$$

On the other hand, by (6.9),

$$n_r + k_2au = m_\mu + (t - k_1a + k_2a)u\ . \tag{6.11}$$

Here $t - k_1a + k_2a > 0$, because $k_2 > k_1$. By (6.6) and (6.11), we now conclude that the word $x^{n_r + k_2au}$ belongs to the language $L(\mathfrak{P}\mathfrak{A}, w_r)$, which contradicts (6.10). Consequently, $L(\mathfrak{P}\mathfrak{A}, w_r)$ is a non-regular language, whence Theorem 7 follows.

6.3. For probabilistic automata satisfying the conditions of Theorem 7, every value of $W(m)$ is a non-regular cut-point. In the following theorem concerning the problem (c) of section 5.1, we weaken the condition related to u_s and show that at least one non-regular cut-point exists.

Theorem 8. *Let* $\mathfrak{PA} = (S, M, \pi_0, F)$ *be a probabilistic automaton over the alphabet* $\{x\}$ *such that* $\pi_0 M(x^m)\pi_F$ *is of the form* (6.3). *Assume that* φ_s *is irrational in degrees and* $\varphi_{s+1}, \ldots, \varphi_{s+p}$ *are rational in degrees. The language* $L(\mathfrak{PA}, \eta)$ *is non-regular for at least one cut-point* η *if* $u_s > z$ *where* z *means the least modulus of the values of the periodic function*

$$S(m) = \sum_{i=s+1}^{s+p} u_i \cos(m\varphi_i + \psi_i) .$$

Proof. Denote the periods of $W(m)$ and $S(m)$ by h and l, respectively. Then there exists a natural number n_r such that $|S(n_r + il)| = z$ $(i = 0, 1, 2, \ldots)$. Denote $w_r = W(n_r)$. Then, for each $i \geqq 0$,

$$|S(n_r + ihl)| = z , \quad W(n_r + ihl) = w_r . \tag{6.12}$$

Theorem 8 is established by proving that the language $L(\mathfrak{PA}, w_r)$ is non-regular. The proof is exactly the same as that of Theorem 7 if one replaces h by hl and remembers the equations (6.12).

As an example we consider the probabilistic automaton

$$\mathfrak{PA} = (\{s_1, \ldots, s_6\}, M, (0, \gamma, \gamma, 1 - 2\gamma, 0, 0), \{s_2, s_3, s_4\})$$

where $0 < \gamma < 1/2$ and

$$M(x) = \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix} \tag{6.13}$$

with

$$M_1 = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 3/4 & 1/4 & 0 \\ 1/4 & 1/2 & 1/4 \end{bmatrix} , \quad M_2 = \begin{bmatrix} 3/8 & 1/4 & 3/8 \\ 1/8 & 1/4 & 5/8 \\ 1/3 & 0 & 2/3 \end{bmatrix} .$$

Note that M_2 is the same matrix as in section 3.6. It follows that, for any $m \geqq 1$,

$$M(x^m) = \begin{bmatrix} M_1^m & 0 \\ 0 & M_2^m \end{bmatrix} .$$

The eigenvalues of $M(x)$ are $\lambda_1 = 1$, $\lambda_2 = 1$, $\lambda_3 = 1/4$, $\lambda_4 = -1/4$, $\lambda_5 = (7 + \sqrt{95}\, i)/48$ and $\lambda_6 = \bar{\lambda}_5$. Hence, $|\lambda_3| = \cdots = |\lambda_4| = 1/4$. Formula (1.5) gives

$$M_1^m = \begin{bmatrix} \frac{3}{5} + \frac{2}{5}\left(-\frac{1}{4}\right)^m & \frac{2}{5} - \frac{2}{5}\left(-\frac{1}{4}\right)^m & 0 \\ \frac{3}{5} - \frac{3}{5}\left(-\frac{1}{4}\right)^m & \frac{2}{5} + \frac{3}{5}\left(-\frac{1}{4}\right)^m & 0 \\ \frac{3}{5} - \left(\frac{1}{4}\right)^m + \frac{2}{5}\left(-\frac{1}{4}\right)^m & \frac{2}{5} - \frac{2}{5}\left(-\frac{1}{4}\right)^m & \left(\frac{1}{4}\right)^m \end{bmatrix} .$$

In the same manner we find that in M_2^m the element $p_{11}^{(m)}$ is of the form

$$p_{11}^{(m)} = 12/37 + |\lambda_5|^m u_5 \cos (m\varphi_5 + \psi_5)$$

where $u_5 > 0$ and $\varphi_5 = \arg \lambda_5$ is irrational in degrees (cf. section 3.6). Now we have

$$\pi_0 M(x^m)\pi_F = \frac{4}{5}\,\gamma + \frac{12}{37}\,(1 - 2\gamma) + \frac{1}{5}\,\gamma\left(-\frac{1}{4}\right)^m + \gamma\left(\frac{1}{4}\right)^m +$$
$$(1 - 2\gamma)u_5\left(\frac{1}{4}\right)^m \cos (m\varphi_5 + \psi_5) .$$

Although $\lambda_1 = 1$ is not simple, we may use Theorem 8, since $\pi_0 M(x^m)\pi_F$ is of the same form as in this theorem. Now $W(m) = 4\gamma/5 + 12(1 - 2\gamma)/37$, $\lambda_s = \lambda_5$, $u_s = (1 - 2\gamma)u_5$ and $S(m) = \gamma + \gamma(-1)^m/5$. The least value of $|S(m)|$ is $z = 4\gamma/5$. The condition $u_s > z$ is satisfied if

$$0 < \gamma < u_5/(4/5 + 2u_5).$$

Thus, the language $L(\mathfrak{P}\mathfrak{A}\,, 4\gamma/5 + 12(1 - 2\gamma)/37)$ is non-regular whenever γ satisfies the above condition.

6.4. If the coefficients of the eigenvalues λ with $0 < |\lambda| < 1$ do not all vanish in formula (6.1), then, for some integers s and q, $1 > |\lambda_s| = \cdots = |\lambda_{s+q}| > 0$ and

$$\pi_0 M(x^m)\pi_F = W(m) + R(m) + S(\lambda_s^m\,, \ldots, \lambda_{s+q}^m) + \varepsilon_1(m)$$

where $R(m)$ and $\varepsilon_1(m)$ are the parts of (6.1) corresponding to eigenvalues λ with $1 > |\lambda| > |\lambda_s|$ and $|\lambda| < |\lambda_s|$, respectively. The following formula, corresponding to (6.3), can now be derived:

$$(6.14) \quad \pi_0 M(x^m)\pi_F = W(m) + R(m) + |\lambda_s|^m \Big(\sum_{i=s}^{s+p} u_i \cos (m\varphi_i + \psi_i) + \varepsilon_2(m)\Big) .$$

Here $u_i > 0$ and $\varepsilon_2(m)$ is a real function tending to zero as m tends to infinity.

The following theorem shows that in some cases the conditions of Theorem 5 are also necessary for the language $L(\mathfrak{P}\mathfrak{A}\,, \eta)$ to be regular for every cut-point.

Theorem 9. *Let* $\mathfrak{PA} = (S, M, \pi_0, F)$ *be a probabilistic automaton over the alphabet* $\{x\}$ *such that* $\pi_0 M(x^m)\pi_F$ *is of the form* (6.14). *A non-regular language can be represented in* $\mathfrak{PA}$ *if* $R(m)$ *vanishes periodically,* φ_s *is irrational in degrees and* $u_s > u_{s+1} + \ldots + u_{s+p}$.

Remark 2. The last condition can be replaced by the condition

$$u_s > \limsup_{m \to \infty} \left| \sum_{i=s+1}^{s+p} u_i \cos(m\varphi_i + \psi_i) \right| .$$

Proof. There exist natural numbers n_r and l such that $R(n_r + il) = 0$ $(i = 0, 1, 2, \ldots)$. Let h be the period of $W(m)$ and $w_r = W(n_r)$. Then, for each $i \geqq 0$,

$$W(n_r + ihl) = w_r, \quad R(n_r + ihl) = 0 .$$

As in the proofs of Theorems 7 and 8, it can be verified that the language $L(\mathfrak{PA}, w_r)$ is infinite but not regular, whence the theorem follows.

