数理解析研究所講究録
229巻 1975年 63-78

系列パターン認識システムとしての確率オートマトンに関する二、三の問題 ――線形空間オートマトンの安定性と近似可能性――

名大 工学部　稲垣 康善
杉野花津江

1. はじめに　もともとオートマトンをシンボル系列の認識機械（recognizer）あるいは受理機械（acceptor）と考え，オートマトンによって受理されるシンボル系列の集合とその性質を議論の対象にする[1]ことによっていわゆる言語とオートマトンの理論が生れたといってもよいだろう．しかしながら，本研究集会のテーマ「時系列パターン認識システムの研究」のような問題への応用が試みられることはどちらかといえば少く，その様な試みを云う場合に従来の有限オートマトンやプッシュダウンオートマトン等のオートマトンの“かたさ”が問題にされ，“柔い”オートマトンを望む意見も多々見受けられた．これは，一つには相似型と計数型の間隙の問題とも受けとることができよう．

一方では，すでに，オートマトンという言葉が研究論文の中に現われ始めた頃に，この様な問題に対して，J. von Neumann

は，有名な論文"Probabilistic logics and the synthesis of reliable organisms from unreliable components"[2]で，確率的論理と神経回路網の信頼性について述べると共に，その第11章の「計数化と多重化に関する一般的注釈」の中で，神経系の計数型特性と相似型特性に注意を向け，相似型手続の可能性として，Sheffer関数の多線論理で実現可能な定数α，凸結合$\alpha\xi+(1-\alpha)\eta$，ならびに$1-\xi\eta$なる三つの基本演算をもつような代数系について述べている．

この信頼性の問題に動機づけられて，Rabinは，有限オートマトンの状態推移規則を確率化して確率オートマトンを定式化した[3]。また，記憶のある通信路モデルとして，Shannonは，通信路の状態sで送信シンボルがσ_iのとき，通信路の状態がs'に推移し，受信シンボルがσ_oである確率$P\{(\sigma_o, s')|(\sigma_i, s)\}$で規定される確率的なシステムを考えている[4]。これを確率的順序機械として，Carlyleが捉えて[5],[6]以後，非定常なマルコフの鎖の理論[7]と共に，確率的有限状態システムとして確率的順序機械の研究が行われてきている．

これらの研究の中から，さらに，一般化された確率オートマトン（Turakainen[8]），線形空間オートマトン（Inagaki et al.[9]）が生れ，さらにZadehのFuzzy集合の概念を用いて，Fuzzyオートマトン（Santos and Wee[10]，Mizumoto et al.[11]，Wechler[12]）

が提案されるにいたっている。

他方，オートマトン理論としては，自然な成り行きとして，確率文法や重みつき文法，言語上の確率測度が考えられるようになり，それに伴って確率的プッシュダウンオートマトンや確率的木オートマトン（probabilistic tree automaton）等が考えられてきている。

この様に，最初はごく現実的な工学的問題に動機づけられて始まった確率オートマトンの理論は，自然にまたオートマトン理論の枠組の中でその内容を豊かにしてきて，最近ではいくつかの応用が考えられるようになってきた。Paz は，詳細にわたって確率オートマトンの成果を紹介 しているその労作(13)の巻末に，応用分野として，

A.情報理論，B.信頼性，C.学習とパターン認識，D.制御，E.その他として，TSS の解析への応用，Markov Chain の関数の問題への応用，D.P. との関係

の様な項目を掲げ，また Fu と Huang の "Stochastic Grammars and Languages" と題する解説(14)の中には，プログラミング言語，コンパイラー，自然言語のモデル化，パターン認識等への応用の可能性の指摘がみられる。

この様な応用の可能性の例がいくつか試みられてきているが，まだ直接に形式を借りるという段階にとどまるか，結

果的にモデルの形式が一致しているという段階にとどまっているものも少くない。しかし，それが全く無意味ということではなく，そこに具体的な応用分野からの新しい問題が生じ，これがオートマトン理論の中での新しい結果を生みだし，互いに理論と応用が作用し合って発展していくことへの第1歩とみることもできよう。

