INFORMATION AND CONTROL 20, 439–479 (1972)

Some Aspects of Linear Space Automata

YASUYOSHI INAGAKI* AND TERUO FUKUMURA

Department of Electrical Engineering, Faculty of Engineering, Nagoya University Furo-cho, Chikusa-ku, Nagoya, Japan

AND

HIROYUKI MATUURA

Nippon Telegraph and Telephone Public Corporation, Tokyo, Japan

Linear space automaton is introduced as a generalization of probabilistic automaton and its various properties are investigated.

Linear space automaton has the abilities equivalent to probabilistic automaton but we can treat the former more easily than the latter because we can make use of properties of the linear space, successfully.

First the solutions are given for the problems of connectivity, state equivalence, reduction and identification of linear space automata. Second, the matrix representation of linear space automaton is investigated and the relations between linear space automaton and probabilistic automaton are shown. Third, we discuss the closure properties of the family of all real functions on a free semigroup Σ^* which are defined by linear space automata and then give a solution to the synthesis problem of linear space automata.

Finally, some considerations are given to the problems of sets of tapes accepted by l.a.'s and also of operations under which the family of all the output functions of l.a.'s is not closed.

1. INTRODUCTION

Since Rabin's (1963) work concerning probabilistic automata (abbreviated p.a.), many researchers such as Paz (1966), Salomaa (1967), Honda and Nasu (1968), Sugino, Inagaki and Fukumura (1968), Turakainen (1968), Carlyle (1968) have concerned with this subject. Despite the above extensive works, there remain some difficulties in the theoretical treatment of p.a. because the set of states of p.a. is a set of stochastic vectors.[1]

* Research supported in part by Matsunaga Science Foundation.

[1] For many other works on this subject, see the exhaustive bibliography of A. Paz (1971).

Hence in this paper we will introduce a new system with a linear space, which will be called linear space automaton (abbreviated l.a.), and discuss its properties as well as its relation to p.a. to have a new view point of p.a.

It will be shown later that many properties of l.a. are easily known and in particular the reduction problem is solved in straightforward way because we can use many well-known properties of the linear space. In order to lay stress on this point we will consider l.a. with linear state space over a general field K in the Sections 2–7.

It will be understood that l.a. is a very natural extension of p.a. and their abilities are equivalent to each other. These facts may give us an idea that l.a. will become a useful tool to treat p.a.

2. Preliminary Definitions

Alphabet and Tape

Let $\Sigma = \{\sigma_i \mid i = 1,..., I\}$ be a nonempty finite set of symbols, which is called an alphabet. A finite sequence consisting of elements of Σ is called a tape, which is denoted by w, z, etc. The length of a tape w, which is denoted by $l(w)$, is defined to be the number of symbols contained in w. The concatenation of tapes w and z is denoted by $w \cdot z$ or simply by wz. The tape with zero length is denoted by ϵ. The set of all tapes generated over Σ is denoted by Σ^*, which contains ϵ.

Linear Space

A finite dimensional linear space over the field K will simply be called "space" and denoted by V, W, etc. The dimension of the space V is denoted by dim V. A subset S of the space V being a space is called a subspace.

Linear Mapping

A linear mapping T from a space V_1 into V_2 is simply called "mapping" and denoted by $T\colon V_1 \to V_2$. When the mapping T transforms $x \in V_1$ to $y \in V_2$, we write this simply as $y = xT$. For an arbitrary subset S_1 of V_1, we will call $S_2 = T(S_1) = \{y \mid y = xT,\ x \in S_1\}$, T-image of S_1 and for an arbitrary subset S_2' of V_2, $S_1' = T^{-1}(S_2') = \{x \mid xT \in S_2'\}$, T-inverse image of S_2', if T^{-1}, the inverse mapping of T, exists. If S_1 is a subspace of V_1, S_2 is a subspace of V_2 and if S_2' is a subspace of V_2, S_1' is also a subspace of V_1.

If T-image V_2' of V_1 is identical to V_2, T is called a mapping from V_1 onto V_2.

Let V_1, V_2, V_3, and V_4 be spaces. For mappings $T_1 : V_1 \to V_2$ and $T_2 : V_2 \to V_3$, the product $T_1T_2 : V_1 \to V_3$ is defined by $xT_1T_2 = (xT_1)\,T_2$ for all $x \in V$. If T_3 is a mapping from V_3 into V_4, then $(T_1T_2)\,T_3 = T_1(T_2T_3) = T_1T_2T_3$ holds, that is, the product of mappings is associative.

3. Linear Space Automaton

In this section the definition of linear space automaton and the relating concepts are introduced.

Definition 3.1. *Linear space automaton* L over a finite alphabet Σ is a system

$$L = \langle V, \{A_i \mid i = 1, 2,..., I\}, u, v\rangle \tag{3.1}$$

where V is the space over the field K, which is called the *state space* of L and of which elements are called the *states* of L. A_i is a mapping from V into V caused by the input symbol $\sigma_i \in \Sigma$. u is an element of V and called the *initial state* of L. v is a linear function from V into K.

The linear space automaton is essentially the same as the generalized automaton defined by Turakainen (1968). But the authors would like to notice that the concept of l.a. is obtained independently of it by the authors (1968).

Definition 3.2. For all $w \in \Sigma^*$, $F(w)$: $V \to V$ is recursively defined as follows and called the *response mapping* of L.

$$\begin{aligned} F(\epsilon) &= E, \\ F(z\sigma_i) &= F(z)\,A_i \qquad \text{for} \quad \forall z \in \Sigma^* \qquad \text{and} \qquad \forall \sigma_i \in \Sigma, \end{aligned} \tag{3.2}$$

where E is the identity mapping from V into V.

Definition 3.3. The function $f(w)$: $\Sigma^* \to K$ is defined by

$$f(w) = uF(w)v \qquad \text{for all } w \in \Sigma^* \tag{3.3}$$

and called the *output function* of V.

From Definition 3.2 and the associative law concerning the product of mappings we have directly the following lemma.

LEMMA 3.1. *For* $\forall w = \sigma_{i_1}\sigma_{i_2} \cdots \sigma_{i_k} \in \Sigma^*$,

$$F(w) = A_{i_1}A_{i_2} \cdots A_{i_k} \tag{3.4}$$

and for $\forall w, \forall z \in \Sigma^*$

$$F(wz) = F(w)F(z). \tag{3.5}$$

4. CONNECTIVITY, DISTINGUISHABILITY AND STATE EQUIVALENCE OF l.a.

DEFINITION 4.1. A subset S of V is called an *invariant subset* of V with respect to (w.r.t.) l.a. L if the following condition holds:

$$x \in S \quad \text{implies} \quad xA_i \in S \qquad \text{for} \quad \forall \sigma_i \in \Sigma. \tag{4.1}$$

Moreover, if S is a subspace, it is called an invariant subspace of V w.r.t. L.

LEMMA 4.1. *Let S be an invariant subspace of V. Then for* $\forall w \in \Sigma^*$,

$$x \in S \quad \textit{implies} \quad xF(w) \in S. \tag{4.2}$$

DEFINITION 4.2. For $x \in V$, the minimal invariant subspace containing x is called the *connected part* of x and denoted by $V(x)$. If the connected part $V(u)$ of the initial state u is identical to V, L is said to be *connected*. And if for any nonzero $x \in V$, its connected part $V(x)$ is identical to V, then L is said *strongly connected*.

We have the following lemma directly from the definition:

LEMMA 4.2. *L is not strongly connected if and only if there exists at least one invariant subspace except* $\{0\}$ *and* V.

DEFINITION 4.3. For $x_1, x_2 \in V$, if there exists $w \in \Sigma^*$ such that

$$x_1F(w)u \neq x_2F(w)u \tag{4.3}$$

x_1 and x_2 are said *distinguishable*. If x_1 and x_2 are not distinguishable then they are said to be *equivalent*.

If all x_1 and x_2 in V such that $x_1 \neq x_2$ are distinguishable, then L is said to be distinguishable.

Now we consider the procedures to obtain $V(x)$ for any given x and to decide whether x_1 and x_2 are distinguishable or not for given x_1 and x_2 in V. These problems can be solved by using well-known techniques for probabilistic automata.[2] Then we describe only the results.

We can assume that $x \neq 0$ without loss of generality because $x = 0$ implies $V(x) = \{0\}$.

Define the sequence of subspaces V_x^k ($k = 1, 2, \ldots$) of V recursively as follows:

$$V_x^1 = \{\alpha x \mid \alpha \in K\}, \tag{4.4}$$

$$V_x^{k+1} = \left\{ x_0 + \sum_{i=1}^{I} x_i A_i \,\middle|\, x_0, x_i \in V_x^k, i = 1, \ldots, I \right\}. \tag{4.5}$$

LEMMA 4.3. *V_x^k is the space spanned by $\{xF(w) \mid w \in \Sigma^*, l(w) \leqslant k - 1\}$ which is the set of all states accessible from the state x with tapes of which length are $(k - 1)$ or less.*

From the definition of V_x^k and Lemma 4.3 we can show that V_x^n is the minimal invariant subspace containing x. Thus we have the following lemma.

LEMMA 4.4. *V_x^n is a connected part $V(x)$ of x.*

Thus V_x^n is an invariant subspace containing x. By Lemma 4.1, we have

$$xF(w) \in V_x^n \qquad \text{for} \quad \forall w \in \Sigma^*. \tag{4.6}$$

Combining this (4.6) and Lemma 4.3, we obtain the following Theorem 4.1.

THEOREM 4.1. *If* l.a. *has n-dimensional state space, then any states accessible from a given state x can be represented by a linear combination of states accessible from x with tapes of which length is $(n - 1)$ or less.*

Moreover, we have the following theorem concerning the state equivalence.

THEOREM 4.2. *For any $x_1, x_2 \in V$, x_1 and x_2 are equivalent if and only if the equation*

$$x_1 F(w) v = x_2 F(w) v \tag{4.13}$$

holds for all tapes with length $(n - 1)$ or less.

[2] See p. 19 of Paz (1971), for example.

5. Equivalence Between l.a.'s

Given two l.a.'s $L_1 = \langle V_1, \{A_i\}, u_1, v_1\rangle$ and $L_2 = \langle V_2, \{B_i\}, u_2, v_2\rangle$, we will denote their response mappings by F_1 and F_2, their output functions by f_1 and f_2.

Definition 5.1. If the equation

$$f_1(w) = f_2(w) \tag{5.1}$$

holds for $\forall w \in \Sigma^*$, the l.a.'s L_1 and L_2 are said to be *equivalent* (or *strongly equivalent*) and this relation will be denoted by $L_1 \equiv^s L_2$.

The relation $\equiv^s$ is obviously an equivalence relation. Another relation, homomorphism between l.a.'s, which is stronger than $\equiv^s$, is useful and defined as follows.

Definition 5.2. A mapping $T: V_1 \to V_2$ is said to be a *homomorphism* from L_1 into L_2 if the following three conditions are satisfied.

$$\text{(i)} \quad xA_iT = xTB_i \qquad \text{for} \quad \forall x \in V_1 \text{ and } \forall \sigma_i \in \Sigma \tag{5.2}$$

$$\text{(ii)} \quad u_1T = u_2 \tag{5.3}$$

$$\text{(iii)} \quad xv_1 = xTv_2 \qquad \text{for} \quad \forall x \in V_1 \tag{5.4}$$

Theorem 5.1. *If there exists a homomorphism from L_1 into L_2 then L_1 is equivalent to L_2, that is, $L_1 \equiv^s L_2$.*

Proof. Let $T: V_1 \to V_2$ be a homomorphism from L_1 into L_2. Using the definition of the response mapping and (5.2), we can prove by induction that $xF_1(w)\,T = xTF_2(w)$ for $\forall x \in V_1$ and $\forall w \in \Sigma^*$. Thus we have $f_1(w) = u_1F_1(w)\,v_1 = u_1F_1(w)\,Tv_2 = u_1TF_2(w)\,v_2 = u_2F_2(w)\,v_2 = f_2(w)$ for $\forall w \in \Sigma^*$. Q.E.D.

Definition 5.3. Let $T: V_1 \to V_2$ be a homomorphism from L_1 *into* L_2.

(i) If T is a mapping from V_1 onto V_2 then T is called a homomorphism from L_1 *onto* L_2.