As an example we consider the probabilistic automaton

$$\mathfrak{PA} = (\{s_1, \ldots, s_6\}, M, (0, 1/4, 1/4, 1/2, 0, 0), \{s_2, s_4\})$$

where $M(x)$ is the matrix (6.13) with

$$M_1 = \begin{bmatrix} 1/4 & 3/4 & 0 \\ 3/4 & 1/4 & 0 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 3/8 & 1/4 & 3/8 \\ 1/8 & 1/4 & 5/8 \\ 1/3 & 0 & 2/3 \end{bmatrix}.$$

Note that M_2 is the same matrix as in the example of section 6.3. The eigenvalues of $M(x)$ are $\lambda_1 = 1$, $\lambda_2 = 1$, $\lambda_3 = 1/2$, $\lambda_4 = -1/2$, $\lambda_5 = (7 + \sqrt{95}\, i)/48$ and $\lambda_6 = \bar{\lambda}_5$. Hence, we obtain (cf. the example of section 6.3)

$$\pi_0 M(x^m)\pi_F = \frac{61}{148} + \frac{1}{8}\left(\left(-\frac{1}{2}\right)^m - \left(\frac{1}{2}\right)^m\right) + \frac{1}{2} u_5 |\lambda_5|^m \cos(m\varphi_5 + \psi_5)$$

where $u_5 > 0$ and $\varphi_5 = \arg \lambda_5$ is irrational in degrees. Although $\lambda_1 = 1$ is not simple, we may use Theorem 9, since $\pi_0 M(x^m)\pi_F$ is of the same form as in this theorem. Here $W(m) = 61/148$, $R(m) = ((-1/2)^m - (1/2)^m)/8$, $u_s = u_5/2$, $\lambda_s = \lambda_5$ and $p = 0$. We find that $R(m)$ vanishes periodically, the period being $l = 2$. The number u_s satisfies the condition of Theorem 9, since $u_s > 0$. Thus, $L(\mathfrak{PA}, 61/148)$ is a non-regular language.

CHAPTER II

PROBABILISTIC AND GENERALIZED AUTOMATA OVER AN ARBITRARY ALPHABET

§ 7. A theorem concerning two-state probabilistic automata

7.1. In this chapter we investigate probabilistic automata over an alphabet I without making any assumption on the number of its elements. We first establish the following

Theorem 10. *Let* $\mathfrak{PA} = (\{s_1, s_2\}, M, (p_1, p_2), F)$ *be a probabilistic automaton over the alphabet* $I = \{x_1, \ldots, x_k\}$. *If the transition matrices*

$$M(x_i) = \begin{bmatrix} 1 - a_i & a_i \\ b_i & 1 - b_i \end{bmatrix} \quad (i = 1, \ldots, k)$$

satisfy the conditions

$$M(x_i) \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{7.1}$$

and

$$a_i b_j = a_j b_i \quad (i, j = 1, \ldots, k), \tag{7.2}$$

then the language $L(\mathfrak{PA}, \eta)$ *is regular for any cut-point* η.

Proof. Without loss of generality, we assume that $F = \{s_1\}$. Let 1 and λ_i be the eigenvalues of $M(x_i)$. Then λ_i is real, because $M(x_i)$ is a stochastic 2×2 matrix. On account of the condition (7.1) we have $|\lambda_i| < 1$ and $a_i + b_i > 0$ $(i = 1, \ldots, k)$. Denote

$$A_i = \frac{1}{a_i + b_i} \begin{bmatrix} b_i & a_i \\ b_i & a_i \end{bmatrix}, \quad B_i = \frac{1}{a_i + b_i} \begin{bmatrix} a_i & -a_i \\ -b_i & b_i \end{bmatrix}.$$

Then

$$M(x_i) = A_i + B_i \lambda_i \quad (i = 1, \ldots, k).$$

From the condition (7.2) it follows that $A_1 = \ldots = A_k$ and $B_1 = \ldots = B_k$. Thus

$$M(x_i) = A_1 + B_1 \lambda_i \quad (i = 1, \ldots, k). \tag{7.3}$$

Clearly $A_1 A_1 = A_1$, $B_1 B_1 = B_1$, $A_1 B_1 = 0$, $B_1 A_1 = 0$. By (7.3), this implies that, for any non-empty word $x_{\nu_1} \cdots x_{\nu_s}$,

$$M(x_{\nu_1} \cdots x_{\nu_s}) = A_1 + B_1 \prod_{i=1}^{s} \lambda_{\nu_i} . \tag{7.4}$$

Denote $a_1/(a_1 + b_1) = a$ and $b_1/(a_1 + b_1) = b$. From the equation (7.4) we now obtain

$$\pi_0 M(x_{\nu_1} \cdots x_{\nu_s}) \pi_F = b + (a - p_2) \prod_{i=1}^{s} \lambda_{\nu_i} . \tag{7.5}$$

Now we divide the eigenvalues $\lambda_1 , \ldots , \lambda_k$ into three classes

$$\{\lambda_{i_1} , \ldots , \lambda_{i_\alpha}\} , \{\lambda_{j_1} , \ldots , \lambda_{j_\beta}\} , \{\lambda_{k_1} , \ldots , \lambda_{k_\gamma}\}$$

where $\alpha , \beta , \gamma \geqq 0$, $\lambda_{i_\nu} = 0$, $\lambda_{j_\nu} > 0$ and $\lambda_{k_\nu} < 0$.

We consider the language $L(\mathfrak{PA} , \eta)$ for different values of η. When doing this, we use formula (7.5) and remember that $|\lambda_i| < 1$ $(i = 1 , \ldots , k)$.

Assume first that $\eta = b$. If $a = p_2$, then $b = p_1$ and $L(\mathfrak{PA} , b)$ is the empty language. If $a > p_2$, then $L(\mathfrak{PA} , b)$ is a language over the alphabet $\{x_{j_1} , \ldots , x_{j_\beta} , x_{k_1} , \ldots , x_{k_\gamma}\}$ containing exactly the words where the total number of letters x_{k_ν} $(\nu = 1 , \ldots , \gamma)$ is even. If $a < p_2$, then the statement of the previous sentence holds with »even» replaced by »odd».

If $\eta > b$, then $L(\mathfrak{PA} , \eta)$ is a finite language. Finally, assume that $\eta < b$. If $a = p_2$, then $b = p_1$ and $L(\mathfrak{PA} , \eta) = W(I)$. If $a \neq p_2$, then $L(\mathfrak{PA} , \eta) = L_1 + L_2$ where L_1 and L_2 are formed as follows. L_1 consists of the words P such that $l(P) > m_0$ where m_0 is a sufficiently large integer, and L_2 contains only words P with $l(P) \leqq m_0$.

From the above considerations it follows that $L(\mathfrak{PA} , \eta)$ is a regular language for any cut-point η.

§ 8. Theorems concerning stochastic languages

8.1. Bukharaev [3] showed that every η-stochastic language $(\eta > 0)$ is also η_1-stochastic for any η_1 such that $0 < \eta_1 < \eta$. The following theorem generalizes this result.

