

層状ランダム回路のパターン分離性について

鳥岡豊士*・川下 満**

The Ability of Separability in a Layered Random Network

Toyoshi TORIOKA, Mituru KAWASHIMO

Abstract

The cortex of the vertebrate cerebellum has a simple and extremely fine structure. The system is organized with the mossy fibers, granule cells, Golgi cells, basket cells, and Purkinje cells. Functionally, the cortex of cerebellum works as a pattern-recognition data processing system. Moreover, this system (network) has two eminent abilities which makes the overlap of two input patterns reduce and makes the learning enhance. The present paper gives a fundamental and theoretical analysis of the ability of the former function and some results of computer-simulated experiment.

1. 緒 言

近年、生物系の種々の機能をつかさどる神経細胞の動作、構造が生理学的な立場からかなり詳しく明らかにされ、研究の著しい発展をもたらしている。しかし多くの神経細胞が複雑に結合され組織化された大規模な神経回路網の動作機能を明らかにしていくには、生理学的立場からはまだほど遠いようである。このため実際の神経回路網を直接に研究する立場を離れて、神経素子のもつ本質をできるだけそこなわないで、単純な神経モデルを構成し、工学的な立場から説明しようとする立場がある。1943年、W.S. McCulloch and W. Pitts は¹⁾、命題論理の立場から脳機能の説明を試みた。1952年に、A. Rapaport²⁾によって多入力、多出力の神経素子を2項分布の立場で解析し、活動度についての2, 3の性質が明らかにされた。更に、1956年には N. Rochester⁴⁾等は、ニューロン回路網内の組織化についてのコンピュータシミュレーションを示している。1959年には、生物系の視覚系のもつパターン分類の能力をもつパーセプトロンが、F. Rosenblatt³⁾によって示された。1960年には、E. R. Caianiello⁵⁾によって、大脳の機能を表わすという大胆な2つの方程式、記憶方程式と神経方程式が示され、以来、学習識別、連想機能⁶⁾をもつモデルが明らかにされてきた。更に最近、L. I. Rozoner^{7)~9)}、

E.M. Harth^{10),11)} et. al., 甘利等^{12)~14)}は、特定の機能の解明ということにとらわれず、神経回路網のもつ基本的動作特性を追求している。特に甘利は、神経素子という微視的立場と回路網全体としての巨視的立場を統計的な立場から結びつける巨視的方程式を導びき、神経回路網の体系化をめざしている。

ここでは、小脳機能をもつという優れたパターン認識機能の解明の手段となる解析を、上記に述べたようにできるだけ実システムのもつ性質を失なわないように工学的モデルを構成しながら展開する。

2. 小脳の組織と機能^{15),16)}

小脳の組織を神経細胞の立場からみるとグラニュール細胞、プルキンエ細胞、ゴルジ細胞、バスケット細胞、星状細胞からなっている。グラニュール細胞は、苔状線維から入力を受け、平均4本の苔状線維とシナップスする。次に、グラニュール細胞は出力に相当するプルキンエ細胞と300~400個のシナップスをする。そして、このグラニュール細胞とプルキンエ細胞を結ぶ線維を平行線維といい、プルキンエ細胞は約20万本の線維とシナップスする。ゴルジ細胞は、グラニュール細胞に抑制性のシナップスを行ない、入力刺激の強弱に関係なく一定数のグラニュール細胞が発火するように制御している。バスケット細胞、星状細胞は平行線

* 工業短期大学部電気工学科

** 工学部電気工科大学院

維から入力を受け、近傍の60~70個のプルキンエ細胞の閾値を制御している。次に小脳の学習機能を考える。苔状線維を伝わって入ってきた刺激パターンは、特定のグラニュール細胞を発火し、ついでプルキンエ細胞を発火する。一方、大脳からはオリブ細胞を通じて、プルキンエ細胞が出力すべき出力パターンの教師となる情報を与える。この結果、プルキンエ細胞は要求される出力を出すように学習される。この学習が反復されるに従い、最終的にはオリブ細胞からの情報なしに、プルキンエ細胞は正しい出力パターンを出すようになり、入力刺激パターンを正しく認識できる