2. 系列パターン認識系の一つの試み

前節で述べた観点から，単純で直接的ではあるが一つの系列パターン認識系をあげ，そこから確率オートマトンの一つの問題を引き出すことにしよう．

まず，系列パターンは，有限長の記号系列であり，この記号系列を構成する記号の個数は有限であると仮定することは自然であろう．この記号の集合を$\Sigma=\{\sigma_1, \sigma_2, \cdots, \sigma_r\}$とする．$\Sigma$から生成される系列パターン（有限記号系列）の集合をΣ^*とする．Σ^*には空系列εをも含まれているものとする．Σ^*の要素を$w=\sigma_{i_1}\cdots\sigma_{i_r}$で表すことにする．

さて，カテゴリーCが与えられたときの系列パターンwの生起確率$P(w|C)$が与えられるものとする．また，$P(C)$：カテゴリーCの生起確率，$P(C|w)$：系列パターンwが与えられたときのCの条件つき確率を約束する．m個のカテゴ

リー C_1, C_2, …, C_m を考えるとき，よく知られているように，

$$P(C_i|w)=\frac{P(w|C_i)P(C_i)}{\sum_{j=1}^{m}P(w|C_j)P(C_j)} \tag{1}$$

いま，$P(C_i)=$一定$=\frac{1}{m}$と仮定すれば，

$$P(C_i|w)=\frac{P(w|C_i)}{\sum_{j=1}^{m}P(w|C_j)} \tag{2}$$

このとき，$\max_i P(C_i|w)$ を与える C_i に w が属するとの判定を行うことにすれば，これは，$\max_i P(w|C_i)$ を与える C_i に w が属するとの判定に等価である。このように考えると必要な機構は，$P(w|C)$ を計算するようなものである。

さて，$P(w|C)$ は，一つの確率言語

$$p_c=\sum_{w\in\Sigma^*}p_c(w)w \;;\; p_c(w)\geqq 0,\ \sum_{w\in\Sigma^*}p_c(w)=1 \tag{3}$$

を定める．ここに，$p_c(w)=P(w|C)$ である。いま，$w\in\Sigma^*$ に対して，

$$D_w p_c=\sum_{w'\in\Sigma^*}p_c(ww')w' \tag{4}$$

なる系列微分演算 D_w を定義する．このとき，

$$\{D_w p_c \mid w\in\Sigma^*\} \tag{5}$$

が張る空間が有限次元の線形空間であれば，式(3)の確率言語を実現するような線形空間オートマトン(l.a.)が構成できる.[(6),(9),(17)]

例えば，カテゴリーCの確率言語 p が次式で与えられたと

する．

$$p = \sum_{w \in \Sigma^*} p(w)\, w \tag{6}$$

ただし，
$$p(w) = \begin{cases} \left(\frac{1}{3}\right)^{n+1}\left(\frac{1}{2}\right)^{m+1} & w = a^n b^n \ (n \geqq 0,\ m > 0) \\ 0 & \text{その他} \end{cases}$$

これは擬似正規表現[18]で表わすと

$$p = \left(\tfrac{1}{3}a\right)^*\left(\tfrac{1}{6}b\right)\left(\tfrac{1}{2}b\right)^* \tag{7}$$

であり，

$$D_a p = \frac{1}{3}p\,,\quad D_b p = \frac{1}{6}\left(\frac{1}{2}b\right)^*,\quad D_{aa} p = \frac{1}{9}p\,,\ D_{ab} p = \frac{1}{18}\left(\frac{1}{2}b\right)^*$$

$$D_{ba} p = 0\,,\quad D_{bb} p = \frac{1}{12}\left(\frac{1}{2}b\right)^*$$

これから，一次独立な擬似正規表現として，p と $D_b p$ をとれば，

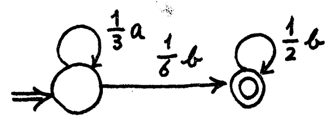
$$p = a(D_a p) + b(D_b p) = \frac{1}{3}a\,p + 1\,b(D_b p)$$

$$D_b p = a(D_{ba} p) + b(D_{bb} p) + \delta(D_b p) = \frac{1}{2}b(D_b p) + \frac{1}{6}\lambda$$

