

3 自己組織化方式による文字認識

(東北大学 電気通信研究所) 野口正一 大泉充郎

[概 略]

ランダム回路網の自己組織化方式を用いたパターン認識の問題は閾素子等による Boolean Function 合成の問題とも多くの共通点を持ち、種々の立場から数多くの研究が続けられ見るべき成果も上がっているが⁽¹⁾ランダム回路網を認識回路網の立場から考えた場合その基本的設計の問題は殆んど未解決のままである。本論文ではこの観点から多層認識回路網を構成するに必要となる各素子問題の連結法、各層の細胞の数、パターンの種類と細胞の数による認識能力の推定等を理論的に考察し、実際この方式を数字パターン認識装置として応用した場合既に求めた理論的考察がどのように合致するか、又各層の結合構成が実際の認識系にどのような影響するかを計算機によるシミュレーションを通して考察してみた。

[はじめに]

最近著しい情報処理の発展に伴い、パターン認識の問題は各方面から取り上げられ多くの見るべき成果をあげてきているが必らずしも十分満足すべきものではない。

パターン認識の問題は従来の電子計算機の方法論とは大分異なりそれ自体認識のためのアルゴリズムが明確とされていない場合が多く、このため系の構成が不完全なものとなる。従来の研究は主として特徴抽出により刺激パターンを符号化し、この情報をもとにある場合には統計的手法から、ある場合には決定論的立場から人間が学習過程を通してアルゴリズムを構成させてきた。

しかしこの方法論では主役は常に人間であり、このため入力情報の量が膨大化、複雑化すれば十分な成果を望めないのは当然である。しかし人間を含めた生物系はそれ自体の中では極めて易々と認識を行つてきている。

即ち入力刺激の膨大な情報を感覚器を通して吸収し、この情報を逐次学習という操作から多くの時間をかけてそれ自体の中に認識に必要なためのアルゴリズムを形成させ、極めて高度の認識を達成させている。

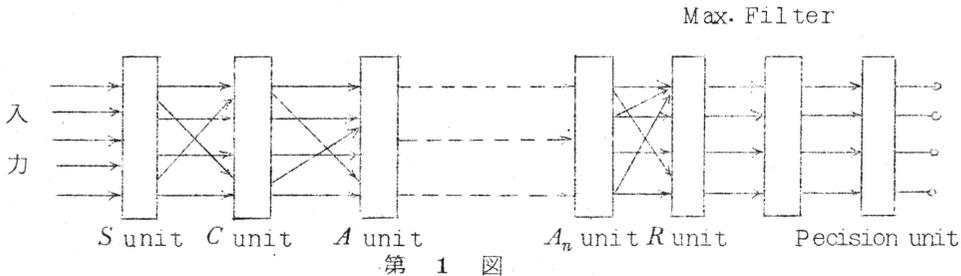
この立場から生物系がもつ認識の手段を工学的な方法論に取り入れれば秀れた認識装置が構成されることは十分に予想される。この立場にたつて研究を行はれてきているのが自己組織化によるパターン認識の研究である。即ち膨大な情報の中から逐次学習を通して認識

に必要なアルゴリズムを形成させ、これを用いて高度の情報処理を行はせんとするものである。

[本 論]

1 認識回路網の一般的構成

考察となる回路網の原理的な構成は第1図にみられるものである。



Sunit 部は入力刺激に対する系の内部えの変換部分で生物系の Sensory の部分に対応する。一般に Sunit の出力はアナログ量でもよく2値の必要はない。Cunit 部は入力のパターンに対し基本的な特徴抽出を行う部分である。例えば図形パターンが対象であれば図形がもつ基本要素即ち線分、斜線、円弧等を検出する回路である。この部分を更に高度化させれば通常の光学系による認識装置の光学 filter に相当する部分となる。

又入力パターンがある論理関数として与えられればある標準展開の各基本項つまり x_i , $x_i x_j, \dots, x_i x_j \dots x_t, \dots$ と考えても良い。Cunit 以後の各 Aunit 部分は Cunit の各要素に対して更に高次の操作を加える層であり、さらに応答素子間の関係から自己組織化を与える系の中心部分である。即ち論理関数の場合なら系が自分自身で構成させたアルゴリズムの中から逐次 $\sum x_i, \sum x_i x_j$, 等を構成させ、図形パターンであれば基本要素から逐次高次の要素を構成させる部分である。一般の回路網はこゝで示したように多くの層から構成されているとは限らず最少単位としては1層の Aunit から構成される。Runit は最終的に入力パターンに対して応答を与える回路であり、入力パターンをどのように対応させるかで多くの変化がある。こゝでは Runit の数を入力パターンのグループに一つずつ対応させ、この出力を逐次入力が増えられた場合比較して最大値検出法によつて決定を行っている。この認識回路網の最小構成単位は Sunit, Aunit, Runit の三層より成るもので屢々実際にも利用され、本論文もこの model を主にして取り扱う。

2 各連合素子間に於ける細胞の数の関係について

既に述べた如く認識回路網では幾つかの連合素子を通して並列的な情報処理が行はれる。今情報の流れを左から右への一方向のものとしこの状態で情報が損はれず伝送されるための必要条件を求めてみる。(勿論パターン自体の情報を整理総合する系では又異なつた議論に

なる。) 今ある \$A\$ 層の細胞の数を \$N\$, そのとれる状態の数を \$\alpha\$, 又 \$(A+1)\$ 層の細胞の数を \$M\$, そのとれる状態数を \$\beta\$ とする. またここで \$N\$ と \$M\$ は十分大きいものとする. 今 \$A\$ 層状態ベクトル \$A = \begin{pmatrix} A_1 \\ \vdots \\ A_N \end{pmatrix}\$ に対し \$(A+1)\$ 層の状態ベクトルを \$B = \begin{pmatrix} B_1 \\ \vdots \\ B_M \end{pmatrix}\$ とすると一般的に次式が成立する.

$$B_i = f_i(A_1, \dots, A_N) \quad (i=1 \sim M) \quad \dots\dots\dots (1)$$

我々の求めることは各状態ベクトル \$A\$ に対し状態ベクトル \$B\$ が unique に与えられる条件を求めることである. この計算の一般的解法は困難であり(1)式が一次不等式で示される場合でも満足すべき解法は見出されていない. そのためここでは情報理論的な統計モデルからこの関係を求めてみる. まず各 \$A_i\$ の生起の確率 \$p(A_i)\$ と(1)式より各 \$p(B_i)\$ の確率は求められるものとし, 各 \$A_i, B_i\$ は夫々独立であると仮定する. 明らかに \$A\$ 層で一つのパターン当りの情報量は \$-N \sum_{i=1}^{\alpha} p(A_i) \log p(A_i)\$ であり, この \$A\$ 層の情報により \$(A+1)\$ 層で発生する情報量は \$-M \sum_{i=1}^{\beta} p(B_i) \log p(B_i)\$ である. この関係から情報が損われずに伝送されるためには少くとも次の関係式が成立しなければならない.

$$-N \sum_{i=1}^{\alpha} p(A_i) \log p(A_i) \leq -M \sum_{i=1}^{\beta} p(B_i) \log p(B_i) \quad \dots\dots\dots (2)$$

勿論この関係は \$N, M\$ が十分大きいことを前提しており有限の \$N, M\$ については更に補正が必要であるがここでは省略する. この関係を閾値素子による簡単な伝送系に応用してみる. (1)式に相当する関係式は次の如く表現される.

$$B_i = \phi \left\{ \sum_{j=1}^N C_{ij} A_j + T \right\} = \phi(x_i) \quad \dots\dots\dots (3)$$

ここで $\phi(x) = 1 \quad x \geq 0$
 $0 \quad x < 0 \quad T$ はある閾値

$$C_{ij} = 1, 0, -1$$

今 \$p(A_j=1) = p_a\$ とし

\$p_r(C_{ij}=1) = p_+, p_r(C_{ij}=-1) = p_-\$ とすると \$x_i\$ は次の平均値 \$m\$ と分散 \$\sigma\$ をもつた正規分布である.

$$m = N p_a (p_+ - p_-) + T \quad \dots\dots\dots (4)$$

$$\sigma = N p_a \{ (p_+ - p_-) - p_a (p_+ - p_-)^2 \}$$

(4)式の関係より簡単に \$p(B_i=1)\$ が次の如く求まる.

$$p(B_i=1) = \frac{1}{2\pi\sigma} \int_{-m}^{\infty} e^{-\frac{x^2}{2\pi\sigma}} dx = \theta_i \quad \dots\dots\dots (5)$$

(2)式と(5)式の関係から各層における \$N\$ と \$M\$ との関係式は次の如くである.

