

Article

On the Entropy Based Associative Memory Model with Higher-Order Correlations

Masahiro Nakagawa

Department of Electrical Engineering, Faculty of Engineering, Nagaoka University of Technology, Kamitomioka 1603-1, Nagaoka, Niigata 940-21, Japan; E-Mail: masanaka@vos.nagaokaut.ac.jp

Received: 2 January 2010 / Accepted: 18 January 2010 / Published: 22 January 2010

Abstract: In this paper, an entropy based associative memory model will be proposed and applied to memory retrievals with an orthogonal learning model so as to compare with the conventional model based on the quadratic Lyapunov functional to be minimized during the retrieval process. In the present approach, the updating dynamics will be constructed on the basis of the entropy minimization strategy which may be reduced asymptotically to the above-mentioned conventional dynamics as a special case ignoring the higher-order correlations. According to the introduction of the entropy functional, one may involve higher-order correlation effects between neurons in a self-contained manner without any heuristic coupling coefficients as in the conventional manner. In fact we shall show such higher order coupling tensors are to be uniquely determined in the framework of the entropy based approach. From numerical results, it will be found that the presently proposed novel approach realizes much larger memory capacity than that of the quadratic Lyapunov functional approach, e.g., associatron.

Keywords: association model; entropy; memory capacity; lyapunov functional

1. Introduction

During the past quarter century, a large number of autoassociative models have been extensively investigated on the basis of the autocorrelation dynamics characterized by the quadratic Lyapunov functional to be minimized. Since the proposals of the pioneering retrieval models by Anderson [1], Kohonen [2], and Nakano [3], some works related to such an autoassociation model of the inter-connected neurons through an autocorrelation matrix were theoretically analyzed by Amari [4], Amit *et al.* [5] and Gardner [6]. So far it has been well appreciated that the storage capacity of the au-

tocorrelation model, or the number of stored pattern vectors, L , to be completely associated vs. the number of neurons N , which is called the relative storage capacity or loading rate and denoted as $\alpha_c = L/N$, is evaluated as $\alpha_c \sim 0.14$ at most for the autocorrelation learning model with the activation function as the signum one ($\text{sgn}(x)$ for the abbreviation) [7,8].

In contrast to the above-mentioned models with monotonous activation functions, neuro-dynamics with a nonmonotonous mapping was recently proposed by Morita [9], Yanai and Amari [10], Shiino and Fukai [11]. They clarified that the nonmonotonous mapping in a neuro-dynamics model possesses a remarkable advantage in the storage capacity, $\alpha_c \sim 0.27-0.4$, superior than the conventional association models with monotonous activation functions, e.g., the signum or sigmoidal function. Therefore activation functions have been considered to be worthwhile of investigation, not only the associative memory models but also learning models in relation with chaos dynamics [12].

In the above-mentioned association models, the dynamics have been restricted to the updating rule on the basis of the quadratic form of the Lyapunov functionals to be minimized through the retrieval process. That is, the nonlinearity of the dynamics results from the nonlinear characteristics of the activation function rather than the updating rule of the internal states derived from the quadratic Lyapunov, or energy, functional form.

From the above-mentioned viewpoint, we shall propose a novel approach based on the entropy defined in terms of the overlaps, which are defined by the inner products between the state vector and the embedded vectors. That is, in the present model the functional to be minimized is defined in terms of the entropy instead of the conventional quadratic functionals. Then it will be found that the higher order dynamics is to be involved in a self-contained manner in the present entropy-based approach. In Section 2 a theoretical framework based on the entropy approach will be described to present the relationship between the present proposal and the conventional model with a quadratic Lyapunov functional to be minimized. Some numerical results will be given in Section 3 and then Section 4 will be devoted to concluding remarks.

2. Theory

Let us consider an associative model with the embedded binary vector $e_i^{(r)} = \pm 1$ ($1 \leq i \leq N, 1 \leq r \leq L$), where N and L are the number of neurons and the number of embedded vectors, respectively, to be retrieved. The states of the neural network are to be characterized in terms of the vector s_i ($1 \leq i \leq N$) and the internal states σ_i ($1 \leq i \leq N$) which are related each other in terms of:

$$s_i = f(\sigma_i) \quad (1 \leq i \leq N) \quad , \quad (1)$$

where $f(\bullet)$ is the activation function of the neuron.

Then we introduce the following entropy which is to be related to the overlaps:

$$I = -\frac{1}{2} \sum_{r=1}^L m^{(r)2} \log\{m^{(r)2}\} \quad , \quad (2)$$

where the overlaps $m^{(r)}$ ($r = 1, 2, \dots, L$) are defined by:

$$m^{(r)} = \sum_{i=1}^N e_i^{\dagger(r)} s_i \quad ; \quad (3)$$

where the covariant vector $e_i^{\dagger(r)}$ is defined in terms of the following orthogonal relation:

$$\sum_{i=1}^N e_i^{\dagger(r)} e_i^{(s)} = \delta_{rs} \quad (1 \leq r, s \leq L) \quad , \tag{4}$$

$$e_i^{\dagger(r)} = \sum_{r'=1}^L a_{rr'} e_i^{(r')} \quad , \tag{5}$$

$$a_{rr'} = (\omega^{-1})_{rr'} \quad , \tag{6}$$

$$\omega_{rr'} = \left(\sum_{i=1}^N e_i^{(r)} e_i^{(r')} \right) \quad , \tag{7}$$

where ω^{-1} denotes the inverse matrix of ω .