よって，初期ベクトル（1，0），出力ベクトル $\begin{pmatrix} 0 \\ \frac{1}{6} \end{pmatrix}$ で，かつ

$$A(a) = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & 0 \end{pmatrix}\,,\qquad A(b) = \begin{pmatrix} 0 & 1 \\ 0 & \frac{1}{2} \end{pmatrix}$$

であるような l.a. が得られる．これは，



で表わされる l.a. に等価である．

さて，この様にして得られた l.a. を用いて図1のような

認識系を考えることができよう．これは，l.a. $L_1, \cdots, L_m$ によって，入力系列パターンwに対する$P\{w \mid C_k\}$, $k=1,\cdots,m$を計算し，その最大値を与えるカテゴリー名C_{i^*}を出力するような系である．

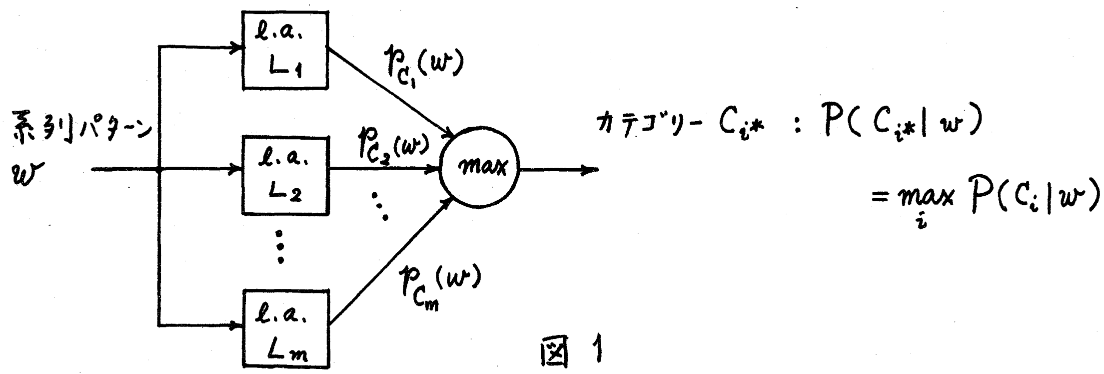


図 1

さて，この様なシステムを考える場合には種々の問題があろうが，ここでは次の様な問題を考えてみる．すなわち，実際にシステムを構成する際にはサンプル系列からl.a.を構成する，すなわち，その推移行列を推定することが必要になろうが，その際に正確にそれを推定することはできず推定誤差を伴う．この誤差の為に，l.a.の出力は$p(w)$の近似値であり，そこに近似誤差δが存在する．従って，この近似誤差δを任意の値におさえるためには，l.a.の推移行列の推定精度はどの程度にすべきか，また，その様な近似が可能である為にはどの様な条件が満たされているべきかという問題がある．しかし，この問題は，サンプル系列からl.a.の推移行列を推定する方法が明らかにされなければ意味をもたない．そこで，

問題をもっと限定して，確率言語 $p(w)$ を規定する l.a. が与えられた時に，この l.a. を δ-近似する問題を考えてみることにしよう。即ち，次節では，一般の l.a.（確率言語を規定する l.a. に限らない）を考えて，l.a. L の出力関数との誤差が δ 以下の出力関数 l.a. L' を決めることができる為の条件を明らかにする。また，この問題は，l.a. L から微小変動した l.a. L' の出力と l.a. L の出力との差を任意の $\delta>0$ で一様におさえるように微小変動 ε をとることができる為の条件を明らかにすることであり，この意味では，l.a. の安定性の問題である。