発火の場合情報を反転して伝送させるからである。ここではこの2つのmodelの認識能力と共にその能力の差も同時に調べてみる。

(1) 学習効果の与え方

各 A unitの素子 a_μ に対し一回の学習量 ΔV_μ として次の如く規定する。与えるパターンを $S_{j,q}$ とすると

$$\Delta V_\mu = a_{\mu,j,q} R_{\mu j} \dots\dots\dots (1)$$

今学習量 V_μ の初期状態を総て0とし、各パターンを1回づつ教育した場合はつき系の能力を推定する。

今考察するパターン全体のグループの数を M とし、各グループに属するパターンの数を l_j ($j=1\sim M$) としよう。しかるとき各パターンを一回づつ学習させた場合 A unitの各 cellの学習量は次の如く与えられる。

$$V_\mu = \sum_{j=1}^M \sum_{q=1}^{l_j} a_{\mu,j,q} R_{\mu j} \dots\dots\dots (2)$$

(2) 判定法

判定法は各 R unitの構成上の問題と大きい関係をもち、その議論も一様ではない。ここでは入力パターンのグループに夫々の R unitの cellの r_i を対応させる。パターンの認識法は今ある入力刺激 $S_{i,p}$ に対し各 R unitの r_i の出力を調べ、この中の最大の r_i を検出してこれに対応するグループを今入ってきたパターンと判定させる。即ち最大値検出法である。

今入力刺激 $S_{i,p}$ を学習修了後に加えると各 r_j の端子に生ずる出力の大きい $R_{i,p,k}$ は次の如く示される。

$$\begin{aligned} R_{i,p,k} &= \sum_{\mu=1}^N \sum_{j=1}^M \sum_{q=1}^{l_j} a_{\mu,i,p} a_{\mu,j,q} \cdot R_{\mu j} \cdot R_{\mu k} \\ &= \sum_{\mu=1}^N K_{i,p,k}^{(\mu)} \dots\dots\dots (3) \end{aligned}$$

ここで N は A unitの cellの数である。

結局アンサンブル平均としての回路網がもつ能力は次の確率 p_c を計算することにより求められる。

$$p_c = p_r \{ (R_{i,p,i} - R_{i,p,k}) > 0 \} \text{ for all } i \text{ and } k (i \neq k) \dots\dots\dots (4)$$

以下次の量を計算する。

$$\begin{aligned} &E(K_{i,p,i} - K_{i,p,k}), \quad V(K_{i,p,i} - K_{i,p,k}) \\ &C\{ (K_{i,p,i} - K_{i,p,k}), (K_{i,p,i} - K_{i,p,l}) \} (k \neq l) \end{aligned}$$

計算の簡単化のため各確率変数 $a_{\mu i,p}$ 等は各 μ について独立で同じ性質を有するものと仮定する。以下 μ についての添字を省略する。

$$\begin{aligned}
 E(K_{i,p,i} - K_{i,p,k}) &= E \left\{ \sum_j^M \sum_q^{\ell_j} (a_{ip} a_{jq} R_{\mu j} R_{\mu i} - a_{ip} a_{jq} R_{\mu j} R_{\mu k}) \right. \\
 &= \sum_q^{\ell_j} E \{ a_{ip} a_{iq} (R_{\mu i})^2 - a_{ip} a_{iq} R_{\mu i} R_{\mu k} \} + \sum_{j \neq i}^M \sum_q^{\ell_j} E (a_{ip} a_{jq} R_{\mu j} R_{\mu i} \\
 &\quad - a_{ip} a_{jq} R_{\mu j} R_{\mu k}) \dots\dots\dots (5)
 \end{aligned}$$

今

$$\frac{(C+C')}{M} = \Delta_1 \dots\dots\dots (6)$$

$$\frac{C(C-1)+C'(C'-1)-2CC'}{M(M-1)} = \Delta_2 \dots\dots\dots (7)$$

$$\left. \begin{aligned}
 p_r(a_{ip}=1) &= p \\
 p_r(a_{ip} \cdot a_{iq}=1) &= p \cdot q \quad (p \neq q) \\
 p_r(a_{ip} \cdot a_{jq}=1) &= p \cdot \alpha \quad (i \neq j) \\
 p_r(a_{ip} \cdot a_{jq} \cdot a_{lq'}) &= p \cdot \alpha^2 \quad (i \neq j \neq l \neq i) \\
 p_r(a_{ip} \cdot a_{iq} \cdot a_{jq'}) &= p q \alpha \quad (p \neq q, i \neq j)
 \end{aligned} \right\} \dots\dots\dots (8)$$

以上 T model

$$\left. \begin{aligned}
 p_r(a_{ip}=1) &= p \\
 p_r(a_{ip} \cdot a_{iq}=1) &= Q \\
 p_r(a_{ip} \cdot a_{jq}=1) &= 0 \\
 p_r(a_{ip} \cdot a_{iq} \cdot a_{jq}=1) &= 0 \quad (i \neq j)
 \end{aligned} \right\} \dots\dots\dots (9)$$

以上 B model とすると

以下 $\ell_i = \ell (i=1 \sim M)$ とすると

$$\begin{aligned}
 E(K_{i,p,i} - K_{i,p,k}) &= (p + (\ell - 1)pq - \ell p\alpha)(\Delta_1 - \Delta_2) \\
 &= \{ p(1 - \alpha) + (\ell - 1)p(q - \alpha) \} (\Delta_1 - \Delta_2) \dots\dots\dots (10)
 \end{aligned}$$

以上 T型

$$\begin{aligned}
 B; E(K_{i,p,i} - K_{i,p,k}) &= (\Delta_1 - \Delta_2) + (\ell - 1)(\Delta_1 - \Delta_2)Q \\
 &= (\Delta_1 - \Delta_2)(1 + (\ell - 1)Q) \dots\dots\dots (11)
 \end{aligned}$$

次に $E(K_{i,p,k} \cdot K_{i,p,k}) (i \neq k)$ を求める.

$$K_{i,p,k} \cdot K_{i,p,k} = \sum_j^M \sum_{j'}^M \sum_q^{\ell} \sum_{q'}^{\ell} (a_{ip})^2 a_{jq} a_{j'q'} R_{\mu j} R_{\mu j'} (R_{\mu k})^2 \dots\dots\dots (12)$$

(1) $j=j'=k$, (2) $j=j' \neq k$, (3) $j=k, j' \neq k$, (4) $j \neq k, j'=k$, (5) $j \neq k, j' \neq k$,

