KŌDAI MATH. SEM. RFP.
21 (1969), 46—57

THE RANDOM NET WHICH HAS BASIC ORGANS REALIZING PARITY BOOLEAN FUNCTIONS

BY YASUICHI HORIBE

This paper is a revised and modified version of the foregoing paper [1] in the sense that the ability of pattern discrimination is much more increased than [1].

1. Brief review of the general concept.

We briefly review the fundamental concept of pattern recognition by random net proposed in [1] (See [1] for detail).

I is the set of n "input points". Denoting by $\pi(A)$ the number of elements belonging to a finite set A, we have $\pi(I)=n$. Any subset $f\subset I$ will be called *(input) pattern* which may be interpreted as a binary 0, 1 sequence of length n. The 2^n possible patterns constitute the *(input) pattern space* F. To say that there are given K categories on F is to say that K probability distributions

$$\mathcal{P}_n=\{P_n^{(1)}, P_n^{(2)}, \cdots, P_n^{(K)}\}$$

are defined on F, considering them to depend on n.

A random net transforms F randomly into another pattern space (output pattern space) G comprising of patterns which are subsets of the set of N output points. Then the random net defines a mapping (assumed deterministic in the present study) φ: $F\to G$. We have thus $\pi(G)\leqq 2^n$.

Given a category $k\in C=\{1, 2, \cdots, K\}$, denote by $Q_{l0}^{(k)}$ the probability that the l-th output point emits signal 0. If the random net has the property that the N output component signals are mutually *independent*, the probability that the corresponding output pattern $\varphi(f)\in G$ is observed, given category k, is given by

$$q_{\varphi(f)}^{(k)}\equiv \prod_{l\in\varphi(f)}(1-Q_{l0}^{(k)})\prod_{l\notin\varphi(f)}Q_{l0}^{(k)}. \tag{1}$$

If we assume a learning mechanism which can estimate the matrix $\mathcal{M}=(Q_{l0}^{(k)})$ and the probabilities $p^{(1)}, p^{(2)}, \cdots, p^{(K)}$ on the category space C, then the *a posteriori* probability method to recognize patterns may be as follows:

(A) An unknown input pattern $f\in F$ is given and the $\varphi(f)\in G$ is observed at the output level.

(B) By (1) $p^{(k)}\cdot q_{\varphi(f)}^{(k)}$, $k=1, 2, \cdots, K$, are calculated.

Received May 9, 1968.

(C) Determine the maximal subset $C(\varphi(f))\subset C$ each of whose categories gives the same maximum value among these K values.

(D) Decide that f belongs to one of categories in $C(\varphi(f))$.

Denoting by $\mathcal{D}[\varphi(f)]$ the decision $C(\varphi(f))$, then we can show that $\mathcal{D}[f]=\mathcal{D}[\varphi(f)]$ if φ is *one-to-one* mapping, where $\mathcal{D}[f]$ is the decision performed directly on the basis of $\mathcal{P}_n$ without using random net.

What was mentioned above suggests that if we could construct the random net such that the output component signals are mutually independent and φ is one-to-one, then we could have an optimal pattern recognition mechanism.

The main object of the present work is to prove that this is effectively possible if the size of the random net which has basic organs realizing parity Boolean functions is properly enlarged.

2. Assumptions on the pattern space structure.

We put the following three basic assumptions on the input patten space F. These assumptions, however, never put any restriction on the generality of pattern space structure. Rather they are not only for convenience in modelling a pattern recognition mechanism, but also they seem to reflect (or at least simplify) some of the intrinsic natures of pattern space for recognition.

Since we always take $n=\pi(I)$ sufficiently large, the pattern variety may be considered to be astronomically enormous.

We call the ratio $\rho(f)\equiv\pi(f)/n$ *stimulation area* (or simply *s-area*) of pattern $f\in F$.

ASSUMPTION 1. For any pattern $f\in F$, $\rho(f)=1/2$.

This is realized, for example, by considering a regulating or normalizing mechanism which regulates any "external pattern" f_e by adding to it an "internal pattern" f_i so as to make $\rho(f_e\cup f_i)=1/2$ (see also [1]).

ASSUMPTION 2. For any pair of distinct patterns $f,f'\in F$, there exist uniform bounds $\underline{b}$ and $\bar{b}$ such that $1\leqq\underline{b}\leqq\pi(f\cap f')\leqq\bar{b}\leqq n/2-1$.

The $\underline{b}$ and $n/2-\bar{b}$ may usually be much greater than 1, since n is large.