3. 線形空間オートマトンの安定性と近似可能性

3.1 諸定義

（定義1）次の性質をみたす $\|x\|$，$\|A\|$ をそれぞれベクトル x，行列 A のノルムと定義する。

$$\begin{cases} x \neq 0 \rightarrow \|x\| > 0 \\ \|\alpha x\| = |\alpha| \|x\| \\ \|x+y\| \leqq \|x\| + \|y\| \end{cases} \tag{8}$$

$$\begin{cases} A \neq 0 \rightarrow \|A\| > 0 \\ \|\alpha A\| = |\alpha| \|A\| \\ \|A+B\| \leqq \|A\| + \|B\| \\ \|A \cdot B\| \leqq \|A\| \cdot \|B\| \end{cases} \tag{9}$$

（定義2）$\|Ax\| \leqq \|A\|\|x\|$ をみたす $\|A\|$，$\|x\|$ は consistent であるという。

（定義3）n 次元空間内の原点 O を内点にもつ凸体を K とす

る．Kに関するxのノルム$\|x\|_K$とAの最小上界$lub_K(A)$を

$$\|x\|_K = glb\{\nu \mid \nu \geqq 0,\ x \in \nu K\} \tag{10}$$

$$lub_K(A) = glb\{\alpha \mid \alpha \geqq 0,\ AK \subset \alpha K\},\quad AK = \{Ak \mid k \in K\} \tag{11}$$

のように定義する．

定義3より，$lub_K(A)$はAのノルムであり，$\|x\|_K$とconsistentである．以下では，$lub_K(A)$を$\|A\|_K$と記す．この行列ノルム$\|A\|_K$に関して次の補題がなり立つ．

（補題1）n次元空間内の原点Oを内点にもつ任意の凸体K, K′に対して，

$$m\|A\|_{K'} \leqq \|A\|_K \leqq M\|A\|_{K'} \tag{12}$$

を満たす正の定数m, Mが存在する．

（定義4） 行列Aの固有値をλ_iとする．$\rho(A) = \max_i |\lambda_i|$を行列$A$のスペクトル半径という．

（補題2）[15] $\rho(A) \leqq \|A\|_K$ (13)

（補題3）[15] $^\forall \varepsilon > 0,\ ^\exists K,\ \|A\|_K \leqq \rho(A) + \varepsilon$ (14)

また，この補題2, 3から，直ちに次を得る．

（補題4）$^\exists K,\ \|A\|_K < 1$ならば$\rho(A) < 1$．逆も成り立つ．

（定義5）線形空間オートマトン(l.a.) Lは

$$L = \langle V, \Sigma, \{A(\sigma_i) \mid \sigma_i \in \Sigma\}, u, v \rangle \tag{15}$$

で与えられるシステムである．ここに，VはLの内部状態の集合でLの状態空間と呼ばれる線形空間，Σは入力記号

の集合，$A(\sigma_i)$ は入力記号 σ_i によって引きおこされる V から V の中への線形写像，$u \in V$ は初期状態，v は V 上で定義される線形実関数である．

（定義6） l.a. $L = \langle V, \Sigma, \{A(\sigma_i)\}, u, v \rangle$ が与えられたとき，

$$^{\forall}\sigma_i \in \Sigma,\ \| A'(\sigma_i) - A(\sigma_i) \|_K < \varepsilon \qquad (16)$$

であるような $\{A'(\sigma_i) \mid \sigma_i \in \Sigma\}$ をもつ任意の l.a. $L' = \langle V, \Sigma, \{A'(\sigma_i)\}, u, v \rangle$ を考える。このとき，任意の $\delta > 0$ に対して，

$$^{\forall}w \in \Sigma^*,\ \| A'(w) - A(w) \|_K < \delta$$

である $\varepsilon > 0$ が存在すれば，l.a. L は強安定であるという。(16)

この強安定（*strongly stable*）の条件は，また l.a. L が L' によって近似可能であることを意味している．

3.2 l.a. が強安定であるための条件

（定理1） すべての $\sigma_i \in \Sigma$ に対して $\|A(\sigma_i)\|_K < 1$ である凸体 K が存在すれば，l.a. L は強安定である．