(ii) If T is a one to one mapping from V_1 into V_2, then T is called an *isomorphism* from L_1 *into* L_2.

(iii) If T is a one to one mapping from V_1 onto V_2, then T is called an *isomorphism* from L_1 *onto* L_2.

The following lemmas are obtained directly from this definition.

LEMMA 5.1. *Consider three* l.a.'*s* L_1, L_2 *and* L_3. *If* T *is an isomorphism (homomorphism) from* L_1 *into (onto)* L_2 *and* T' *is an isomorphism (homomorphism) from* L_2 *into (onto)* L_3, *then the product* TT' *of mappings is an isomorphism (homomorphism) from* L_1 *into (onto)* L_3.

LEMMA 5.2. *If* T *is an isomorphism from* L_1 *onto* L_2 *then the inverse mapping* T^{-1} *of* T *is also an isomorphism from* L_2 *onto* L_1.

Here we would like to state another theorem concerning the equivalence of l.a.'s. As this theorem is well known for probabilistic automata, then it can be proved in a way similar to them.[3]

THEOREM 5.2. *Assume that* l.a.'*s* L_1 *and* L_2 *have* n_1 *and* n_2 *dimensional state spaces, respectively. Then* $L_1 \equiv^s L_2$ *if and only if* $u_1F_1(w)\, v_1 = u_2F_2(w)\, v_2$ *holds for all tapes* w *with length* $(n_1 + n_2 - 1)$ *or less.*

6. REDUCTION OF l.a.

DEFINITION 6.1. An l.a. L is said to be *reducible* if there exists any one which is equivalent to L and of which state space has the smaller dimension than that of L. Otherwise it is said *irreducible*.

We will call it reduction of L to obtain the irreducible l.a. equivalent to L. In the following of this section, the reduction of l.a. L defined by Definition 3.1 is considered.

Denote the connected part $V(u)$ of u of the l.a. L by W. As W is an invariant subspace of V, we have a mapping

$$B_i : W \to W \tag{6.1}$$

by restricting the domain of $A_i : V \to V$ to W. That is, for all $x \in W$, we can define xB_i by

$$xB_i = xA_i \in W. \tag{6.2}$$

And, by v^c we denote the function obtained by restricting the domain V of the linear function v to W. That is, v^c is the function such that

$$xv^c = xv \tag{6.3}$$

holds for all $x \in W$.

[3] See Theorem 2.7 on p. 25 of Paz (1971), for example.

DEFINITION 6.2. A new l.a. L^c is defined by

$$L^c = \langle W, \{B_i\}, u, v^c \rangle \tag{6.4}$$

and L^c is called the *connected part* of L.

LEMMA 6.1. *The* l.a. L^c *is a connected* l.a. *which is equivalent to* L.

Proof. The identity mapping E_W from W to W can be considered to be a mapping from W into V. Using (6.2), (6.3) and the fact that the initial states of L^c and L are the same, we can easily show that the mapping E_W is a homomorphism from L^c into L. Thus, by Theorem 5.1, $L^c \equiv^s L$.

Next, the connectivity of L^c is easily shown from the fact that W is the connected part of u. Q.E.D.

EXAMPLE 6.1. Let the field K be the real field R and assume that $\Sigma = \{\sigma_1\}$. Then, we obtain the connected part $L_1{}^c$, of L_1 defined by

$$L_1 = \langle V, \{A_1\}, u, v \rangle,$$

where V is the three-dimensional linear space generated by the basis $\{e_1, e_2, e_3\}$ over the real field R, that is, $V = R^3$. The mapping $A_i : V \to V$ is the one defined by

$$e_1A_1 = \tfrac{1}{2}e_1 + \tfrac{1}{2}e_2, \qquad e_2A_1 = -\tfrac{1}{2}e_1 + \tfrac{1}{2}e_2, \qquad e_3A_1 = e_3.$$

The initial state $u = e_1$, and the linear real function $v\colon V \to R$ is the one such that

$$e_1v = 1, \qquad e_2v = e_3v = 0.$$

Hence, it is easily known that the connected part W of u is the space spanned by $\{e_1, e_2\}$. Thus, the connected part $L_1{}^c$ of L_1 is obtained as follows:

$$L_1{}^c = \langle W, \{B_1\}, u, v^c \rangle,$$

where W is the space spanned by $\{e_1, e_2\}$, B_1 is the mapping such that

$$e_1B_1 = \tfrac{1}{2}e_1 + \tfrac{1}{2}e_2, \ \ e_2B_1 = -\tfrac{1}{2}e_1 + \tfrac{1}{2}e_2 \ \text{ and } \ u = e_1, \ \ e_1v^c = 1, \ \ e_2v^c = 0.$$

Now, assume that $\dim V = n$ and denote tapes with length $(n - 1)$ or less by $z_1, z_2, \ldots, z_m$, where their order may be arbitrary except $z_1 = \epsilon$ (null string) and the value of m is determined by

$$m = \sum_{k=0}^{n-1} I^k = (1 - I^n)/(1 - I). \tag{6.5}$$

Here, for $x \in V$, we define $y \in K^m$ as follows:

$$y = (xF(z_1)v, xF(z_2)v, \ldots, xF(z_m)v). \tag{6.6}$$

That is, y is the m-dimensional row vector obtained by arranging one after another the outputs of L in the state x, to which $z_1, z_2, \ldots, z_m$ are applied. Obviously, the correspondence from x to y is a linear mapping and written by

$$y = xT. \tag{6.7}$$

Thus, T is a mapping from V into K^m.

LEMMA 6.2. *For* $\forall x_1, \forall x_2 \in V$,

(i) x_1 *and* x_2 *are equivalent if and only if*

$$x_1 T = x_2 T \tag{6.8}$$

holds.

(ii) *If* (6.8) *holds, then for* $\forall \sigma_i \in \Sigma$

$$x_1 A_i T = x_2 A_i T. \tag{6.9}$$

Proof. (1) By the definition of T, the statement (i) is not more than an alternative expression of Theorem 4.2.

(2) From (i), if (6.8) holds then x_1 and x_2 are equivalent and thus, $x_1 F(w)v = x_2 F(w)v$ holds for $\forall w \in \Sigma^*$. In particular, by putting $w = \sigma_i z_l$ $(1 \leqslant l \leqslant m)$ we have

$$x_1 A_i F(z_l)v = x_2 A_i F(z_l)v \qquad (l = 1, \ldots, m). \tag{6.10}$$

Thus (6.10) means (6.9). Q.E.D.

Now, let Z be the image of W by T and define $y \in Z$ by

$$y = xT \tag{6.11}$$

for $x \in W$. For this y we define yC_i by

$$yC_i = xA_i T. \tag{6.12}$$

This C_i is a mapping from Z into Z. In fact, by (ii) of Lemma 6.2, the right hand side of (6.11) is uniquely determined for y; in other words, it is determined independently of the selection of x which satisfies (6.11). Thus, C_i is known to be a unique correspondence from Z into K^m. Obviously, this correspondence is linear. Therefore, C_i is a linear mapping from Z into K^m. Finally, W is the invariant subspace of L, and so, if $x \in W$ then $xA_i \in W$. Thus, $yC_i = xA_iT \in Z$.

Next, for the vector

$$y = xT = (xF(z_1)v, xF(z_2)v, \ldots, xF(z_m)v) \tag{6.13}$$

a linear function v^r from Z into K is defined by

$$yv^r = xF(z_1)v = xF(\epsilon)v = xv. \tag{6.14}$$

That is, yv^r is the first component of y. Finally, the $u^r \in Z$ is defined by

$$u^r = uT. \tag{6.15}$$

Using the Z, $\{C_i\}$, u^r, v^r defined above, we may construct a new l.a. L^r as is given in the following definition:

DEFINITION 6.3. For a given l.a. L we define l.a. L^r by

$$L^r = \langle Z, \{C_i\}, u^r, v^r \rangle. \tag{6.16}$$

LEMMA 6.3. *The mapping T_1 obtained by restricting the domain V of $T: V \to K^m$ to W is a mapping from W onto Z and also a homomorphism from L^c onto L^r.*

Proof. As Z is the image of W by T, it is evident that T_1 is a mapping from W onto Z.

Next, putting $y = xT$ for $x \in W$, from (6.11) and (6.12), we have

$$xA_iT = yC_i = xTC_i\,. \tag{6.17}$$

Thus, for B_i and T_1 obtained from A_i and T by restricting their domain V to W, the following equation holds:

$$xB_iT = xT_1C_i \qquad \text{for} \quad \forall x \in W. \tag{6.18}$$

From (6.15), we have

$$u^r = uT_1 \tag{6.19}$$

and from (6.13) and (6.14) we have

$$xv^c = xv = yv^r = xTv^r = xT_1v^r \qquad \text{for} \quad \forall x \in W, \tag{6.20}$$

where $y = xT_1$. Thus, T_1 is a homomorphism from L^c onto L^r. Q.E.D.

THEOREM 6.1. *The* l.a. L^r *is the irreducible* l.a. *equivalent to* L.

Proof. By Lemma 6.1, L and L^c are equivalent and by Lemma 6.3 and Theorem 5.1 L^c and L^r are equivalent. Thus, L^r is equivalent to L.

Next, we show the irreducibility of L^r. Assume that $L' = \langle V', \{A_i'\}, u', v' \rangle$ is equivalent to L^r and is also equivalent to L. Let the response mapping of L' be F' and define $T': V' \to K^m$ by the following equation:

$$x'T' = (x'F'(z_1)v', x'F'(z_2)v', \ldots, x'F'(z_m)v') \qquad \text{for} \quad x' \in V'. \tag{6.21}$$

From the above assumption that L' and L are equivalent, we have

$$uF(w)F(z_l)v = u'F'(w)F'(z_l)v' \qquad \text{for} \quad \forall w \in \Sigma^* \text{ and } l = 1, 2, \ldots, m. \tag{6.22}$$

Therefore, by the definition of T and T' the equation

$$uF(w)T = u'F'(w)T' \tag{6.23}$$

holds.

On the other hand, combining Lemmas 4.3 and 4.4, we know that W is the space spanned by

$$\{uF(z_1), uF(z_2), \ldots, uF(z_m)\} \tag{6.24}$$

and Z is the image of W by T. Thus, Z is the space spanned by

$$\{uF(z_1)T, uF(z_2)T, \ldots, uF(z_m)T\}. \tag{6.25}$$

Here, let Z' be the image of V' by T' and since the right hand side of (6.23) is an element of Z' for $\forall w \in \Sigma^*$, any element of (6.25) is contained in Z'. Therefore, the space Z is a subspace of Z', that is, $Z \subseteq Z'$.

Hence, recalling that we have denoted the image of V' by Z', we obtain

$$\dim Z \leqslant \dim Z' \leqslant \dim V'. \tag{6.26}$$

Thus, we have known that L^r is irreducible. Q.E.D.

THEOREM 6.2. *If we know all values of output function* $f(w)$ *of* L *for all tapes* w *such that* $0 \leqslant l(w) \leqslant 2n - 1$, *then we can construct* L^r, *where* n *is the dimension of the state space of* L.

Proof. As described in the proof of Theorem 6.1, the state space Z of L^r is the space spanned by (6.25). Therefore, these row vectors can be determined by the values of output function $f(w)$ of L for all tapes w with length $(2n - 2)$ or less.

In order to determine $C_j : Z \to Z$, it is sufficient to know

$$uF(z_i)\, TC_j = uF(z_i)\, A_j T = uF(z_i \sigma_j)T \qquad \text{for all } \sigma_j \in \Sigma \text{ and } i = 1, 2,..., m. \tag{6.27}$$

These can be obtained from values of $f(w)$ of L for tapes w such that $0 \leqslant l(w) \leqslant 2n - 1$. Determination of u^r and v^r is easy. Q.E.D.