ASSUMPTION 3. For any category $k\in C$, $\zeta_n^{(k)}\equiv\sum_{f\in F}\{P_n^{(k)}(f)\}^2\to 0$ as $n\to\infty$, i.e. $1-\zeta_n^{(k)}=\sum_{f\neq f'}P_n^{(k)}(f)P_n^{(k)}(f')\to 1$, $n\to\infty$ (this convergence is considered to be very rapid perhaps of order α^{-n}, $\alpha>1$).

It is natural that the patterns f^* to be recognized are "imbedded" in various random (or noisy) patterns f_r, both of them constitute our patterns f, i.e. $f=f^*\cup f_r$, and furthermore f^* itself has tremendously various forms. Thus once a pattern has occurred, then *exactly the same pattern* will almost never occur in future, when n becomes large. This consideration, though intuitive, leads to the assumption 3.

3. A class of basic organs usable in the random net.

Every basic organ contained in the random net is to realize a certain Boolean function. The basic organ has, therefore, finite input lines receiving binary stimuli $\sigma_1, \sigma_2, \cdots, \sigma_s$ $(=0$ or $1)$ and one output line emitting a binary signal σ_0 $(=0$ or $1)$, realizing a map: Φ: $\{0,1\}^s \rightarrow \{0,1\}$.

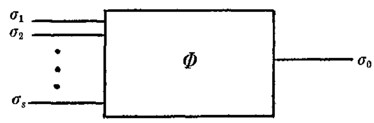

FIG. 1. Basic organ

A Boolean function $\Phi(\sigma_1, \sigma_2, \cdots, \sigma_s)$ is called *symmetric* if the function value remains invariant under any permutation of variables $\sigma_1, \sigma_2, \cdots, \sigma_s$. A class of basic organs which will be used to construct the random net in the present study is a subclass of basic organs realizing symmetric Boolean functions. That is a class of organs which realize *parity functions* for s odd with $s \geqq 3$.

The parity Boolean function is defined as:

$$\Phi(\sigma_1, \sigma_2, \cdots, \sigma_s) = \sigma_1 \oplus \sigma_2 \oplus \cdots \oplus \sigma_s,$$

where the operation $\oplus$ means "exclusive or", i.e. $0\oplus 0=0$, $0\oplus 1=1$, $1\oplus 0=1$, $1\oplus 1=0$. Let us denote this function (also the basic organ realizing it) of $s=2r+1$ variables by Φ_{2r+1}, $r \geqq 1$.

If the basic organ Φ_{2r+1} is constructed in *canonic form* only with (3-input) majority elements, the number of elements needed is shown to be $2(r+1)^2-1$, but if with various threshold elements, the minimum number of elements needed in canonic form construction is shown to be $1+[\log_2(2r+1)]$, where $[a]$ is the integer part of the real number a. See for deiail [2].

The most significant role of the organ Φ_{2r+1} in the pattern recognizing random net is in its probabilistic logic property. That is, when $2r+1$ input lines are stimulated mutually independently each with probability ρ, the probability $u_{2r+1}(\rho)$ that the output line is stimulated is given by

$$\begin{aligned} u_{2r+1}(\rho) &= \sum_{k=0}^{r} \binom{2r+1}{2k+1} \rho^{2k+1}(1-\rho)^{2(r-k)} \\ &= \frac{1}{2} u_{2r+1}(\rho) + \frac{1}{2}\left\{ u_{2r+1}(\rho) - \sum_{k=0}^{2r+1} \binom{2r+1}{k} \rho^k (1-\rho)^{2r+1-k} \right\} + \frac{1}{2} \\ &= \frac{1}{2}\left\{ \sum_{k=0}^{r} \binom{2r+1}{2k+1} \rho^{2k+1}\{-(1-\rho)\}^{2(r-k)} + \sum_{k=0}^{r} \binom{2r+1}{2k} \rho^{2k}\{-(1-\rho)\}^{2(r-k)+1} \right\} + \frac{1}{2} \\ &= \frac{1}{2}\{\rho-(1-\rho)\}^{2r+1} + \frac{1}{2} = \frac{1}{2}\{1-(1-2\rho)^{2r+1}\}. \end{aligned}$$

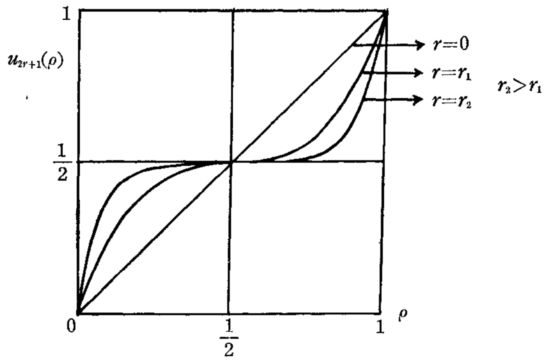

FIG. 2.

