2004 International Symposium on Nonlinear
Theory and its Applications (NOLTA2004)
Fukuoka, Japan, Nov. 29 - Dec. 3, 2004

The Associative Ability of Homogeneous Neural Networks

Noritaka Shigei†, Hiromi Miyajima† and Shuji Yatsuki†

†Department of Electrical and Electronics Engineering, Kagoshima University
1-21-40, Korimoto, Kagoshima 890-0065 Japan
Email: shigei@eee.kagoshima-u.ac.jp

Abstract—This paper proposes homogeneous neural networks (HNNs). They are new associative memory systems that realize shift-invariant properties, that is, they can associate not only the memorized pattern but also its shifted ones. The transition property of HNNs is analyzed by the statistical method. We show that autocorrelation model of HNNs cannot memorize over the number, $\frac{\binom{m}{k}}{2m\log m}$, of patterns, where m is the number of neurons and k is the order of connections.

1. Introduction

Many studies have been done with neural networks, which may be mutual connected, multi-layered and self-organized models. The models have been applied to various applications such as pattern recognition, associative memory and combinatorial optimization problems[1]-[3]. Associative memory is one of the well-studied fields of neural networks. Many associative memory models processing static and sequential patterns have been studied such as autocorrelation associative memory [4]-[17]. Specifically, higher order neural networks are known as a generalized model whose potential is represented as the weighted sum of products for input variables and are more effective in associative memory than the conventional model. However, when the pattern is memorized in the conventional associative memory models, its shifted patterns are not recalled. In order to perform it, neural networks with homogeneous structure are desired to propose like cellular automata[18].

In this paper, we propose homogeneous neural networks (HNNs), that is, each neuron of them has the identical weights. It is shown that HNNs are possible to perform associative memory, but the condition that $k \geq 2$ is needed. Further, we show that autocorrelation HNNs cannot memorize over the number, $\frac{\binom{m}{k}}{2m\log m}$, of patterns.

2. Higher order NNs and associative memory

Each neuron has m inputs. The state of the i-th neuron, the output, is represented by a function f of a potential u_i. The potential u_i increases in proportion to the weighted sum $\sum_{[L_k]} w_{i,[L_k]} x_{l_1} \cdots x_{l_k}$ of all combinations of products of k pieces of input variables x_j, $1 \leq j \leq m$ and a threshold θ_i, where $[L_k] = l_1 \cdots l_k$ and $\sum_{[L_k]}$ is defined as $\sum_{l_1=1}^{m-k+1} \sum_{l_2=l_1+1}^{m-k+2} \cdots \sum_{l_k=l_{k-1}+1}^{m}$ to exclude the overlapping of variables. Then, the output of the i-th neuron, z_i, is determined by $u_i = \sum_{k=1}^{r} \sum_{[L_k]} w_{i,[L_k]} x_{l_1} \cdots x_{l_k} - \theta_i$ and $z_i = f(u_i)$. The neural element is called a higher order neuron with the order r. In this paper, we will consider higher order neurons with only the order k. And threshold values θ_i are taken to 0. Then, the potential u_i is as follows: $u_i = \sum_{[L_k]} w_{i,[L_k]} x_{l_1} \cdots x_{l_k}$. When $k = 1$, the potential is represented as follows: $u_i = \sum_{l_1=1}^{m} w_{i,l_1} x_{l_1}$. P pairs of memory patterns $\{X^{(s)}, Z^{(s)}\}$ for $s = 1, \cdots, P$ are memorized in the networks, where $X^{(s)}$ and $Z^{(s)}$ are m and n dimensional vectors, respectively, as follows: $X^{(s)} = (x_1^{(s)}, \cdots, x_m^{(s)})^T$, $Z^{(s)} = (z_1^{(s)}, \cdots, z_n^{(s)})^T$, where T represents the transposition of a vector and each element takes 1 or -1.