EXAMPLE 6.2. Let us reduce l.a. of Example 6.1. This l.a. has $I = 1$, $\dim V = 3$ and so $m = 3$. Hence putting $z_1 = \epsilon$, $z_2 = \sigma_1$, $z_3 = \sigma_1\sigma_1$, the space Z spanned by $uF(z_1)T$, $uF(z_2)T$ and $uF(z_3)T$ is obtained as follows:

$$\begin{aligned} uF(z_1)T &= (uF(z_1)\, F(z_1)v,\ uF(z_1)\, F(z_2)v,\ uF(z_1)\, F(z_3)v) \\ &= (e_1 v,\ e_1 A_1 v,\ e_1 A_1 A_1 v) \\ &= (1, \tfrac{1}{2}, 0), \\ uF(z_2)T &= (\tfrac{1}{2}, 0, -\tfrac{1}{4}), \\ uF(z_3)T &= (0, -\tfrac{1}{4}, -\tfrac{1}{4}). \end{aligned}$$

The vectors $uF(z_1)T$ and $uF(z_2)T$ are linearly independent and $uF(z_3)T$ can be represented by a linear combination of the other two, i.e.,

$$uF(z_3)T = -\tfrac{1}{2} uF(z_1)T + uF(z_2)T.$$

Hence, we choose the basis $\{y_1, y_2\}$ of Z so that

$$y_1 = uF(z_1)T, \qquad y_2 = uF(z_2)T.$$

The mapping $C_1 : Z \to Z$ is determined as follows:

$$\begin{aligned} y_1 C_1 &= uF(z_1)\, TC_1 = uF(z_1)\, A_1 T = uF(z_2)T = y_2, \\ y_2 C_1 &= uF(z_2)\, TC_1 = uF(z_2)\, A_1 T = uF(z_3)T = -\tfrac{1}{2} y_1 + y_2. \end{aligned}$$

Finally, the linear real function v^r over Z is determined as

$$y_1 v^r = 1, \qquad y_2 v^r = \tfrac{1}{2}$$

and the initial state u^r is also determined as $u^r = y_1$. Here, l.a. $L^r = \langle Z, \{C_1\}, u^r, v^r \rangle$ obtained in the above discussion is the irreducible l.a. equivalent to L.

COROLLARY 6.1. *The* l.a. *L is irreducible if and only if L is connected and distinguishable.*

Proof. From Theorem 6.1, L is irreducible if and only if

$$\dim V = \dim Z. \tag{6.28}$$

Since Z is the image of the subspace W of V by T, (6.28) holds if and only if W coincides with V and T is one to one mapping. Therefore, combining Definition 4.2 of connectivity and (i) of Lemma 6.2 this corollary results.

Q.E.D.

COROLLARY 6.2. *If two* l.a.'*s* L_1 *and* L_2 *are equivalent and both are irreducible, there exists an isomorphism from* L_1 *onto* L_2.

Proof. The assumption of the corollary implies that output functions of L_1 and L_2 are the same and the dimensions of their state space are equal. Therefore, it follows from Theorem 6.2 that $L_1{}^r$ of L_1 and $L_2{}^r$ of L_2 are the same one.

From Corollary 6.1, L_1 is known to be connected and distinguishable. Therefore, T is a one-to-one mapping from V onto Z. It is easily shown that this T is an isomorphism from L_1 onto $L_1{}^r$. Similarly, it is known that there exists an isomorphism T' from L_2 onto $L_2{}^r$.

Hence, if we denote the inverse mapping of T' by T'^{-1} then TT'^{-1} is obviously an isomorphism from L_1 onto L_2. Q.E.D.

COROLLARY 6.3. *If* l.a.'*s* L_1 *and* L_2 *are equivalent and* L_2 *is irreducible then there exists a homomorphism from the connected part* $L_1{}^c$ *of* L_1 *onto* L_2. *Particularly, if* L_1 *is connected then a homomorphism exists from* L_1 *onto* L_2.

Proof. From Lemma 6.3 there exists a homomorphism T_1 from $L_1{}^c$ onto $L_1{}^r$. Since $L_1{}^r$ is irreducible, it follows from Corollary 6.2 that there exists an isomorphism T_2 from $L_1{}^r$ onto L_2. Hence, T_1T_2 is a homomorphism from L_1 onto L_2. Q.E.D.

7. l.a.'s REPRESENTED BY MATRICES

7.1. *Notations Relating to Vector and Matrix*

In the following sections we consider mainly row vector spaces and matrices; here we refer to some notations relating to them.

A mapping $T: K^m \to K^n$ is the matrix with m rows and n columns. These row vectors will be denoted by $a_1, a_2, \ldots, a_m$. That is, the mapping T is represented by

$$T = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, \qquad a_1, a_2, \ldots, a_m \in K^n.$$

Hence, the number of linearly independent vectors in $a_1, a_2, \ldots, a_m$ is called the rank of T and denoted by rank T.

Let us denote the transposes of vector x and matrix A by x^t and A^t, respectively.

The direct sum of vectors

$$x = (\xi_1, \xi_2, \ldots, \xi_m) \in K^m \qquad \text{and} \qquad y = (\eta_1, \eta_2, \ldots, \eta_n) \in K^n$$

is denoted by $x \oplus y$ and defined by

$$x \oplus y = (\xi_1, \xi_2, \ldots, \xi_m, \eta_1, \eta_2, \ldots, \eta_n) \in K^{m+n}. \tag{7.1}$$

The direct sum of $m \times m$ matrix A and $n \times n$ matrix B is denoted by $A \oplus B$ and defined by

$$A \oplus B = \begin{pmatrix} A & O_{m,n} \\ O_{n,m} & B \end{pmatrix}, \tag{7.2}$$

where $O_{m,n}$ and $O_{n,m}$ are $m \times n$ and $n \times m$ zero matrices, respectively.

Next, the direct product $x \otimes y$ of x and y is defined by

$$x \otimes y = (\xi_1\eta_1, \xi_1\eta_2, \ldots, \xi_1\eta_n, \ldots, \xi_m\eta_1, \xi_m\eta_2, \ldots, \xi_m\eta_n). \tag{7.3}$$

And the direct product of $A = (\alpha_{ij})$ and B is

$$A \otimes B = \begin{pmatrix} \alpha_{11}B & \alpha_{12}B & \cdots & \alpha_{1m}B \\ \alpha_{21}B & \alpha_{22}B & \cdots & \alpha_{2m}B \\ \alpha_{m1}B & \alpha_{m2}B & \cdots & \alpha_{mm}B \end{pmatrix}. \tag{7.4}$$

7.2. *Matrix Representation of* l.a. *and Invariant Subspace*

Let us consider l.a. of Definition 3.1. Let us denote a finite subset of the state space V of L by

$$S = \{a_1, a_2, \ldots, a_r\}. \tag{7.5}$$

Here we assume that the space spanned by S contains the initial state u

of L and it is an invariant subspace of V. That is, u can be represented by

$$u = \sum_{k=1}^{r} \theta_k a_k \,; \qquad \theta_k \in K, \quad a_k \in S \tag{7.6}$$

and for any $a_k \in S$, $a_k A_i$ can be represented by

$$a_k A_i = \sum_{j=1}^{r} \alpha_{kj}^{(i)} a_j \,; \qquad \alpha_{kj}^{(i)} \in K, \quad a_j \in S. \tag{7.7}$$

And we also assume that the linear function v is a mapping from V into K, i.e., can be represented by

$$a_k v = \eta_k \in K \qquad \text{for all } a_k \in S. \tag{7.8}$$

Here we define l.a. L' of which state space is K^r as follows.

$$L' = \langle K^r, \{A_i'\}, u', v' \rangle, \tag{7.9}$$

where $A_i' = (\alpha_{kj}^{(i)})$ is the $r \times r$ square matrix whose (k, j) entry is $\alpha_{kj}^{(i)}$ determined by (7.7) and u' and v' are defined as follows:

$$u' = (\theta_1, \theta_2, ..., \theta_r), \qquad v' = (\eta_1, \eta_2, ..., \eta_r)^t.$$

Moreover, we define a mapping $T: K^r \to V$ by

$$yT = \sum_{j=1}^{r} \xi_j a_j \qquad \text{for} \quad y = (\xi_1, ..., \xi_r) \in K^r \tag{7.10}$$

LEMMA 7.1. (i) *The mapping $T: K^r \to V$ is a homomorphism from L' into L.*

(ii) *If the space spanned by S coincides with V, then T is a homomorphism from L' onto L.*

(iii) *If the elements of S are linearly independent, then T is an isomorphism from L' into L.*

(iv) *If S is a basis of V, then T is an isomorphism from L' onto L.*

Proof. First, from the definitions we can prove that for $\forall y \in K^r$ and $\forall \sigma_i \in \Sigma$,

$$yA_i'T = yTA_i \tag{7.11}$$

holds. From the definitions of u', v' and T, we also know that $u'T = u$ and $yv' = yTv$. Thus we know that T is a homomorphism from L' into L and have verified (i) of this lemma.

It is obvious that if the space spanned by S coincides with V, then T is a mapping from K^r onto V and that if elements of S are linearly independent, T is a one-to-one mapping. With these discussions and (i) of this lemma, we may obtain (ii)–(iv). Q.E.D.

As known from the above discussions, for any l.a. we can construct the one with the row vector space as its state space, which is equivalent to it. Such l.a.'s will be called l.a. *represented by matrices* to distinguish them from the others. If we take the set S of (7.5) as the basis of V, the space spanned by S is V and obviously the invariant subspace containing the initial state u.

Here we notice that if the elements of S are not linearly independent, the expressions of (7.6) and (7.7) are not unique and so u' and A_i' are not uniquely determined.

Now, we proceed to consider the invariant subspace of l.a. represented by matrices. We consider l.a. $L = \langle K^n, \{A_i\}, u, v\rangle$, where A_i's are $n \times n$ square matrices, $u \in K^n$ and $v^t \in K^n$.

Let V be the invariant subspace of K^n of L, its dimension be n_1, and its basis be $S_1 = \{e_1, e_2, \ldots, e_{n_1}\}$. Hence, as is well known, we can select $S_2 = \{e_{n_1+1}, e_{n_1+2}, \ldots, e_n\}$ so that $S = S_1 \cup S_2 = \{e_1, e_2, \ldots, e_n\}$ becomes the basis of K^n. Thus, for $\forall \sigma_i \in \Sigma$ and $\forall e_k \in S$, $e_k A_i$ can be uniquely represented by

$$e_k A_i = \sum_{j=1}^{n} \beta_{kj}^{(i)} e_j; \qquad k = 1, 2, \ldots, n. \tag{7.12}$$

On the other hand, since S_1 have been assumed to be the basis of the invariant subspace, for e_k $(k = 1, 2, \ldots, n_1)$, $e_k A_i$ can be represented by a linear combination of elements of S_1. Then, because $\beta_{kj}^{(i)}$'s are the uniquely determined values, we have

$$\beta_{kj}^{(i)} = 0; \qquad k = 1, 2, \ldots, n_1, \quad j = n_1 + 1, \ldots, n. \tag{7.13}$$

Here, denoting (k, j) element of the matrix A_i by $\alpha_{kj}^{(i)}$ and putting $e_k = (\xi_{k1}, \xi_{k2}, \ldots, \xi_{kn})$; $k = 1, 2, \ldots, n$, $e_k A_i$ can be represented by

$$e_k A_i = (\zeta_1, \zeta_2, \ldots, \zeta_n), \tag{7.14}$$

where

$$\zeta_r = \sum_{j=1}^{n} \xi_{kj} \alpha_{jr}^{(i)}. \tag{7.15}$$

Concerning ζ_r, we also have from (7.12)

$$\zeta_r = \sum_{j=1}^{n} \beta_{kj}^{(i)} \xi_{jr} . \tag{7.16}$$

Comparing (7.15) and (7.16), we have

$$\sum_{j=1}^{n} \xi_{kj} \alpha_{jr}^{(i)} = \sum_{j=1}^{n} \beta_{kj}^{(i)} \xi_{jr} . \tag{7.17}$$

Thus, if we denote $n \times n$ square matrix having $\beta_{kj}^{(i)}$ as its (k, j) entry by B_i and $n \times n$ square matrix having ξ_{kj} as its (k, j) entry by Q, (7.17) means $QA_i = B_i Q$. Since S is the basis of K^n and ξ_{kj} has been defined as the j-th component of $e_k \in S$, Q is nonsingular and so there exists its inverse Q^{-1}. Thus, we have

$$QA_i Q^{-1} = B_i . \tag{7.18}$$

Here, we notice that from (7.13) B_i has the form of

$$B_i = \begin{array}{c} \begin{array}{cc} \overbrace{\hphantom{B_i'}}^{n_1} & \overbrace{\hphantom{D_i}}^{n - n_1} \end{array} \\ \left(\begin{array}{c:c} B_i' & O \\ \hdashline C_i & D_i \end{array} \right) \end{array} \begin{array}{l} \} n_1 \\ \} n - n_1 \end{array} , \tag{7.19}$$

where B_i' is the $n_1 \times n_1$ square matrix. Thus, if l.a. L has an n_1-dimensional invariant subspace, then there exists a nonsingular square matrix Q such that (7.18) holds.