ようになる。ここで、苔状線維が直接プルキンエ細胞にシナプスされないで、グラニュール細胞を介してシナプスされているのは、二つのオーバーラップした刺激パターンを次元を上げることにより、オーバーラップの少ない認識しやすいパターンに分離するためである。更に、ゴルジ細胞は、先にのべたように入力刺激の強弱にかかわらず、一定数のグラニュール細胞が発火するように抑制し、システムの感度を上げるとともに、プルキンエ細胞が小入力刺激パターンでも、苔状線維からの入力を処理できるようにしている。Fig. 1に小脳の概略を示す。

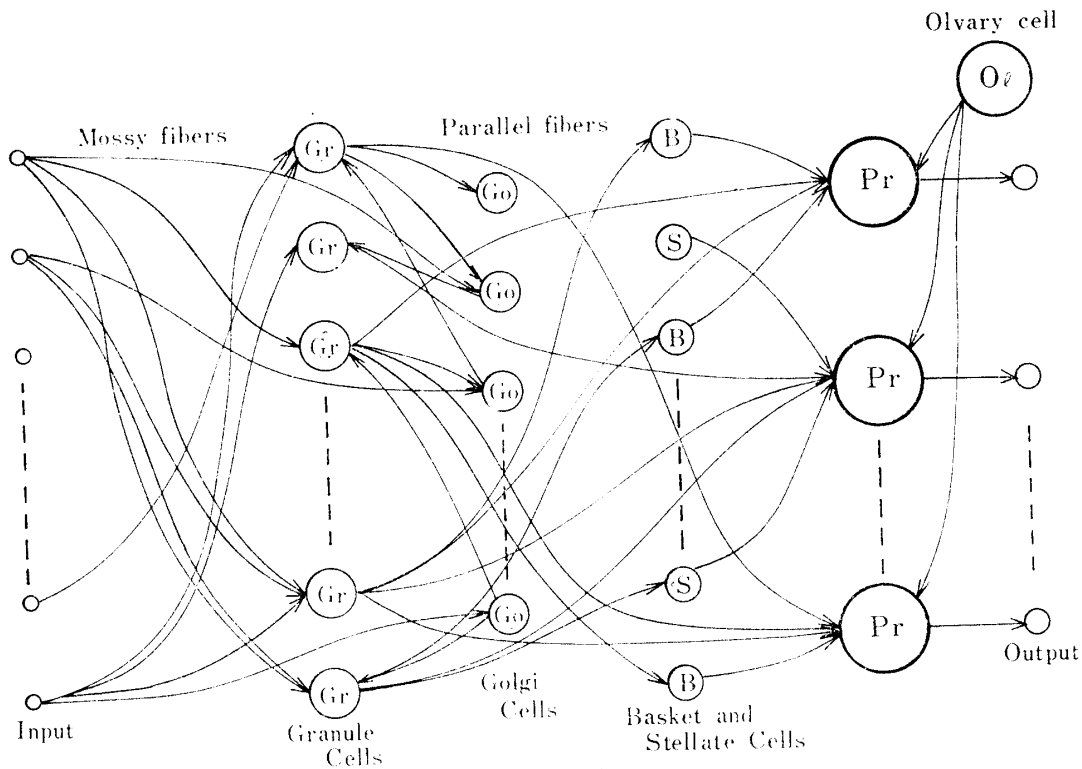


Fig. 1 Generalized Structure of Cerebellar Cortex

3. 層状ランダムネットと記号の説明

小脳のシステム、機能が示されたが、ここでは、学習機能、活動度を一定にする機能等についてはふれず、入力層とグラニュール細胞層間での次元拡大によって行なわれるパターン分離の能力についてのみ解析を行なう。更に、入力層とグラニュール層との結線数を平均4本とせず、ここではn本であるとして解析を行なう。この際、グラニュール細胞を適当な閾値をもつ閾値素子として考えて行なう。図1におけるn次元入力パターンXとYを

$$X = (x_1, x_2, \dots, x_n) \dots\dots\dots(1)$$

$$Y = (y_1, y_2, \dots, y_n) \dots\dots\dots(2)$$

$$x_i \text{ or } y_i = \begin{cases} 1 (\text{興奮状態}) \\ -1 (\text{興奮していない状態}) \end{cases}$$