We consider a two-layered network consisting of the input layer and the output layer as shown in Fig.1. In the conventional neural networks, the weight w_{ij} is determined by the correlation learning as follows[14]:

$$w_{i,[L_k]} = \frac{1}{\binom{m}{k}} \sum_{s=1}^{P} z_i^{(s)} x_{l_1}^{(s)} \cdots x_{l_k}^{(s)} \tag{1}$$

When the memorized pattern $X^{(s)}$ or one with the noise $X^{(s)\prime}$ is input in this case, the corresponding pattern $Z^{(s)}$ is output (See Fig.2). However, when the shifted pattern shift$(X^{(s)}, \alpha)$ of $X^{(s)}$ is input, the pattern $Z^{(s)}$ is not recalled, where shift$(X^{(s)}, \alpha)$ = $(x^{(s)}_{(1+\alpha) \bmod m+1}, x^{(s)}_{(2+\alpha) \bmod m+1}, \cdots, x^{(s)}_{(m+\alpha) \bmod m+1})^T$ for $\alpha \in \{0, \cdots, m-2\}$. Therefore, we propose a new learning method that when the memorized $X^{(s)}$ or its shifted patterns is input, the desired pattern $Z^{(s)}$ is recalled (Fig.3).

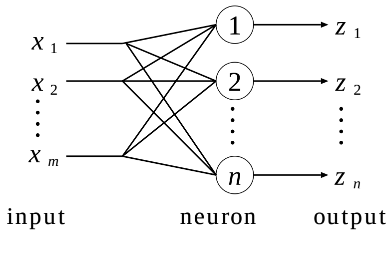


Figure 1: A two-layered neural network.

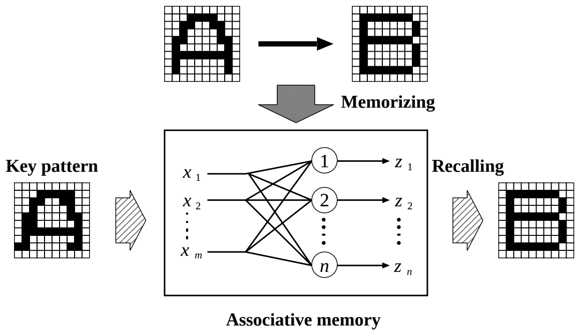
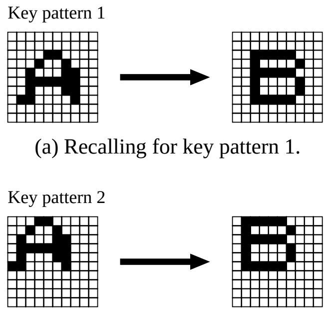


Figure 2: Associative memory.



(a) Recalling for key pattern 1.

(b) Recalling for key pattern 2.

Figure 3: The recalling in the proposed model.

First, the weight w is defined as follows:

$$w_{[\Gamma_k]} = \frac{1}{\binom{m}{k}} \sum_{j=1}^{m} \sum_{s=1}^{P} z_j^{(s)} x_{j+\gamma_1}^{(s)} \cdots x_{j+\gamma_k}^{(s)}, \tag{2}$$

where $[\Gamma_k] = \gamma_1 \cdots \gamma_k$ and $\gamma_i \in \{1, \cdots, n\}$. The neural networks with the weights defined by Eq.(2) are called homogeneous neural networks (HNNs).

Example 1.

For $k = 1, 2$, the following weights are obtained:

$$\begin{aligned} w_{[\Gamma_1]} &= w_\alpha = \tfrac{1}{m} \textstyle\sum_j \sum_s z_j^{(s)} x_{j+\alpha}^{(s)} \\ w_{[\Gamma_2]} &= w_{\alpha,\beta} = \tfrac{1}{\binom{m}{2}} \textstyle\sum_j \sum_s z_j^{(s)} x_{j+\alpha}^{(s)} x_{j+\beta}^{(s)}, \end{aligned}$$

where $\alpha, \beta \in \{1, \cdots, n\}$. □

Remark that the Eq.(2) does not include the suffix i as compared with the Eq.(1). It means that the weights are homogeneous and do not depend on the places of the neurons. When the r-th pattern $X^{(r)}$ is input, the transition property is defined as follows:

$$u_i = \sum_{[\Gamma_k]} \left(\frac{1}{\binom{m}{k}} \sum_j \sum_s z_j^{(s)} x_{j+\gamma_1}^{(s)} \cdots x_{j+\gamma_k}^{(s)} \right) \cdot x_{i+\gamma_1}^{(r)} \cdots x_{i+\gamma_k}^{(r)} \tag{3}$$

$$z_i = f(u_i) \tag{4}$$

$$f(u) = \mathrm{sgn}(u) = \begin{cases} 1 & u > 0 \\ -1 & u \leq 0, \end{cases} \tag{5}$$

where $\sum_{[\Gamma_k]}$ is defined as $\sum_{\gamma_1=1}^{m-k} \sum_{\gamma_2=\gamma_1+1}^{m-k+1} \cdots \sum_{\gamma_k=\gamma_{k-1}+1}^{m-1}$.

3. The associative ability of HNNs

This chapter describes the abilities of heteroassociative ($Z^{(s)} \neq X^{(s)}$) and autoassociative ($Z^{(s)} = X^{(s)}$) memories for HNNs. In the following, we will evaluate the probability that each neuron outputs correctly, in other words, the output of the i-th neuron is $z_i^{(r)}$ for an input pattern $X^{(r)}$. Let $r = 1$ without loss of generality. It is assumed that each element of memory patterns, $x_j^{(s)}$ and $z_i^{(s)}$, takes the value 1 or -1 with a probability $\frac{1}{2}$, independently to each other. Further, it is assumed that m and P are sufficiently large.

3.1. The ability of heteroassociative memory

The potential u_i for an input pattern $X^{(1)}$ is given by

$$u_i = \sum_{[\Gamma_k]} w_{[\Gamma_k]} x_{i+\gamma_1}^{(1)} \cdots x_{i+\gamma_k}^{(1)} = z_i^{(1)} + \frac{\sum\sum_{j\neq i \text{ or } s\neq 1} z_i^{(s)} \left(\sum_{[\Gamma_k]} x_{j+\gamma_1}^{(s)} x_{i+\gamma_1}^{(1)} \cdots x_{j+\gamma_k}^{(s)} x_{i+\gamma_k}^{(1)} \right)}{\binom{m}{k}}. \tag{6}$$

Let h and h_1 be defined as follows:

$$h_1 = \sum_{[\Gamma_k]} x_{j+\gamma_1}^{(s)} x_{i+\gamma_1}^{(1)} \cdots x_{j+\gamma_k}^{(s)} x_{i+\gamma_k}^{(1)}, \tag{7}$$

$$h = \frac{\sum\sum_{j\neq i \text{ or } s\neq 1} z_i^{(s)} h_1}{\binom{m}{k}}. \tag{8}$$

The term h is called the interference one. Let the expectation of h_1 be denoted by $\mathrm{E}[h_1]$. Then, we have

$$\begin{aligned} \mathrm{E}[h_1] &= \mathrm{E}[\sum_{[\Gamma_k]} x_{j+\gamma_1}^{(s)} x_{i+\gamma_1}^{(1)} \cdots x_{j+\gamma_k}^{(s)} x_{i+\gamma_k}^{(1)}] \\ &= \sum_{[\Gamma_k]} \mathrm{E}[x_{j+\gamma_1}^{(s)} x_{i+\gamma_1}^{(1)} \cdots x_{j+\gamma_k}^{(s)} x_{i+\gamma_k}^{(1)}] \\ &= 0, \end{aligned} \tag{9}$$

because of $\mathrm{Prob}(x_j^{(s)} = \pm 1) = \frac{1}{2}$. Let the variance of h_1 be denoted by $\mathrm{Var}[h_1]$. We have