Then the entropy defined by Equation (2) can be minimized by the following condition:

$$|m^{(r)}| = \delta_{rs} \quad (1 \leq r, s \leq L) \quad , \tag{8}$$

and:

$$\sum_{r=1}^L m^{(r)2} = 1 \quad . \tag{9}$$

That is, regarding $m^{(r)2}$ ($1 \leq r \leq L$) as the probability distribution in Equation (2), a target pattern may be retrieved by minimizing the entropy I with respect to $m^{(r)}$ or the state vector s_i to achieve the retrieval of a target pattern in which the Equation (8) and Equation (9) are to be satisfied. Therefore the entropy function may be considered to be a functional to be minimized during the retrieval process of the auto-association model instead of the conventional quadratic Lyapunov, *i.e.* energy functional, E :

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_{ij} s_i^{\dagger} s_j \quad , \tag{10}$$

where s_i^{\dagger} is the covariant vector defined by:

$$s_i^{\dagger} = \sum_{r=1}^L \sum_{j=1}^N e_i^{\dagger(r)} e_j^{(r)} s_j \quad , \tag{11}$$

and the connection matrix w_{ij} is defined in terms of:

$$w_{ij} = \sum_{r=1}^L e_i^{(r)} e_j^{\dagger(r)} \quad . \tag{12}$$

Substituting Equation (12) into Equation (10), one may readily find:

$$E = -\frac{1}{2} \sum_{r=1}^L m^{(r)2} \quad . \tag{13}$$

According to the steepest descent approach in the discrete time model, the updating rule of the internal states $s_i = f(\sigma_i) = \text{sgn}(\sigma_i)$, may be defined by:

$$\sigma_i(t+1) = -\eta \frac{\partial I}{\partial s_i^{\dagger}} \quad (1 \leq i \leq N) \quad , \tag{14}$$

where $\eta (> 0)$ is a coefficient. Substituting Equation (2) and Equation (3) into Equation (14) and noting the following relation with the aid of Equation (11):

$$m^{(r)} = \sum_{i=1}^N e_i^{\dagger(r)} s_i = \sum_{i=1}^N e_i^{(r)} s_i^{\dagger} \tag{15}$$

one may readily derive the following relation:

$$\begin{aligned} \sigma_i(t+1) &= -\eta \frac{\partial I}{\partial s_i^{\dagger}} = +\frac{1}{2} \eta \frac{\partial}{\partial s_i^{\dagger}} \sum_{r=1}^L m^{(r)2} \log\{m^{(r)2}\} \\ &= \frac{1}{2} \eta \frac{\partial}{\partial s_i^{\dagger}} \sum_{r=1}^L \left\{ \sum_{j=1}^N e_j^{(r)} s_j^{\dagger}(t) \right\}^2 \log \left[\left\{ \sum_{j=1}^N e_j^{(r)} s_j^{\dagger}(t) \right\}^2 \right] \\ &= \eta \sum_{r=1}^L e_i^{(r)} \left\{ \sum_{j=1}^N e_j^{(r)} s_j^{\dagger}(t) \right\} \left[1 + \log \left[\left\{ \sum_{j=1}^N e_j^{(r)} s_j^{\dagger}(t) \right\}^2 \right] \right] \\ &= \eta \sum_{r=1}^L e_i^{(r)} m^{(r)} \left[1 + \log\{m^{(r)2}\} \right] . \end{aligned} \tag{16}$$

Generalizing somewhat the above dynamics, we propose the following dynamic rule for the internal states in order to unify the conventional quadratic dynamics as well as the presently proposed entropy approach as mentioned below:

$$\begin{aligned} \sigma_i(t+1) &= \eta \sum_{r=1}^L e_i^{(r)} \left\{ \sum_{j=1}^N e_j^{\square r} s_j(t) \right\} \left[1 + \log \left\{ (1-\alpha) + \alpha \left\{ \sum_{j=1}^N e_j^r s_j(t) \right\}^2 \right\} \right] \\ &= \eta \sum_{r=1}^L e_i^{(r)} \{m^{(r)}(t)\} \left[1 + \log \left\{ (1-\alpha) + \alpha \{m^{(r)}(t)\}^2 \right\} \right]. \end{aligned} \tag{17}$$

In the above expression α ($0 < \alpha < 1$) is considered to be a control parameter of the present model as follows. First, in the limit of $\alpha \rightarrow 0$, the above dynamics will be reduced to the conventional auto-correlation dynamics:

$$\begin{aligned} \sigma_i(t+1) &= \eta \lim_{\alpha \rightarrow 0} \sum_{r=1}^L e_i^{(r)} \{m^{(r)}(t)\} \left[1 + \log \left\{ (1-\alpha) + \alpha \{m^{(r)}(t)\}^2 \right\} \right] \\ &= \eta \sum_{r=1}^L e_i^{(r)} m^{(r)}(t) = -\eta \sum_{r=1}^L e_i^{(r)} \sum_{j=1}^N e_j^{\dagger(r)} s_j(t) \\ &= \eta \sum_{j=1}^N w_{ij} s_j(t) . \end{aligned} \tag{18}$$

On the other hand, Equation (17) results in Equation (16) in the case of $\alpha \rightarrow 1$. Therefore one may control the dynamics between the autocorrelation ($\alpha \rightarrow 0$) and the entropy based approach ($\alpha \rightarrow 1$) on the basis of the presently proposed generalized approach defined by Equation (17).