とする。次に、グラニュール細胞出力Z、閾値Hおよび入力層と出力層のシナプス結合の重み行列を

$$Z = (z_1, z_2, \dots, z_N) \dots\dots\dots(3)$$

$$H = (h_1, h_2, \dots, h_N) \dots\dots\dots(4)$$

$$W = \begin{pmatrix} w_{11}, w_{12}, \dots, w_{1n} \\ w_{21}, w_{22}, \dots, w_{2n} \\ \vdots \quad \vdots \quad \quad \quad \vdots \\ w_{N1}, w_{N2}, \dots, w_{Nn} \end{pmatrix} \dots\dots\dots(5)$$

N: グラニュール層の次元

とすると、各入力パターン X, Y に対する出力は、

$$\left. \begin{aligned} z_X &= \text{sgn}(W \cdot X^T - HT) \\ z_Y &= \text{sgn}(W \cdot Y^T - HT) \end{aligned} \right\} \dots\dots\dots (6)$$

と表わされ、要素は

$$\left. \begin{aligned} z_{X_i} &= \text{sgn} \left(\sum_{j=1}^n w_{ij} x_j - h_i \right) \\ z_{Y_i} &= \text{sgn} \left(\sum_{j=1}^n w_{ij} y_j - h_i \right) \end{aligned} \right\} \dots\dots\dots (7)$$

となる。更に以上の説明より、入力パターン X, Y が分離されている様子を表わすため距離 d_s をハミング距離を用いて

$$d_s = \frac{1}{2n} \sum_{i=1}^n |x_i - y_i| \dots\dots\dots (8)$$

と定義する。また入力パターンの情報は、興奮状態の入力細胞からのみ伝えられる。そして後の解析では、これらの細胞の割合が重要となるので、パターン X, Y の興奮状態にある細胞数の割合を S_x, S_y 、入力パターン X と Y に共通に興奮する細胞数の割合を k とすると、興奮状態にある細胞間でのハミング距離 d は、

$$d = \frac{S_x + S_y - 2k}{S_x + S_y - k} \dots\dots\dots (9)$$

と表わされる。また d_H は、

$$d_H = S_x + S_y - 2k \dots\dots\dots (8)'$$

となる。これらの関係を図2に示す。

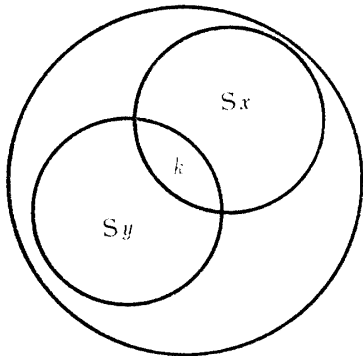


Fig. 2 Illustration of S_x, S_y, k

4. 層状ランダムネットの基本特性

オーバーラップの大きい入力パターンをオーバーラップの小さいパターンに変換するには、興奮状態にある細胞数と共通に興奮する細胞数の割合 $k/S_x, k/S_y$ がグラニューール細胞層で小さくなるようなランダム回路を構成すればよい。このようになる条件を示す基本的な関係を導くため、ここでは、 $S_x = S_y = S$ と仮定し、(7)式の sgn を取ったものを u_i, v_i とすると、

$$\left. \begin{aligned} u_i &= \sum_{j=1}^n w_{ij} x_j - h_i \\ v_i &= \sum_{j=1}^n w_{ij} y_j - h_i \end{aligned} \right\} \dots\dots\dots (10)$$