Conversely, assume that there exists a nonsingular $n \times n$ square matrix Q such that QA_iQ^{-1} is in the form of the right hand side of (7.19). Hence if we denote (k, j) entry of Q by ξ_{kj} and (k, j) entry of QA_iQ^{-1} by $\beta_{kj}^{(i)}$, then we know that (7.17) holds. Here, if we put $e_k = (\xi_{k1}, \xi_{k2}, \ldots, \xi_{kn})$; $k = 1, 2, \ldots, n$, it follows from the nonsingularity of Q that the set $\{e_1, e_2, \ldots, e_n\}$ is the basis of K^n. Since we have assumed that QA_iQ^{-1} is in the form of the right hand side of (7.19), (7.13) holds. Thus, from (7.17), we know that e_kA_i can be represented by a linear combination of $\{e_1, e_2, \ldots, e_{n_1}\}$ for $k = 1, 2, \ldots, n_1$. That is, the n_1-dimensional subspace spanned by S_1 is an invariant subspace of K^n of L.

From these discussions, we have the following theorem:

THEOREM 7.1. l.a. $L = \langle K^n, \{A_i\}, u, v\rangle$ *has an* n_1*-dimensional invariant subspace if and only if there exists a nonsingular matrix* Q *such that, for all* $\sigma_i \in \Sigma$, QA_iQ^{-1} *is in the form of the right hand side of* (7.19).

Here, if we have Q such that QA_iQ^{-1} is in the form of the right hand side of (7.19) for $\forall\sigma_i \in \Sigma$, from l.a. L we can construct a new l.a. $L' = \langle K^n, \{B_i\}, u', v'\rangle$, where $B_i = QA_iQ^{-1}$, $u' = uQ^{-1}$, $v' = Qv$. Hence, Q^{-1}: $K^n \to K^n$ is an isomorphism from L onto L'. In a sense, the construction of such l.a. L' may be considered as the reduction of l.a. But, it is a rather difficult problem left open for decision whether or not there exists any invariant subspace except $\{0\}$ and K^n, that is, whether l.a. L is strongly connected or not.

7.3. *Dual* l.a. *and Properties of Connectivity and Distinguishability*

For l.a. $L = \langle K^n, \{A_i\}, u, v\rangle$ we define $L^t = \langle K^n, \{A_i^t\}, v^t, u^t\rangle$. This L^t will be called the *transpose* of L or *dual* l.a. It is clear from the definition that $(L^t)^t = L$.

LEMMA 7.2. *Denoting the response mappings of* L *and* L^t *by* F *and* G, *respectively, and their output functions by* f *and* g, *respectively, for* $\forall w \in \Sigma^*$ *we have*

$$G(w) = (F(w^R))^t, \tag{7.20}$$

$$g(w) = f(w^R) \tag{7.21}$$

where w^R *denotes the reverse of a tape* w, *which is recursively defined by* $\epsilon^R = \epsilon$ *and* $(w\sigma)^R = \sigma w^R$.

Next, the dual properties of connectivity and distinguishability are summarized in the form of a theorem.

THEOREM 7.2. (i) *L is connected if and only if L^t is distinguishable.*

(ii) *L is distinguishable if and only if L^t is connected.*

(iii) *L is irreducible if and only if L^t is irreducible.*

(iv) *L is strongly connected if and only if L^t is strongly connected.*

Proof. (i) Assume that L is connected. Hence, for arbitrary elements x_1 and x_2 such that $x_1 \neq x_2$, we put $x = x_1 - x_2$. From the assumption of connectivity of L, x can be represented by

$$x = \alpha_1 uF(z_1) + \alpha_2 uF(z_2) + \cdots + \alpha_m uF(z_m); \qquad \alpha_i \ (i = 1,..., m) \in K, \tag{7.22}$$

where $z_1, z_2,..., z_m$ are the tapes with length $(n - 1)$ or less, which have been defined in Section 6. Since we have assumed that $x_1 \neq x_2$ and so $x \neq 0$, we have

$$xx^t \neq 0. \tag{7.23}$$

Substituting (7.22) to (7.23), we have $\alpha_1 x\{uF(z_1)\}^t + \cdots + \alpha_m x\{uF(z_m)\}^t \neq 0$. Thus, there exists at least an integer $k (1 \leqslant k \leqslant m)$ such that $x\{uF(z_k)\}^t = xG(z_k{}^R)\, u^t \neq 0$ and therefore $x_1 G(z_k{}^R)\, u^t \neq x_2 G(z_k{}^R)\, u^t$. Thus, L^t is distinguishable.

Next assume that L is not connected. Then, if W denotes the connected part of the initial state u of L, then $W \subsetneq K^n$. Hence, there exists a nonzero element a in K^n which is orthogonal to all elements in W. For such an element $a \in K^n$ and all $w \in \Sigma^*$, we have $aG(w)\, u^t = a\{uF(w^R)\}^t = uF(w^R)\, a^t$. Since W is the connected part of u and so $uF(w^R) \in W$ and since the element a is orthogonal to any x in W, i.e., $xa^t = 0$ for all $x \in W$, we have $uF(w^R)\, a^t = 0$. Thus, $aG(w)\, u^t = 0$ holds for all $w \in \Sigma^*$. Therefore, a and 0 can not be distinguished by L^t.

(ii) Obvious from (i) since $L = (L^t)^t$.

(iii) Obvious from (i), (ii) and Corollary 6.1.

(iv) Let V denote a subspace of K^n and define the annihilator $V^\perp$ of V by $V^\perp = \{y \mid yx^t = 0, \text{ for } \forall x \in V\}$. Here, if V is an invariant subspace of L, then $V^\perp$ is also an invariant subspace of L^t. Because, for $\forall y \in V^\perp$, $yA_i{}^t x^t = y(xA_i)^t$. Therefore, noticing $xA_i \in V$ for $\forall x \in V$, we have $yA_i{}^t x^t = y(xA_i)^t = 0$. This means that $yA_i{}^t \in V^\perp$ and so $V^\perp$ is an invariant subspace of L^t.

Since $\dim V + \dim V^\perp = n$, if V is neither $\{0\}$ nor K^n then so $V^\perp$.

Accordingly, it follows from Lemma 4.2 that if L is not strongly connected then so L^t.

The inverse of this also holds since $L = (L^t)^t$. Q.E.D.

One more property is described concerning the duality between L and L^t.

THEOREM 7.3. *Consider two* l.a.'s, $L_1 = \langle K^{n_1}, \{A_i\}, u_1, v_1 \rangle$ *and* $L_2 = \langle K^{n_2}, \{B_i\}, u_2, v_2 \rangle$. *Let T be a mapping from K^{n_1} into K^{n_2}. Then,*

(i) *T is a homomorphism from L_1 into L_2 if and only if T^t is a homomorphism from $L_2{}^t$ into $L_1{}^t$.*

(ii) *T is a homomorphism from L_1 onto L_2 if and only if T^t is an isomorphism from $L_2{}^t$ into $L_1{}^t$.*

(iii) *T is an isomorphism from L_1 onto L_2 if and only if T^t is an isomorphism from $L_2{}^t$ onto $L_1{}^t$.*

Proof. The necessary and sufficient condition that T is a homomorphism from L_1 into L_2 is that

$$A_iT = TB_i, \qquad u_1T = u_2 \qquad \text{and} \qquad v_1 = Tv_2 \tag{7.24}$$

hold. Applying the transpose operation to the both sides of these equations, we have

$$B_i^tT^t = T^tA_i^t, \qquad v_2^tT^t = v_1^t \qquad \text{and} \qquad u_2^t = T^tu_1^t. \tag{7.25}$$

Conversely, if (7.25) holds, then (7.24) holds. Thus, we have confirmed the validity of (i) of this theorem.

The statements (ii) and (iii) are obvious from (i). Q.E.D.

7.4. *Eigenvalues of Response Mapping*

In this section we will make some consideration on eigenvalues of response mappings of l.a. L.

LEMMA 7.3. *Let A be an $m \times m$ square matrix and B be an $n \times n$ square matrix. Assume that an $m \times n$ matrix T satisfies*

$$AT = TB. \tag{7.26}$$

Then, if rank $T = r$ ($\leqslant m, n$), A and B hold at least r eigenvalues in common.

Since this lemma can easily be proved by using well-known properties of linear space, the proof is omitted.

LEMMA 7.4. *For* l.a.*'s L_1 and L_2 of Theorem* 7.3, *if there exists a homomorphism T with rank r, then for $\forall w \in \Sigma^*$ the response mappings $F_1(w)$ and $F_2(w)$ of L_1 and L_2 hold at least r eigenvalues in common.*

Proof. If T is a homomorphism from L_1 into L_2, for $\forall w \in \Sigma^*$,

$$yF_1(w)T = yTF_2(w) \qquad \text{for} \quad y \in K^{n_1}. \tag{7.27}$$

Thus,

$$F_1(w)T = TF_2(w). \tag{7.28}$$

If we combine (7.28) and Lemma 7.3, this lemma results. Q.E.D.

LEMMA 7.5. *Let L_1, L_2 and T be the same as Lemma* 7.4. *Then,*

(i) *If T is a homomorphism from L_1 onto L_2, then the eigenvalues of $F_1(w)$ are contained in those of $F_2(w)$.*

(ii) *If T is an isomorphism from L_1 into L_2, then the eigenvalues of $F_1(w)$ are contained in those of $F_2(w)$.*

(iii) *If T is an isomorphism from L_1 onto L_2, then the eigenvalues of $F_1(w)$ are identical to those of $F_2(w)$.*

Proof. (i) Since rank $T = r = n_2$, $F_1(w)$ and $F_2(w)$ hold n_2 eigenvalues in common. On the other hand, $F_2(w)$ has just n_2 eigenvalues. Therefore, the eigenvalues of $F_2(w)$ are contained in those of $F_1(w)$. The (ii) and (iii) are justified in a similar way. Q.E.D.

THEOREM 7.4. *Let L_1 and L_2 be equivalent and L_2 be irreducible. Then, for $\forall w \in \Sigma^*$, the eigenvalues of the response mapping $F_2(w)$ of L_2 are contained in those of $F_1(w)$ of L_1.*

Proof. Let $V(u_1)$ denote the connected part of the initial state u_1 of L_1. Assume that dim $V(u_1) = r$ and $S = \{a_1', \ldots, a_r'\}$ is the basis of $V(u_1)$. Here we construct $L_1' = \langle K^r, \{A_i'\}, u_1', v_1' \rangle$ by the same way as in the case of the l.a. of (7.9). Hence, the mapping $T_1 : K^r \to K^{n_1}$, which is defined similarly to (7.10), is an isomorphism from L_1' into L_1. This results from (iii) of Lemma 7.1, since the elements of S_1 are linearly independent. Thus, by (ii) of Lemma 7.5, the eigenvalues of the response mapping $F_1'(w)$ of L_1' are contained in those of $F_1(w)$ for all $w \in \Sigma^*$. It is obvious from the construction procedure of L_1' that L_1' is the connected part of L_1. Therefore, it follows from Corollary 6.3 that there exists a homomorphism from L_1' onto L_2. Hence, from (i) of Lemma 7.5, the eigenvalues of $F_2(w)$ are contained in those of $F_1'(w)$. Thus the eigenvalues of $F_2(w)$ are contained in those of $F_1(w)$. Q.E.D.

8. l.a. AND PROBABILISTIC AUTOMATA

For the convenience of describing relations between l.a. and probabilistic automaton we will consider only l.a.'s defined over the real field R instead of K in this and the following sections. Thus, if we write simply l.a., it will mean l.a. defined over R.

Under the above assumption, A_i is a mapping from the linear space over R into itself and $f(w)$ is a real function over Σ^*. Hence, similarly to the case of p.a., the set of tapes $T(L)$ which are accepted by l.a. L may be defined as follows.