Now it seems to be worthwhile to see the higher-order correlation in Equation (17) expanding the right-hand-side of Equation (17) as follows.

$$\begin{aligned}
 \sigma_i(t+1) &= \eta \sum_{r=1}^L e_i^{(r)} m^{(r)}(t) \left[1 + \log \left\{ (1-\alpha) + \alpha \{m^{(r)}(t)\}^2 \right\} \right] \\
 &= \eta \sum_{r=1}^L e_i^{(r)} m^{(r)}(t) \left[1 + \log(1-\alpha) + \log \left\{ 1 + \beta m^{(r)}(t)^2 \right\} \right] \\
 &= \eta \sum_{r=1}^L e_i^{(r)} m^{(r)}(t) \left[\left\{ 1 + \log(1-\alpha) \right\} + \sum_{n=1}^{\infty} \frac{(-1)^{n-1} \beta^n \{m^{(r)}(t)\}^{2n}}{n} \right]
 \end{aligned}
 \tag{19}$$

where β is defined by:

$$\beta = \frac{\alpha}{1-\alpha} . \tag{20}$$

Substituting Equation (3) into Equation (19), one may eventually derive the following up-dating rule for the internal state, *i.e.*:

$$\sigma_i(t+1) = \eta \sum_{n=0}^{\infty} \sum_{j_1=1}^N \sum_{j_2=1}^N \cdots \sum_{j_{2n+1}=1}^N w_{ij_1 \cdots j_{2n+1}}^{(2n+1)} \left\{ \prod_{k=1}^{2n+1} s_{j_k}(t) \right\} , \tag{21}$$

where $w_{ij_1 \cdots j_n}^{(2n+1)}$ ($0 \leq n < \infty$) are the connection weight tensors between neurons involving such higher-order correlations as $n \geq 1$ and are to be expressed by means of $e_i^{(r)}$ and $e_i^{\dagger(r)}$ comparing Equation (19) and Equation (21). Of course the lowest order connection weight $w_{ij_1}^{(1)}$ in Equation (21) corresponds to w_{ij_1} in Equation (12), *i.e.*:

$$\begin{aligned}
 w_{ij_1}^{(1)} &= w_{ij_1} \{1 + \log(1-\alpha)\} \\
 &= \sum_{r=1}^L e_i^{(r)} e_{j_1}^{\dagger(r)} \{1 + \log(1-\alpha)\} .
 \end{aligned}
 \tag{22}$$

Thus the lowest correlation is reduced to the conventional quadratic framework expressed in terms of Equation (10) and Equation (12) as $\alpha \rightarrow 0$. Furthermore, for the higher-order connection tensors appearing in Equation (10c), one may readily obtain the following results:

$$w_{ij_i}^{(2n+1)} = \frac{(-1)^{n-1} \beta^n}{n} \sum_{r=1}^L e_i^{(r)} e_{j_1}^{\dagger(r)} e_{j_2}^{\dagger(r)} \cdots e_{j_{2n+2}}^{\dagger(r)} . \quad (1 \leq n \leq \infty) \tag{23}$$

It should be borne in mind here that all of the connection tensors, *i.e.* $w_{ij_1 \cdots j_n}^{(2n+1)}$ ($0 \leq n < \infty$) are to be uniquely determined in terms of the embedding vectors $e_i^{(r)}$ ($1 \leq i \leq N, 1 \leq r \leq L$) and $e_i^{\dagger(r)}$ ($1 \leq i \leq N, 1 \leq r \leq L$), which may be related to each other according to Equation (5) to Equation (7). Thus the present approach substantially includes the higher-order correlations beyond the conventional approach defined by Equation (11), in which the correlation between neurons is restricted up to the second-order contribution corresponding to the quadratic Lyapunov functional given by Equation (10). For practical association of the stored patterns, the connection tensors $w_{ij_1 \cdots j_n}^{(n)}$ ($1 \leq n < \infty$) defined by Equation (21) have to be utilised instead of the embedded vectors, *i.e.* $e_i^{(r)}$ and $e_i^{\dagger(r)}$ ($1 \leq r \leq L$).