$$\begin{aligned} \mathrm{Var}[h_1] &= \mathrm{E}\left[\left(\sum_{[\Gamma_k]} x_{j+\gamma_1}^{(s)} x_{i+\gamma_1}^{(1)} \cdots x_{j+\gamma_k}^{(s)} x_{i+\gamma_k}^{(1)} \right)^2 \right] \\ &= \binom{m}{k} + \sum_{[\Gamma_k]} \sum_{[\Gamma_k' \neq \Gamma_k]} \mathrm{E}[(x_{j+\gamma_1}^{(s)} x_{i+\gamma_1}^{(1)} \cdots x_{j+\gamma_k}^{(s)} x_{i+\gamma_k}^{(1)}) \\ &\quad \cdot (x_{j+\gamma_1'}^{(s)} x_{i+\gamma_1'}^{(1)} \cdots x_{j+\gamma_k'}^{(s)} x_{i+\gamma_k'}^{(1)})] = \binom{m}{k}, \end{aligned} \tag{10}$$

where $\sum_{[\Gamma_k]} \sum_{[\Gamma_k' \neq \Gamma_k]}$ means the sum for all combinations of $\gamma_1, \cdots, \gamma_k$ and $\gamma_1', \cdots, \gamma_k'$ such that $\gamma_\alpha' \neq \gamma_\alpha$ for at least one α. In the Eq.(8), $z_i^{(s)}$ is 1 or -1 independently of h_1.

Table 1: Associative probabilities by HNNs with the order k: $m = 100, P = 1505$.

k	2	3	4	5
$\frac{mP}{\binom{m}{k}}$	30.4	0.931	0.038	0.002
Prob	0.572	0.85	≈ 1	1

Table 2: Capacity of associative memory by HNNs with the order k: Prob$> 0.99, m = 100$.

k	2	3	4	5
P	9	292	7,059	135,520

Hence, by the Central Limit Theorem, h is regarded as the normal distribution with mean 0 and variance $\frac{m(P-1)}{\binom{m}{k}}$ or $\frac{(m-1)P}{\binom{m}{k}} \approx \frac{mP}{\binom{m}{k}}$. When $z_i^{(1)} = 1$, the output of the i-th neuron is correct for $u_i > 0$, and when $z_i^{(1)} = -1$, the output is correct for $u_i \leq 0$. Therefore, the probability that each neuron outputs correctly, is given by

$$\begin{aligned} \mathrm{Prob}(z_i = z_i^{(1)}) &= \mathrm{Prob}(h \leq 1) = \mathrm{Prob}(h > -1) \\ &= \Phi\left(\sqrt{\frac{\binom{m}{k}}{mP}}\right), \end{aligned} \quad (11)$$

where $\Phi(u)$ is the error integral function defined by $\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} \exp(-\frac{s^2}{2}) ds$. By the Eq.(11), the number of memory patterns memorized, is in proportion to the total number of weights. The Fig.4 shows theoretical and experimental results. In the case of heteroassociative memory, the results in numerical simulations are in fairly general agreement with the theoretical ones. If $k = 1$, then $\mathrm{Prob}(z_i = z_i^{(1)}) = \Phi(\sqrt{\frac{1}{P}})$. It means that it is impossible to perform associative memory by using HNNs for $k = 1$.

Let us show some properties using the Eq.(11).

Example 2.

In the case where $m = 100$ and $P = 1505$, the probabilities of the Eq.(11) are shown in Table 1. □

Example 3.

The numbers of memory patterns satisfying $\mathrm{Prob}(z_i = z_i^{(1)}) > 0.99$ at $m = 100$, are shown in Table 2. □

3.2. Autoassociative ability

In this section, we will consider autoassociative memory, the case where $Z^{(s)} = X^{(s)}$. Let us evaluate $\mathrm{Prob}(z_i = x_i^{(1)})$ in the following. The weight $w_{[\Gamma_k]}$ is determined by

$$w_{[\Gamma_k]} = \frac{1}{\binom{m}{k}} \sum_{j=1}^{m} \sum_{s=1}^{P} x_j^{(s)} x_{j+\gamma_1}^{(s)} \cdots x_{j+\gamma_k}^{(s)}. \quad (12)$$