$j \neq j'$ の夫々の場合に別けて(12)式を計算すると

$$\begin{aligned}
 T; E(K_{i,p,k} \cdot K_{i,p,k}) &= \sum_q^{\ell} E \{ (a_{ip})^2 (a_{kq})^2 (R_{\mu k})^4 \} \\
 &+ \sum_{\substack{q \\ q \neq q'}}^{\ell} \sum_{\substack{q' \\ q' \neq q}}^{\ell} E \{ (a_{ip})^2 (a_{kq}) (a_{kq'}) (R_{\mu k})^4 \} + \sum_{\substack{j=1 \\ j \neq k \\ j \neq i}}^M \sum_{\substack{q \\ q \neq q'}}^{\ell} \sum_{\substack{q' \\ q' \neq q}}^{\ell} E \{ (a_{ip})^2 (a_{jq}) (a_{jq'}) (R_{\mu j})^2 (R_{\mu k})^2 \} \\
 &+ \sum_{\substack{j=1 \\ j \neq k, j \neq i}}^M \sum_{\substack{q \\ q \neq q'}}^{\ell} E \{ (a_{ip})^2 (a_{jq})^2 (R_{\mu j})^2 (R_{\mu k})^2 \} + \sum_{\substack{q \\ q \neq q'}}^{\ell} \sum_{\substack{q' \\ q' \neq q}}^{\ell} E \{ (a_{ip})^2 (a_{iq}) (a_{iq'}) (R_{\mu i})^2 (R_{\mu k})^2 \} \\
 &+ 2 \sum_{\substack{j=1 \\ j \neq k \\ j \neq i}}^M \sum_{\substack{q \\ q \neq q'}}^{\ell} \sum_{\substack{q' \\ q' \neq q}}^{\ell} E \{ (a_{ip})^2 (a_{jq}) (a_{kq'}) (R_{\mu j}) (R_{\mu k})^3 \} + 2 \sum_{\substack{q \\ q \neq q'}}^{\ell} \sum_{\substack{q' \\ q' \neq q}}^{\ell} E \{ (a_{ip})^2 (a_{iq}) (a_{kq'}) (R_{\mu i}) (R_{\mu k})^3 \} \\
 &+ \sum_{\substack{j \\ j \neq k \\ j \neq i}}^{\ell} \sum_{\substack{j' \\ j' \neq k \\ j' \neq i}}^{\ell} \sum_{\substack{q \\ q \neq q'}}^{\ell} E \{ (a_{ip})^2 (a_{jq}) (a_{j'q'}) R_{\mu j} \cdot R_{\mu j'} \cdot (R_{\mu k})^2 \} \\
 &+ 2 \sum_{\substack{j=1 \\ j \neq k, j \neq i}}^M \sum_{\substack{q \\ q \neq q'}}^{\ell} \sum_{\substack{q' \\ q' \neq q}}^{\ell} E \{ (a_{ip})^2 (a_{jq}) (a_{iq'}) R_{\mu j} \cdot R_{\mu i} (R_{\mu k})^2 \} \\
 &= p\alpha \ell \Delta_1 (1 + \alpha(\ell - 1)q) + p \Delta_3 \{ 1 + 3q(\ell - 1) + q^2(\ell - 1)(\ell - 2) \} \\
 &+ 2p\alpha \ell \Delta_2 \{ (M - 2)\alpha \ell + 1 + (\ell - 1)q \} + p\alpha \ell \Delta_4 (M - 2) \{ (M - 3)\alpha \ell + 2 + 2(\ell - 1)q \} \\
 &\simeq p q \alpha \ell^2 \Delta_1 + (M - 1) \ell^2 p \cdot q \cdot \alpha \Delta_3 + 2(M - 1) \ell^2 p \alpha^2 \Delta_2 + (M - 1)(M - 2) \ell^2 p \alpha^2 \Delta_4 \\
 & \dots \dots \dots (13)
 \end{aligned}$$

ここで

$$\Delta_3 \simeq \frac{C(C-1) + C'(C'-1) + 2CC'}{M(M-1)} \dots \dots \dots (14)$$

$$\Delta_4 \simeq \frac{C\{(C-1)(C-2) + C'(C'-1) - 2C'(C-1)\} + C'\{C(C-1) + (C'+1)(C'-2) - 2C(C'-1)\}}{M(M-1)(M-2)} \dots \dots \dots (15)$$

全く同様にして B model について計算すると

$$B; E(K_{i,p,k} \cdot K_{i,p,k}) \simeq \ell^2 Q \Delta_1 + (M - 1) \ell^2 Q \Delta_3 \dots \dots \dots (16)$$

全く同様にして

$E(K_{i,p,i} \cdot K_{i,p,i})$ を計算する.

$$\begin{aligned}
 T; E(K_{i,p,i} \cdot K_{i,p,i}) &\simeq p q^2 \ell^2 \Delta_1 + (M - 1) \ell^2 p q \alpha \Delta_3 \\
 &+ 2(M - 1) p q \cdot \alpha \ell^2 \Delta_2 + (M - 1)(M - 2) p \alpha^2 \ell^2 \Delta_4 \dots \dots \dots (17)
 \end{aligned}$$

$$B; E(K_{i,p,i} \cdot K_{i,p,i}) = Q' \ell^2 \Delta_1 + (M - 1) \ell^2 Q \Delta_3 \dots \dots \dots (18)$$

ここで

$$Q' = E(a_{ip} \cdot a_{iq} \cdot a_{iq'} = 1) \dots \dots \dots (19)$$

$p \neq q \neq q' \neq p$

次に

$E(K_{i,p,k} \cdot K_{i,p,l})$ を求める.

$$\begin{aligned}
 E(K_{i,p,k} \cdot K_{i,p,l}) &= \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{kq})(a_{lq'}) (R_{\mu k} \cdot R_{\mu l})^2 \} \\
 &+ \sum_{j=1}^M \sum_q \sum_{q'} E \{ (a_{ip})^2 \cdot (a_{kq}) \cdot (a_{j'q'}) \cdot (R_{\mu k})^2 \cdot (R_{\mu l} \cdot R_{\mu j'}) \} \\
 &\quad \substack{j' \neq k \\ j' \neq l} \\
 &+ \sum_q \sum_{q'} E \{ (a_{ip})^2 \cdot (a_{lq})(a_{kq'}) \cdot (R_{\mu l} \cdot R_{\mu k})^2 \} \\
 &+ \sum_{j=1}^M \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{lq})(a_{j'q'}) \cdot (R_{\mu l})^2 (R_{\mu k}) \cdot (R_{\mu j'}) \} \\
 &\quad \substack{j' \neq l \\ j' \neq k} \\
 &+ \sum_{j=1}^M \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{jq}) \cdot (a_{lq'}) \cdot (R_{\mu l})^2 \cdot (R_{\mu k}) \cdot (R_{\mu j}) \} \\
 &\quad \substack{j \neq l \\ j \neq k} \\
 &+ \sum_{j=1}^M \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{jq})(a_{kq'}) \cdot (R_{\mu k})^2 \cdot (R_{\mu l} \cdot R_{\mu j}) \} \\
 &\quad \substack{j \neq l \\ j \neq k} \\
 &+ \sum_{j \neq l} \sum_{j' \neq l} \sum_{q \neq k} \sum_{q' \neq k} E \{ (a_{ip})^2 \cdot (a_{jq})(a_{j'q'}) R_{\mu k} \cdot R_{\mu l} \cdot R_{\mu j} \cdot R_{\mu j'} \} \\
 &+ \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{kq})(a_{kq'}) (R_{\mu k})^3 (R_{\mu l}) \} + \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{lq} \cdot a_{lq'}) (R_{\mu k}) \cdot (R_{\mu l})^3 \} \\
 &+ \sum_{j=1}^M \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{iq})(a_{iq'}) (R_{\mu k}) \cdot (R_{\mu l}) \cdot (R_{\mu j})^2 \} \dots \dots \dots (20) \\
 &\quad \substack{j \neq k \\ j \neq l}
 \end{aligned}$$

以下各 T, B model について概略の結果のみを示そう.