（定理2） l.a. L が強安定ならば，すべての $\sigma_i \in \Sigma$ に対して $\rho(A(\sigma_i)) < 1$ ．

（定理3） l.a. L が強安定ならば $\lim_{l(w)\to\infty} \|A(w)\|_K = 0$. 逆も成り立つ．

（証明） 強安定であることから，任意の凸体 K に対して，

$$\| A'(w) - A(w) \|_K < \delta' \quad \text{for all} \quad w \in \Sigma^* \qquad (17)$$

が成り立つ．この式（17）は，$A'(\sigma_i) - A(\sigma_i) = E_i$, $\|E_i\|_K < \varepsilon$

を満たす任意の E_i に対して成り立つから，$A'(\sigma_i)=A(\sigma_i)E_i$，$E_i=\mu I$，$0<\mu<1$ に対しても成り立つ．そこで，$w=\sigma_{i_1}\cdots\sigma_{i_k}$ とすれば，

$$A'(w)=A'(\sigma_{i_1})\cdot\cdots\cdot A'(\sigma_{i_k})=A(\sigma_{i_1})\mu I\cdot\cdots\cdot A(\sigma_{i_k})\mu I$$
$$=\mu^k A(w) \tag{18}$$

に注意すれば，

$$\|A'(w)-A(w)\|_K=|\mu^k-1|\,\|A(w)\|_K=(1-\mu^k)\|A(w)\|_K$$

を得る．従って，式(17) より，

$$(1-\mu^k)\|A(w)\|_K<\delta' \tag{19}$$

ここで，$k\geqq 1$ に対して $0<1-\mu^k<1$ であるから，

$$\|A(w)\|_K<\delta'/(1-\mu^k) \tag{20}$$

さて，$\delta'/(1-\mu^k)\leqq\delta$ $(\delta'<\delta)$ とすると

$$\mu^k\leqq(\delta-\delta')/\delta\equiv\delta''\quad(0<\delta''<1)$$
$$k\geqq\log\delta''/\log\mu>0\quad(\because\log\delta''<0,\ \log\mu<0)$$

であることに注意して，N を $N\geqq\log\delta''/\log\mu$ を満たす最小の正整数にとれば，$w\in\Sigma^N\Sigma^*$ なるすべての w に対して，$\|A(w)\|_K<\delta$ がなり立つことが知られる。

逆に，すべての $w\in\Sigma^N\Sigma^*$ と任意の $\delta>0$ に対して $\|A(w)\|_K<\delta$ なる $N>0$ が存在するならば，l.a. L は強安定であることを示す．

$w\in\Sigma^+$ に対して，$l(w)\leq N$ ならば，$\|A'(w)-A(w)\|_K<\delta'$

である $\varepsilon' > 0$ が存在することは明らかである．

$l(w) \geqq N$ のとき，　$w = x_0 x_1 \cdots x_n$，$l(x_0) \leqq N-1$，$l(x_i) = N$ $(i=1,2,\cdots,n)$ とおく．

$$\|A'(w) - A(w)\|_K \leqq \|A'(x_0) - A(x_0)\|_K \|A(x_1)\|_K \cdots \|A(x_n)\|_K$$
$$+ \|A'(x_0)\|_K \|A'(x_1) - A(x_1)\|_K \|A(x_2)\|_K \cdots \|A(x_n)\|_K$$
$$+ \cdots\cdots$$
$$+ \|A'(x_0)\|_K \cdots \|A'(x_{n-1})\|_K \|A'(x_n) - A(x_n)\|_K$$

ここで，仮定より，$^\exists N > 0$，$l(w) \geqq N$ なる w に対して $\|A(w)\|_K < \delta$ であるから，$\|A(x_i)\|_K < \delta$ $(i=1,2,\cdots,n)$．また，$\|A'(x_i) - A(x_i)\|_K < \delta'$ なる $\varepsilon' > 0$ が存在する．$(i=0,\cdots,n)$．さらに，$\|A'(x_0)\|_K < \alpha$ （α は有限），$\|A'(x_i)\| < \delta''$ $(i=1,2,\cdots,n)$ なる $\varepsilon'' > 0$ が存在することも明らか．したがって，$\varepsilon = \min(\varepsilon', \varepsilon'')$ とすれば，$l(x_i) \leqq N$ なる x_i と $\varepsilon > 0$ に対して，

$$\|A'(x_i) - A(x_i)\|_K < \delta'$$
$$\|A'(x_0)\|_K < \delta$$
$$\|A'(x_i)\|_K < \delta''$$