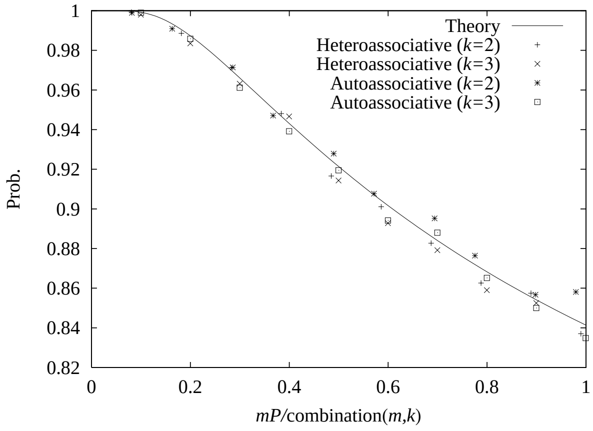


Figure 4: Associative probabilities by HNNs with the order k (theoretical and experimental results).

Table 3: Memory capacity of HNNs.

		k		
		2	3	4
m	100	5	175	4,257
	500	20	3,332	414,029
	1,000	36	12,027	2,997,871

The potential u_i for an input pattern $X^{(1)}$ is given by

$$\begin{aligned} u_i = & \sum_{[\Gamma_k]} w_{[\Gamma_k]} x_{i+\gamma_1}^{(1)} \cdots x_{i+\gamma_k}^{(1)} = x_i^{(1)} + \\ & \frac{x_i^{(1)}}{\binom{m}{k}} \sum_{j \neq i} \sum_{[\Gamma_k]} x_j^{(1)} x_{j+\gamma_1}^{(1)} \cdots x_{j+\gamma_{k-1}}^{(1)} x_{i+\gamma_1}^{(1)} \cdots x_{i+\gamma_k}^{(1)} + \\ & \frac{1}{\binom{m}{k}} \sum_{s \neq 1} \sum_{j} \sum_{[\Gamma_k]} x_j^{(s)} x_{j+\gamma_1}^{(s)} \cdots x_{j+\gamma_k}^{(s)} x_{i+\gamma_1}^{(1)} \cdots x_{i+\gamma_k}^{(1)}, \end{aligned} \quad (13)$$

Using the same method as 3.1, we can calculate the expectations and the variances of the second and third terms, which are regarded as the normal distribution. Further, we can calculate the covariance of the second and third terms. By their results, the interference term is regarded as the normal distribution.

Therefore, we have

$$\mathrm{Prob}(z_i = x_i^{(1)}) = \Phi\left(\sqrt{\frac{\binom{m}{k}}{mP}}\right). \quad (14)$$

As a result, it is shown that the result of autoassociative memory is the same as one of heteroassociative memory. Fig.4 shows theoretical and experimental results. The results in numerical simulations are in fairly general agreement with the theoretical ones.

Autoassociative memory is used as dynamical systems. There is some definition of memory capacity of autoassociative memory. In this paper, we define the memory capacity P_c as the maximum number of memory patterns that each memory pattern is an equilibrium state in the network with over a probability p_e which is close to 1. In other words, memory capacity P_c is defined as the maximum number of memory patterns P which satisfies the following equation:

$$\left(\mathrm{Prob}(z_i = x_i^{(1)})\right)^m \geq p_e \tag{15}$$

Eq.(15) can be approximated as

$$\mathrm{Prob}(z_i = x_i^{(1)}) \geq 1 - \frac{1 - p_e}{m}. \tag{16}$$

So, we have

$$\Phi\left(\sqrt{\frac{\binom{m}{k}}{mP_c}}\right) = 1 - \frac{1 - p_e}{m}. \tag{17}$$

The left hand of Eq.(17) can be approximated as

$$\Phi\left(\sqrt{\frac{\binom{m}{k}}{mP_c}}\right) = 1 - \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\binom{m}{k}}{2mP_c}\right) / \frac{\binom{m}{k}}{2mP_c}, \tag{18}$$

because, when $m \to \infty$, then $\frac{mP_c}{\binom{m}{k}} \to 0$. By using assumption of m, which is sufficiently large, Eq.(17) can be expanded as follows:

$$\frac{\binom{m}{k}}{2mP_c} \approx \log m. \tag{19}$$

Therefore, HNNs cannot memorized over the following P_c:

$$P_c = \frac{\binom{m}{k}}{2m \log m}. \tag{20}$$

The result shows the memory capacity of autoassociative memory for HNNs.