Theorem 11. *Every stochastic language is η_1-stochastic for any η_1 such that $0 < \eta_1 < 1$.*

Proof. Let $L = L(\mathfrak{PA} , \eta)$, where $\mathfrak{PA} = (\{s_1 , \ldots , s_n\} , M , \pi_0 , F)$ is a probabilistic automaton over the alphabet I. Let η_1 be fixed, $0 < \eta_1 < 1$. By omitting the trivial case $\eta_1 = \eta$, we may assume that $\eta_1 \neq \eta$. Consider the probabilistic automaton $\mathfrak{P}\mathfrak{A}_1 = (\{s_1, \ldots, s_n, s_{n+1}\}, M_1, \pi_1, F_1)$ where, for each $x \in I$,

$$M_1(x) = \begin{bmatrix} M(x) & 0 \\ 0 & 1 \end{bmatrix} \tag{8.1}$$

and π_1 as well as F_1 are defined as follows. If $\eta_1 < \eta$, then $\pi_1 = ((\eta_1/\eta)\pi_0, 1 - \eta_1/\eta)$ and $F_1 = F$. If $\eta_1 > \eta$, then $\pi_1 = (p\pi_0, 1 - p)$ and $F_1 = F \cup \{s_{n+1}\}$ where $p = (1 - \eta_1)/(1 - \eta)$. From (8.1) it follows that, for any $P \in W(I)$,

$$M_1(P) = \begin{bmatrix} M(P) & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus, for any word $P \in W(I)$,

$$\pi_1 M_1(P) \pi_{F_1} = \begin{cases} (\eta_1/\eta)\pi_0 M(P)\pi_F & \text{if } \eta_1 < \eta, \\ p\pi_0 M(P)\pi_F + (1 - p) & \text{if } \eta_1 > \eta. \end{cases} \tag{8.2}$$

From (8.2) and the choice of p it follows that in both cases $L = L(\mathfrak{P}\mathfrak{A}_1, \eta_1)$, which implies our theorem.

According to Theorem 11, every stochastic language is, for example, $\frac{1}{2}$-stochastic. The restriction $\eta_1 > 0$ is essential, because every 0-stochastic language is regular. Conversely, every regular language is 0-stochastic, because every finite deterministic automaton can be rewritten as a probabilistic automaton where the initial distribution and the transition matrices consist of 0's and 1's only.

8.2. PAZ [12] constructed a probabilistic automaton where the intersection of a stochastic language and a regular language can be represented. In the following theorem we construct a more economical probabilistic automaton where both the sum and the intersection of such languages can be represented.

Theorem 12. *The sum $L_1 + L_2$ and the intersection $L_1 \cap L_2$ of a stochastic language L_1 and a regular language L_2 (over the alphabet I) are both stochastic languages.*

Proof. Let $L_1 = L(\mathfrak{P}\mathfrak{A}_1, \eta)$, where

$$\mathfrak{P}\mathfrak{A}_1 = (\{s_1, \ldots, s_n\}, M_1, (p_1, \ldots, p_n), F_1)$$

is a probabilistic automaton over I. Let $L_2 = L(\mathfrak{P}\mathfrak{A}_2, 0)$, where

$$\mathfrak{P}\mathfrak{A}_2 = (\{s_{n+1}, \ldots, s_{n+m}\}, M_2, (1, 0, \ldots, 0), F_2)$$

and each row in each $M_2(x)$ is a co-ordinate vector. Thus, $\mathfrak{P}\mathfrak{A}_2$ is a finite deterministic automaton, where L_2 is represented, rewritten as a probabilistic automaton.

Consider the $(m + n)$-state probabilistic automaton

$$\mathfrak{P}\mathfrak{A} = (\{s_1 , \ldots , s_n , s_{n+1} , \ldots , s_{n+m}\} , M , \pi_0 , F)$$

over I where, for each $x \in I$,

$$M(x) = \begin{bmatrix} M_1(x) & 0 \\ 0 & M_2(x) \end{bmatrix} ,$$

$\pi_0 = (\frac{1}{2}p_1 , \ldots , \frac{1}{2}p_n , \frac{1}{2} , 0 , \ldots , 0)$ and $F = F_1 \cup F_2$. Consequently, for each $P \in W(I)$,

$$M(P) = \begin{bmatrix} M_1(P) & 0 \\ 0 & M_2(P) \end{bmatrix} .$$

Since, in addition,

$$\pi_F = \begin{bmatrix} \pi_{F_1} \\ \pi_{F_2} \end{bmatrix} ,$$

the following equation holds for any word $P \in W(I)$:

(8.3) $\quad \pi_0 M(P)\pi_F = \frac{1}{2}(p_1 , \ldots , p_n)M_1(P)\pi_{F_1} + \frac{1}{2}(1 , 0 , \ldots , 0)M_2(P)\pi_{F_2} .$

Since $(1 , 0 , \ldots , 0)M_2(P)\pi_{F_2}$ equals 1 or 0, depending on whether $P \in L_2$ or $P \notin L_2$, (8.3) implies the equations

$$L_1 + L_2 = L(\mathfrak{P}\mathfrak{A} , \tfrac{1}{2}\eta)$$

and

$$L_1 \cap L_2 = L(\mathfrak{P}\mathfrak{A} , \tfrac{1}{2}(\eta + 1)) .$$

This proves the theorem.

As an immediate corollary we obtain the following

Theorem 13. *If L_1 is a stochastic language and L_2 a regular language, then the language $L_1 - L_2$ is stochastic. Let L' be a language obtained from a language L by adding or removing finitely many words. Then L' is stochastic if and only if L is stochastic.*

The *left derivative* of a language L with respect to a word P is defined by $\partial_P L = \{Q \mid PQ \in L\}$. Bukharaev [3] has proved the following result:

Lemma 8. *All left derivatives of a stochastic language are stochastic languages. Conversely, if there exists an integer k such that all left derivatives of a language L with respect to words of length k are stochastic languages, then L is a stochastic language.*

Our next theorem follows from this lemma, Theorem 12 and the equations

$$\partial_P(L_1 + L_2) = \partial_P L_1 + \partial_P L_2 \,, \quad \partial_P(L_1 \cap L_2) = (\partial_P L_1) \cap (\partial_P L_2) \,.$$

Theorem 14. *Let L_1 and L_2 be stochastic languages. If there exists an integer k such that, for each word P of length k, either $\partial_P L_1$ or $\partial_P L_2$ is a regular language, then $L_1 + L_2$ and $L_1 \cap L_2$ are stochastic languages.*

8.3. For a probabilistic automaton $\mathfrak{PA} = (S\,, M\,, \pi_0\,, F)$ and a cut-point η, we denote

$$L(\mathfrak{PA}\,, \eta\,, =) = \{P \mid \pi_0 M(P)\pi_F = \eta\} \,.$$

Now the following result concerning closure under complementation can be established.

Theorem 15. *Assume that $L = L(\mathfrak{PA}\,, \eta)$ is a stochastic language and that the language $L(\mathfrak{PA}\,, \eta\,, =)$ is regular. Then the complement of L is a stochastic language.*

Proof. Let $\mathfrak{PA} = (S\,, M\,, \pi_0\,, F)$. Denote by $\bar{F}$ the complement of F with respect to S. Then

$$\begin{aligned} \bar{L} &= \{P \mid \pi_0 M(P)\pi_F \leqq \eta\} \\ &= \{P \mid \pi_0 M(P)\pi_{\bar{F}} > 1 - \eta\} + L(\mathfrak{PA}\,, \eta\,, =) \,, \end{aligned}$$

whence Theorem 15 follows, by Theorem 12.

The hypothesis of the previous theorem is not satisfied for all stochastic languages; STARKE [18] constructed a probabilistic automaton over the alphabet $\{0\,, 1\}$ for which

$$L(\mathfrak{PA}\,, \eta\,, =) = \{0^n 1 0^n 1 \mid n \geqq 1\} + 011 \,. \tag{8.4}$$

This language is not regular.

We return to the closure problems in § 10.

§ 9. Generalized probabilistic automata

9.1. PAGE [11] has considered probabilistic automata for which the components of π_F as well as the cut-point η are allowed to be arbitrary real numbers. An open problem has been whether the family of representable languages remains unaltered in this generalization. In what follows, we solve this problem.

Definition 2. A *generalized probabilistic automaton* over the alphabet I is an ordered quadruple $\mathfrak{GPA} = (S , M , \pi_0 , f_0)$, where S, M and π_0 are as in Definition 1 and f_0 is an n-dimensional column vector with real components (the *final vector*). If f_0 is a stochastic vector, then $\mathfrak{GPA}$ is a generalized probabilistic automaton with a *final distribution.*

The domain of M is extended from I to $W(I)$ in the same way as before.