よって，

$$\|A'(w) - A(w)\|_K < \delta'\delta^n + \alpha\delta'\delta^{n-1} + \alpha\delta''\delta'\delta^{n-2} + \cdots + \alpha\delta''^{n-1}\delta'$$
$$< \delta_M^{n+1} + n\alpha\delta_M^n \quad (\delta_M = \mathrm{Max}(\delta, \delta', \delta'')).$$

ここで，$\delta_M < 1$ ととることができるから，

$$\delta_M^{n+1} + n\alpha\delta_M^n \longrightarrow 0. \qquad \text{（証明終り）}$$

（定理4） $\lim_{l(w)\to\infty} \|A(w)\|_K = 0$ ならば $\lim_{l(w)\to\infty} \rho(A(w)) = 0$．また，逆も成り立つ．

（定理5） $\lim_{l(w)\to\infty} \|A(w)\|_K = 0$ ならば，$^\forall w \in \Sigma^+, \lim_{m\to\infty} \|A^m(w)\|_K = 0$.

（定理6） $^\forall w \in \Sigma^+, \lim_{m\to\infty} \|A^m(w)\|_K = 0$ ならば，$^\forall w \in \Sigma^+, \rho(A(w)) < 1$．また逆も成り立つ．

以上の結果をまとめて図2に示す．

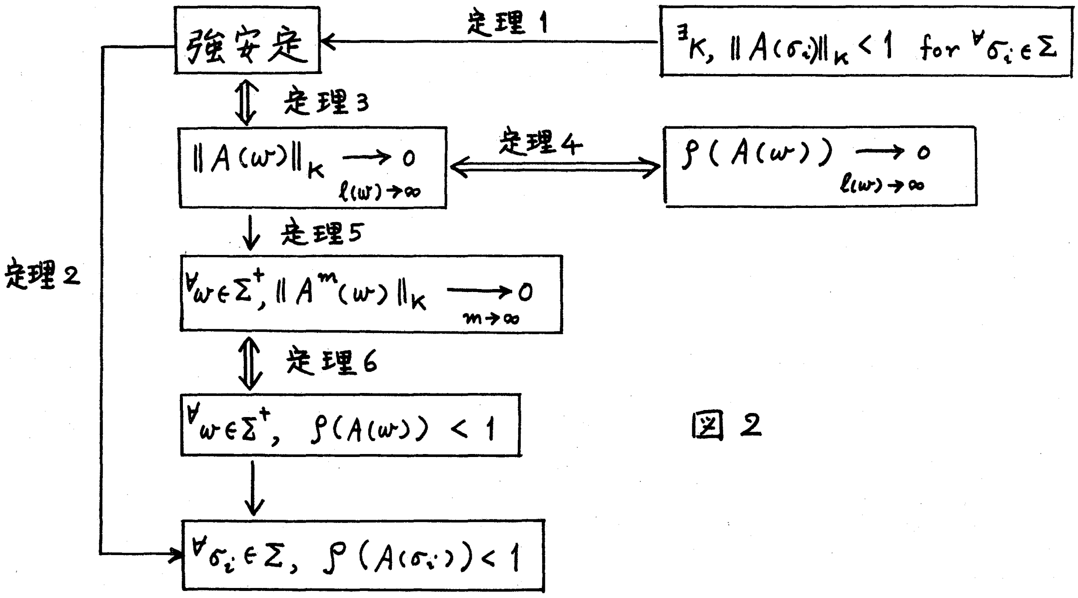


図2

3.3 強安定な l.a. の例

3.2節で示した条件から，次の様な性質をもつ l.a. は強安定であることが知られる．

(i) $A(\sigma_k) = (a_{ij}^{(k)}), i,j = 1, \cdots, n$ として，$k = 1, \cdots, r$ に対して，