Example 4.

Let us compute P_c for m and k. The results are shown in Table 3. □

4. Conclusion

This paper proposed homogeneous neural networks and analyzed the associative ability of them. HNNs were possible to perform associative memory, but the condition that $k \geq 2$ was needed. Further, the transition property of HNNs was analyzed. It was shown that the memory capacity of autocorrelation model is $\frac{\binom{m}{k}}{2m \log m}$.

References

[1] J. Hertz, A. Krogh and R. G. Palmer, Introduction to the Theory of Neural Computation, Perseus Books Publishing, 1991.

[2] N.K. Kasabov, Foundations of Neural Networks, Fuzzy Systems and Knowledge Engineering, The MIT Press, 1996.

[3] K. Mehrotra, C.K. Mohan and S. Ranka, Elements of Artificial Neural Networks, The MIT Press, 1997.

[4] S. Amari and K. Maginu, "Statistical Neurodynamics of Associative Memory," Neural Networks, Vol.1, pp.63–73, 1988.

[5] S. Amari, "Mathematical Foundations of Neurocomputing," Proceedings of the IEEE, Vol.79, No.9, pp.1443–1463, 1990.

[6] R. Meir and E. Domany, "Exact Solution of a Layered Neural Network Memory," Physical Rev. Letter, Vol.59, pp.359–362, 1987.

[7] M. Okada, "Notions of Associative Memory and Sparse Coding," Neural Networks, Vol.9, No.8, pp.1429–1458, 1996.

[8] M. Oda and H. Miyajima, "Autoassociative Memory Using Refractory Period of Neurons and Its On-line Learning," IEEE Proc. of the ECS, pp.623–626, 2001.

[9] D. Amit, H. Gutfreund and H. Sompolinsky, "Storing Infinite Numbers of Patterns in a Spin-Glass Model of Neural Networks," Physical Review Letters, Vol.55, No.14, pp.1530–1533, 1985.

[10] T. Kohonen, "Correlation Matrix Memories," IEEE Trans. on Computers, C-21, pp.353-359, 1972.

[11] B. Kosko, "Adaptive Bilateral Associative Memory," Appl. Optics, Vol.26, pp.4947–4960, 1987.

[12] S. Yoshizawa, M. Morita and S. Amari, "Capacity of Associative Memory using a Nonmonotonic Neuron Model," Neural Networks, Vol.6, pp.167–176, 1993.

[13] M. Morita, "Memory and Learning of Sequential Patterns by Nonmonotone Neural Networks," Neural Networks, Vol.9, No.8, pp.1447–1489, 1996.

[14] S. Yatsuki and H. Miyajima, "Associative Ability of Higher Order Neural Networks," in Proc. ICNN'97, Vol.2, pp.1299-1304, 1997.

[15] S. Yatsuki and H. Miyajima, "Statistical Dynamics of Associative Memory for Higher Order Neural Networks," IEEE Proc. ISCAS 2000, Vol.3, pp.670–673, 2000.

[16] L.F. Abbott and Y. Arian, "Storage Capacity of Generalized Networks," Physical Review A, Vol.36, No.10, pp.5091–5094, 1987.

[17] H. Miyajima, N. Shigei and Y. Hamakawa, "Higher Order Differential Correlation Associative Memory of Sequential Patterns," Proc. IJCNN 2004, 2004 (in print).

[18] S. Wolfram, "Universality and Complexity in Cellular Automata," Physica D, 10, pp.1–35, 1984.