DEFINITION 8.1.[4] The *set of tapes accepted by* l.a. is defined by

$$T(L) = \{w \in \Sigma^* \mid f(w) > 0\}. \tag{8.1}$$

If for two l.a.'s L_1 and L_2,

$$T(L_1) = T(L_2) \tag{8.2}$$

holds, that is,

$$f_1(w) > 0 \qquad \text{if only if} \quad f_2(w) > 0, \tag{8.3}$$

then L_1 and L_2 are said to be weakly equivalent.

8.1. l.a.*'s with Bounded Output Functions and Extension of Rabin's Theorem*

As the first step of describing the relations between l.a.'s and p.a.'s we intend to extend the Rabin's result to the case of l.a., which states that the tape set accepted by a p.a. with isolated cutpoint is a regular set.

In this section we consider l.a. L with the linear space V over the real field R.

$$L = \langle V, \{A_i \mid i = 1, 2,..., \Gamma\}, u, v\rangle \tag{8.4}$$

DEFINITION 8.2. A real function $\rho(x)$ over the linear space V is called the norm of V, if the following conditions are satisfied

(i) $$0 \leqslant \rho(x) < \infty \qquad \text{for} \quad \forall x \in V, \tag{8.5}$$

(ii) $$\rho(x) = 0 \qquad \text{if and only if} \quad x = 0, \tag{8.6}$$

(iii) $$\rho(x_1 + x_2) \leqslant \rho(x_1) + \rho(x_2) \qquad \text{for} \quad \forall x_1, \forall x_2 \in V, \tag{8.7}$$

(iv) $$\rho(\alpha x) = |\alpha| \rho(x) \qquad \text{for} \quad \forall x \in V, \quad \forall \alpha \in R. \tag{8.8}$$

LEMMA 8.1. *If there exists a norm $\rho(x)$ such that the following condition*

$$\rho(xA_i) \leqslant \rho(x) \qquad \textit{for} \quad \forall x \in V \quad \textit{and} \quad \forall \sigma_i \in \Sigma \tag{8.9}$$

holds, then the output function $f(w)$ of L is bounded, that is, there exists a constant positive number M such that

$$|f(w)| \leqslant M \qquad \textit{for} \quad \forall w \in \Sigma^*. \tag{8.10}$$

Proof. Using (8.9) iteratively, we obtain

$$\rho(xF(w)) \leqslant \rho(x) \qquad \text{for} \quad \forall x \in V \quad \text{and} \quad \forall w \in \Sigma^*. \tag{8.11}$$

Especially, putting $x = u$ we have $\rho(uF(w)) \leqslant \rho(u)$.

[4] We consider only 0 as the cutpoint since by using the result of Honda and Nasu (1968) we can easily prove that the families of tape sets accepted by l.a.'s with arbitrary cutpoints and with the one fixed to 0 are the same.

On the other hand, it is well known that, for a linear real function v defined on the linear space V over R and for a norm ρ of V, there exists a constant $a > 0$ such that $|xv| < a\rho(x)$ for any $x \in V$. Thus, the output function $f(w) = uF(w)v$ is bounded. Q.E.D.

LEMMA 8.2. *If* l.a. L *is irreducible and its output function is bounded, there exists a norm satisfying* (8.9).

Proof. For $x \in V$, let us define $\rho(x)$ by

$$\rho(x) = \sup\{|xF(w)v| \mid w \in \Sigma^*\}. \tag{8.12}$$

Then this $\rho(x)$ is the norm satisfying (8.9). In fact, this is proved as follows.

First, it is obvious that ρ satisfies (8.7) and (8.8). Here, we will prove that (8.5) and (8.6) hold. From the definition of $\rho(x)$ of (8.12), we know $0 \leqslant \rho(x)$ for $\forall x \in V$ and $\rho(0) = 0$.

Thus, for the proof that (8.5), (8.6) hold and then ρ is a norm, it is sufficient only to show that

$$x \neq 0 \quad \text{implies} \quad \rho(x) \neq 0 \tag{8.13}$$

and

$$\rho(x) < \infty \qquad \text{for} \quad \forall x \in V. \tag{8.14}$$

As we have assumed that l.a. L is irreducible by Corollary 6.1, it is connected and distinguishable. As L is distinguishable, 0 and $x(\neq 0)$ are distinguishable, that is, there exists at least a tape $w \in \Sigma^*$ such that $xF(w)v \neq 0F(w)v = 0$. Thus, from the definition of $\rho(x)$, we know that (8.13) holds.

Next, if $x \in V$ can be represented by $x = uF(z)$ for some tape $z \in \Sigma^*$, then $\rho(x) = \sup\{|uF(z)F(w)v| \mid w \in \Sigma^*\} = \sup\{|uF(zw)v| \mid w \in \Sigma^*\}$. Here, as f is bounded, we have $\rho(x) \leqslant M < \infty$ from (8.10). As L is connected, $\forall x \in V$ can be represented by $x = \sum_{i=1}^{m} \alpha_i uF(z_i)$, $\alpha_i \in R$, where z_i's ($i = 1,..., m$) are the tapes defined in Section 6. On the other hand, as we have known that ρ satisfies (8.7) and (8.8), we obtain

$$\rho(x) \leqslant \sum_{i=1}^{m} |\alpha_i| \rho(uF(z_i)) \leqslant \left(\sum_{i=1}^{m} |\alpha_i|\right) M < \infty.$$

Thus, we have shown that (8.14) holds and that ρ is a norm.

Finally, it is obvious from (8.12) that (8.9) holds. Q.E.D.

Now we extend Rabin's result to the case of l.a. with the bounded output function f. We assume that 0 is the isolated cutpoint of f, that is, there exists $\delta > 0$ such that

$$\delta \leqslant |f(w)| \qquad \text{for} \quad \forall w \in \Sigma^*. \tag{8.15}$$[5]

Without loss of generality, we also assume that L is irreducible. Therefore, from Lemma 8.2 we can define the norm $\rho(x)$ satisfying (8.9).

Now, let us consider the right invariant equivalence relation $\equiv_{T(L)}$ over over Σ^* generated by the tape set

$$T(L) = \{w \in \Sigma^* \mid f(w) > 0\} = \{w \in \Sigma^* \mid \delta \leqslant f(w) \leqslant M\}. \tag{8.16}$$

That is, the relation $\equiv_{T(L)}$ is defined as follows; for $\forall w_1, \forall w_2 \in \Sigma^*$, $w_1 \equiv_{T(L)} w_2$ if and only if, for $\forall z \in \Sigma^*$, $w_1 z \in T(L)$ implies $w_2 z \in T(L)$ and conversely. We prove in the following that the relation $\equiv_{T(L)}$ has a finite index.

Assume that two tapes w_1 and w_2 do not satisfy the relation $\equiv_{T(L)}$, then there exists at least a tape $z \in \Sigma^*$ such that either of the following conditions:

$$\text{(i)} \qquad w_1 z \in T(L) \quad \text{and} \quad w_2 z \notin T(L) \tag{8.17}$$

$$\text{(ii)} \qquad w_1 z \notin T(L) \quad \text{and} \quad w_2 z \in T(L) \tag{8.18}$$

holds. From (8.15), it is seen that (8.17) implies

$$f(w_2 z) \leqslant -\delta < 0 < \delta \leqslant f(w_1 z) \tag{8.19}$$

and (8.18) implies

$$f(w_1 z) \leqslant -\delta < 0 < \delta \leqslant f(w_2 z). \tag{8.20}$$

Thus, whichever condition holds, we have

$$|f(w_1 z) - f(w_2 z)| \geqslant 2\delta. \tag{8.21}$$

Therefore,

$$\begin{aligned} \rho(uF(w_1) - uF(w_2)) &= \sup\{|(uF(w_1) - uF(w_2))\, F(w)v| \mid w \in \Sigma^*\} \\ &= \sup\{|f(w_1 w) - f(w_2 w)| \mid w \in \Sigma^*\} \\ &\geqslant 2\delta. \end{aligned} \tag{8.22}$$

On the other hand, from (8.10) we have

$$\rho(uF(w_1)) \leqslant M, \qquad \rho(uF(w_2)) \leqslant M. \tag{8.23}$$

[5] See footnote 4.

Now, let $\{w_1, w_2, ..., w_K\}$ be a set of tapes in which any pair of distinct tapes does not satisfy the relation $\equiv_{T(L)}$. Putting $x_i = uF(w_i)$; $i = 1, 2, ..., K$, it follows from (8.28) that

$$\rho(x_j - x_k) \geqslant 2\delta \qquad \text{for} \quad \forall j, \quad \forall k, \quad 1 \leqslant j \neq k \leqslant K, \tag{8.24}$$

and

$$\rho(x_j) \leqslant M \qquad \text{for} \quad \forall j, \quad 1 \leqslant j \leqslant K \tag{8.25}$$

hold.

Now, let us define the pseudo sphere g_j with the center x_j, and the radius δ by $g_j = \{x \mid \rho(x - x_j) < \delta\}$. Then, from (8.24), we have

$$g_i \cap g_k = \varnothing \text{ (empty set)} \qquad \forall j, k, \quad j \neq k. \tag{8.26}$$

Let the pseudo sphere G of which center is the origin and radius is $(M + \delta)$ be $G = \{x \mid \rho(x) \leqslant M + \delta\}$. Then, from (8.25) and the definition of g_j we obtain $g_j \subseteq G$ for $j = 1, 2, ..., K$. Thus,

$$\bigcup_{j=1}^{K} g_j \subseteq G. \tag{8.27}$$

Now, assume that dim $V = n$. By introducing an appropriate basis to the space V, we can determine the one-to-one mapping T from V to R^n. Here, for a subset S of V, we define its volume by the one of the image of S by the mapping T. Then, regardless the selection of the basis, the following lemma holds.

LEMMA 8.3. *Let the volume of the pseudo sphere with radius* 1 *be denoted by* a. *Then,* $0 < a < \infty$. *And the volume of a pseudo sphere with radius* r *is* ar^n, *which is determined only by the radius.*

From this lemma and (8.26) and (8.27), we have $Ka\delta^n \leqslant a(M + \delta)^n$, and then $K \leqslant (M + \delta)^n/\delta^n$. Thus, we have known that the right invariant equivalence relation $\equiv_{T(L)}$ has finite index.

The above discussions prove the following theorem:

THEOREM 8.1. *Let* L *be a linear space automaton with n-dimensional state space. If its output function is bounded and the cutpoint* 0 *is isolated with respect to* L, *that is, if there exist* δ *and* M *such that* $\delta \leqslant |f(w)| \leqslant M$ *for* $\forall w \in \Sigma^*$ *and* $0 < \delta \leqslant M < \infty$, *then the tape set* $T(L)$ *accepted by* L *is the regular one* *and there exists a deterministic finite state automaton with* N *or less states which accepts* $T(L)$, *where*

$$N = (M + \delta)^n/\delta^n. \tag{8.28}$$

Here we notice that Theorem 8.1 also holds even if l.a. L is not irreducible.

8.2. l.a. *and Modified Probabilistic Automata*

The main object of this section is that we introduce the modified probabilistic automata as the special l.a.'s and show that the tape acceptance abilities of l.a.'s and modified probabilistic automata are equivalent. Using Honda and Nasu's (1968) results, we can easily show that the abilities of the modified probabilistic automata and the probabilistic automata defined by Rabin (1963) are equivalent. Hence, we know that l.a.'s have the same ability as the probabilistic automata. Furthermore, we show the conditions that a given l.a. should satisfy so that it has the strongly equivalent modified p.a.

DEFINITION 8.3. Let $L = \langle R^n, \{A_i\}, u, v\rangle$ be l.a. of which state space is the n-dimensional row vector space R^n over R. If each A_i is a stochastic matrix and u is an n-dimensional stochastic row vector, then it will be called *modified p.a.*[6]

THEOREM 8.2. *For an arbitrary* l.a. L, *there exists a modified* p.a. L_p *which is weakly equivalent to* L. *That is, there exists a modified* p.a. L_p *such that* $T(L) = T(L_p)$.

By Lemma 7.1, we have known that for an arbitrary l.a. there exists an equivalent one whose state space is a row vector space. Thus, it is sufficient to consider l.a. of Definition 8.3.