For any real number η, the language represented in $\mathfrak{GPA}$ with the cut-point η is defined to be the set

$$L(\mathfrak{GPA} , \eta) = \{P \in W(I) \mid \pi_0 M(P) f_0 > \eta\} .$$

Lemma 9. *Assume that* $L = L(\mathfrak{GPA} , \eta)$ *where* $\mathfrak{GPA} = (S , M , \pi_0 , f_0)$. *For any real number* c, $L = L(\mathfrak{GPA}_1 , \eta + c)$ *where* $\mathfrak{GPA}_1 = (S , M , \pi_0 , f_1)$ *and* $f_1 = f_0 + (c , \ldots , c)^T$.

Proof. For any word $P \in W(I)$,

$$\pi_0 M(P) f_1 = \pi_0 M(P) f_0 + c ,$$

because $\pi_0 M(P)$ is a stochastic vector. This proves the lemma.

Theorem 16. *A language* L *can be represented in a generalized probabilistic automaton (generalized probabilistic automaton with a final distribution) if and only if it can be represented in a probabilistic automaton.*

Proof. To establish the »if»-part, we assume that $L = L(\mathfrak{PA} , \eta)$, where $\mathfrak{PA} = (S , M , \pi_0 , F)$ is a probabilistic automaton over I. Denote by k the number of the elements of F $(k > 0)$. Then the generalized probabilistic automaton $\mathfrak{GPA} = (S , M , \pi_0 , k^{-1}\pi_F)$ has a final distribution and $L(\mathfrak{PA} , \eta) = L(\mathfrak{GPA} , \eta/k)$.

To prove the »only if»-part, we assume that L can be represented in a generalized probabilistic automaton. By choosing c large enough in

Lemma 9, we may assume that $L = L(\mathfrak{G}\mathfrak{P}\mathfrak{A}, \eta)$ where $\mathfrak{G}\mathfrak{P}\mathfrak{A} = (\{s_1, \ldots, s_n\}, M, \pi_0, f_0)$ and the components of f_0 are positive. Denote

$$\pi_0 = (p_1, \ldots, p_n), \quad f_0 = (q_1, \ldots, q_n)^T$$

and $q = q_1 + \ldots + q_n$. Thus $q_i > 0$ $(i = 1, \ldots, n)$ and $q > 0$. Let $q_i' = q_i/q$ for each $i = 1, \ldots, n$. Consider the probabilistic automaton

$$\mathfrak{P}\mathfrak{A} = (\{s_1, \ldots, s_n, s_{n+1}, \ldots, s_{n^2}\}, M_1, \pi_1, F_1)$$

where

$$\pi_1 = n^{-1}(\pi_0, \ldots, \pi_0) \quad (\pi_0 \text{ occurs } n \text{ times})$$

and

$$F_1 = \{s_1, s_{n+2}, s_{2n+3}, \ldots, s_{n^2}\}$$

and, for each $x \in I$,

$$M_1(x) = \begin{bmatrix} q_1'M(x) & q_2'M(x) \ldots q_n'M(x) \\ q_1'M(x) & q_2'M(x) \ldots q_n'M(x) \\ \cdot & \cdot \quad \ldots \quad \cdot \\ q_1'M(x) & q_2'M(x) \ldots q_n'M(x) \end{bmatrix}. \tag{9.1}$$

Clearly, $M_1(x)$ is a stochastic $n^2 \times n^2$ matrix and π_1 is an n^2-dimensional stochastic row vector.

Let $P \in W(I)$ be an arbitrary non-empty word. From the construction of $M_1(x)$ it follows that (9.1) holds if x is replaced by the word P. Using this result, it can be verified that

$$\pi_1 M_1(P) \pi_{F_1} = \sum_{i=1}^{n} \sum_{j=1}^{n} p_i q_j' p_{ij}(P) \tag{9.2}$$

where $p_{ij}(P)$ denotes the (i, j)th element of the matrix $M(P)$. On the other hand,

$$\pi_0 M(P) f_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} p_i q_j p_{ij}(P). \tag{9.3}$$

Combining (9.2) and (9.3), we obtain

$$\pi_1 M_1(P) \pi_{F_1} = q^{-1} \pi_0 M(P) f_0 .$$

Consequently, the language $L(\mathfrak{P}\mathfrak{A}, \eta/q)$ contains exactly the same non-empty words as the language $L(\mathfrak{G}\mathfrak{P}\mathfrak{A}, \eta)$, whence our theorem follows, by Theorem 13.

Thus, a final vector can always be reduced to a vector whose components are 0's and 1's only. Note that the initial distribution can always be reduced to a co-ordinate vector (cf. [17]).

As an immediate corollary of Lemma 9 we obtain

Theorem 17. *If* $L = L(\mathfrak{GPA}, \eta)$, *then, for any real number* η_1, *there exists a generalized probabilistic automaton* $\mathfrak{GPA}_1$ *such that* $L = L(\mathfrak{GPA}_1, \eta_1)$.

Thus, every language of the form $L(\mathfrak{GPA}, \eta)$ can be represented in some $\mathfrak{GPA}_1$ with the cut-point 0 (cf. Theorem 11).

More general automata will be considered in § 11.

§ 10. Closure properties of a subfamily of stochastic languages

10.1. Let $\mathfrak{GPA}_1 = (S_1, M_1, \pi_1, f_1)$ and $\mathfrak{GPA}_2 = (S_2, M_2, \pi_2, f_2)$ be generalized probabilistic automata over I with n and m states, respectively. Denote $\pi_1 = (p_1, \ldots, p_n)$, $f_1 = (q_1, \ldots, q_n)^T$ and, for each $x \in I$, $M_1(x) = [p_{ij}(x)]$. Consider the mn-state generalized probabilistic automaton $\mathfrak{GPA} = (S_1 \times S_2, M, \pi_0, f_0)$ where

$$\pi_0 = (p_1\pi_2, p_2\pi_2, \ldots, p_n\pi_2), \quad f_0 = \begin{bmatrix} q_1 f_2 \\ \cdot \\ \cdot \\ \cdot \\ q_n f_2 \end{bmatrix}$$

and, for each $x \in I$,

$$M(x) = M_1(x) \otimes M_2(x) = \begin{bmatrix} p_{11}(x)M_2(x) & p_{12}(x)M_2(x) & \ldots & p_{1n}(x)M_2(x) \\ p_{21}(x)M_2(x) & p_{22}(x)M_2(x) & \ldots & p_{2n}(x)M_2(x) \\ \cdot & \cdot & \ldots & \cdot \\ p_{n1}(x)M_2(x) & p_{n2}(x)M_2(x) & \ldots & p_{nn}(x)M_2(x) \end{bmatrix}.$$

Thus, $M(x)$ is the Kronecker product of $M_1(x)$ and $M_2(x)$ (cf. [7], p. 97, [12]). From the construction of the matrices $M(x)$ it follows that, for any non-empty word $P \in W(I)$, $M(P) = M_1(P) \otimes M_2(P)$. Clearly, this also holds for $P = \Lambda$. By a direct computation, we now verify that, for any word $P \in W(I)$,

$$\pi_0 M(P) f_0 = (\pi_1 M_1(P) f_1)(\pi_2 M_2(P) f_2). \tag{10.1}$$

Thus, we have established the following

Lemma 10. *For two generalized probabilistic automata* $\mathfrak{GPA}_1 = (S_1, M_1, \pi_1, f_1)$ *and* $\mathfrak{GPA}_2 = (S_2, M_2, \pi_2, f_2)$ *over* I, *there exists a generalized probabilistic automaton* $\mathfrak{GPA} = (S, M, \pi_0, f_0)$ *such that* (10.1) *holds for any word* $P \in W(I)$.

In what follows, we also need the following

Lemma 11. *For two generalized probabilistic automata* $\mathfrak{GPA}_1 = (S_1, M_1, \pi_1, f_1)$ *and* $\mathfrak{GPA}_2 = (S_2, M_2, \pi_2, f_2)$ *over* I, *there exists a generalized probabilistic automaton* $\mathfrak{GPA} = (S, M, \pi_0, f_0)$ *such that, for any word* $P \in W(I)$,

$$\pi_0 M(P) f_0 = \pi_1 M_1(P) f_1 + \pi_2 M_2(P) f_2 .$$

Proof. By defining

$$\pi_0 = \tfrac{1}{2}(\pi_1, \pi_2), \quad f_0 = \begin{bmatrix} 2f_1 \\ 2f_2 \end{bmatrix}$$

and, for each $x \in I$, $M(x)$ as in the proof of Theorem 12, we obtain the desired automaton.