今

$$\begin{aligned}
 d_5 \simeq & \frac{C(C-1)(C-2)(C-3) + 6C(C-1)C'(C'-1) + C'(C'-1)(C'-2)(C'-3)}{M(M-1)(M-2)(M-3)} \\
 & - \frac{4C(C-1)(C-2)C' - 4C'C(C-1)(C-2)}{\dots \dots \dots} \dots \dots \dots (21)
 \end{aligned}$$

とおく.

$$\begin{aligned}
 T; E(K_{i,p,k} \cdot K_{i,p,l}) &\simeq p\alpha^2 \ell^2 \Delta_3 + (M-2) \ell^2 p\alpha^2 \Delta_4 \\
 &+ \ell^2 p\alpha^2 \Delta_3 + 2(M-2) \ell^2 p\alpha^2 \Delta_4 + (M-2) \ell^2 p\alpha^2 \Delta_4 \\
 &+ (M-2)(M-3) \ell^2 p\alpha^2 \Delta_5 + pqa\ell^2 \Delta_2 + pqa\ell^2 \Delta_2 + (M-2)pqa\ell^2 \Delta_4
 \end{aligned}$$

$$\begin{aligned} \simeq & 2p\alpha^2\ell^2\Delta_3 + 4(M-2)\ell^2p\alpha^2\Delta_4 + (M-2)(M-3)\ell^2p\alpha^2\Delta_5 + 2pq\alpha\ell^2\Delta_2 \\ & + (M-2)pq\alpha\ell^2\Delta_4 \dots\dots\dots (22) \end{aligned}$$

$$\begin{aligned} B; E(K_{i,p,k} \cdot K_{i,p,\ell}) & \simeq \ell^2Q\Delta_2 + \ell^2Q\Delta_2 + (M-2)Q\ell^2\Delta_4 \\ & = 2\ell^2Q\Delta_2 + (M-2)Q\ell^2\Delta_4 \dots\dots\dots (23) \end{aligned}$$

最後 (C)

$E(K_{i,p,i} \cdot K_{i,p,\ell})$ を計算する. ($i \neq \ell$)

$$\begin{aligned} E(K_{i,p,i} \cdot K_{i,p,\ell}) & = \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{iq}) (a_{lq'}) (R_{\mu i} \cdot R_{\mu \ell})^2 \} \\ & + \sum_{\substack{j \neq i \\ j' \neq \ell}} \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{iq}) (a_{jq'}) \cdot (R_{\mu i})^2 (R_{\mu \ell}) (R_{\mu j'}) \} \\ & + \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{lq}) (a_{iq'}) \cdot (R_{\mu \ell} \cdot R_{\mu i})^2 \} \\ & + 2 \sum_{\substack{j=1 \\ j \neq \ell \\ j \neq i}}^M \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{lq'}) (a_{jq}) (R_{\mu \ell})^2 (R_{\mu i}) (R_{\mu j}) \} \\ & + \sum_{\substack{j=1 \\ j \neq \ell \\ j \neq i}}^M \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{jq}) (a_{iq'}) \cdot (R_{\mu i})^2 (R_{\mu \ell}) (R_{\mu j}) \} \\ & + \sum_{\substack{j \\ j \neq i \\ j \neq \ell \\ i \neq j'}} \sum_{q'} \sum_{q''} E \{ (a_{ip})^2 (a_{jq}) (a_{jq'}) R_{\mu i} \cdot R_{\mu \ell} R_{\mu j} R_{\mu j'} \} \\ & + \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{iq}) (a_{iq'}) (R_{\mu i})^3 (R_{\mu \ell}) \} \\ & + \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{lq}) (a_{lq'}) (R_{\mu \ell})^3 (R_{\mu i}) \} \\ & + \sum_{\substack{j \\ j \neq i \\ j \neq \ell}} \sum_q \sum_{q'} E \{ (a_{ip})^2 (a_{jq}) (a_{jq'}) (R_{\mu i}) (R_{\mu \ell}) (R_{\mu j})^2 \} \dots\dots\dots (24) \end{aligned}$$

以下 T, B model (C) について概略値の結果を示すと

$$\begin{aligned} T; E(K_{i,p,i} \cdot K_{i,p,\ell}) & \simeq pq\alpha\Delta_3\ell^2 + (M-2)\ell^2pq\alpha\Delta_4 \\ & + pq\alpha\Delta_3\ell^2 + 2(M-2)\ell^2pq\alpha\Delta_4 + (M-2)pq\alpha\ell^2\Delta_4 \\ & + (M-2)(M-3)p\alpha^2\ell^2\Delta_5 + pq^2\ell^2\Delta_2 + pq\alpha\ell^2\Delta_2 \\ & + (M-2)pq\alpha\ell^2\Delta_4 = 2pq\alpha\ell^2\Delta_3 + 4(M-2)\ell^2pq\alpha\Delta_4 \\ & + (M-2)(M-3)p\alpha^2\ell^2\Delta_5 + 2pq\alpha\ell^2\Delta_2 + (M-2)pq\alpha\ell^2\Delta_4 \dots\dots\dots (25) \end{aligned}$$

$$\begin{aligned}
 B; E(K_{i,p,i} \cdot K_{i,p,l}) &\simeq l^2 Q_{A_2} + l^2 Q_{A_2} + (M-2) l^2 Q_{A_4} \\
 &= 2 l^2 Q_{A_2} + (M-2) l^2 Q_{A_4} \dots\dots\dots (26)
 \end{aligned}$$

今 M, C, C', r を十分大きいものとし

$$\frac{C}{M} = \frac{C'}{M} = \frac{\delta}{2} \dots\dots\dots (27)$$

とおくと

$$\left. \begin{aligned}
 A_1 = \delta, \quad A_2 = -\frac{2C}{M(M-1)} \simeq -\frac{\delta}{M} \\
 A_3 = \delta^2, \quad A_4 \simeq -\frac{\delta^2}{M} \quad A_5 \simeq 3 \left(\frac{\delta}{M}\right)^2
 \end{aligned} \right\} \dots\dots\dots (28)$$

以上の計算より

$$\left. \begin{aligned}
 E(K_{i,p,i} - K_{i,p,k}) &\simeq \{p(1-\alpha) + (\ell-1)p(q-\alpha)\} \delta && \text{for } T \\
 &\simeq (1 + (\ell-1)Q) \delta && \text{for } B
 \end{aligned} \right\} \dots\dots (29)$$

$$\left. \begin{aligned}
 E(K_{i,p,k} \cdot K_{i,p,k}) &\simeq pq\alpha l^2 \delta_1 + Mpq\alpha l^2 \delta^2 - 2 l^2 p\alpha^2 \delta \\
 - l^2 Mpq\alpha^2 \delta^2 &\simeq Mpa l^2 \delta^2 (q-\alpha) && \text{for } T \text{ model} \\
 &\simeq M l^2 Q \delta^2 && \text{for } B \text{ model}
 \end{aligned} \right\} \dots\dots (30)$$

$$\left. \begin{aligned}
 E(K_{i,p,i} \cdot K_{i,p,i}) &\simeq Mpa l^2 \delta^2 (q-\alpha) && \text{for } T \text{ model} \\
 &\simeq M l^2 Q \delta^2 && \text{for } B \text{ model}
 \end{aligned} \right\} \dots\dots (31)$$

$$\left. \begin{aligned}
 E(K_{i,p,k} \cdot K_{i,p,l}) &\simeq 3p\alpha^2 l^2 \delta^2 - 3pq\alpha l^2 \delta^2 \\
 &= -3pa l^2 \delta^2 (q-\alpha) && \text{for } T \text{ model} \\
 &\simeq -Q l^2 \delta^2 && \text{for } B \text{ model}
 \end{aligned} \right\} \dots\dots (32)$$

$$\left. \begin{aligned}
 E(K_{i,p,i} \cdot K_{i,p,l}) &\simeq -3pq\alpha l^2 \delta^2 + 3p\alpha^2 l^2 \delta^2 \\
 &= -3pa l^2 \delta^2 (q-\alpha) && \text{for } T \text{ model} \\
 &\simeq -\delta^2 Q l^2 && \text{for } B \text{ model}
 \end{aligned} \right\}$$