$x_{j's}, y_{j's}$ はランダム変数で独立であるので、 n が十分大きいときは、 u_i, v_i は平均値、分散

$$\left. \begin{aligned} m_{ui} &= E(u_i) = n\sigma w - h \\ m_{vi} &= E(v_i) = n\sigma w - h \end{aligned} \right\} \dots\dots\dots (11)$$

$$\left. \begin{aligned} \sigma_{ui}^2 &= Var(u_i) = n\sigma w^2 + \sigma h^2 \\ \sigma_{vi}^2 &= Var(v_i) = n\sigma w^2 + \sigma h^2 \end{aligned} \right\} \dots\dots\dots (12)$$

$$\sigma_{uiv_i} = Var(u_i v_i) = n(1 - 2d_H)\sigma w^2 + \sigma h^2$$

w, h : w_{ij}, h_i の分散, 平均値

の正規分布に従うことが分る。従って、2変数の密度関数 $f(u_i, v_i)$ は、

$$f(u_i, v_i) = \frac{1}{2\pi\sqrt{1 - \rho^2}\sigma_{ui}\sigma_{vi}} \cdot e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(u_i - m_{ui})^2}{\sigma_{ui}^2} - \frac{2(u_i - m_{ui})(v_i - m_{vi})}{\sigma_{ui}\sigma_{vi}} + \frac{(v_i - m_{vi})^2}{\sigma_{vi}^2} \right\}}$$

(13)

ただし、

$$\rho = \frac{n(1 - 2d_H)\sigma w^2 + \sigma h^2}{n\sigma w^2 + \sigma h^2}$$

となる。このときのグラニューール細胞側でのハミング距離は

$$\begin{aligned} d_g &= \frac{1}{2N} \sum_{i=1}^N |z_{Xi} - z_{Yi}| = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} |z_{Xi} - z_{Yi}| \\ &= \frac{1}{N} \sum_{i=1}^N r_i \end{aligned} \dots\dots\dots (14)$$

ただし、 $r_i = \frac{1}{2} |z_{Xi} - z_{Yi}|$

となり、1つのグラニューール細胞に関する変換後のハミング距離の関係は

$$P\{r_i > 0\} = \text{Prob}\{u_i v_i < 0\} \dots\dots\dots (15)$$

と求められる。ここで、 $m_{ui} = m_{vi} = 0$ とすれば、

$$\text{Prob}\{u_i v_i < 0\} = \Psi(d_H) = \frac{2}{\pi} \sin^{-1} \sqrt{Ad_H} \quad (16)$$

ただし、 $A = \frac{n\sigma w^2}{n\sigma w^2 + \sigma h^2}$

となる¹⁴⁾。従って入力側でハミング距離 d_H をもつパターンは変換後で、それ以上のハミング距離をもつ確率 $P(d_H)$ は

$$P(d_H) = \sqrt{\frac{1 - \Psi(d_H)}{2\pi}} \int_{d_H - \Psi(d_H)}^{\frac{1 - \Psi(d_H)}{\sigma_{d_H}}} \frac{1}{\sigma_{d_H}} e^{-\frac{x^2}{2}} dx \dots\dots (17)$$

と表わされる。一方、入力側の興奮細胞の割合 S と

共通に興奮する割合の k の比を t とするとき、変換後のこの値は、変換後の各素子の興奮確率 φ が、 $m_{ni} = m_{vi} = 0$ とすれば、

$$\text{Prob}(u_i > 0) = \varphi(S) = \frac{1}{\pi} \left(\sin^{-1} \sqrt{\frac{1+\rho}{2}} + \sin^{-1} \sqrt{\frac{1-\rho}{2}} \right) = \frac{1}{2} \quad \dots\dots\dots(18)$$

となり、一方共通に発火する確率 $\varphi(k)$ は

$$\text{Prob}(u_i > 0, v_i > 0) = \varphi(k) = \frac{1}{\pi} \sin^{-1} \sqrt{\frac{1+\rho}{2}} \quad \dots\dots\dots(19)$$