4. The net construction by random connections.

We now state how the random net is constructed with the basic organs Φ_{2r+1}. Although a plan to use organs of various r in the net construction would be possible, r is fixed arbitrarily in the present study.

Take N_1 basic organs. Each input line of each organ is connected completely randomly to one of n input points. Then we obtain a random net $\mathfrak{N}(0,1)$ between the layer $\mathcal{L}_0$ consisting of n input points and the layer $\mathcal{L}_1$ consisting of N_1 basic organs. Take next N_2 basic organs. Each input lines of each of these organs is randomly connected to one of N_1 output lines of layer $\mathcal{L}_1$, obtaining a random net $\mathfrak{N}(0,2)$ between $\mathcal{L}_0$ and the layer $\mathcal{L}_2$ consisting of N_2 basic organs. Continue this construction until we reach the layer $\mathcal{L}_L$ to obtain the net $\mathfrak{N}(0,L)$, which may be considered as the net comprising of L subrandom bets $\mathfrak{N}(0,1), \mathfrak{N}(1,2), \cdots, \mathfrak{N}(L-1,L)$ in series. The random net $\mathfrak{N}(0,L)$ thus constructed has n input points and N_L output lines on the whole.

5. Maintainability of stimulating area.

The random net $\mathfrak{N}(0,\lambda)$, $1\leqq\lambda\leqq L$, transforms the input pattern space F into the pattern space G_λ which consists of all possible binary patterns of length N_λ that can be observed at the layer $\mathcal{L}_\lambda$. Consider especially the net $\mathfrak{N}(0,1)$ and the corresponding pattern space G_1. When an input pattern f ($\rho(f)=1/2$) in F is presented to the input layer $\mathcal{L}_0$, its corresponding transformed pattern $g_1 \in G_1$ has $\pi(g_1)$ which is a realization of a random variable obeying normal distribution $\boldsymbol{N}((1/2)N_1, (1/2)(1-1/2)N_1)$ with mean $(1/2)N_1$ and the variance $(1/2)(1-1/2)N_1$, if n, hence N_1, is taken large (note that we must take $n\leqq N_1\leqq N_2\leqq\cdots$.), since $u_{2r+1}(1/2)=1/2$ and the random connection operation may be regarded as a Bernoulli trial of coin tossing. Hence the s-area $\rho(g_1)=\pi(g_1)/N_1$ of g_1 is a realization of a random variable obeying $\boldsymbol{N}(1/2, 1/4N_1)$. Let us write this last statement as

$$\rho(g_1)\mathscr{R}\boldsymbol{N}(1/2, 1/4N_1).$$

Generally, if the random net $\mathfrak{N}(0, \lambda)$ transforms $f\in F$ into $g_\lambda \in G_\lambda$, $1\leqq\lambda\leqq L$, then

$$\rho(g_{\lambda+1})\mathscr{R}\boldsymbol{N}\left(u_{2r+1}(\rho(g_\lambda)), u_{2r+1}(\rho(g_\lambda))(1-u_{2r+1}(\rho(g_\lambda)))\cdot\frac{1}{N_{\lambda+1}}\right).$$

Varying f in F we see that the s-area $\rho_{\lambda+1}$ of patterns $\in G_{\lambda+1}$ observable at the layer $\mathcal{L}_{\lambda+1}$ is a variable having (approximate) density

$$D(\rho_{\lambda+1})=\int_{-\infty}^{\infty}\boldsymbol{n}(\rho_{\lambda+1};\rho_\lambda)\,d\rho_\lambda\int_{-\infty}^{\infty}\boldsymbol{n}(\rho_\lambda;\rho_{\lambda-1})\,d\rho_{\lambda-1}\cdots\int_{-\infty}^{\infty}\boldsymbol{n}(\rho_2;\rho_1)\boldsymbol{n}(\rho_1)d\rho_1,$$