以上の計算より我々は T model については

$$V(K_{i,p,i} - K_{i,p,k}), \quad C\{(K_{i,p,i} - K_{i,p,k}), (K_{i,p,i} - K_{i,p,l})\}$$

の値を夫々次の如く求めることができる。

$$V(K_{i,p,i} - K_{i,p,k}) \simeq 2Mpa l^2 \delta^2 (q-\alpha) \dots\dots\dots (33)$$

$$C\{(K_{i,p,i} - K_{i,p,k}), (K_{i,p,i} - K_{i,p,l})\} \simeq Mpa l^2 \delta^2 (q-\alpha) \dots\dots\dots (34)$$

但し以上の近似は

$$M\delta \gg 1 \dots\dots\dots (35)$$

と仮定している。

同様に B model については

$$V(K_{i,p,i} - K_{i,p,k}) \cong 2M\ell^2 Q \delta^2 \quad \dots\dots\dots (36)$$

$$C\{(K_{i,p,i} - K_{i,p,k}), (K_{i,p,i} - K_{i,p,\ell})\} \cong M\ell^2 Q \delta^2 \quad \dots\dots\dots (37)$$

次の確率ベクトルを定義しよう.

$$[\mathbf{K}] = [(K_{i,p,i} - K_{i,p,1}), (K_{i,p,i} - K_{i,p,2}) \dots (K_{i,p,i} - K_{i,p,M})] \quad \dots\dots\dots (38)$$

しかるとき $[\mathbf{K}]$ の平均値及び共分散行列の各要素の値は夫々 (29) 式 (33) ~ (37) 式で示される.

さて既に設定した如く各 μ についての独立性から各確率ベクトル $[\mathbf{K}]$ を N 個選んで

$$[\mathbf{x}] = \sum_{\mu=1}^N [\mathbf{K}^{\mu}] \quad \dots\dots\dots (39)$$

を構成させる. しかる時 $[\mathbf{x}]$ は次の平均値, 共分散行列をもつ $(M-1)$ 次元の正規分布である.

$$E[\mathbf{x}] = (a \dots a) \cdot \mathbf{E} \cdot N \quad \dots\dots\dots (40)$$

ここで

$$\left. \begin{aligned} a &= \{p(1-\alpha) + (\ell-1)p(q-\alpha)\} \delta && \text{for } T \\ &= \{1 + (\ell-1)Q\} \delta && \text{for } B \end{aligned} \right\} \quad \dots\dots\dots (41)$$

\mathbf{E} = unit matrix

$$\text{con}[\mathbf{x}]^{-1} = N \begin{bmatrix} 2b & b & \dots & b \\ b & 2b & \dots & b \\ \vdots & \vdots & & \vdots \\ b & b & \dots & 2b \end{bmatrix} \quad \dots\dots\dots (42)$$

ここで

$$\left. \begin{aligned} b &= M\ell^2 p \alpha (q-\alpha) \delta^2, && \text{for } T \text{ model} \\ &= M\ell^2 Q \delta^2 && \text{for } B \text{ model} \end{aligned} \right\} \quad \dots\dots\dots (43)$$

(42) 式の固有値 $\lambda_1, \dots, \lambda_{(M-1)}$ は簡単に計算され次の如くなる.

$$\left. \begin{aligned} \lambda_1 &= \{2b + (M-2)b\} \cdot N \\ \lambda_2 &= \lambda_3 = \dots = \lambda_{M-1} = b \cdot N \end{aligned} \right\} \quad \dots\dots\dots (44)$$

以上の準備から (4) 式の確率 p_c は次の如く求めることができる.

$$p_c \cong \left(\frac{1}{2\pi}\right)^{M-1} \cdot \frac{1}{\{N M b\}^{\frac{1}{2}} \{b \cdot N\}^{\frac{M-2}{2}}} \int_{-Na}^{\infty} e^{-\frac{x^2}{2NMb}} dx \left\{ \int_{-Na}^{\infty} e^{-\frac{x^2}{2Nb}} dx \right\}^{(M-2)} \quad \dots\dots\dots (45)$$

さて系の認識能力を定める測度として次の量 D を定義する.

$$D = \frac{\text{平均値}}{\text{偏差}} \dots\dots\dots (46)$$

今 N が十分大きいとすると (46) 式は次の如くなる。

$$D \approx \frac{Na}{\sqrt{NMb}} = \sqrt{\frac{N}{M}} \cdot \sqrt{\frac{a^2}{b}} \dots\dots\dots (47)$$

この議論から次のことが証明できる。

$$\lceil N \text{ の大いさを } N \geq \sigma^2 \frac{b}{a^2} \dots\dots\dots (48)$$

より大きくとれば十分の精度でパターンを認識することができる。ここで σ^2 は適当な定数

今この考察を T model と B model に適用し両者の識別能力を検討してみる。必要なのは $\sqrt{\frac{a^2}{b}}$ の値である。

$$\left. \begin{aligned} \sqrt{\frac{a^2}{b}} &\approx \frac{p(q-\alpha)}{Mp\alpha(q-\alpha)} = \sqrt{\frac{(q-\alpha)}{M\alpha}} && \text{for } T \text{ model} \\ &\approx \sqrt{\frac{Q}{M}} && \text{for } B \text{ model} \end{aligned} \right\} \dots\dots\dots (49)$$

既に与えた q, Q の定義からも判る如くこれは同一グループに属するパターンの類似性を平均的に与えるものである。これから $q=Q < 1$ とすれば (49) 式より

$$\frac{(q-\alpha)}{\alpha} \geq 1 \dots\dots\dots (50)$$

の関係で T model と B model の差が示される。一般の場合には $\alpha \approx \frac{1}{2}, \frac{1}{2} < q < 1$ の関係で示され今 $q = \frac{3}{4}$ とすると略 B model は T model より半分の素子で同一の能力を与えることが分る。

今迄の考察は総て強制学習法のもとの認識能力であつた。しかし実際上の系の能力は誤差訂正学習法と呼ばれる学習のもとの定められる場合が多い。この誤差学習法による能力決定の議論も進んでいるが^{(2),(3)} ここでは省略する。ただ系の能力は当然のことながら強制学習法の場合に比較して増加し、しかもこの能力は強制学習法による能力から推定され、又強制学習法による系の能力の関係は誤差訂正学習法に於ける学習回数として表現されることを附記しておく。

4 計算機によるシミュレーション^{(4),(5)}

今迄の議論は強制学習法のもとの十分大きい M, N を想定して求めたものある。計算機によるシミュレーションを行なう場合 M, N を十分大きくすることは計算機の能力の点から不可能なので比較的小さい M, N の場合につき、今迄考察した基本的な関係を数字パターンを対象にして調べてみる。シミュレーションの方法は2大別され最初の場合は主として N, C, C' の数が認識に及ぼす影響、 T model と B model の差異等について調べる。後者はさら

に認識装置として構成される場合各文字図形の基本的な topology の選出が系の能力にどのように影響するかを調べる。

シミュレーション (1)