The detailed proof is omitted since Turakainen (1968) gave the proof for l.a. of Definition 8.3. But the more concise and more straightforward proof was given by the authors (1968) independently of him, although these two proofs are essentially same. Turakainen (1968) also proved that the tape acceptance abilities of the generalized automata (l.a.'s), the generalized p.a.'s (modified p.a.'s) and p.a.'s are same.

It should be noticed that for a general l.a. there does not always exist a modified p.a. which is strongly equivalent to it, because the output function of any modified p.a. is bounded but that of l.a. is not necessarily bounded.

[6] The modified p.a. corresponds to the generalized probabilistic automaton of Turakainen (1968).

Hence, in what follows, we consider on what conditions a given l.a. has the equivalent modified p.a.

Let $L = \langle V, \{A_i\}, u, v\rangle$ be l.a. with the linear space over the real field R. Let $S = \{a_1, a_2, ..., a_r\}$ be a finite subset of V. Let us denote the convex polyhedron of which extreme points are all elements of S by Conv(S), that is,

$$\mathrm{Conv}(S) = \left\{x \,\middle|\, x = \sum_{i=1}^{r} \zeta_i a_i\,, \zeta_i \geqslant 0, \sum_{i=1}^{r} \zeta_i = 1\right\}. \tag{8.29}$$

LEMMA 8.4. *If there exists S satisfying the following conditions (i) and (ii), then* l.a. L *has an equivalent modified* p.a.

(i) $u \in \mathrm{Conv}\,(S)$,

(ii) Conv(S) is an invariant subset of V.

Proof. Since the conditions (i) and (ii) hold, l.a. L' defined by (7.9) with $K = R$ can be reconstructed to be a modified p.a. That is, from the condition (i), the right hand side of (7.6) can be represented by a convex combination of elements of S [7]. Moreover, from the condition (ii), for any $a_k \in S$, $a_k A_i$ can be represented by a convex combination of elements of S. Thus, u' of L' of (7.9) and $\alpha_{kj}^{(i)}$ defined by (7.7) can be determined so that u' may be a stochastic vector and A_i' having $\alpha_{kj}^{(i)}$ as its (k, j) entry may be a stochastic matrix. Hence, l.a. L' is a modified p.a. and equivalent to L by Lemma 7.1. Q.E.D.

If l.a. L is irreducible, the inverse of Lemma 8.4 also holds. This will be shown in the following lemma.

LEMMA 8.5. *If* l.a. L *is irreducible and has an equivalent modified* p.a., *then there exists a subset S satisfying the conditions* (i) *and* (ii) *of Lemma* 8.4.

Proof. Let the modified p.a. equivalent to L be

$$L_p = \langle R^m, \{P_i\}, u_p\,, v_p\rangle \tag{8.30}$$

and denote the set of all m-dimensional stochastic row vectors by Ω^m. Ω^m is the convex polyhedron with m extreme points, which are the row vectors having 0 elements except only one 1 element.

[7] But we should notice that, if the elements of S are not linearly independent, $\theta_1, \theta_2, ..., \theta_r$ in the right hand side of (7.6) are not uniquely determined as already mentioned and so there may exist representations not to be convex combination.

Since P_i is stochastic matrix, $x \in \Omega^m$ implies $xP_i \in \Omega^m$. This means that Ω^m is an invariant subset of L_p.

Here, according to Definition 6.2, the connected part $L_p{}^c$ of L_p is defined by $L_p{}^c = \langle W_p, \{B_i{}^p\}, u_p, v_p \rangle$. Since $x \in W_p \cap \Omega^m$ implies $xB_i{}^p = xP_i \in W_p \cap \Omega^m$, $W_p \cap \Omega^m$ is an invariant subspace of $L_p{}^c$. Because the intersection of any subspace V of R^n and the convex polyhedron Conv (S), where S is a finite subset of R^n, is also a convex polyhedron[8], $W_p \cap \Omega^m$ is a convex polyhedron. That is, there exists a finite subset S_1 such that $W_p \cap \Omega^m =$ Conv (S_1).

On the other hand, since L is irreducible and equivalent to L_p, from Lemma 6.3 there exists a homomorphism from $L_p{}^c$ onto L. If the image of S_1 by T_1 is denoted by S_2, the image of Conv (S_1) by T_1 is obviously Conv (S_2). As Conv (S_1) is an invariant subset of $L_p{}^c$ and contains u_p, it is obvious that Conv (S_2) is an invariant subset of L and contains u. Thus, S_2 satisfies the conditions (i) and (ii) of Lemma 8.4. Q.E.D.

Combining Lemmas 8.4 and 8.5, the following theorem results.

THEOREM 8.3. *If* l.a. L *is irreducible, the necessary and sufficient condition that there exists a modified* p.a. *equivalent to* L *is that there exists* S *satisfying the conditions* (i) *and* (ii) *of* Lemma 8.4.

Here, we notice that even if a given l.a.'s output function is bounded, it does not always have a strongly equivalent modified p.a. This is shown by the following example.

EXAMPLE 8.1. Let $\Sigma = \{\sigma\}$ and consider $L = \langle R^2, \{A\}, u, v \rangle$, where

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

and θ/π is not rational, and $u = (1, 0)$, $v = (0, 1)^t$.

The output function of this l.a. is $f(w) = \sin l\theta$, for $w = \sigma^l \in \Sigma^*$. Obviously, this function is bounded. On the other hand, the output functions of l.a.'s with the state space of which dimensions are 1 or less must be represented in the form $g(w) = \alpha \cdot \beta^l$, $\alpha, \beta \in R$. Thus, l.a. L is irreducible.

Now, we know from Theorem 7.4 that the set of eigenvalues of the response function of l.a.'s equivalent to L must contain the eigenvalues of A for $\sigma \in \Sigma$. But eigenvalues of A are $e^{\pm j\theta}$ and θ/π is not rational. Therefore these eigen-

[8] For example, refer to Valentine (1964).

values can not be ones of the stochastic matrix. Thus, there does not exist any modified p.a. equivalent to L.

As shown in the above example, even if the output function of l.a. is bounded, it does not always have the equivalent modified p.a. But if the output function is bounded, the following corollary holds.

COROLLARY 8.1. *If* l.a. $L = \langle V, \{A_i\}, u, v\rangle$ *has the output function* $f(w)$ *which is bounded, for any* ϵ *such that* $0 < \epsilon < 1$ *there exists a modified* p.a. *with the output function*

$$g(w) = \epsilon^{l(w)} f(w) \tag{8.31}$$

where $l(w)$ *is the length of* w.

Proof. We can assume without loss of generality that L is irreducible. Hence, by Lemma 8.2, there exists the norm $\rho(x)$ of V satisfying the condition, $\rho(xA_i) \leqslant \rho(x)$ for $\forall x \in V$. Here, let us define l.a. L' by $L' = \langle V, \{B_i\}, u, v\rangle$, where $B_i = \epsilon A_i$. Clearly, the output function of this l.a. L' is $g(w)$ of (8.31).

Considering the properties of $\rho(x)$, we have

$$\rho(xB_i) \leqslant \epsilon\rho(x) \qquad \text{for} \quad \forall x \in V. \tag{8.32}$$

Here, put

$$a = \rho(u) \qquad \text{and} \qquad b = \frac{1}{\epsilon} a. \tag{8.33}$$

If $a = 0$ then $u = 0$ and $f(w)$ is identically equal to zero. In such case this corollary is trivial. Thus we assume that $a \neq 0$. Then, $0 < a < b$. Hence, there exists a convex polyhedron Conv (S) such that[9]

$$\{x \mid \rho(x) \leqslant a\} \subseteq \mathrm{Conv}\,(S) \subseteq \{x \mid \rho(x) \leqslant b\}. \tag{8.34}$$

By (8.32), (8.33) and (8.34), we know that

$$\begin{gathered} x \in \mathrm{Conv}\,(S) \quad \text{implies} \quad \rho(x) \leqslant b, \\ \rho(x) \leqslant b \quad \text{implies} \quad \rho(xB_i) \leqslant \epsilon b = a, \\ \rho(xB_i) \leqslant a \quad \text{implies} \quad xB_i \in \mathrm{Conv}\,(S). \end{gathered}$$

Therefore, Conv (S) is an invariant subset of L'.

On the other hand, from (8.33) and (8.34), we know that $u \in \mathrm{Conv}\,(S)$. Thus, it follows from Lemma 8.4 that there exists a modified p.a. strongly equivalent to L'. Q.E.D.

[9] This statement is proved in Klee (1959)'s Corollary 6.4, p. 104.

9. Construction of l.a. and Quasi-Regular Function[10]

Since for any given l.a. we can obtain the equivalent l.a. with the state space which is the row vector space, we may consider the latter l.a.'s in this section. Moreover, for the simplicity of description, we will denote l.a. by

$$L = \langle \{A_i\}, u, v \rangle \tag{9.1}$$

leaving out the symbol of the state space, where A_i's, u and v are the square matrices, the row vector and the column vector, respectively, and these dimensions are the same.

Definition 9.1. The real function f over Σ^* is called *quasi-regular function* if there exists l.a. having f as its output function. Let us denote the family of all the quasi-regular functions by $\mathscr{L}$.

For the construction of l.a.'s, it is convenient to know under what operations the family $\mathscr{L}$ is closed. Hence we define some operations on $\mathscr{L}$ and will show that $\mathscr{L}$ is closed under these operations.

Definition 9.2. Let f and g be real functions over Σ^*. For f and g, let us define new real functions over Σ^*:

(i) *sum* of f and g; $(f+g)(w) = f(w) + g(w)$, (9.2)

(ii) *product* of f and g; $f \cdot g(w) = f(w) \cdot g(w)$, (9.3)

(iii) *convolution* of f and g; $f \circ g(w) = \sum_{w_1 \cdot w_2 = w} f(w_1) \cdot g(w_2)$, (9.4)

where $\sum_{w_1 \cdot w_2 = w}$ means the summation over all the pairs w_1, w_2 such that $w_1 \cdot w_2 = w$.

Next, let us define the operations for l.a.'s corresponding to the above defined operations on $\mathscr{L}$.

Definition 9.3.[11] Let two l.a.'s L_1 and L_2 be $L_1 = \langle \{A_i\}, u_1, v_1 \rangle$ and $L_2 = \langle \{B_i\}, u_2, v_2 \rangle$. Then, we define new l.a.'s using L_1 and L_2:

(i) *direct sum* of L_1 and L_2:

$$L_1 \oplus L_2 = \langle \{A_i \oplus B_i\}, u_1 \oplus u_2, v_1 \oplus v_2 \rangle, \tag{9.5}$$

[10] The discussions which will be developed in this section is also warranted in the case of general field K instead of the real field R.

[11] As to operations $\oplus$ and $\otimes$, refer to the Section 7.1.

(ii) *direct product* of L_1 and L_2 :

$$L_1 \otimes L_2 = \langle \{A_i \otimes B_i\}, u_1 \otimes u_2, v_1 \otimes v_2 \rangle, \tag{9.6}$$

(iii) *convolution* of L_1 and L_2 :

$$L_1 \circ L_2 = \langle \{C_i\}, u, v \rangle, \tag{9.7}$$

where

$$C_i = \begin{pmatrix} A_i & A_i D \\ 0 & B_i \end{pmatrix}, \qquad D = v_1 u_2$$

$u = u_1 \oplus u_1 v_1 u_2$ and $v = 0 \oplus v_2$.

LEMMA 9.1. *Denote the response mappings of L_1 and L_2 by F_1 and F_2, respectively. Then the response mapping F of $L_1 \circ L_2$ is given by*

$$F = \begin{pmatrix} F_1 & F_{12} \\ 0 & F_2 \end{pmatrix}, \tag{9.8}$$

where F_{12} is determined by

$$F_{12}(w) = -F_1(\epsilon)\, DF_2(w) + \sum_{w_1 \cdot w_2 = w} F_1(w_1)\, DF_2(w_2) \qquad \text{for} \quad \forall w \in \Sigma^*. \tag{9.9}$$

Proof. We can prove the lemma by induction. First, for $w = \epsilon$, we have

$$F(\epsilon) = \begin{pmatrix} F_1(\epsilon) & F_{12}(\epsilon) \\ 0 & F_2(\epsilon) \end{pmatrix} = E \qquad \text{(unit matrix)}. \tag{9.10}$$

Next, assume that (9.8) holds for $w = z$. Thus, for $\forall \sigma_i \in \Sigma$, we have

$$F(z\sigma_i) = F(z)\, C_i = \begin{pmatrix} F_1(z) & F_{12}(z) \\ 0 & F_2(z) \end{pmatrix} \begin{pmatrix} A_i & A_i D \\ 0 & B_i \end{pmatrix}. \tag{9.11}$$

Hence, it follows from (9.9) and the definitions of F_1 and F_2 that

$$F(z\sigma_i) = \begin{pmatrix} F_1(z\sigma_i) & F_{12}(z\sigma_i) \\ 0 & F_2(z\sigma_i) \end{pmatrix}. \tag{9.12}$$

Q.E.D.