3. Results

The embedded vectors are set to the binary random vectors as follows:

$$e_i^{(r)} = \text{sgn}(z_i^{(r)}) \quad (1 \leq r \leq L) \quad , \quad (24)$$

where $z_i^{(r)}$ ($1 \leq i \leq N$, $1 \leq r \leq L$) are the zero-mean pseudo-random numbers between -1 and +1. For simplicity, the activation function, Equation (1), is set to:

$$s_i = f(\sigma_i) = \text{sgn}(\sigma_i), \quad (25)$$

where $\text{sgn}(\bullet)$ denotes the signum function defined by:

$$\text{sgn}(x) = \begin{cases} -1 & (x < 0) \\ 0 & (x = 0) \\ +1 & (x > 0) \end{cases} \quad (26)$$

The initial vector $s_i(0)$ ($1 \leq i \leq N$) is set to:

$$s_i(0) = \begin{cases} -e_i^{(s)} & (1 \leq i \leq H_d) \\ +e_i^{(s)} & (H_d + 1 \leq i \leq N) \end{cases} \quad , \quad (27)$$

where $e_i^{(s)}$ is a target pattern to be retrieved and H_d is the Hamming distance between the initial vector $s_i(0)$ and the target vector $e_i^{(s)}$. The retrieval is successful if:

$$m^{(s)}(t) = \sum_{i=1}^N e_i^{\dagger(s)} s_i(t) \quad (28)$$

results in 1 for $t \geq 1$, in which the system may be in a steady state such that:

$$s_i(t+1) = s_i(t) \quad , \quad (29)$$

$$\sigma_i(t+1) = \sigma_i(t) \quad . \quad (30)$$

To see the retrieval ability of the present model, the success rate S_r is defined as the rate of the success for 1,000 trials with the different embedded vector sets $e_i^{(r)}$ ($1 \leq i \leq N$, $1 \leq r \leq L$). To control from the autocorrelation dynamics after the initial state ($t \sim 1$) to the entropy based dynamics ($t \sim T_{\max}$), the parameter α in Equation (17) was simply controlled by:

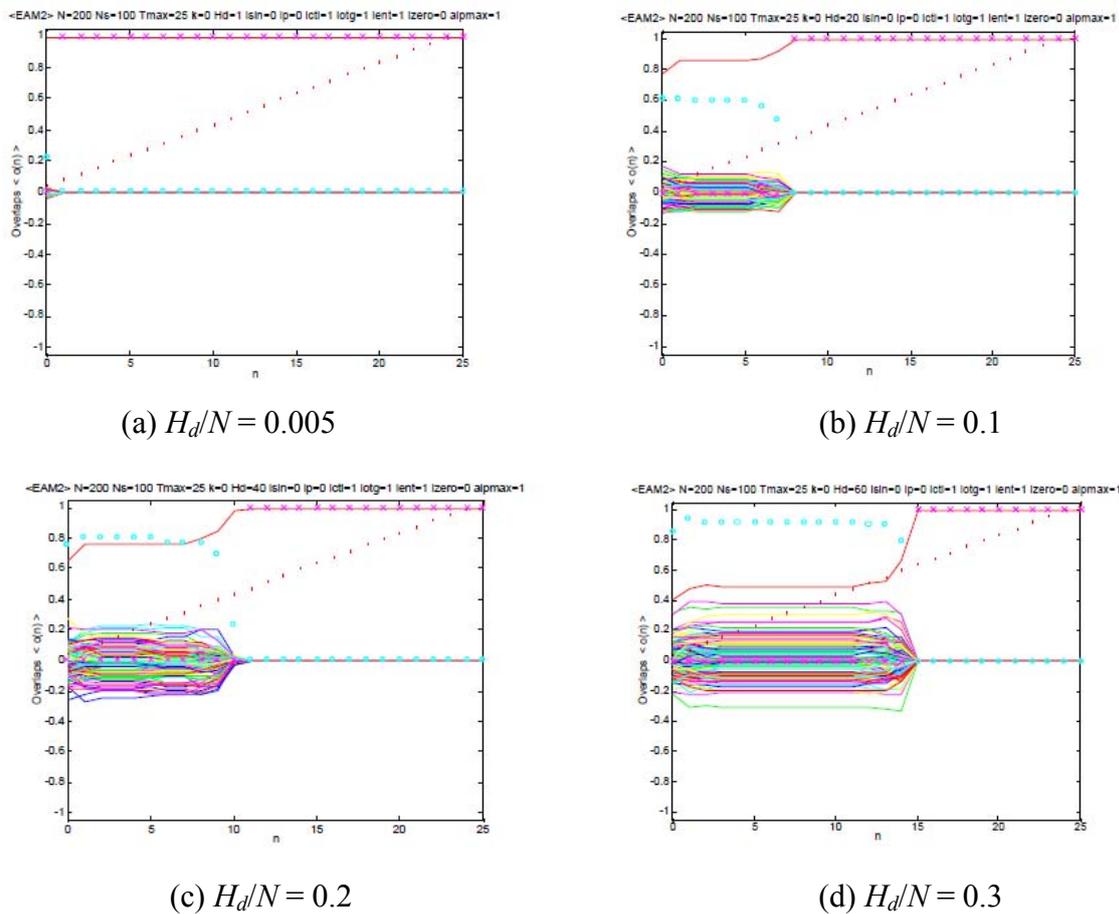
$$\alpha = \frac{t}{T_{\max}} \alpha_{\max} \quad (0 \leq t \leq T_{\max}) \quad (31)$$

where T_{\max} and α_{\max} are the maximum values of the iterations of the updating according to Equation (17) and α , respectively.