回路網の構成は最も基本的な S, A, R unit からなる三層とする。

各層及び結合の構成

- (1) S unit の cell ; 0, 1 表示の 40 ビット
- (2) A unit の cell ; 100 及び 200
- (3) R unit の数 ; 8 個
- (4) 教育法 ; 誤差訂正学習法
- (5) S unit より A unit への結合法 ; A unit の cell に S unit よりランダムに正結合 2, 負結合 1
- (6) A unit の閾値 ; 1 又は 2
- (7) A unit の cell の model ; T model 又は B model
- (8) A unit より R unit への結合法 ; R unit の cell に A unit より (i) 20 個正結合, (ii) 40 個正結合, 40 個負結合

次に実際の計算機に於ける A unit の割りつけを第 2 図, 使用した各文字を第 3 図に, 又動作の概略の flow chart を第 4 図に示す。尚ここでの雑音は一様乱数表より求めたものである。次に実際のシミュレーションの結果を示す。第 5 図と第 6 図は A unit から R unit への結線法とし常に正結合にのみ限定し, A unit の cell を 100 個と 200 個の場合につき比較したものである。 R unit の cell への入力第 5 図の場合は 20 個, 第 6 図の場合は 40 個である。この実験結果より判る如く A unit の cell の数が倍に増加すれば学習回数は約半減し, 系の能力は予想した如く増加している。第 5 図と第 7 図は結線法にさらに負結合の効果を加えた場合その能力がどのように増加するかを示したものである。今迄の実験は総て T model を対象にしていたが B model との対応を調べるため第 7 図の構成を B model にしその結果を第 8 図に示す。この結果から B model は T model より著しく能力が増加していることが知れる。尚本計算に用いた計算機は東北大学 SENAC-1 である。

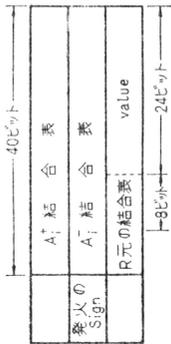


図 2

The allocation of the memory for the each cell connection in A unit.

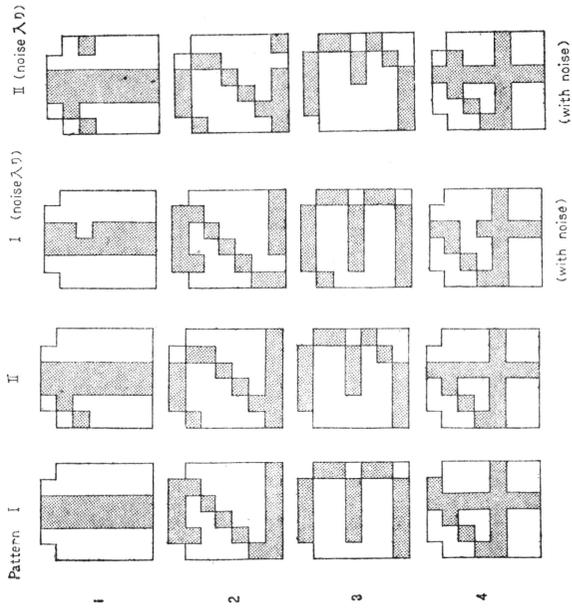


図 3

The patterns used for the computer simulations.

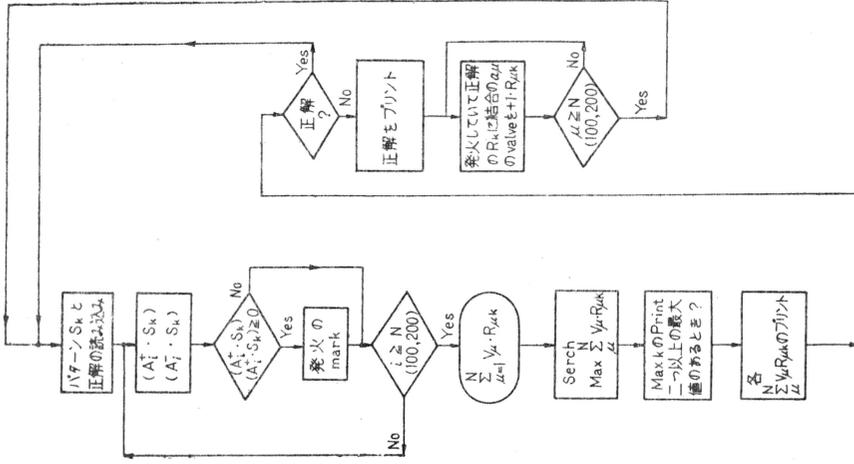
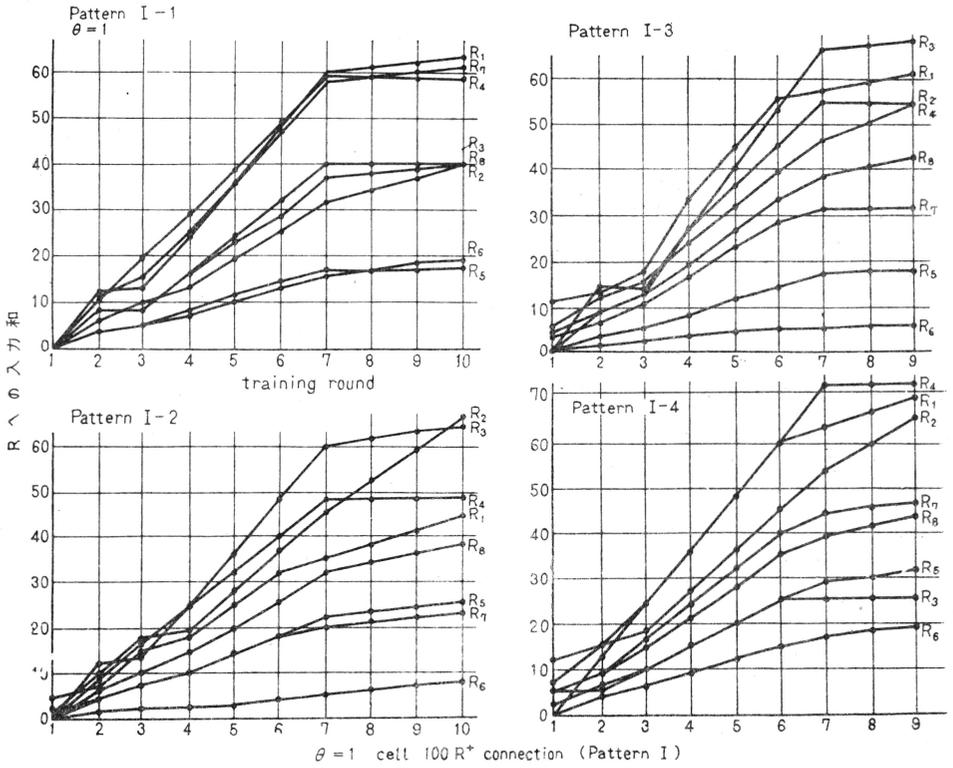
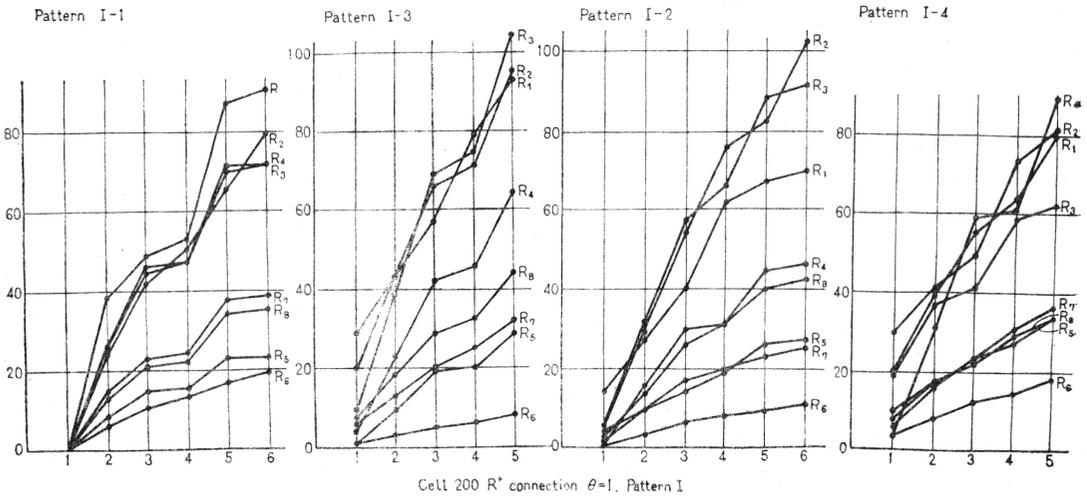


図 4 The flow chart of the computer simulation.