COROLLARY 9.1. *Let the output functions of L_1 and L_2 be f_1 and f_2, respectively. Then, the output function of $L_1 \circ L_2$ is $f_1 \circ f_2$.*

Proof. Denoting the output function of $L_1 \circ L_2$ by f, we have

$$f = uFv = (u_1 \oplus u_1 v_1 u_2) \begin{pmatrix} F_1 & F_{12} \\ 0 & F_2 \end{pmatrix} (0 \oplus v_2)$$
$$= \{u_1 F_1 \oplus (u_1 F_{12} + u_1 v_1 u_2 F_2)\}(0 \oplus v_2)$$
$$= u_1 F_{12} v_2 + u_1 v_1 u_2 F_2 v_2 \tag{9.13}$$

from Lemma 9.1 and the definition of u and v. Rewriting (9.13) for an arbitrary tape $w \in \Sigma^*$ by using (9.9), we obtain

$$f(w) = -u_1 F_1(\epsilon)\, v_1 u_2 F_2(w)\, v_2 + \sum_{w_1 \cdot w_2 = w} u_1 F_1(w_1)\, v_1 u_2 F_2(w_2)\, v_2$$
$$+ u_1 v_1 u_2 F_2(w)\, v_2$$
$$= \sum_{w_1 \cdot w_2 = w} u_1 F_1(w_1)\, v_1 u_2 F_2(w_2)\, v_2 = f_1 \cdot f_2(w). \tag{9.14}$$

Q.E.D.

THEOREM 9.1. *If $f_1, f_2 \in \mathscr{L}$ then*

$$f_1 + f_2 \in \mathscr{L}, \qquad f_1 \cdot f_2 \in \mathscr{L} \qquad \text{and} \qquad f_1 \circ f_2 \in \mathscr{L}.$$

Proof. We can easily prove that the output functions of $L_1 \oplus L_2$ and $L_1 \otimes L_2$ are $f_1 + f_2$ and $f_1 \cdot f_2$, respectively. Corollary 9.1 gives the proof for $f_1 \circ f_2$. Q.E.D.

DEFINITION 9.4. For $f \in \mathscr{L}$ we define three unary operations as follows:

(i) *multiplication by constant*: $(\alpha f)(w) = \alpha f(w)$ (9.15)

where $\alpha \in R$.

(ii) *dagger*: $f^+ = f + f \circ f + f \circ f \circ f + \cdots$ (9.16)[12]

This is defined only when $f(\epsilon) = 0$.

(iii) *reverse*: $f^R(w) = f(w^R)$ (9.17)

LEMMA 9.2. *If $f(\epsilon) = 0$ then $f^+(w)$ is the summation of*

$$f(w_1) \cdot f(w_2) \cdot \cdots \cdot f(w_k) \tag{9.18}$$

[12] Since the operation of convolution satisfies the associative law, $(f \circ g) \circ h = f \circ (g \circ h)$ and this is denoted simply by $f \circ g \circ h$.

over all k's $(1 \leqslant k \leqslant l(w))$ *and all the k-tuples* $w_1, w_2, \ldots, w_k$ *such that* $w_1 w_2 \cdots w_k = w$.

Proof. Put

$$g_k = \overbrace{f \circ f \circ \cdots \circ f}^{k},$$

then, from the definition of convolution, $g_k(w)$ is the summation of (9.18) over all k-tuples $w_1, w_2, \ldots, w_k$ such that their concatenation is equal to w. Since $f(\epsilon) = 0$, we know that $g_k(w) = 0$ for $k \geqslant l(w) + 1$. On the other hand, f^+ is represented by $f^+ = g_1 + g_2 + \cdots + g_k + \cdots$. Thus, this lemma holds.

Q.E.D.

DEFINITION 9.5. We define new l.a.'s from l.a. L of (9.1) as follows:

(i) *multiplication of L by constant*: $\alpha L = \langle \{A_i\}, u, \alpha v \rangle$ (9.19)

where $\alpha \in R$.

(ii) *dagger* of L: $L^+ = \langle \{A_i(E + vu)\}, u, v \rangle$. (9.20)

This L^+ is defined only when $uv = 0$.

(iii) *transpose* of L: $L^t = \langle \{A_i^t\}, v^t, u^t \rangle$. (9.21)

LEMMA 9.3. *If we denote the output function of L by f, then the output functions of* αL *and* L^t *are* αf *and* f^R, *respectively. And if* $f(\epsilon) = 0$, *then the output function of* L^+ *is* f^+.

Proof. As to αL, it is obvious, and as to L^t, it has been already proved in Lemma 7.2 of Section 7.3. Here we will give the proof for L^+. Denoting the output function of L^+ by g, we have

$$g(w) = uA_{i_1}(E + vu)\, A_{i_2}(E + vu) \cdots A_{i_l}(E + vu)v \tag{9.22}$$

for $w = \sigma_{i_1}\sigma_{i_2} \cdots \sigma_{i_l} \in \Sigma^*$. Since $uv = 0$ and $(E + vu)v = v$, we obtain

$$g(w) = uA_{i_1}(E + vu)\, A_{i_2}(E + vu) \cdots A_{i_{l-1}}(E + vu)\, A_{i_l}v. \tag{9.23}$$

Expanding the right hand side of (9.23), this term is represented by the summation of (9.18) over all k's and k-tuples $w_1, w_2, \ldots, w_k$ such that their concatenation is equal to w. Thus, by Lemma 9.2, we know that $g = f^+$.

Q.E.D.

This lemma gives the following theorem.

THEOREM 9.2. *If $f \in \mathscr{L}$ then $\alpha f \in \mathscr{L}$ and $f^R \in \mathscr{L}$. Moreover, if $f(\epsilon) = 0$, then $f^+ \in \mathscr{L}$.*

Now, for a subset S of Σ^*, we denote the characteristic function of S by χ_S, that is,

$$\chi_S(w) = \begin{cases} 1 & \text{for} \quad w \in S \\ 0 & \text{for} \quad w \notin S \end{cases} \tag{9.24}$$

THEOREM 9.3. *Let f be a real function over Σ^*. The necessary and sufficient condition that $f \in \mathscr{L}$ is that f can be obtained from $\chi_{\{\epsilon\}}$, $\chi_{\{\sigma_1\}}$,..., $\chi_{\{\sigma_I\}}$ by finite numbers of applications of sum, convolution, dagger and multiplication by constant. In other words, $\mathscr{L}$ is the minimal family of real functions over Σ^* which is closed under the operations of sum, convolution, dagger and multiplication by constant.*

Proof. The necessity is proved by Sugino, Inagaki and Fukumura (1968). The sufficiency can be shown as follows. In fact, there exist l.a.'s of which output functions are $\chi_{\{\epsilon\}}$, $\chi_{\{\sigma_1\}}$,..., $\chi_{\{\sigma_I\}}$. For example, l.a. whose output function is equal to $\chi_{\{\sigma_1\}}$ can be obtained by determining A_i's, u and v as follows:

$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad A_2 = A_3 = \cdots = A_I = 0,$$

$$u = (1, 0), \qquad v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus, $\chi_{\{\epsilon\}}$, $\chi_{\{\sigma_1\}}$,..., $\chi_{\{\sigma_I\}} \in \mathscr{L}$.

This fact and Theorems 9.1 and 9.2 establish the sufficiency. Q.E.D.

THEOREM 9.4. *$\chi_S \in \mathscr{L}$ if and only if S is regular.*

Proof. Since any finite automaton can be considered as l.a. by representing it by use of matrices, the characteristic function χ_S of the regular set S is contained in $\mathscr{L}$.

Conversely, assume that $\chi_S \in \mathscr{L}$. Then, $f = \chi_S - \frac{1}{2}\chi_{\Sigma^*} \in \mathscr{L}$. It can be easily known that this function f is to be such that

$$f(w) = \begin{cases} \frac{1}{2} & \text{for} \quad w \in S \\ -\frac{1}{2} & \text{for} \quad w \notin S. \end{cases}$$

Thus, f is bounded and the cutpoint 0 is isolated. Therefore, by Theorem 8.1, S is regular. Q.E.D.

The above discussions give a solution of the problem to specify the quasi-regular function, that is, to answer the question what real functions over Σ^* can be realized by l.a.'s. The algebraic properties of $\mathscr{L}$ are detailed in Inagaki, Sugino and Fukumura (1970).

Here, we present some examples of the quasi-regular functions.

EXAMPLE 9.1. Examples of the quasi-regular functions;

(1) The function giving the length $l(w)$ of a tape w. This can be realized by

$$A_1 = A_2 = \cdots = A_I = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad u = (1, 0), \qquad v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

(2) The function $f_S(w)$ giving the number of elements which are contained in a given regular set S and which also are subwords of tape w.

If the set S is regular, then $\chi_S \in \mathscr{L}$. Therefore, $f = \chi_{\Sigma^*} \circ \chi_S \circ \chi_{\Sigma^*}$ is an element of $\mathscr{L}$ and then the desired function. Here we should notice that the function $l(w)$ of (1) is a special case of $f_S(w)$ and $l(w)$ may be represented by $l = \chi_{\Sigma^*} \circ \chi_\Sigma \circ \chi_{\Sigma^*}$.

Next let us refer to output functions of modified p.a.'s and denote the set of all the output functions of modified p.a.'s by $\mathscr{P}'$.

THEOREM 9.5. *The set $\mathscr{P}'$ is closed under the operations of sum* $(+)$, *product* $(\cdot)$, *multiplication by constant and reverse.*

Proof. As for product and multiplication by constant, the theorem is obvious. As for sum, by using $\frac{1}{2}u_1 \oplus \frac{1}{2}u_2$ and $2v_1 \oplus 2v_2$ instead of $u_1 \oplus u_2$ and $v_1 \oplus v_2$ in (9.5), respectively, we know that there exists a modified p.a. of which output function is the sum of two modified p.a.'s functions. Finally, as for reverse, refer to Nasu and Honda (1968). Q.E.D.

Remark 9.1. We should notice that the set $\mathscr{P}'$ is not closed under operations of $\circ$ and $^+$. In fact, χ_Σ and χ_{Σ^*} are in $\mathscr{P}'$ but the function l cited in Example 9.1 (2), which is represented by $l = \chi_{\Sigma^*} \circ \chi_\Sigma \circ \chi_{\Sigma^*}$ is not in $\mathscr{P}'$ because it is not bounded. As for the operation $^+$, the characteristic function of $\Sigma^* - \epsilon$ is in $\mathscr{P}'$ but the function obtained by applying the operation $^+$ to it is not bounded.

10. An Application of Quasi-Regular Functions and Sets of Tapes Accepted by l.a.'s

In this section, we discuss some properties of the family of languages accepted by l.a.'s by using the quasi-regular functions.

Theorem 10.1. *Let S be an arbitrary regular set and L be an arbitrary* l.a. *Then there exist* l.a.'s *L' and L'' such that $S \cup T(L) = T(L')$ and $S \cap T(L) = T(L'')$, i.e., the family of languages accepted by* l.a.'s *is closed under the operations of union and intersection with the regular set.*

Proof. Although this theorem is obvious from Theorem 44 of Paz (1966), we prove here this theorem by using quasi-regular functions.

Denoting the output function of L by f, we have

$$S \cup T(L) = \{w \mid g_1(w) > 0\},$$

where g_1 is represented by $g_1 = \chi_S + (\chi_{\Sigma^*} - \chi_S) \cdot f$ and this is in $\mathscr{L}$. Next, defining g_2 by $g_2 = f \cdot \chi_S$ we know $S \cap T(L) = \{w \mid g_2(w) > 0\}$. Q.E.D.