と求められる。従がって、変換後の t を $R(t)$ とすれば、出力共通発火率は

$$R(t) = \frac{\varphi(k)}{\varphi(S)} = \frac{2}{\pi} \sin^{-1} \sqrt{\frac{1+\rho}{2}} \quad \dots\dots\dots(20)$$

と求められる。従がって、全グラニューール素子について $R(t) < t$ となる確率を考えれば、

$$\chi_N(t) = \frac{1}{\sqrt{2\pi}} \int_{-R(t)/N}^{t-R(t)/N} e^{-\frac{x^2}{2}} dx \quad \dots\dots\dots(21)$$

となる。従がって、パターンが入力層—出力側で次元拡大によって分離されるためには、必要情報を担う興奮素子における t の割合が小さくなる確率(21)の値が N の増大とともに大きくなることはいえよいことが分る。このことは、全体でのパターンのハミング距離が変換後、確率 $P(d_H)$ で大きくなることを意味する。(18式から(21)式の導出は付録で示す。

5. パターンの分離性

前節で層状 ランダム 回路網の特性を明らかにしたが、ここでは入力パターンが出力側で分離される条件を明らかにする。このためまず入力側のパターンに対する条件を先に確す。(9)式と(8)式において、入力素子全体に対するハミング距離 d_H は、興奮素子のみでのハミング距離 d より小さいはずであるから、

$$d_H < d \quad \dots\dots\dots(22)$$

従がって、(22)式に(9)、(8)を代入すると、

$$k \leq S \leq \frac{k+1}{2} \quad (\because S_x = S_y) \quad \dots\dots\dots(23)$$

なる式が求まり、入力素子興奮割合と共通に興奮する割合との関係がもとまる。また、(17)より $A=1$ とすれば、層状ランダム回路網によって入力パターン間のハミング距離が出力側で拡大される範囲は

$$\psi(d) > d \quad \text{すなわち} \quad d < \frac{1}{2} \quad \dots\dots\dots(24)$$

と求まり、かつ $\frac{1}{2}$ が極限の状態であることが分る。そこでこの範囲にある入力パターンの範囲を求めると

$$S < k + \frac{1}{4} \quad \dots\dots\dots(25)$$

となる。そして、(23)、(25)を満足する領域では、相関数 $\rho \geq 0$ となり、それ以外では $\rho < 0$ となることが分る。これを図(3)に示す。次に、入力パターンが出力側

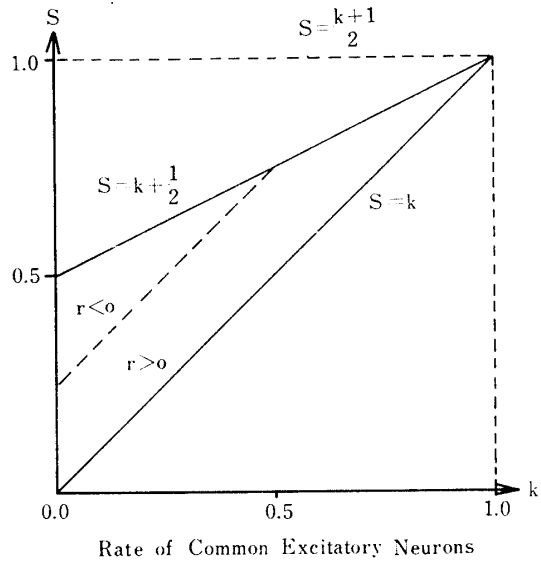


Fig. 3 The area of admissible input patterns

で分離されるためには、(21)式であたえられる分離の度合を示す確率 χ_N の積分の上限が正になる必要がある。故に、

$$t - R(t) > 0 \quad \dots\dots\dots(26)$$

(26)式を代入して

$$\rho < 2 \left\{ \sin \left(\frac{\pi}{2} t \right) \right\}^2 - 1 \quad \dots\dots\dots(27)$$