5

The training curve for the pattern (I), when $\theta_1=1$ and A unit cell=100, with all positive connectors.



6

The training curve for the pattern (I) when $\theta_1=1$ and A unit cell=200 with all positive connections.

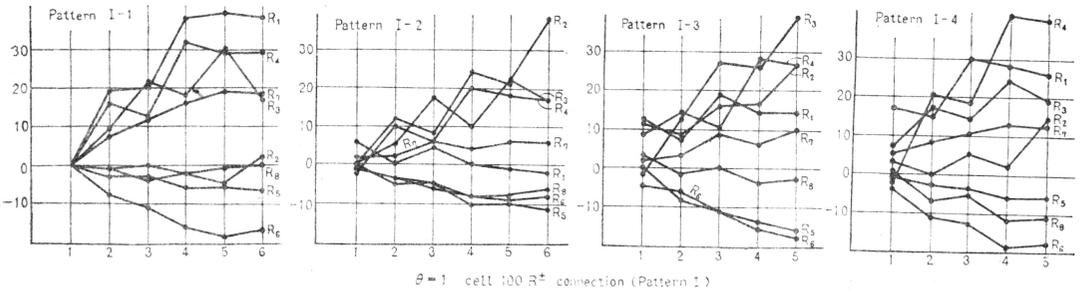


图 7

The training curve for the pattern (I) when $\theta_1=1$ and A unit cell=100 with both positive and negative connections.

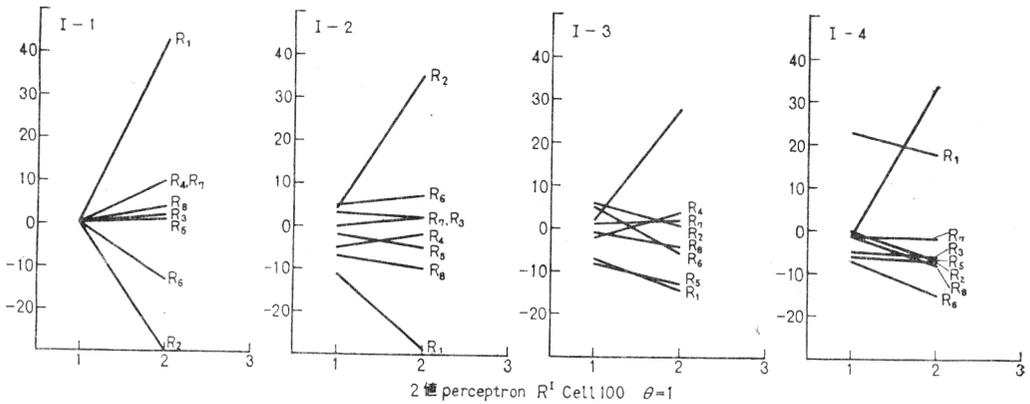


图 8

The training curve for the pattern (I) when $\theta_1=1$ and A unit cell=100 in the binary state network.

シミュレーション (Ⅱ)

このシミュレーションは(I)の規模を更に増加させ積極的に文字認識装置として考察を加えたものである。このため今迄の如く単純な三層の場合と別に新たに文字パターンの基本的なトポロジーを抽出する特徴抽出回路を有する四層認識装置を設定し、この回路が認識の上でどのような働きをもつかを吟味する。又このような基本的特徴の抽出回路は実際の生物に於いても存在することが生理学的な研究からも認められている。系の model の設定は次の如くである。

(I) 三層の場合

(i) S unitの数; 01表示の72 bit (8×9)(ii) A unitの数; 300(iii) R unitの cellの数; 10

(iv) 教育の方法; 誤差訂正法

(v) S unitより A unitへの結合法; A unitの各 cellに S よりランダムに正結合4, 負結合2(vi) A unitの閾値; 2(vii) A unitより R unitへの結合法; R unitの各 cellに A unitよりランダムに正結合30, 負結合30

(viii) 判定法; 最大値検出法

(II) 四層の場合

この block-diagram は第9図にみられる通りである。構成法の大要は次の如くである。

(i) 特徴抽出回路の各要素; 第10図の如くである。

即ち S unit 72 bits中に4 bit で構成される基本特徴素子を考え夫々(a)~(j)迄とし、夫々の数を全体の構成要素の数で配分し全部で300個とする。内訳は(a)=25, (b)=23, (c)=23, (d)=29, (e)=24, (f)=24, (g)=17, (h)=17, (i)=24, (j)=(94)である。

以下の構成は三層の場合と同様である。

使用したパターンを第11図, 12図に示す。

今三層回路網を用いてパターンIのみ学習させた様子を第13図に示す。単一のパターンのみでは四層回路網を用いても略同様な傾向を示し大した変化は認められない。しかしパターンIとパターンIIを同時に教育させると四層回路網は極めて秀れた動作特性を示し学習回数が著しく少なくなる。今この変化の様子を文字4について示したものが第14図, 第15図であり、全体の結果は第1表にまとめて示されている。

以上の考察からも判る如く特徴抽出回路の設定は系の認識能力に多くの影響を与える。

尚第1表で示される S/N の定義は次式により与えられる。

$$S/N = \frac{(M-1)R_i}{\sum_{j=1}^M R_j (j \neq i)} \quad \text{for max } R_i (i=1 \sim M) \quad \dots\dots\dots (1)$$

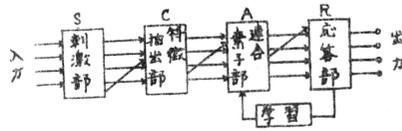
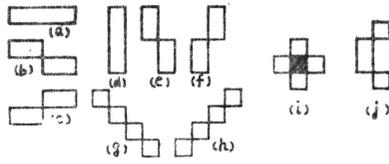


图 9



特征素子 特征素子
图 10 Characteristic elements.

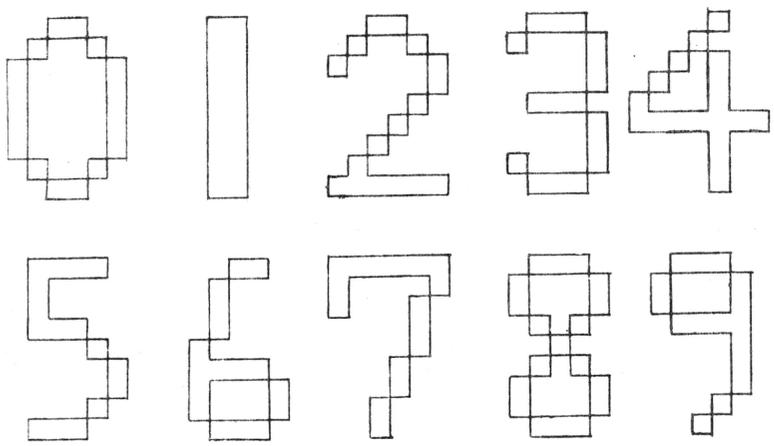


图 11 Pattern I.