where $\boldsymbol{n}(\rho_1)$ is the normal density of ρ_1 obeying $\boldsymbol{N}(1/2, 1/4N_1)$ and $\boldsymbol{n}(\rho_{\lambda+1};\rho_\lambda)$ is the normal density function of $\rho_{\lambda+1}(\rho_\lambda)$ obeying $\boldsymbol{N}(u_{2r+1}(\rho_\lambda), u_{2r+1}(\rho_\lambda)(1-u_{2r+1}(\rho_\lambda))/N_{\lambda+1})$, $1\leqq\lambda\leqq L-1$. The mean $\int\rho_\lambda D(\rho_\lambda)d\rho_\lambda$ of ρ_λ is obviously 1/2 from the symmetricity of $u_{2r+1}(\rho)$ with respect to the point (1/2, 1/2). But the variance of ρ_λ is not readily obtained because of the difficulty for exactly calculating $D(\rho_\lambda)$. An approximate upper bound of the variance is, however, obtained by the following argument. Since the concentration of s-areas around 1/2 mainly concerns us, it is sufficient to consider the problem only in the neighbourhood of 1/2. In

$$\rho_{\lambda+1}(\rho_\lambda)\mathscr{R}\boldsymbol{N}\left((u_{2r+1}(\rho_\lambda), u_{2r+1}(\rho_\lambda)(1-u_{2r+1}(\rho_\lambda))\frac{1}{N_{\lambda+1}}\right)$$

we estimate the variance larger, denoting ([1])

$$\rho_{\lambda+1}(\rho_\lambda)\mathscr{R}^*\boldsymbol{N}\left(u_{2r+1}(\rho_\lambda), \frac{1}{4}\cdot\frac{1}{N_{\lambda+1}}\right),$$

since $0\leqq u_{2r+1}(\rho_\lambda)(1-u_{2r+1}(\rho_\lambda))\leqq 1/4$. And furthermore estimate larger the deviation of the mean from 1/2, denoting $\rho_{\lambda+1}(\rho_\lambda)\mathscr{R}^{**}\boldsymbol{N}(a(r)(\rho_\lambda-1/2)+1/2, 1/4N_{\lambda+1})$, since in the neighbourhood of 1/2, $|u_{2r+1}(\rho_\lambda)-1/2|\leqq|a(r)(\rho_\lambda-1/2)|$ for suitable $a(r)$ depending on r such that $0<a(r)\ll 1$. Then we easily see that, if we put $\rho_\lambda\mathscr{R}^*\boldsymbol{N}(\mu_\lambda, d_\lambda^2)$ and $\rho_{\lambda+1}\mathscr{R}^*\boldsymbol{N}(\mu_{\lambda+1}, d_{\lambda+1}^2)$, then

$$\begin{cases}\mu_{\lambda+1}=a(r)\mu_\lambda+\dfrac{1}{2}(1-a(r))=\dfrac{1}{2}+a(r)\left(\mu_\lambda-\dfrac{1}{2}\right),\\[2ex] d_{\lambda+1}^2=[a(r)]^2d_\lambda^2+\dfrac{1}{4N_{\lambda+1}}, \qquad 1\leqq\lambda\leqq L-1,\end{cases}$$

since the lemma 1 of [1] (i.e., if the variable x obeys $\boldsymbol{N}(m,\sigma^2)$ and the variable $y(x)$ for fixed x obeys $\boldsymbol{N}(ax+b,\delta^2)$, a, b constants, then $y(x)$ averaged by x obeys $\boldsymbol{N}(am+b,(a\sigma)^2+\delta^2)$ is applicable to the two distributions $\boldsymbol{N}(\mu_\lambda, d_\lambda^2)$ and $\boldsymbol{N}(a(r)(\rho_\lambda-1/2)+1/2, 1/4N_{\lambda+1})$. The initial conditions $\mu_1=1/2$, $d_1^2=1/4N_1$ give us

$$\mu_L=\frac{1}{2},$$

$$d_L^2=\frac{[a(r)]^{2(L-1)}}{4N_1}+\frac{[a(r)]^{2(L-2)}}{4N_2}+\cdots+\frac{[a(r)]^2}{4N_{L-1}}+\frac{1}{4N_L}$$

$$\leqq\frac{1}{4N_1}\{1+[a(r)]^2+[a(r)]^4+\cdots+[a(r)]^{2(L-1)}\}$$

$$\leqq\frac{1}{4N_1(1-[a(r)]^2)}.$$

Remark that $a(r+1)<a(r)$, and that $a(r)\to 0$ when r becomes large. And also note that $N_1=N_1(n)\to\infty$ as $n\to\infty$ ($N_1(n)$ will be remarked in section 9.). Thus we have

LEMMA 1. *The random net* $\mathfrak{N}(0,L)$ *maintains the s-area of any input pattern to be* 1/2 *with the standard deviation* d_L *from* 1/2 *having the following bound*:

$$d_L^2\leqq\sum_{\lambda=1}^{L}\frac{[a(r)]^{2(L-\lambda)}}{4N_\lambda}\leqq\frac{1}{4N_1(1-[a(r)]^2)},$$

where $0<a(r+1)<a(r)\ll 1$ *and* $a(r)\to 0$, $r\to\infty$ *and* $N_1=N_1(n)\to\infty$, $n\to\infty$.