Choosing $N = 200$, $\eta = 1$, $T_{\max} = 25$, $L/N = 0.5$ and $\alpha_{\max} = 1$, we first present an example of the dynamics of the overlaps in Figures 1(a)–(d) (entropy based approach) and Figures 2(a)–(d) (associatron), in which the abscissa and the ordinate are for the retrieval steps after the initial states and the overlaps derived from Equation (16), respectively. Therein the cross symbols (\times) and the open circles (o) represent the success of retrievals, in which Equation (8) and Equation (9) are satisfied, and the entropy defined by Equation (2), respectively, for a retrieval process. In addition the time dependence of the parameter α/α_{\max} defined by Equation (31) is depicted as dots (.). In Figures 1(a)–(d) after a transient state,

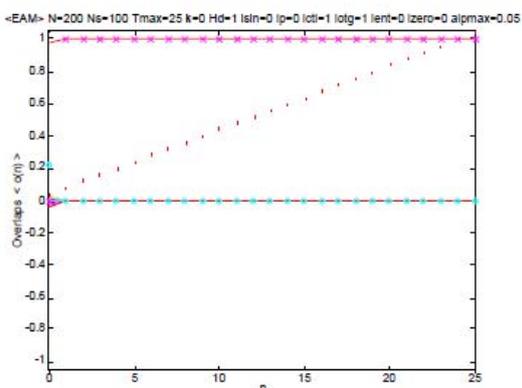
it may be confirmed that the complete association corresponding to the conditions, Equation (8) and Equation (9), can be achieved, even for such a relatively large Hamming distance of the initial vector from a target vector as $H_d/N = 0.1-0.15$. On the other hand, in Figures 2(a)–(d), a trapping in a local minimum is found to be inevitable for $L/N = 0.5$ ($\gg 0.14$ which is the relative storage capacity for the autocorrelation model as discussed by Amari and Maginu [8] (see Concluding remarks), in which Equation (8) and Equation (9) cannot be achieved even for $H_d/N \rightarrow 0$ with $L/N > 0.5$. In addition one may see that the retrieval cannot be achieved beyond $H_d/N = 0.05$ as in Figures 2(c) and (d). From these results one may certainly confirm the advantage of our approach beyond the conventional models based on the quadratic Lyapunov (energy) functionals.

Figure 1. The time dependence of overlaps of the present entropy based model defined by Equation (17).

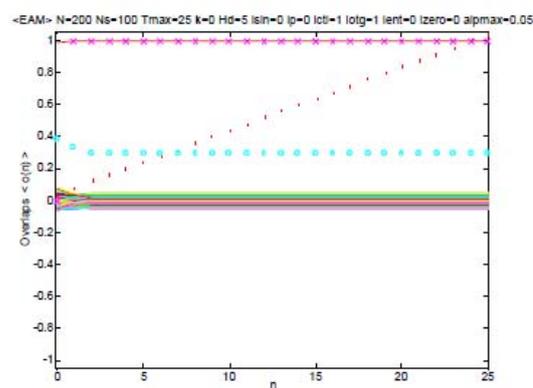


Then we shall present the dependence of the success rate S_r on the loading rate L/N are depicted in Figure 3 for various Hamming distances H_d with $N = 100$. For comparison, the corresponding results of the associatron model with $\alpha \sim 0$, *i.e.* Equation (11), are shown in Figure 4. Comparing between Figures 3 and 4, it is found that the present approach may achieve a relatively larger memory capacity beyond the conventional autocorrelation strategy. Therefore the presently proposed nonlinear dynamics with the higher-order correlations involved in Equation (17) or Equation (21) based on the entropy functional to be minimized has a great advantage for the storage capacity beyond the conventional one based on Equation (10) and Equation (18).

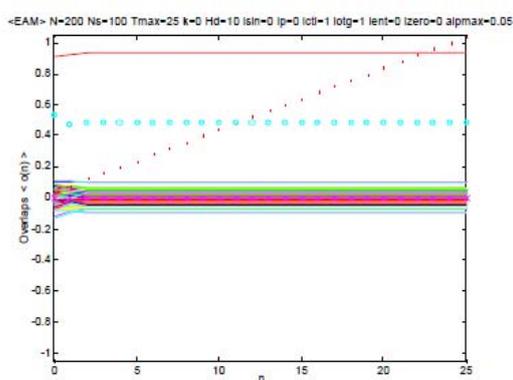
Figure 2. The time dependence of overlaps of the associatron defined by Equation (18).