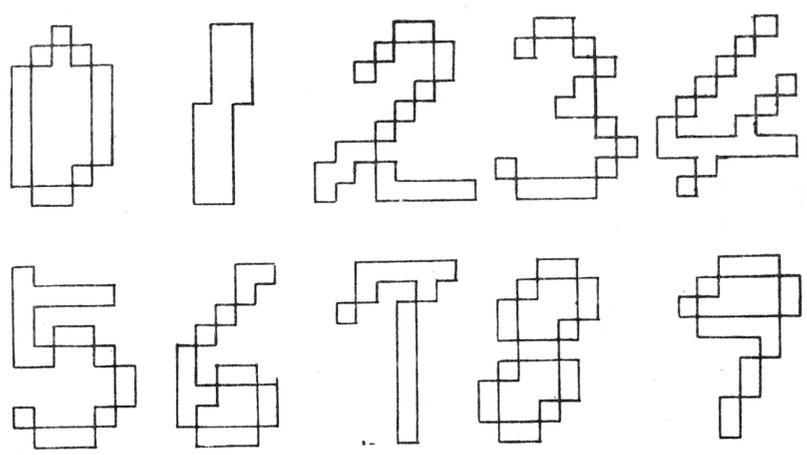


图 12 Pattern II

	S/N					最低学習回数				
	パターン I			I と II 重合		パターン I			I と II 重合	
0	2.2	1.5	4.7	2.4	2.0	2	2	2	14	11
1	2.0	2.4	3.1	2.4	2.5	3	2	2	13	4
2	2.5	1.8	4.2	1.7	1.7	2	2	2	13	5
3	1.3	1.4	4.4	2.2	1.5	2	2	2	10	8
4	2.0	2.6	4.7	1.5	2.2	2	2	2	14	4
5	1.6	1.7	5.3	3.0	1.8	2	2	2	9	4
6	2.0	2.3	4.8	1.8	1.9	2	2	2	11	4
7	1.8	1.5	4.2	1.9	1.8	2	2	2	9	8
8	1.4	1.8	2.7	2.0	2.0	2	2	2	8	7
9	1.5	2.3	12.8	2.2	2.9	2	2	2	11	5
	3層	4層	2値	3層	4層	3層	4層	2値	3層	4層

第1表 比較 Comparison

[結 語]

本論文では自己組織化方式による認識のために必要となる基礎理論と計算機シミュレーションの結果を報告した。まず回路網構成に必要な各層間の素子の数量関係を情報理論的な立場から検討を加えた。これにより将来更に多層化した回路網を構成させる場合各層結合のための必要条件を設定することができる。次に系の認識能力を強制学習法による設定のもとで特にパターンの数が多い場合種々の回路網のパラメータのもとで推定した。又この結論の重要な点として2値的認識回路網は通常的方式のものより認識能力の点で約倍以上秀れていることを立証できた。本論文では触れなかつたが逐次誤差訂正学習法の効果は強制学習法のもつ能力と極めて密接な関係にあり、後者の推定値から一般の回路網の能力が推定できる。これにより我々は比較的単純な認識回路網については設計に必要な方法論を確立したことになる。次に計算機によるシミュレーションについては今迄の理論的な考察を検討する意味と実際にこの装置をパターン認識装置として用いた場合どのような識別能力をもつか、又特徴抽出に相当する操作が系の能力にどのような影響を与えるかについて考察するために行つたものである。この実験より既に求めた理論的検討との一致を確めることができたとともに実際の認識装置としても極めて秀れた装置として利用できることを証明することができた。しかしこの装置の構成はあく迄計算機上のモデルであり、実際の hardware の立場にたてば多くの問題点を有することは事実である。

例えば系の構成に不可決である各 A unit の Value の大いさを十分の精度で広い領域にわたり non-destructive の方法で読み出し、しかも簡単に Value を制御できる素子を安く構成させることはその一つの重要な問題点である。しかしこれらの問題も将来発展する

する電子工学により解決されるであろう。

さてここで取り扱われた自己組織化回路網の研究はまだまだその端初についたに過ぎない。即ち実際の生物系のmodelでは情報処理を行なう層は何重にも多層化され、しかもそれらが forward 又は back ward の Coupling を通して複雑にからみ合っている。今後の問題はこれらの複雑化した回路網の情報処理能力を理論的に求めるとともに更に大型の電子計算機を用いて大規模なシミュレーションを行つて実験的にもこれらの回路網の性質を十分よく説明する必要がある。

終りに色々と御討論、御援助を頂いた本学計算センター高橋理助教授に厚く御礼申し上げます。

参 考 文 献

- (1) F. Rosenblatt ; Principle of Neurodynamics (Spartan Books 1962)
- (2) 野口, 大泉 ; 認識回路網の性質について ; 通信学会専門研究会資料 ; 64' 6
- (3) S. Noguchi et al ; "The Pattern Recognition by organizing the random network" I.C.M.C.I. I-11 1964
- (4) 野口, 他 ; "パーセプトロンに関する一考察", 通信学会情報と制御の研究 (38. 第3号)
- (5) 鈴木, 野口 ; ランダム回路網によるパターン認識 ; 東北大学電通談話会記録 Vol. 33
June 1964

本 PDF ファイルは 1965 年発行の「第 6 回プログラミング—シンポジウム報告集」をスキャンし、項目ごとに整理して、情報処理学会電子図書館「情報学広場」に掲載するものです。

この出版物は情報処理学会への著作権譲渡がなされていませんが、情報処理学会公式 Web サイトの https://www.ipsj.or.jp/topics/Past_reports.html に下記「過去のプログラミング・シンポジウム報告集の利用許諾について」を掲載して、権利者の検索をおこないました。そのうえで同意をいただいたもの、お申し出のなかったものを掲載しています。

過去のプログラミング・シンポジウム報告集の利用許諾について

情報処理学会発行の出版物著作権は平成 12 年から情報処理学会著作権規程に従い、学会に帰属することになっています。

プログラミング・シンポジウムの報告集は、情報処理学会と設立の事情が異なるため、この改訂がシンポジウム内部で徹底しておらず、情報処理学会の他の出版物が情報学広場(=情報処理学会電子図書館)で公開されているにも拘らず、古い報告集には公開されていないものが少からずありました。

プログラミング・シンポジウムは昭和 59 年に情報処理学会の一部門になりましたが、それ以前の報告集も含め、この度学会の他の出版物と同様の扱いにしたいと考えます。過去のすべての報告集の論文について、著作権者(論文を執筆された故人の相続人)を探し出して利用許諾に関する同意を頂くことは困難ですので、一定期間の権利者搜索の努力をしたうえで、著作権者が見つからない場合も論文を情報学広場に掲載させていただきたいと思えます。その後、著作権者が発見され、情報学広場への掲載の継続に同意が得られなかった場合には、当該論文については、掲載を停止致します。

この措置にご意見のある方は、プログラミング・シンポジウムの辻尚史運営委員長(tsuji@math.s.chiba-u.ac.jp)までお申し出ください。

加えて、著作権者について情報をお持ちの方は事務局まで情報をお寄せくださいますようお願い申し上げます。

期間：2020 年 12 月 18 日～2021 年 3 月 19 日

掲載日：2020 年 12 月 18 日

プログラミング・シンポジウム委員会

情報処理学会著作権規程

<https://www.ipsj.or.jp/copyright/ronbun/copyright.html>