By the above result we may assume in the sections to follow that the s-area of any input pattern is maintained to be precisely 1/2 in order not to make the argument inessentially complicated.

6. Changes of intersecting stimulating area.

We shall consider another important factor of the random net, namely the intersecting stimulating area defined below.

We again focus our attention on the subrandom net $\mathfrak{N}(0,1)$ and two fixed distinct input patterns f and f' in F. Define by $\sigma(f\cap f')=\pi(f\cap f')/n$ the *intersecting stimulating area, simply i.s-area,* of these two patterns. Suppose $\sigma(f\cap f')=\sigma$. Then $0<\sigma<1/2=\rho(f)=\rho(f')$. The problem is how this i.s-area changes as these two patterns are transformed by the net $\mathfrak{N}(0,1)$. For this it is sufficient to investigate the probability that a basic organ in $\mathcal{L}_1$ is stimulated (output 1) by both f and f' when its $2r+1$ input lines are randomly connected to input points.

Denote by α, $\bar{\alpha}$, β, β', the number of input lines of the basic organ which are connected to the points in sets $f\cap f'$ $I-(f\cup f')$, $f-(f\cap f')$, $f'-(f\cap f')$, respectively. Note here that $\alpha+\bar{\alpha}+\beta+\beta'=2r+1$ and $\pi(f\cap f')/n=\pi(I-(f\cup f'))/n=\sigma$, $\pi(f-(f\cap f'))/n=\pi(f'-(f\cap f'))/n=1/2-\sigma$.

We easily see that the basic organ is stimulated by both f and f' if and only if $\alpha+\beta=\text{odd}$ and $\alpha+\beta'=\text{odd}$. Therefore $\alpha+\bar{\alpha}$ must be odd and $\beta+\beta'$ must be even (including 0), since $\alpha+\bar{\alpha}+\beta+\beta'=2r+1=\text{odd}$. The probability that $\alpha+\bar{\alpha}=2k+1$ and $\beta+\beta'=2r+1-(2k+1)=2(r-k)$, $0\leqq k\leqq r$, is

$$\binom{2r+1}{2k+1}\sigma^{2k+1}\left(\frac{1}{2}-\sigma\right)^{2(r-k)}.$$

Fix $k=0,1,\cdots,r$. The number of ways of choosing α input lines and $\bar{\alpha}$ input lines out of $2k+1$ input lines such that α is odd $\bar{\alpha}$ is even is

$$\sum_{a=0}^{k}\binom{2k+1}{2a+1}.$$

But in this case β and β' must be even. The number of ways of choosing β input lines and β' input lines out of $2(r-k)$ input lines such that β and β' are both even is

$$\sum_{a=0}^{r-k}\binom{2(r-k)}{2a}.$$

Hence

$$\sum_{a=0}^{k}\binom{2k+1}{2a+1}\cdot\sum_{a=0}^{r-k}\binom{2(r-k)}{2a}$$

ways.

Similarly for the case $\alpha=\text{even}$, $\bar{\alpha}=\text{odd}$, $\beta=\text{odd}$, $\beta'=\text{odd}$, we have

$$\sum_{a=0}^{k}\binom{2k+1}{2a}\cdot\sum_{a=0}^{r-k-1}\binom{2(r-k)}{2a+1}$$

ways. Therefore for fixed k, there are

$$\sum_{a=0}^{k}\binom{2k+1}{2a+1}\cdot\sum_{a=0}^{r-k}\binom{2(r-k)}{2a}+\sum_{a=0}^{k}\binom{2k-1}{2a}\sum_{a=0}^{r-k-1}\binom{2(r-k)}{2a+1}=\frac{2^{2r+1}}{2}$$

ways of connection patterns in all, since the expansions of

$$(1-1)^{2k+1},\qquad (1-1)^{2(r-k)},\qquad (1+1)^{2k+1},\qquad \text{and}\qquad (1+1)^{2(r-k)},$$

show that

$$\sum_{a=0}^{k}\binom{2k+1}{2a+1}=\sum_{a=0}^{k}\binom{2k+1}{2a}=\frac{1}{2}\cdot 2^{2k+1},$$

$$\sum_{a=0}^{r-k-1}\binom{2(r-k)}{2a+1}=\sum_{a=0}^{r-k}\binom{2(r-k)}{2a}=\frac{1}{2}\cdot 2^{2(r-k)}.$$