が求まる。従がって、入力側のパターンが分離されるためには、(27)式の関係が成り立つ必要があり、パターン分離性の条件式となる。従がって、 $t - R(t)$ がわずかでも正になるならグラニューール細胞層の次元拡大によって χ_N の確率値が1となる確率 P は

$$\lim_{N \rightarrow \infty} P\{1 - \chi_N < \epsilon\} = 1 \quad (\because \epsilon > 0) \quad \dots\dots\dots(28)$$

となる。図(4)、図(5)に相関係数 ρ 入力 の 共通発火率 t 、出力の 共通発火率 $R(t)$ の関係を示す。図(6)に出力素子数(ニューロン数) N に対する確率 $P(d_H)$ 、 χ_{40} の変化の様子を一定のパラメータの場合について示す。

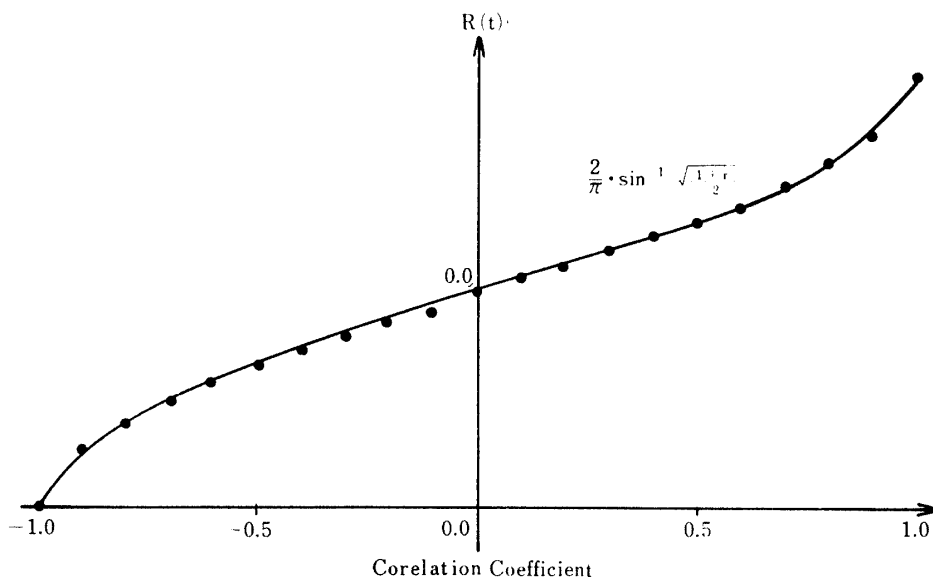


Fig. 4 The relation between ρ and $R(t)$

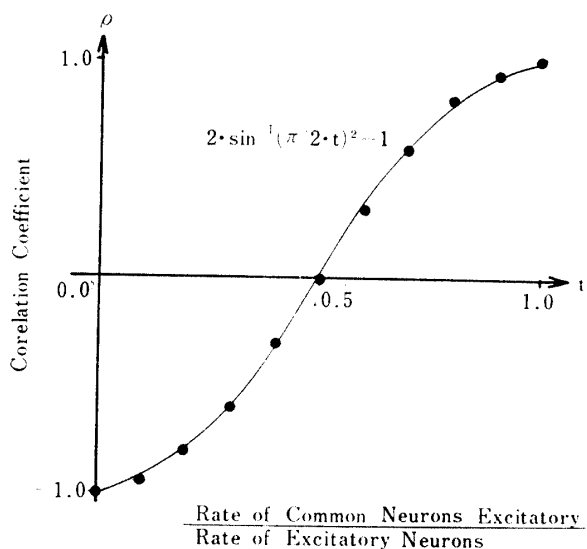


Fig. 5 The relation between t and ρ

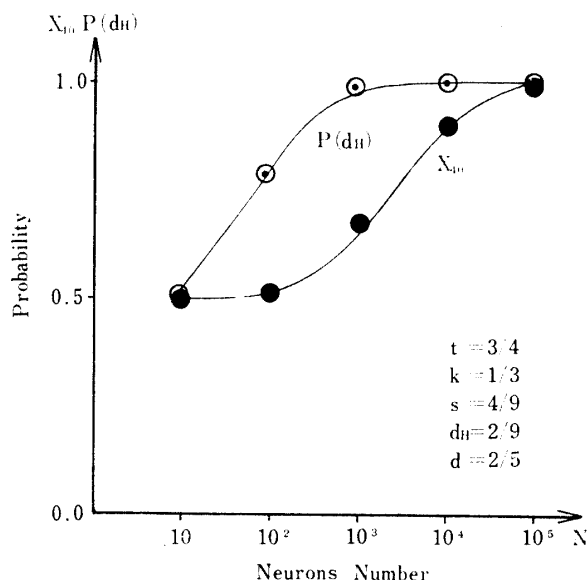


Fig. 6 The relation between $\rho(d_H)$ or X_{40} and N

6. シミュレーション実験

以上の解析が正しいことをたしかめるため、入力次元8，出力次元40の層状ランダム回路網を組みシミュレーション実験を行なった。実験としては

(1) $d_H=0.250$, $S=0.5$, $t=0.75$

(2) $d_H=0.750$, $S=0.5$, $t=0.572$

なる関係をもつパターンを用いて行なった。この結果、**実験(1)**については確率1.0でパターンが分離され、**実験(2)**については確率0.478で分離が行なわれた。**実験(1)**では(20)式が満足され、当然分離されることが分り、式(21)から求めた X_{40} の確率0.9とかなり一致する。一方、**実験(2)**では(20)式が満足されず、

分離されるよりもむしろ重なりあってくることが分り、理論式(21)からも確率 $X_{40}=0.24$ となり、実験結果もその傾向を示していることが分る。

7. 結 論

以上の解析、及びシミュレーション実験の結果より、次のことが明らかになった。

- (1) パターンが分離される条件(20)を満足する t は、図(5)より t が小さくなるほど満足されにくくなり、分離機能がなくなることが明らかになった。
- (2) シミュレーション実験の結果は、小杉らが行なった結果と $t > R(t)$ が満足されるとき良

く一致する。

(3) 入力次元を出力側で拡大することにより、分離作用が強調されることになる。

今後の問題としては、小脳がもつグラニューール細胞層の発火率を1定にする機能と、分離機能との関係、および入力側とグラニューール細胞層の接続数が小脳におけるように少ない(約 $n=4$) の場合の解析を行なう必要があると思われる。最後に、日頃御世話になっている東京大学甘利助教授に文献を頂いただいたことを付記するとともに、深く感謝する。

参 考 文 献

- 1) W.S. McCulloch and W. Pitt: "A Logical Calculus of the Ideas Immanent in Neurons Activity" Bull. Math. Biophys, 1943, 5, 115-133 (1943)