10.2. For any generalized probabilistic automaton $\mathfrak{GPA} = (S, M, \pi_0, f_0)$ over I, we denote

$$L(\mathfrak{GPA}, \eta, \neq) = \{P \mid \pi_0 M(P) f_0 \neq \eta\}$$

and, respectively,

$$L(\mathfrak{GPA}, \eta, =) = \{P \mid \pi_0 M(P) f_0 = \eta\} .$$

Let $\mathcal{L}(\neq)$ be the family consisting of the languages L over I such that, for some $\mathfrak{GPA}$ and some η, $L = L(\mathfrak{GPA}, \eta, \neq)$. The family $\mathcal{L}(=)$ is defined analogously.

In what follows, we also use the notations $\mathcal{L}_r$ and $\mathcal{L}_s$ to mean, respectively, the family of regular languages and the family of stochastic languages over I.

Lemma 9 and Theorem 16 are useful tools in constructions involving probabilistic automata. We shall use them in the following considerations.

Theorem 18. *The family* $\mathcal{L}_r$ *is a proper subset of the family* $\mathcal{L}(\neq)$. *The family* $\mathcal{L}(\neq)$ *is a subset of the family* $\mathcal{L}_s$.

Proof. For each $L \in \mathcal{L}_r$, there exists a probabilistic automaton $\mathfrak{PA} = (S, M, \pi_0, F)$ such that $L = \{P \mid \pi_0 M(P) \pi_F \neq 0\}$. Thus, $\mathcal{L}_r$ is a subset of $\mathcal{L}(\neq)$. The first part of the theorem now follows from the fact that the complement of the language (8.4) is not regular.

To establish the last part, we choose an arbitrary language L from $\mathcal{L}(\neq)$. From the proof of Lemma 9 it follows that, for some $\mathfrak{G}\mathfrak{P}\mathfrak{A}_1 = (S_1 , M_1 , \pi_1 , f_1)$ over I, $L = L(\mathfrak{G}\mathfrak{P}\mathfrak{A}_1 , 0 , \neq)$. By Lemma 10, there exists a generalized probabilistic automaton $\mathfrak{G}\mathfrak{P}\mathfrak{A} = (S , M , \pi_0 , f_0)$ such that, for any word $P \in W(I)$,

$$\pi_0 M(P) f_0 = (\pi_1 M_1(P) f_1)^2 .$$

This implies that $L = L(\mathfrak{G}\mathfrak{P}\mathfrak{A} , 0)$, whence the last part of Theorem 18 follows, by Theorem 16.

Theorem 19. *The family $\mathcal{L}(\neq)$ is closed under sum and intersection.*

Proof. Let $L_1 \in \mathcal{L}(\neq)$ and $L_2 \in \mathcal{L}(\neq)$ be arbitrarily fixed. By the proof of Lemma 9, we may assume that $L_i = L(\mathfrak{G}\mathfrak{P}\mathfrak{A}_i , 0 , \neq)$ where $\mathfrak{G}\mathfrak{P}\mathfrak{A}_i = (S_i , M_i , \pi_i , f_i)$ $(i = 1 , 2)$. By Lemmas 10 and 11, there exists a generalized probabilistic automaton $\mathfrak{G}\mathfrak{P}\mathfrak{A} = (S , M , \pi_0 , f_0)$ such that, for each word $P \in W(I)$,

$$\pi_0 M(P) f_0 = (\pi_1 M_1(P) f_1)^2 + (\pi_2 M_2(P) f_2)^2 .$$

This implies that $L_1 + L_2 = L(\mathfrak{G}\mathfrak{P}\mathfrak{A} , 0 , \neq)$, whence the closure under sum follows.

To establish the closure under intersection, let L_1 and L_2 be as above. By Lemma 10, there exists a generalized probabilistic automaton $\mathfrak{G}\mathfrak{P}\mathfrak{A} = (S , M , \pi_0 , f_0)$ such that, for each word $P \in W(I)$, the equation (10.1) holds. Consequently, $L_1 \cap L_2 = L(\mathfrak{G}\mathfrak{P}\mathfrak{A} , 0 , \neq)$, whence the closure under intersection follows.

Theorem 20. *If $L_1 \in \mathcal{L}(\neq)$ and L_2 is a stochastic language, then* $L_1 \cap L_2$ is a stochastic language.

Proof. We may assume that $L_1 = L(\mathfrak{G}\mathfrak{P}\mathfrak{A}_1 , 0 , \neq)$ and $L_2 = L(\mathfrak{G}\mathfrak{P}\mathfrak{A}_2 , 0)$ where $\mathfrak{G}\mathfrak{P}\mathfrak{A}_i = (S_i , M_i , \pi_i , f_i)$ $(i = 1 , 2)$. By Lemma 10, there exists a generalized probabilistic automaton $\mathfrak{G}\mathfrak{P}\mathfrak{A} = (S , M , \pi_0 , f_0)$ such that, for each word $P \in W(I)$,

$$\pi_0 M(P) f_0 = (\pi_1 M_1(P) f_1)^2 (\pi_2 M_2(P) f_2) .$$

This implies that $L_1 \cap L_2 = L(\mathfrak{G}\mathfrak{P}\mathfrak{A} , 0)$, whence the theorem follows, by Theorem 16.

We do not know whether or not the family $\mathcal{L}(\neq)$ is a proper subfamily of $\mathcal{L}_s$.

10.3. We consider for a moment the family $\mathcal{L}(=)$. By (8.4), $\mathcal{L}_r$ is a proper subfamily of $\mathcal{L}(=)$. By means of Lemmas 10 and 11 we see that $\mathcal{L}(=)$ is closed under sum and intersection. Every language of $\mathcal{L}(=)$ is the complement of a language of $\mathcal{L}(\neq)$. Consequently, if there exists a language $L \in \mathcal{L}(=)$ which is not stochastic, then $\mathcal{L}_s$ is not closed under complementation. It can be verified that, for example, the language $\{0^n 1^n \mid n \geqq 1\}$ belongs to $\mathcal{L}(=)$. However, we do not know whether or not it is stochastic.

The problem of the closure under complementation is closely related to the question whether or not the family of representable languages remains unaltered if the sign $>$ is replaced by the sign $\geqq$ in the definition of a language represented in a generalized probabilistic automaton. Namely, $\mathcal{L}_s = \mathcal{L}(\geqq)$ if and only if $\mathcal{L}_s$ is closed under complementation. The definition of the family $\mathcal{L}(\geqq)$ is analogous to that of $\mathcal{L}(\neq)$.

§ 11. Generalized automata

11.1. As we have seen, the generalization of π_F is not essential as far as the family of representable languages is concerned. In what follows, we show that this holds even if the elements of π_0 and of the matrices $M(x)$ are allowed to be arbitrary real numbers.

Definition 3. A *generalized automaton* over the alphabet I is an ordered quadruple $\mathfrak{GA} = (S, M, \pi_0, f_0)$, where $S = \{s_1, \ldots, s_n\}$ is a finite non-empty set (the set of internal states), M is a mapping of I into the set of $n \times n$ matrices with real elements, π_0 is an n-dimensional row vector with real components (the *initial vector*) and f_0 is an n-dimensional column vector with real components (the *final vector*).

The domain of M is extended from I to $W(I)$ in the same way as before.

For any real number η, the language represented in $\mathfrak{GA}$ with the cut-point η is defined to be the set

$$L(\mathfrak{GA}, \eta) = \{P \in W(I) \mid \pi_0 M(P) f_0 > \eta\}.$$

A language L is called a $\mathfrak{GA}$-language if and only if, for some $\mathfrak{GA}$ and η, $L = L(\mathfrak{GA}, \eta)$.