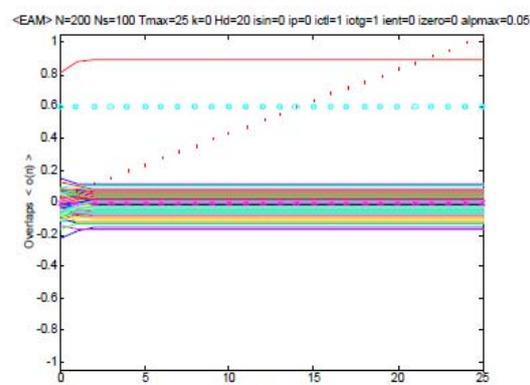
(a) $H_d/N = 0.005$



(b) $H_d/N = 0.025$

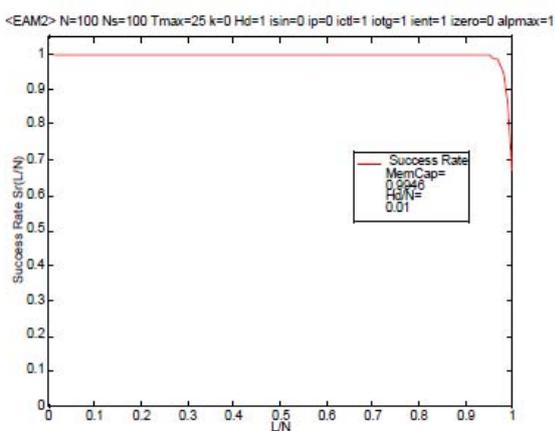


(c) $H_d/N = 0.05$

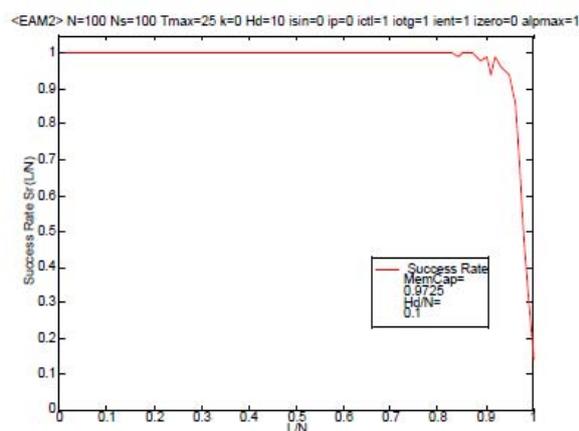


(d) $H_d/N = 0.1$

Figure 3. The dependence of the success rate on the loading rate $\alpha = L/N$ of autoassociation model based on Equation (17) (entropy based approach).



(a) $H_d/N = 0.01$



(b) $H_d/N = 0.1$

Figure 3. Cont.

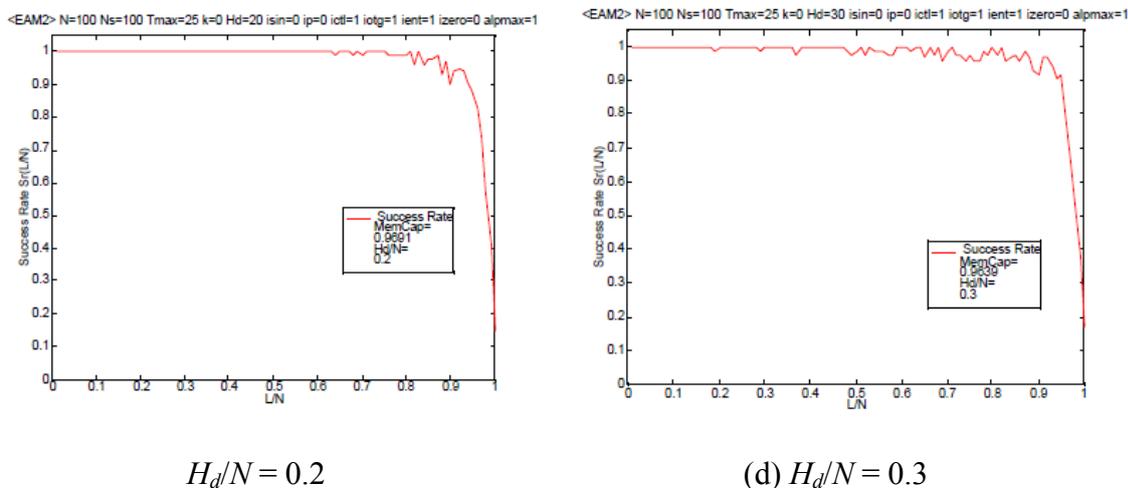
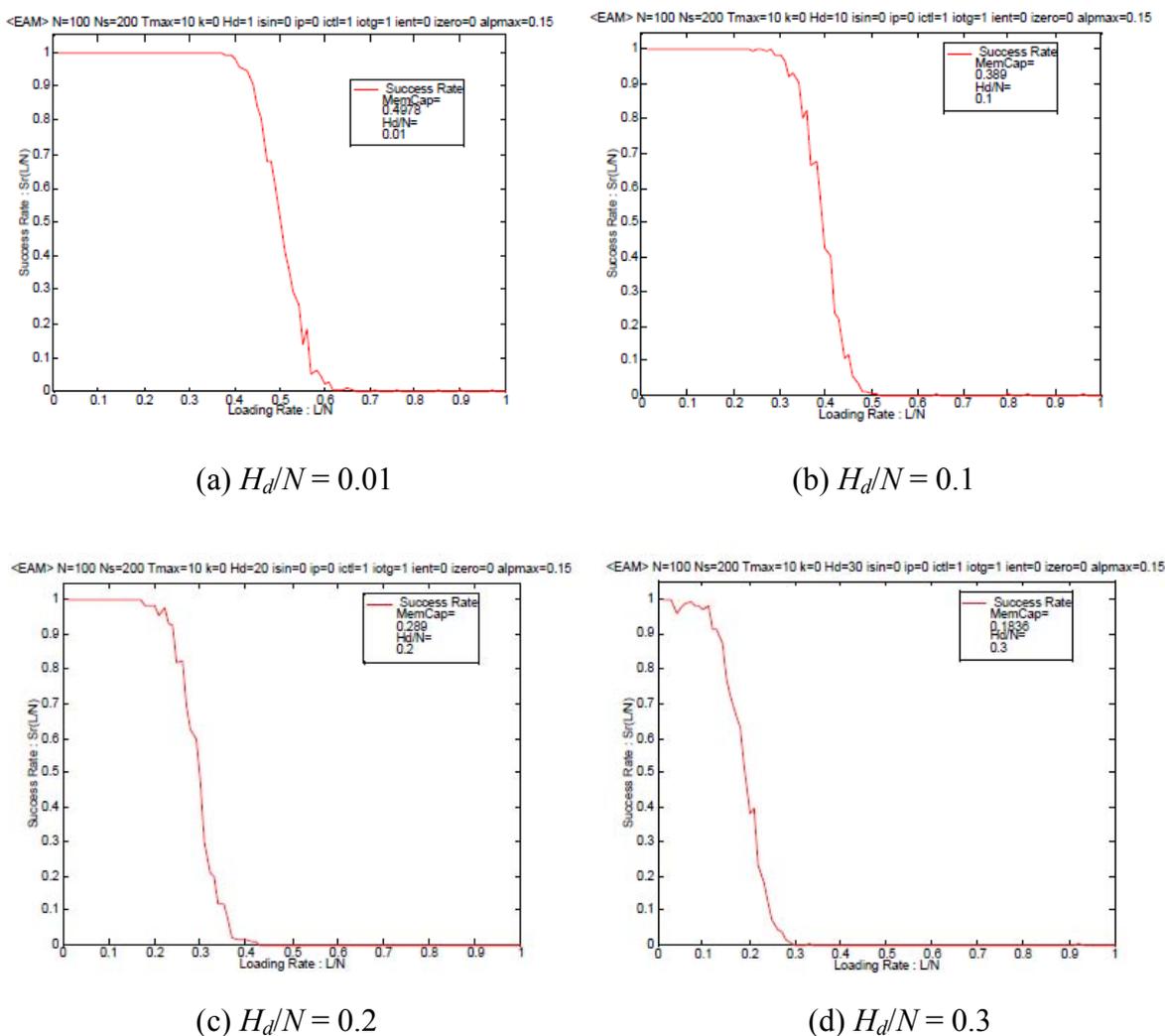