- 2) A. Rapoport: "Ignition Phenomena in Random Nets", Bull. Math. Biophys. Volume 14, (1952)
- 3) F. Rosenblatt: "Principle of Neurodynamics" Spartan (1962)
- 4) N. Rochester, et. al.: "The nets a cell assembly theory of the action of the brain", Using a large digital Computer, IRE, IT-2, 80 (Sep. 1955)
- 5) E.R. Caianello: "Outline of a Theory of Thought-Process and Thinking Machine", J. Theoret. Biol (1961) 2, 204-235
- 6) Poul B. post: "A Lifelike Model for Association Relevance", Proc. Interntional Joint. Conf. on Artificial Intelligence (1969) 271-290, Washinton May 7, 9
- 7) L.I. Rozoner: "Random Logical net, I" Automation and Remote Control, May 5, 137 (1969)
- 8) L.I. Rozoner: "Random Logical net, II" Automation and Remote Control, June 6, 99 (1969)
- 9) L.I. Rozoner: "Random Logical net, III", Automata and Remoto Control, July 7, 127, (1969)
- 10) E.M. Harth, et al: "Brain Function and Neural Dynamics", J. Theoret, Biol (1970), 25, 93-120
- 11) P.A. Anninos, et. al: "Dynamics of Neural Structures", J. Thoret, Biol, (1970), 26, 121-148
- 12) 甘利俊一: "アナログ神経素子よりなるランダム回路の特性", 171-178, 電子通信学会誌, Vol. 55-D No. 3
- 13) 甘利俊一: "興奮および抑制性神経素子よりなるランダム回路", Vol. 55-D, No. 3, 179-18,
- 14) A menthod of Statistical Neurodynamics", Kybernetik 14, 201-215 (1974)
- 15) D. Marr: "A Theory of Cerebellar Cortex", J. Pysiol. 202, 437-470 (1969)
- 16) J.S. Albus: "Theory of Cerebellar Function", Math. Bioscience 10, 25-61
- 17) 小杉, 内藤: "神経回路のパターン分類能力に関する一考察", 電子通信学会資料, MBE 74-4, 37-48 (1974)