11.2. In the following four lemmas we show how, for any given generalized automaton, a generalized probabilistic automaton representing the same language can be constructed.

Lemma 12. *Every* $\mathfrak{GA}$*-language* L *can be represented in a generalized automaton* $\mathfrak{GA}_1 = (S_1 , M_1 , \pi_1 , f_1)$ *where, for each* $x \in I$, *the row and column sums of* $M_1(x)$ *equal zero.*

Proof. Let $L = L(\mathfrak{GA} , \eta)$, where $\mathfrak{GA} = (S , M , \pi_0 , f_0)$ is an n-state generalized automaton. Clearly, for each $x \in I$, there exist real numbers $\delta_1(x) , \ldots , \delta_n(x)$, $\gamma_0(x) , \ldots , \gamma_n(x)$ such that in the matrix

$$(11.1) \qquad M_1(x) = \left[\begin{array}{c|ccc|c} 0 & 0 & \cdots & 0 & 0 \\ \hline \delta_1(x) & & & & 0 \\ \cdot & & & & \cdot \\ \cdot & & M(x) & & \cdot \\ \cdot & & & & \cdot \\ \delta_n(x) & & & & 0 \\ \hline \gamma_0(x) & \gamma_1(x) & \ldots & \gamma_n(x) & 0 \end{array}\right]$$

every row and column sum equals 0. From this construction it follows that, for each non-empty word $P \in W(I)$ and for some real numbers $\delta_1(P) , \ldots , \delta_n(P)$, $\gamma_0(P) , \ldots , \gamma_n(P)$, $M_1(P)$ is obtained from (11.1) by replacing x by P. Consequently, if we define $\pi_1 = (0 , \pi_0 , 0)$ and $f_1 = (0 , f_0^T , 0)^T$, then, for the $(n + 2)$-state generalized automaton $\mathfrak{GA}_1 = (S_1 , M_1 , \pi_1 , f_1)$, the equation

$$\pi_1 M_1(P) f_1 = \pi_0 M(P) f_0$$

holds whenever $P \in W(I)$. This implies that $L = L(\mathfrak{GA}_1 , \eta)$, whence the lemma follows.

Lemma 13. *Every* $\mathfrak{GA}$*-language* L *can be represented in a generalized automaton* $\mathfrak{GA}_1 = (S_1 , M_1 , \pi_1 , f_1)$ *where, for each* $x \in I$, *the elements of the matrix* $M_1(x)$ *are non-negative.*

Proof. By Lemma 12, we may assume that $L = L(\mathfrak{GA} , \eta)$ for an n-state generalized automaton $\mathfrak{GA} = (S , M , \pi_0 , f_0)$ such that the row and column sums of the matrices $M(x)$ $(x \in I)$ equal 0.

For any real number a, denote by $N(a)$ the $n \times n$ matrix whose elements equal a. Let $\delta > 0$ be so large that, for each $x \in I$, the elements of the matrix $M_2(x) = M(x) + N(\delta)$ are non-negative. By the assumption concerning the matrices $M(x)$ $(x \in I)$, both $M(x)N(a)$ and $N(a)M(x)$ are zero matrices. This implies that, for any $x \in I$ and $y \in I$,

$$M_2(xy) = M_2(x)M_2(y) = M(xy) + N(n\delta^2) . \tag{11.2}$$

It is easy to verify that the row and column sums of $M(xy)$ equal 0. Let $P \in W(I)$ be an arbitrary non-empty word. Proceeding inductively, we infer from (11.2) that

$$M_2(P) = M(P) + N(n^{l(P)-1}\, \delta^{l(P)}) . \tag{11.3}$$

Let A be the 2×2 matrix whose rows equal $(0\,,\,1)$. Consider the $(2n+2)$-state generalized automaton $\mathfrak{GA}_1 = (S_1\,,\,M_1\,,\,\pi_1\,,f_1)$ where

$$\pi_1 = (\pi_0\,,\,\pi_0\,,\,\pi_0 f_0\,,\,0)\,,\quad f_1 = (f_0^T\,,\,-f_0^T\,,\,1\,,\,0)^T$$

and, for each $x \in I$,

$$M_1(x) = \begin{bmatrix} M_2(x) & 0 & 0 \\ 0 & N(\delta) & 0 \\ 0 & 0 & A \end{bmatrix} .$$

Consequently, for any non-empty word $P \in W(I)$,

$$M_1(P) = \begin{bmatrix} M_2(P) & 0 & 0 \\ 0 & N(n^{l(P)-1}\, \delta^{l(P)}) & 0 \\ 0 & 0 & A \end{bmatrix} . \tag{11.4}$$

Formulas (11.3) and (11.4) together with the definition of $M_1(A)$ now imply that, for any word $P \in W(I)$,

$$\pi_1 M_1(P) f_1 = \pi_0 M(P) f_0 .$$

Thus $L = L(\mathfrak{GA}_1\,,\,\eta)$, whence the lemma follows.

Lemma 14. *Every $\mathfrak{GA}$-language L can be represented in a generalized automaton $\mathfrak{GA}_1 = (S_1\,,\,M_1\,,\,\pi_1\,,f_1)$ where the matrices $M_1(x)$ $(x \in I)$ are stochastic.*

Proof. By Lemma 13, we may assume that $L = L(\mathfrak{GA}\,,\,\eta)$ for an n-state generalized automaton $\mathfrak{GA} = (S\,,\,M\,,\,\pi_0\,,f_0)$ where the elements of the matrices $M(x)$ $(x \in I)$ are non-negative.

Let $\delta > 1$ be a real number larger than the largest row sum in the matrices $M(x)$ $(x \in I)$. For each $x \in I$, there exist real numbers $\delta_i(x)$, $0 \leqq \delta_i(x) \leqq 1$, $(i = 1, \ldots, n)$ such that the matrix

$$M_2(x) = \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline \delta_1(x) & \\ \cdot & \\ \cdot & \delta^{-1}M(x) \\ \cdot & \\ \delta_n(x) & \end{array}\right]$$

is stochastic. From this construction it follows that, for any non-empty word $P \in W(I)$ and for some real numbers $\delta_1(P), \ldots, \delta_n(P)$,

$$M_2(P) = \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline \delta_1(P) & \\ \cdot & \\ \cdot & \delta^{-l(P)}M(P) \\ \cdot & \\ \delta_n(P) & \end{array}\right].$$

Consider the $(n + 3)$-state generalized automaton $\mathfrak{G}\mathfrak{A}_1 = (S_1, M_1, \pi_1, f_1)$ where

$$\pi_1 = (0, \pi_0, \eta, 0), \quad f_1 = (0, f_0^T, -1, 0)^T$$

and, for each $x \in I$,

$$M_1(x) = \left[\begin{array}{c|cc} M_2(x) & & 0 \\ \hline & \delta^{-1} & 1 - \delta^{-1} \\ 0 & 0 & 1 \end{array}\right].$$

It follows that, for any non-empty word $P \in W(I)$,

$$M_1(P) = \left[\begin{array}{c|cc} M_2(P) & & 0 \\ \hline & \delta^{-l(P)} & 1 - \delta^{-l(P)} \\ 0 & 0 & 1 \end{array}\right].$$

We now conclude that, for any word $P \in W(I)$,

$$\pi_1 M_1(P) f_1 = \delta^{-l(P)}(\pi_0 M(P) f_0 - \eta).$$

Thus $L = L(\mathfrak{G}\mathfrak{A}_1, 0)$, whence the lemma follows.

Lemma 15. *Every* $\mathfrak{G}\mathfrak{A}$*-language* L *can be represented in a generalized probabilistic automaton.*

Proof. By Lemma 14, we may assume that $L = L(\mathfrak{G}\mathfrak{A}, \eta)$ for an n-state generalized automaton $\mathfrak{G}\mathfrak{A} = (S, M, \pi_0, f_0)$ where the matrices $M(x)$ $(x \in I)$ are stochastic.