Figure 4. The dependence of the success rate on the loading rate $\alpha = L/N$ of autoassociation model based on Equation (18) (associatron).



The depression of the success rate at $L/N \sim 1$ in Figure 3 may be considered to result from the fact such that:

$$w_{ij} = \sum_{r=1}^L e_i^{(r)} e_j^{\dagger(r)} = \delta_{ij} \quad (L = N) \quad , \quad (32)$$

where Equation (18) reads:

$$\begin{aligned} \sigma_i(t+1) &= \eta \sum_{j=1}^N w_{ij} s_j(t) \\ &= \eta s_i(t) \quad . \end{aligned} \quad (33)$$

Thus, noting that $s_i(t) = \text{sgn}(\sigma_i(t))$ and $\eta > 0$, one has:

$$\begin{aligned} \sigma_i(t+1) &= \eta s_i(t) \\ &= \eta s_i(0) \quad . \quad (t \geq 0) \end{aligned} \quad (34)$$

4. Concluding Remarks

In the present paper, we have proposed an entropy based association model instead of the conventional autocorrelation dynamics. From numerical results, it was found that a large memory capacity may be achieved on the basis of the entropy approach. This advantage of the association property of the present model is considered to result from the fact such that the present dynamics to update the internal state Equation (17) assures that the entropy, Equation (2) is minimized under the conditions, Equation (8) and Equation (9), which corresponds to the successful retrieval of a target pattern.

To conclude this work, we shall show the dependence of the storage capacity, which is defined as the area covered in terms of the success rate curves as shown in Figures 3 and Figure 4, on the Hamming distance in Figure 5. Therein one may see again the great advantage of the present model based on the entropy functional to be minimized beyond the conventional quadratic form. In fact one may realize the considerably larger storage capacity in the present model in comparison with the associatron over $H_d/N \sim 0.5$. The memory retrievals for the associatron become troublesome near $H_d/N = 0.5$ as seen in Figure 5 since the directional cosine between the initial vector and a target pattern eventually vanishes therein. Remarkably, even in such a case, the present model attains a remarkably large memory capacity because of the higher-order correlations involved in Equation (17) or Equation (21), as expected from Figure 3.

As a future problem, it seems to be worthwhile to involve a chaotic dynamics in the present model introducing a periodic activation function such as sinusoidal one and to extend the autocorrelation model replacing $e_i^{\dagger(r)}$ by $e_i^{(r)}/N$ in the present approach, in which the connection matrix w_{ij} and the overlaps $m^{(r)}$ read:

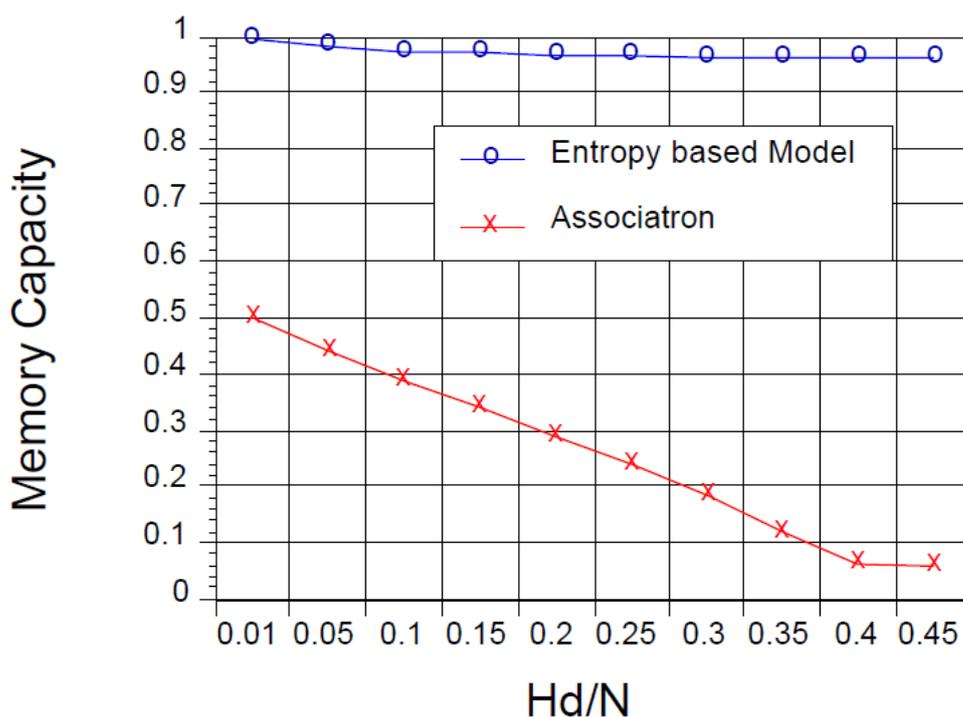
$$w_{ij} = \frac{1}{N} \sum_{r=1}^L e_i^{(r)} e_j^{(r)} \quad , \quad (35)$$

and:

$$m^{(r)}(t) = \frac{1}{N} \sum_{i=1}^N e_i^{(r)} s_i(t) \quad , \quad (36)$$

respectively, corresponding to Equation (12) and Equation (15). The entropy based approach with Equation (20), *i.e.* autocorrelation dynamics, is now in progress in the relation with chaos dynamics [12] and will be reported elsewhere as a separated paper and to be compared with the previous works [13,14] in the near future. Furthermore it seems to be worthwhile to examine the truncation effects of the expansion tensors as in Equation (21), which was not directly derived in our previous work [15], for practical applications related to the hardware implementation.

Figure 5. The dependence of the storage capacity on the Hamming distance. Here symbols o and x are for the entropy based approach and the associatron, respectively.



Acknowledgments

This work was supported in part by the 21st Century COE (Center of Excellence) Program " Global Renaissance by Green Energy Revolution" and the Grant-in-Aid for Science Research (15300070) from the Ministry of Education Culture, Sports, Science and Technology of Japan.

References

1. Anderson, J.A. A simple neural network generating interactive memory. *Math. Biosci.* **1972**, *14*, 197–220.
2. Kohonen, T. Correlation matrix memories. *IEEE Trans. Comput.* **1972**, *C-21*, 353–359.
3. Nakano, K. Associatron—a model of associative memory. *IEEE Trans.* **1972**, *SMC-2*, 381–388.
4. Amari, S. Neural Theory of association and concept formation. *Biol. Cybern.* **1977**, *26*, 175–185.
5. Amit, D.J.; Gutfreund, H.; Sompolinsky, H. Storing infinite numbers of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.* **1985**, *55*, 1530–1533.

6. Gardner, E. Structure of metastable states in the Hopfield model. *J. Phys. A: Math. Gen.* **1986**, *19*, 1047–1052.
7. Kohonen, T.; Ruohonen, M. Representation of associated pairs by matrix operators, *IEEE Trans.* **1973**, *22*, 701–702.
8. Amari, S.; Maginu, K. Statistical Neurodynamics of associative memory. *Neural Networks* **1988**, *1*, 63–73.
9. Morita, M. Associative memory with nonmonotone dynamics. *Neural Networks* **1993**, *6*, 115–126.
10. Yanai, H.-F.; Amari, S. Auto-associative Memory with two-stage dynamics of non-monotonic neurons. *IEEE Trans. Neural Networks* **1996**, *7*, 803–815.
11. Okada, M.; Shiino, M.; Fukai, T. Random and systematic dilutions of synaptic connections in a neural network with a nonmonotonic response functions. *Phys. Rev. E* **1998**, *57*, 2095–2103.
12. Nakagawa, M. *Chaos and Fractals in Engineering*; World Scientific Inc.: Singapore, 1999.
13. Yatsuki, S.; Miyajima, H. Associative ability of higher order neural networks. *Neural Networks* **1997**, *2*, 1299–1304.
14. Gorban, A.N.; Mirkes, Y.M.; Wunsch, D.C. Higher order orthogonal tensor networks: Information capacity and reliability. *Proc. Neural Networks* **1997**, *2*, 1311–1314.
15. Nakagawa, M. Entropy based associative model. *Lect. Notes Comput. Sci.* **2006**, *4232/2006*, 397–406.

© 2010 by the authors; licensee Molecular Diversity Preservation International, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).