付 録

本文(13)式は、平均値が0で分散が等しいゆえ

$$f(u_i, v_i) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma^2} e^{-\frac{1}{2(1-\rho^2)\sigma^2}(u_i^2 - 2\rho u_i v_i + v_i^2)} \dots\dots\dots(付1)$$

となる。従がって、 $u_i > 0$ となる確率 $\text{Prob}(u_i > 0)$ は、

$$\text{Prob}(u_i > 0) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma^2} \int_{-\infty}^{\infty} \int_0^{\infty} f(u_i, v_i) du_i dv_i \dots\dots\dots(付2)$$

となる。ここで

$$\left. \begin{aligned} u_i &= \sqrt{2(1-\rho^2)}\sigma \cos\theta \\ v_i &= \sqrt{2(1-\rho^2)}\sigma \sin\theta \end{aligned} \right\} \dots\dots\dots(付3)$$

と変数変換し積分すれば、

$$\text{Prob}(u_i > 0) = \frac{\sqrt{1-\rho^2}}{2\pi} \int_{-\pi/2}^{\pi/2} \int_0^{\infty} r \cdot e^{-r^2(1-2\rho\sin\theta\cos\theta)} dr d\theta = \frac{1}{2} \dots\dots\dots(付4)$$

となる。(19)式は同様の考え方により

$$\begin{aligned} \text{Prob}(u_i > 0, v_i > 0) &= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\pi/2} \int_0^{\infty} r e^{-r^2(1-2\rho\sin\theta\cos\theta)} dr d\theta \\ &= \frac{1}{\pi} \sin^{-1} \sqrt{\frac{1+\rho}{2}} \dots\dots\dots(付5) \end{aligned}$$

となる。従がって、(20)式 $R(t)$ は

$$R(t) = \frac{\text{Prob}(u_i > 0, v_i > 0)}{\text{Prob}(u_i > 0)} = \frac{2}{\pi} \sin^{-1} \sqrt{\frac{1+\rho}{2}} \dots\dots\dots(付6)$$

となる。次に、(21)式の $R(t) < t$ となる確率 χ_N は2項分布より、

$$\chi_N = \sum_{x=0}^{t-N} {}^N C_x R(t)^x \cdot (1-R(t))^{N-x} \dots\dots\dots(付7)$$

と表わされる。ここでは、素子数 N は十分であるので、この式を正規分布で近似すると、平均値 m_R 、分散 σ_R^2 となり、

$$\begin{aligned} m_R &= N \cdot R(t) \dots\dots\dots(付8) \\ \sigma_R^2 &= N \cdot R(t) (1-R(t)) \\ \chi_N &= \frac{1}{\sqrt{2\pi}\sigma_R} \int_0^{t-N} e^{-\frac{(x-m_R)^2}{2\sigma_R^2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\frac{R(t)}{\sigma_R} \cdot N}^{\frac{t-R(t)}{\sigma_R} \cdot N} e^{-t^2/2} dt \dots\dots\dots(付9) \end{aligned}$$

ともとまる。