Theorem 10.2. *Let f and g be real functions over Σ^* which assume only nonnegative values, and define*

$$S_1 = \{w \in \Sigma^* \mid f(w) > 0\} \qquad \textit{and} \qquad S_2 = \{w \in \Sigma^* \mid g(w) > 0\}.$$

Then $f + g$, $f \cdot g$, $f \circ g$, f^+ and f^R are also nonnegative-valued real functions and the sets of tapes for which these functions assume positive values are $S_1 \cup S_2$, $S_1 \cap S_2$, $S_1 \cdot S_2$, $S_1 \cdot S_1^*$ and S_1^R respectively, where f^+ can be taken into consideration only when $f(\epsilon) = 0$.

The proof is obvious and omitted.

Corollary 10.1. *The family of tape sets defined by $S = \{w \mid f(w) \neq 0\}$ for some $f \in \mathscr{L}$ is closed under the operations $\cup$, $\cap$, $\cdot$, $*$ and reverse ^{R}.*

Proof. If $f \in \mathscr{L}$, then $f \cdot f \in \mathscr{L}$. Therefore, $S = \{w \mid f(w) \neq 0\} = \{w \mid ff(w) > 0\}$. Thus S is the set of tapes for which the nonnegative real function ff assumes positive values. Therefore, Theorem 10.2 asserts that this theorem holds for $\cap$, $\cup$, $\cdot$ and ^{R}.

Here we prove it for the operation $*$. Put $g = ff \in \mathscr{L}$ and $g' = g - g(\epsilon) \cdot \chi_{\{\epsilon\}}$. g' is a nonnegative real function and $g'(\epsilon) = 0$ and hence

$$\{w \mid g'(w) > 0\} = \{w \mid g(w) > 0\} - \{\epsilon\} = S - \{\epsilon\}.$$

Thus, $h = g'^{+} + \chi_{\{\epsilon\}}$ is also a nonnegative real function and $\{w \mid h(w) > 0\} = \{w \mid g'^{+} + \chi_{\{\epsilon\}}(w) > 0\} = S \cdot S^* \cup \{\epsilon\} = S^*$. Q.E.D.

EXAMPLE 10.1. Examples of constructions of l.a.'s accepting the given tape sets:

(1) Put $\Sigma = \{\sigma_1, \sigma_2\}$ and let S be the set of tapes w which contains more σ_1's than σ_2's. Let us construct l.a. which accepts S. In order to do this, it is sufficient to construct l.a. having output function $g = 2^{l_1} - 2^{l_2}$. Such l.a. certainly exists. In fact, l.a. L with the output function g may be obtained as follows. $L = \langle\{A_1, A_2\}, u, v\rangle$, where

$$A_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \qquad u = (1, 1), \qquad v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

(2) The set $\{\sigma_1{}^l\sigma_2{}^l \mid l \geqslant 0\}$ is accepted by l.a. of which output function is given by

$$f = \chi_{\sigma_1{}^*\sigma_2{}^*}[1 - (2^{l_1} - 2^{l_2})^2],$$

where l_1 and l_2 is the functions giving the numbers of σ_1 and σ_2 in the tape, respectively.

(3) The sets $\{\sigma_1{}^l\sigma_2{}^l\sigma_1{}^l \mid l \geqslant 0\}$ and $\{\sigma_1{}^l\sigma_2{}^l\sigma_1^{2l} \mid l \geqslant 0\}$ are also acceptable for some l.a.'s.

11. OPERATIONS UNDER WHICH $\mathscr{L}$ IS NOT CLOSED

DEFINITION 11.1. Let f and g be real functions over Σ^*. We define two functions $f \vee g$ and $f \wedge g$ as follows:

$$f \vee g(w) = \max\{f(w), g(w)\}, \tag{11.1}$$

$$f \wedge g(w) = \min\{f(w), g(w)\}. \tag{11.2}$$

THEOREM 11.1 *$\mathscr{L}$ is not closed under operations $\vee$ and $\wedge$.*

Proof. Putting $\Sigma = \{\sigma_1, \sigma_2\}$, we denote the number of σ_1 contained in $w \in \Sigma^*$ by $l_1(w)$ and the number of σ_2 contained in w by $l_2(w)$. Here we consider the following functions f and g.

$$f = (\tfrac{1}{2})^{l_1} - (\tfrac{1}{2})^{l_2}, \tag{11.3}$$

$$g = 0. \tag{11.4}$$

Since g is the characteristic function of the empty set, which is a regular set, we know from Theorem 9.4 that $g \in \mathscr{L}$.

Next f is the output function of L_p defined as follows:

$$L_p = \langle\{P_1, P_2\}, u, v\rangle, \tag{11.5}$$

where

$$P_1 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix},$$

$$u = (\tfrac{1}{2}, \tfrac{1}{2}, 0), \qquad v = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}.$$

Thus, $f \in \mathscr{L}$.

Here we assume that there exists l.a. L of which output function is $f \vee g$, and that this L has n-dimensional state space and is defined by

$$L = \langle R^n, \{A_1, A_2\}, u, v\rangle. \tag{11.6}$$

The output of L for $w = \sigma_2^{l_2}\sigma_1^{l_1} \in \Sigma^*$ is given by

$$uA_2^{l_2}A_1^{l_1}v = f \vee g(w) = \max\{(\tfrac{1}{2})^{l_1} - (\tfrac{1}{2})^{l_2}, 0\}. \tag{11.7}$$

Now, from this L, we can construct a new l.a. L' with the set of input symbol, $\Sigma' = \{\sigma_1\}$ having only one element, as follows:

$$L' = \langle R^n, \{B_1\}, u', v'\rangle, \tag{11.8}$$

where $B_1 = A_1$, $u' = uA_2{}^n$, $v' = v$. Hence the output of L' for $w' = \sigma_1{}^l$ is given by

$$u'B_1{}^lv' = uA_2{}^nA_1{}^lv = \max\{(\tfrac{1}{2})^l - (\tfrac{1}{2})^n, 0\}. \tag{11.9}$$

Therefore we know that

$$u'B_1{}^lv' \begin{cases} = 0 & \text{for} \quad 0 \leqslant l \leqslant n, \\ > 0 & \text{for} \quad l \geqslant n+1. \end{cases} \tag{11.10}$$

On the other hand, the dimension of the state space of L' is n and from (11.10), the outputs of L' is equal to 0 for all tapes with length n or less. Therefore it follows from Theorem 4.2 that u' and 0 and not distinguishable.

That is, $u'B_1{}^l v' = 0$ for all $l \geqslant 0$. But this contradicts to (11.10). Therefore, $f \vee g \notin \mathscr{L}$.

In the same way we can prove that $f \wedge g \notin \mathscr{L}$. Q.E.D.

Thus we have known that $\mathscr{L}$ is not closed under the operations $\vee$ and $\wedge$. But we can prove that if f and g are in $\mathscr{L}$ and bounded there exists the sequence of elements of $\mathscr{L}$ uniformly converging to $f \vee g$ and $f \wedge g$, respectively. For this end, we prove the following lemma.

LEMMA 11.1. *Putting $0 \leqslant a \leqslant b \leqslant 1$, we define the sequence $\{c_n\}$ by*

$$c_1 = ab \qquad \textit{and} \qquad c_{n+1} = c_n + (a - c_n)(b - c_n).$$

Then the sequence $\{c_n\}$ converges to $\min\{a, b\}$ at the rate not slower than the rate at which the following sequence $\{\theta_n\}$ converges to 0.

$$\theta_1 = \tfrac{1}{4}, \qquad \theta_{n+1} = \theta_n - \theta_n{}^2.$$

The proof is omitted since it is easily done.

From Lemma 11.1, we obtain the interesting corollary.

COROLLARY 11.1. *Let f and g be real functions over Σ^* and assume that f and g satisfy the conditions that $0 \leqslant f(w) \leqslant 1$ and $0 \leqslant g(w) \leqslant 1$ for all $w \in \Sigma^*$. Define the sequence $h_1, h_2, \ldots, h_n, \ldots$ of real functions over Σ^* as follows:*

$$h_1 = fg, \qquad h_{n+1} = h_n + (f - h_n)(g - h_n).$$

Then the sequence $\{h_n\}$ converges uniformly to $f \wedge g$.

THEOREM 11.2. *If $f, g \in \mathscr{L}$ are bounded, then there exist sequences of elements of $\mathscr{L}$, which are bounded, converging uniformly to $f \vee g$ and $f \wedge g$, respectively.*

Proof. Assume that $|f(w)| \leqslant M$ and $|g(w)| \leqslant M$ for all $w \in \Sigma^*$, and *define* f' and g' as follows:

$$f'(w) = \frac{1}{2M} f(w) + \frac{1}{2}, \qquad g'(w) = \frac{1}{2M} g(w) + \frac{1}{2}. \tag{11.11}$$

Then, we may easily verify that $f' \in \mathscr{L}$ and $g' \in \mathscr{L}$, and $0 \leqslant f'(w) \leqslant 1$ and $0 \leqslant g'(w) \leqslant 1$ for all $w \in \Sigma^*$.

Here, we define the sequence $\{h_n'\}$ by

$$h_1' = f'g' \qquad \text{and} \qquad h'_{n+1} = h_n' + (f' - h_n')(g' - h_n'). \quad (11.12)$$

From Corollary 11.1, this sequence $\{h_n'\}$ converges uniformly to $f' \wedge g'$. On the other hand, as $\mathscr{L}$ is closed under operations of sum, product and multiplication by constant, h_n' $(n = 1, 2,...)$ are contained in $\mathscr{L}$.

By defining the sequence $\{h_n\}$ as $h_n = 2M(h_n' - \frac{1}{2})$, we see that $\{h_n\}$ converges uniformly to $f \wedge g$ and h_n $(n = 1, 2,...)$ are the elements of $\mathscr{L}$.

Concerning $f \vee g$, we can also verify the statement of this theorem in the same way. Q.E.D.

COROLLARY 11.2 *The family $\mathscr{P}'$ of the all output functions of the modified* p.a.'*s is not closed under the operations $\vee$ and $\wedge$. But if f, $g \in \mathscr{P}'$, then there exist the sequences of elements of $\mathscr{P}'$ converging to $f \vee g$ and $f \wedge g$, respectively.*

ACKNOWLEDGMENTS

The authors would like to thank Prof. N. Honda of Tohoku University for his helpful discussions and encouragements. They are also grateful to their colleagues and Prof. K. Ikegaya of Nagoya University for their encouragements. They thank the referee for his valuable comments.

RECEIVED: April 9, 1971

REFERENCES

CARLYLE, J. W. (1969), Stochastic finite-state system theory, *in* "System Theory," Zadeh and Polack eds., McGraw-Hill, New York.

HONDA, N. AND NASU, M. (1968), Fuzzy events realized by finite probabilistic automata, *Information and Control* **12**, 284–303.

INAGAKI, Y., SUGINO, K., AND FUKUMURA, T. (1970), Algebraic properties of quasi-regular expression and linear space automata, *The Transactions of the Institute of Electronics and Communication Engineers of Japan*, Vol. 53-C, 5, pp. 309–316 (in Japanese).

KLEE, V. (1959), Some characterization of convex polyhedra, *Acta Math.* **102**, 79–107.

MATUURA, H., INAGAKI, Y., AND FUKUMURA, T. (1968), Generalization of automaton and its analysis, *Papers of Technical Group on Automata of the Institute of Electronics and Communication Engineers of Japan* (January, 1968) A67-45 (in Japanese).

PAZ, A. (1966), Some aspects of probabilistic automata, *Information and Control* **10**, 215–219.

PAZ, A. (1971), "Introduction to Probabilistic Automata," Academic Press, New York.

RABIN, M. O. (1963), Probabilistic automata, *Information and Control* **6**, 230–245.

SALOMAA, A. (1967), On *m*-adic probabilistic automata, *Information and Control* **10**, 215–219.

SUGINO, K., INAGAKI, Y., AND FUKUMURA, T. (1968), An analysis of probabilistic automaton by its state characteristic equation, *The Transactions of the Institute of Electronics and Communication Engineers of Japan*, Vol. 51-C, 1, pp. 29–36 (in Japanese).

TURAKAINEN, P. (1968), On probabilistic automaton and their generalizations, *Ann. Acad. Sci. Fennicae Ser. A*, **429**, 1–53.

VALENTINA, F. A. (1964), "Convex Sets," p. 143, McGraw-Hill, New York.