Let $\delta > 0$ be so large that, for $\pi_0 = (p_1, \ldots, p_n)$, the components of the vector $(p_1 + \delta, \ldots, p_n + \delta)$ are positive. Denote $\gamma = p_1 + \ldots + p_n + 2n\delta$. Thus $\gamma > 0$. Consider the $2n$-state generalized probabilistic automaton $\mathfrak{G}\mathfrak{P}\mathfrak{A} = (S_1, M_1, \pi_1, f_1)$ where

$$\pi_1 = \gamma^{-1}(p_1 + \delta, \ldots, p_n + \delta, \delta, \ldots, \delta), \quad f_1 = \begin{bmatrix} f_0 \\ -f_0 \end{bmatrix}$$

and, for each $x \in I$,

$$M_1(x) = \begin{bmatrix} M(x) & 0 \\ 0 & M(x) \end{bmatrix}.$$

We conclude that, for any word $P \in W(I)$,

$$\pi_1 M_1(P) f_1 = \gamma^{-1} \pi_0 M(P) f_0 .$$

Thus $L = L(\mathfrak{G}\mathfrak{P}\mathfrak{A}, \eta/\gamma)$, whence the lemma follows.

11.3. As an immediate consequence of Lemma 15 and Theorem 16 we obtain

Theorem 21. *A language* L *can be represented in a generalized automaton if and only if it can be represented in a probabilistic automaton, i.e., if and only if it is a stochastic language.*

Thus, the family of languages representable in generalized automata equals the family $\mathcal{L}_s$.

By the *mirror image* of a language L, in symbols $\mathrm{mi}(L)$, we mean the language obtained from L by writing all words backwards. By means of Theorem 21, we now establish the following

Theorem 22. *A language* L *is stochastic if and only if the mirror image of* L *is stochastic.*

Proof. Since $\mathrm{mi}(\mathrm{mi}(L)) = L$, it suffices to prove that if L is a stochastic language, so is $\mathrm{mi}(L)$.

Let $L = L(\mathfrak{P}\mathfrak{A}, \eta)$, where $\mathfrak{P}\mathfrak{A} = (S, M, \pi_0, F)$ is a probabilistic automaton. Consider the generalized automaton $\mathfrak{GA} = (S, M_1, \pi_1, f_1)$ where $\pi_1 = \pi_F^T$, $f_1 = \pi_0^T$ and, for each $x \in I$, $M_1(x) = M(x)^T$. For each word $P \in W(I)$, we now obtain

$$\pi_1 M_1(\mathrm{mi}(P)) f_1 = \pi_0 M(P) \pi_F ,$$

which implies that $\mathrm{mi}\,(L) = L(\mathfrak{GA}, \eta)$. Theorem 22 now follows from Theorem 21.

The *right derivative* of a language L with respect to a word P is defined by

$$\bar{\partial}_P L = \{Q \mid QP \in L\} .$$

Clearly, $\mathrm{mi}\,(\bar{\partial}_P L)$ is the left derivative of $\mathrm{mi}\,(L)$ with respect to the word $\mathrm{mi}\,(P)$. By Theorem 22 and Lemma 8, we now obtain the following

Theorem 23. *All right derivatives of a stochastic language are stochastic languages. Conversely, if there exists an integer k such that all right derivatives of a language L with respect to words of length k are stochastic languages, then L is a stochastic language.*

We give another proof for the first part of this theorem. Let $L = L(\mathfrak{PA}, \eta)$, where $\mathfrak{PA} = (S, M, \pi_0, F)$. Let P be an arbitrary word. The language $\bar{\partial}_P L$ is represented in the generalized probabilistic automaton $\mathfrak{GPA} = (S, M, \pi_0, M(P)\pi_F)$. By Theorem 16, we conclude that $\bar{\partial}_P L$ is a stochastic language.

Theorem 23 implies that Theorem 14 holds for right derivatives, too.

§ 12. Realizability of mappings

12.1. Let V_n be the set of n-dimensional stochastic row vectors and Z a mapping of $W(I)$ into V_n. Then Z is said to be *realizable* by a probabilistic automaton if and only if, for some $\mathfrak{PA} = (S, M, \pi_0, F)$ over I,

$$Z(P) = \pi_0 M(P) \quad \text{for all} \quad P \in W(I) .$$

Clearly, Z is realized by a probabilistic automaton if and only if, for each $x \in I$, there exists a stochastic $n \times n$ matrix $M(x)$ such that

$$Z(Px) = Z(P)M(x) \quad \text{for all} \quad P \in W(I) .$$

If r is the maximal number of linearly independent vectors $Z(P)$, and $Z(P_1), \ldots, Z(P_r)$ are linearly independent, then we say that these vectors form a *basis* of Z.

SALOMAA [16] has proved that the following two conditions are both necessary and sufficient for a mapping $Z : W(I) \to V_n$ to be realizable by a probabilistic automaton:

(i) *For all* $P \in W(I)$ *and all* $x \in I$, *if* $Z(P) = \sum_{i=1}^{k} \alpha_i Z(Q_i)$ then $Z(Px) = \sum_{i=1}^{k} \alpha_i Z(Q_i x)$.

(ii) *If* $Z(Q_1), \ldots, Z(Q_t)$, *where* $1 \leqq t \leqq n$, *are linearly independent and* $x \in I$, *then there exist n-dimensional row vectors* $Z_{t+1}, \ldots, Z_n$, $U_{t+1}, \ldots, U_n$ *such that the matrix*

$$\begin{bmatrix} Z(Q_1) \\ \cdot \\ \cdot \\ \cdot \\ Z(Q_t) \\ Z_{t+1} \\ \cdot \\ \cdot \\ \cdot \\ Z_n \end{bmatrix}^{-1} \begin{bmatrix} Z(Q_1 x) \\ \cdot \\ \cdot \\ \cdot \\ Z(Q_t x) \\ U_{t+1} \\ \cdot \\ \cdot \\ \cdot \\ U_n \end{bmatrix}$$

is stochastic.

Another criterion can be found in [2] and [3].

Remark 3. The existence of vectors Z_i, U_i $(i = t+1, \ldots, n)$ for some basis of Z only is sufficient in the condition (ii).

12.2. Using the conditions (i) and (ii), we establish the following criterion.

Theorem 24. *Let* $Z(P_1), \ldots, Z(P_r)$ *form a basis of a mapping* $Z : W(I) \to V_n$. *Then* Z *is realizable by a probabilistic automaton if and only if the following two conditions are satisfied:*

(i)′ *For all* $P \in W(I)$ *and all* $x \in I$, *if* $Z(P) = \sum_{i=1}^{r} \alpha_i Z(P_i)$ *then* $Z(Px) = \sum_{i=1}^{r} \alpha_i Z(P_i x)$.

(ii)′ *For all* $x \in I$, *there exist n-dimensional row vectors* $Z_{r+1}, \ldots, Z_n$, $U_{r+1}, \ldots, U_n$ *such that*

$$(12.1) \qquad A(x) = \begin{bmatrix} Z(P_1) \\ \cdot \\ \cdot \\ \cdot \\ Z(P_r) \\ Z_{r+1} \\ \cdot \\ \cdot \\ \cdot \\ Z_n \end{bmatrix}^{-1} \begin{bmatrix} Z(P_1x) \\ \cdot \\ \cdot \\ \cdot \\ Z(P_rx) \\ U_{r+1} \\ \cdot \\ \cdot \\ \cdot \\ U_n \end{bmatrix}$$

is a matrix with non-negative elements and, furthermore, the sum of the components of Z_i *equals the sum of the components of* U_i $(i = r+1, \ldots, n)$.