The probability to be obtained is, therefore,

$$\sum_{k=0}^{r} \frac{2^{2r+1}}{2}\binom{2r+1}{2k+1}\sigma^{2k+1}\left(\frac{1}{2}-\sigma\right)^{2(r-k)}$$

$$=\frac{1}{2}\sum_{k=0}^{r}\binom{2r+1}{2k+1}(2\sigma)^{2k+1}(1-2\sigma)^{2(r-k)}$$

$$=\frac{1}{4}\{1-(1-4\sigma)^{2r+1}\},$$

where the last equality has been shown (put $\rho=2\sigma$) in section 5. Put

$$v_{2r+1}(\sigma)=\frac{1}{4}\{1-(1-4\sigma)^{2r+1}\}.$$

By a similar argument to that of section 5, if f, f' correspond to $g_1, g_1' \in G_1$, respectively, then i. s-area of g_1 and g_1' satisfies

$$\sigma(g_1 \cap g_1')=\frac{\pi(g_1 \cap g_1')}{n}\cdot \mathcal{R}\boldsymbol{N}\left(v_{2r+1}(\sigma),\, v_{2r+1}(\sigma)(1-v_{2r+1}(\sigma))\frac{1}{N_1}\right).$$

When the net $\mathfrak{N}(0, \lambda)$ transforms f and f' into $g_\lambda, g_\lambda' \in G_\lambda$, respectively, then i. s-area $\sigma_\lambda=\sigma(g_\lambda \cap g_\lambda')$ of g_λ and g_λ' changes *in the mean* recursively as follows:

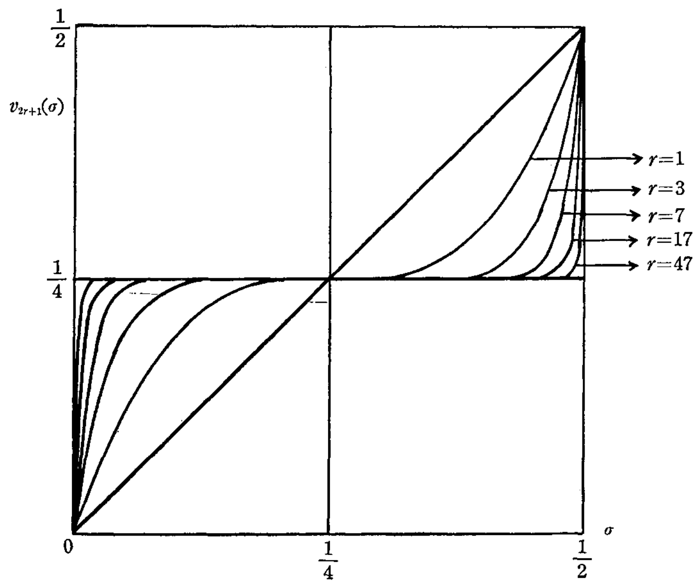

FIG. 3.

$$\sigma_{\lambda+1}=v_{2r+1}(\sigma_\lambda), \qquad \sigma_0=\sigma.$$

The variance argument is quite analogous with section 5 and the function $v_{2r+1}(\sigma)$ has the property to nicely "attract" i. s-areas to the point 1/4. We therefore omit the detail and conclude:

LEMMA 2. *The i. s-area $\sigma=\sigma(f\cap f')$ of distinct input patterns f and f' in F changes in the mean through the random net $\mathfrak{N}(0, L)$ with arbitrarily small deviations (if n taken large) as follows*: $\sigma_{\lambda+1}=1/4\{1-(1-4\sigma_\lambda)^{2r+1}\}$, $0\leqq\lambda\leqq L-1$, *where* $\sigma_0=\sigma$, *and we have* $\sigma_L\to 1/4$ *as* $L\to\infty$, *the limit not depending on the choice of f and f'.*

How the i. s-area will be changed for various r may be seen by Fig. 3 below.