$$\sum_{j=1}^{n} |a_{ij}^{(k)}| < 1 \quad \text{for } i = 1, \cdots, n.$$

(ii) $A(\sigma_k), k = 1, \cdots, r$ が，対角要素の絶対値が1より小さい同じ型の三角行列．

さらに,

(iii) $A(\sigma_k)$, $k=1,\cdots,r$ が交換可能ならば,

$$l.a.\text{が強安定である} \Leftrightarrow \rho(A(\sigma_k))<1 \quad \text{for } k=1,\cdots,r$$

(iv) $A(\sigma_k)$, $k=1,\cdots,r$ が同じ凸体Kに属するクラスMの行列であるならば, 即ち, $^{\exists}K$, $\rho(A(\sigma_k))=\|A(\sigma_k)\|_K$ for $k=1,\cdots,r$ ならば,(15)

$$l.a.\text{が強安定である} \Leftrightarrow \rho(A(\sigma_k))<1 \quad \text{for } k=1,\cdots,r$$

が成立することが知られる.

4. むすび

系列パターン認識系の問題に直接的な l.a. の応用を考えてみて, そこから l.a. の安定性と近似可能性の問題を取り上げ, 種々の性質を明らかにした. 主な結果は図2にまとめられている.

文献

(1) Rabin, M.O. and Scott, D. (1959), Finite automata and their decision problem, in " Sequential Machines " (E.F.Moore ed.), Addison-Wesley, Reading, Massachusetts (1964).

(2) von Neumann, J. (1956), Probabilistic logics and the synthesis of reliable organisms from unreliable components, Ann. Math. Studies, 34, 43-98.

(3) Rabin, M.O. (1963), Probabilistic automata, Inf. and Cont'l, 6, 3, p.230.

(4) Shannon, C.E. and Weaver, W. (1948), " The Mathematical Theory of Communication ", Univ. of Illinois Press, Urbana, Illinois.

(5) Carlyle, J.W. (1963), On the external probability structure of finite state channels, Inf. and Cont'l, 7, 167-175.

(6) Carlyle, J.W. (1969),Stochastic finite-state system theory, in " System Theory " (L. Zadeh and E. Polak eds.), p.387, McGraw-Hill.

(7) Hajnal, J. (1958), Weak ergodicity in nonhomogeneous Markov chains, Proc. Cambridge Philos. Soc., 54, 233-246.

(8) Turakainen, P. (1968), On probabilistic automata and their generalization, Ann. Acad. Sci. Fennicae, Ser.A.I., no.429.

(9) Inagaki, Y., Fukumura, T. and Matuura, H. (1972), Some aspects of linear space automata, Inf. and Cont'l, 20, 439-479.

(10) Santos, E.S. and Wee, W.G. (1968), General formulation of sequential machines, Inf. and Cont'l, 12, 5-10.

(11) Mizumoto, M., Toyoda, J. and Tanaka, K. (1969), Some considerations on fuzzy automata, J. Comput. System Sci., 3, 111-124.

(12) Wechler, W. and Dimitrov, V.(1974), R-fuzzy automata, Information Processing 74, 657-660.

(13) Paz, A. (1971), " Introduction to Probabilistic Automata ", Academic Press, New York and London.

(14) Fu, K.S. and Huang, T., Stochastic grammars and languages, International J. of Comput. and Information Sci., 1, 135-170.

(15) Householder, A.S. (1964), " The Theory of Matrices in Numerical Analysis ", Blaisdell Pub. Com.

(16) 杉野，稲垣，福村（1974），線形空間オートマトンの安定性について，
昭和49年度電子通信学会全国大会 1523.

(17) Paz, A. and Robinovich, M. (1974), Linear automata approximation problem, IEEE Trans. on Computers, C-23, 249-255.

(18) 稲垣，杉野，福村（1970），擬似正規表現の代数的性質と線形空間オートマトン，電子通信学会論文誌，53-C，309－316.