Proof. Assume first that Z is realizable by a probabilistic automaton. Then the conditions (i) and (ii) are satisfied. This implies that (i)′ is satisfied. By the condition (ii), there exist row vectors Z_i, U_i $(i = r+1, \ldots, n)$ such that (12.1) is a stochastic matrix. Thus, its elements are non-negative. Furthermore,

$$(12.2) \qquad \begin{bmatrix} Z(P_1x) \\ \cdot \\ \cdot \\ \cdot \\ Z(P_rx) \\ U_{r+1} \\ \cdot \\ \cdot \\ \cdot \\ U_n \end{bmatrix} = \begin{bmatrix} Z(P_1) \\ \cdot \\ \cdot \\ \cdot \\ Z(P_r) \\ Z_{r+1} \\ \cdot \\ \cdot \\ \cdot \\ Z_n \end{bmatrix} A(x) .$$

Since $A(x)$ is a stochastic matrix, it follows that the sum of the components of U_i equals the sum of the components of Z_i $(i = r+1, \ldots, n)$. Thus, also the condition (ii)′ is satisfied.

Conversely, assume that (i)′ and (ii)′ are satisfied. To prove that (i) is satisfied, let $P \in W(I)$ and $x \in I$ be arbitrarily fixed and let $Z(P) = \sum_{i=1}^{k} \beta_i Z(Q_i)$. Now $Z(Q_i) = \sum_{j=1}^{r} \beta_{ij} Z(P_j)$, so that

$$Z(P) = \sum_{j=1}^{r} \Big(\sum_{i=1}^{k} \beta_i \beta_{ij} \Big) Z(P_j) \,.$$

By the condition (i)′, we now have

$$Z(Px) = \sum_{i=1}^{k} \sum_{j=1}^{r} \beta_i \beta_{ij} Z(P_j x) \,, \quad Z(Q_i x) = \sum_{j=1}^{r} \beta_{ij} Z(P_j x) \,,$$

so that $Z(Px) = \sum_{i=1}^{k} \beta_i Z(Q_i x)$. Thus, the condition (i) is satisfied.

From the condition (ii)′ we obtain the equation (12.2) where the elements of $A(x)$ are non-negative. Denote in (12.2) the first factor of the matrix on the right by $[z_{ij}]$. The row sums of $A(x)$ satisfy the equations

$$z_{i1} x_1 + \ldots + z_{in} x_n = \begin{cases} 1 & \text{for } i = 1 \,, \ldots , r \,, \\ a_i & \text{for } i = r+1 \,, \ldots , n \,, \end{cases}$$

where a_i denotes the sum of the components of U_i. For this inhomogenous system of n equations, the determinant of coefficients is different from zero. Thus, the solution is uniquely determined. Since $x_1 = \ldots = x_n = 1$ is a solution, we conclude that the row sums of $A(x)$ equal 1. Consequently, $A(x)$ is a stochastic matrix. We have thus verified that, for the basis $Z(P_1) \,, \ldots , Z(P_r)$, the condition (ii) is satisfied. From Remark 3 it now follows that Z is realizable by a probabilistic automaton. The proof is thus complete.

Using Theorem 24, it can be verified that the condition (ii)′ may be replaced by the following condition:

(ii)″ *If* $Z_{r+1} \,, \ldots , Z_n$ *are n-dimensional stochastic row vectors and, for each* $i = 1 \,, \ldots , n$,

$$\sum_{j=1}^{r} \alpha_{ij} Z(P_j) + \sum_{j=r+1}^{n} \alpha_{ij} Z_j$$

is a stochastic vector, then, for each $x \in I$, *there exist n-dimensional row vectors* $U_{r+1} \,, \ldots , U_n$ *independent on* i *such that*

$$\sum_{j=1}^{r} \alpha_{ij} Z(P_j x) + \sum_{j=r+1}^{n} \alpha_{ij} U_j$$

is a stochastic vector for each $i = 1 \,, \ldots , n$.

If $r = n$, then (ii)$''$ gets the following simpler form:

If $\sum_{i=1}^{n} \alpha_i Z(P_i)$ *is a stochastic vector and* $x \in I$, *then* $\sum_{i=1}^{n} \alpha_i Z(P_i x)$ *is a stochastic vector.*

12.3. Finally, we consider the realizability of mappings $Z : W(I) \rightarrow W_n$ by generalized automata, where W_n denotes the set of all n-dimensional row vectors with real components. The notion of the realizability by generalized automata is defined as in section 12.1. As we show in the following theorem, only the condition (i)$'$ is now needed.

Theorem 25. *Let* $Z(P_1), \ldots, Z(P_r)$ *form a basis of a mapping* $Z : W(I) \rightarrow W_n$. *Then* Z *is realizable by a generalized automaton if and only if the following condition is satisfied:*

For all $P \in W(I)$ *and all* $x \in I$, *if* $Z(P) = \sum_{i=1}^{r} \alpha_i Z(P_i)$ *then* $Z(Px) = \sum_{i=1}^{r} \alpha_i Z(P_i x)$.

Proof. The »only if»-part follows from the distributive law of matrix multiplication. For the »if»-part, assume that the condition is satisfied. We choose vectors Z_i, $U_i \in W_n$ $(i = r+1, \ldots, n)$ such that $Z(P_1), \ldots, Z(P_r)$, $Z_{r+1}, \ldots, Z_n$ are linearly independent, and form the matrix (12.1). Then

$$Z(P_i x) = Z(P_i)A(x) \quad (i = 1, \ldots, r) . \tag{12.3}$$

Let $P \in W(I)$ be an arbitrary word. Since $Z(P_1), \ldots, Z(P_r)$ form a basis of Z, we have

$$Z(P) = \sum_{i=1}^{r} \alpha_i Z(P_i) .$$

Our assumption now implies that

$$Z(Px) = \sum_{i=1}^{r} \alpha_i Z(P_i x) .$$

Consequently, by (12.3), $Z(Px) = Z(P)A(x)$. The proof is thus complete.

University of Turku
Turku, Finland

References

[1] Бухараев, Р. Г.: Некоторые эквивалентности в теории вероятностных автоматов. - Уч. зап. Казанск. ун-т, 124 (1964), 45 – 65.

[2] —»— Критерий представимости событий в конечных вероятностных автоматах. - Докл. Акад. Наук СССР 164 (1965), 289 – 291.

[3] —»— Критерий представимости событий в конечных вероятностных автоматах, manuscript to be published.

[4] Carlyle, J. W.: Reduced forms for stochastic sequential machines. - J. Math. Anal. Appl. 7 (1963), 167 – 175.

[5] Feller, W.: An introduction to probability theory and its applications, Vol. 1. 3rd ed., John Wiley and Sons, Inc., New York (1968).

[6] Gantmacher, F. R.: The theory of matrices, Vol. 1, 2. Chelsea Publishing Company, New York (1960).

[7] Halmos, P.: Finite-dimensional vector spaces. 2nd ed., D. Van Nostrand Company, Inc., Princeton, New Jersey (1958).

[8] Kleene, S. C.: Representation of events in nerve nets and finite automata. - Automata Studies, Princeton (1956), 3 – 41.

[9] Niven, I.: Irrational numbers. Math. Association of America, The Carus Math. Monographs, distributed by John Wiley and Sons, Inc., New York(1956).

[10] Olmsted, J. M. H.: Rational values of trigonometric functions. - Amer. Math. Monthly 52 (1945), 507 – 508.

[11] Page, C. V.: Equivalences between probabilistic and deterministic sequential machines. - Information and Control 9 (1966), 469 – 520.

[12] Paz, A.: Some aspects of probabilistic automata. - Information and Control 9 (1966), 26 – 60.

[13] Rabin, M. O.: Probabilistic automata. - Information and Control 6 (1963), 230 – 245.

[14] Salomaa, A.: Theorems on the representation of events in Moore-automata. - Ann. Univ. Turku. Ser. A I 69 (1964).

[15] —»— On probabilistic automata with one input letter. - Ann. Univ. Turku. Ser. A I 85 (1965).

[16] —»— On events represented by probabilistic automata of different types. - Canad. J. Math. 20 (1968), 242 – 251.

[17] —»— Theory of automata. Pergamon Press, to appear.

[18] Starke, P. H.: Stochastische Ereignisse und Wortmengen. - Z. Math. Logik Grundlagen Math. 12 (1966), 61 – 68.

[19] Turakainen, P.: On non-regular events representable in probabilistic automata with one input letter. - Ann. Univ. Turku. Ser. A I 90 (1966).

Printed August 1968