7. A note on one-to-one correspondence between pattern spaces.

We remark here the obvious fact that there exists one-to-one correspondence between the input space F and the output pattern space G_L of the layer $\mathcal{L}_L$. The random net $\mathfrak{N}(0, L)$, as was noted in section 1, may be considered as a mapping φ: $F\to G_L$. It is our requirement for the random net system that this φ should be one-to-one in order to ascertain the optimal *a posteriori* probability method for recognizing patterns.

Take arbitrary distinct patterns f and f' in F, then $\sigma(f\cap f')\leqq 1/2-1/n<1/2$. The corresponding patterns g_L, g'_L in G_L cannot be identical, i.e. cannot be $\sigma(g_L\cap g'_L)=1/2$, by lemma 2. Hence φ is one-to-one with arbitrarily high probability in the sense that, if n and L are taken large, the "variance problem" encountered in sections 5 and 6 will be solved and σ_L will be sufficiently near 1/4 even if $\sigma(f\cap f')=1/2-1/n$.

LEMMA 3. *The input pattern space F and the corresponding pattern spacc G_L can be made one-to-one with aıbitrarily high probability if n and L are taken large.*

8. Statistical independence evaluation.

We shall show in this section that, when a category is given, then the components constituting patterns in G_L become mutually independent if n and L are taken large. Since the argument, however, is almost same as that given in [1], we shall only outline the main points without a detailed proof.

Suppose that an arbitrary category is given, i.e. a probability distribution $P_n\in\mathcal{P}_n=\{P_n^{(1)}, P_n^{(2)}, \cdots, P_n^{(K)}\}$ is given on the input pattern space F. The net symbol $\mathfrak{N}(0, L)$, in this section, will also be meant the set of all possible random nets constructible between $\mathcal{L}_0$ and $\mathcal{L}_L$. The set $\mathfrak{N}(0, L)$ is called the net space, which has a natural probability measure induced by the random connection operation. Then the product space $F\otimes\mathfrak{N}(0, L)$ has the product probability measure of two measures: P_n on F and the above measure on $\mathfrak{N}(0, L)$. Thus the output value $x(=0, 1)$ of any basic organ in $\mathcal{L}_L$ is a random variable defined on this product space.

Now take s $(2\leqq s\leqq N_L)$ basic organs from $\mathcal{L}_L$, and denote their output random variables by $x_1, x_2, \cdots, x_s$. We denote, furthermore, by E the expectation over the pattern space F for a fixed net, and by $\mathcal{E}$ that over the net space $\mathfrak{N}(0, L)$ for a fixed pattern in F. Then we can prove that

$$\mathcal{E}[|U^{(s)}_{1,2,\cdots,s}(n, L)|]\leqq\sqrt{2(\sigma^s_{M,L}-\sigma^s_{m,L})\zeta_n},$$

where

$$U^{(s)}_{1,2,\cdots,s}(n, L)=E[x_1x_2\cdots x_s]-E[x_1]E[x_2]\cdots E[x_s],$$

$$\sigma_{M,L}=\max_{f\neq f'}\sigma_L(f\cap f'),$$

$$\sigma_{m,L}=\min_{f\neq f'}\sigma_L(f\cap f'),$$

and

$$\zeta_n=\sum_f P^2_n(f).$$

By the result of lemma 2 we have rapid convergences: $\sigma_{M,L}\to 1/4$, $\sigma_{m,L}\to 1/4$, when L becomes large, and by the assumption on the pattern space F (see section 2), we have also a rapid convergence $\zeta_n\to 0$, when n becomes large. Thus $\mathcal{E}[|U^{(s)}_{1,2,\cdots,s}(n, L)|]\to 0$, $n\to\infty$, $L\to\infty$.

Now the measure P_n defines a joint probability distribution $p_{n,L}(x_1=\varepsilon_1, \cdots, x_s=\varepsilon_s)$, $\varepsilon_i=0, 1$ and s-th product distribution $p^{(1)}_{n,L}(x_1=\varepsilon_1)\cdots p^{(s)}_{n,L}(x_s=\varepsilon_s)$, $\varepsilon_i=0, 1$. Denote by $\varphi_{n,L}(t_1, \cdots, t_s)$, $\psi_{n,L}(t_1, \cdots, t_s)$ the s-dimensional characteristic functions for these two distributions, respectively. Then we can show that

$$\mathcal{E}|\varphi_{n,L}(t_1, \cdots, t_s)-\psi_{n,L}(t_1, \cdots, t_s)|<\alpha_{n,L}\cdot e^T$$

for arbitrary but fixed bounded domain $|t_1|+\cdots+|t_s|\leqq T$, where

$$\alpha_{n,L}=\max_{1\leqq i_1,\cdots,i_\nu\leqq s}\mathcal{E}[|U^{(\nu)}_{i_1,\cdots,i_\nu}(n, L)|].$$

Since $\alpha_{n,L}\to 0$ as $n, L\to\infty$, we have

$$\mathcal{E}|\varphi_{n,L}(t_1, \cdots, t_s)-\psi_{n,L}(t_1, \cdots, t_s)|\to 0, \qquad n, L\to\infty,$$

which is valid for arbitrary s and arbitrary P_n in $\mathcal{P}_n$.

LEMMA 4. *For a given category the output component signals constituting output patterns in G_L become mutually independent as n and L are taken large.*

We may summarize all the preceding results in

THEOREM. *The random net with basic organs each realizing a parity function Φ_{2r+1}, if n the number of input points and L the number of layers are taken large, becomes pattern recognizable in the sense of section 1.*

9. On number of basic organs.

Since we in a first step like the present study are mainly concerned with an approximate size of the random net, we shall calculate very rough number of basic organs needed in the net.

From the results obtained so far, it is now easy to see an approximate relation among three variables, n, r, and L, in the pattern recognizable random net. Suppose $(h_1/n)\leqq\sigma(f\cap f')\leqq 1/2-h_2/n$ for any pairs $f\neq f'$ in F, and put $h=\min\{h_1, h_2\}$. Then $1\leqq h$. From the recursion formula for i.s-area in section 9, we have $1-4\sigma_{\lambda+1} =(1-4\sigma_\lambda)^{2r+1}$. Hence $\tau_{\lambda+1}=\tau_\lambda^{2r+1}$, $\tau_\lambda=1-4\sigma_\lambda$. Put $\tau_0=1-4\sigma_0=1-4h/n$, then $\tau_L=\tau_0^{(2r+1)^L}\approx e^{-4h(2r+1)^L/n}$. Assume here that h is of order such that $e^{-4h}\approx 0$, which is quite plausible when n is large. For the random net to become pattern recognizable in the sense of the Theorem given earlier, it is, therefore, sufficient (at least practically) to take $(2r+1)^L/n=1$. Hence

$$L\approx\frac{\log n}{\log(2r+1)}=\log_{(2r+1)}n.$$

On the other hand it might be possible to take $n=N_1=\cdots=N_L$ if n is sufficiently large. But in practical situations it seems, perhaps, to be reasonable to take $cn=N_1=\cdots=N_L$, where c is greater than 1 but not necessary large, since $d_L\leqq 1/2\sqrt{N_1(1-[\alpha(r)]^2)}$ in lemma 1.

If we admit the above estimation, then the number of basic organs may be approximately $cnL\approx cn\log n/\log(2r+1)=cn\log_{(2r+1)}n$. As was noted in section 3, the number of threshold elements (or formal neurons) needed to realize a basic organ may be of order $c'\log(2r+1)$, $1<c'<2$. Then the total number of threshold elements in the random net is approximately of order

$$c''n\log n,$$

where c'' is, for example, $1\ll c''<10$. It is interesting to note that this number does not depend on r and L.

For instance take $n=2\times10^3$, $r=7$ (see Fig. 3), and $c''=5$, then $L=3$, $c''n\log n \approx 10^5$. If $n=10^4$, $L=4$ and $c''n\log n\approx 5\times10^5$.

10. Random net as a lossless channel.

We conclude the present study with one remark.

The random net $\mathfrak{N}(0, L)$ may be regarded as an information transmitting channel with sending signal space F and the receiving signal space G_L. This channel characterized by φ: $F\to G_L$ has been assumed in the present work to be deterministic. But there certainly exists the possibility that the channel will be noisy, due to, for example, malfunctioning of basic organs contained in the net. In such a situation, errors in recognition might become not to be neglected. There arises, therefore, the reliability problem in this sense. On one hand even in the case that φ is probabilistic, if the channel is lossless, then the recognition ability is never weakened [1]. Thus the reliability guaranteeing the losslessness might be one of the basic problems in the pattern recognition by random net.

The author is very much indebted to Prof. K. Kunisawa in preparing this paper.

REFERENCES

[1] HORIBE, Y., Pattern recognition by random net. Kōdai Math. Sem. Rep. **20** (1968), 355–373.

[2] LEWIS, P. M., AND C. L. COATES, Threshold logic. John Wiley & Sons, Inc. (1967).

DEPARTMENT OF MATHEMATICS,
TOKYO INSTITUTE OF TECHNOLOGY.