ELSEVIER

Physica A 218 (1995) 291-297

PHYSICA A

The stability of the 2^n-element number neural network models

Z.X. Chen, J.W. Shuai, R.T. Liu, B. X. Wu

Department of Physics, Xiamen University, Xiamen 361005, P.R. China

Received 1 January 1995

Abstract

In this paper, we defined a energy function for the 2^n-element number neural network models with $n < 3$. It is proved that the memorized patterns are local minima of the energy function. In the case of random purely sequential dynamics, the energy of the network decreases monotonically with time and so the nonlinear 2^n-element number neural network with $n < 3$ must end up in a state of equilibrium, which is stable against changes in the state of any single neuron.

1. Introduction

To process multistate gray or color patterns intelligently, such as contern-addressable storage, parallel distributed process and associative recognition, some multistate neural network models have been suggested and discussed. Kanter generalized the Ising neuron to the Potts neuron and so developed the Pott-glass models of neural networks [1]. Using the property of plane rotation of the complex numbers, the discrete-state phasor neural network model [2] is suggested by Noest. By generalizing the neuron state ± 1 to other integers, Rieger proposed the Q-state neural network model [3]. Shuai introduced the 2^n-element number [4] into the neural networks and developed the 2^n-element number neural networks [5,6]. The 2^n-element number neural network models can be applied to recognize the 4, 16, 256 levels gray or color patterns that are widely used in computers. The encoding schemes of the 2^n-element number neuron can be set up in exactly the same way as the computer encoding schemes.

For the nonlinear Hopfield neural network model [7,8], after introducing an important concept, i.e. the energy function, into the network, an unambiguous criterion is obtained for the study of stability. In the case of random purely sequential dynamics the Hopfield model must reach a stationary point that corresponds to a local minimum of the energy

Elsevier Science B.V.
SSDI 0378-4371(95)00128-X

function. The network thus always ends up in a state of equilibrium, which is stable against changes in the state of any single neuron. In the case of fully parallel dynamics the network can either reach a steady state or permanently cycle between two different configurations [9,10]. In this paper, the energy function and the stability of the 2^n-element number neural network with $n < 3$ are discussed. The energy of the networks is defined in Section 2. Section 3 is devoted to the proof that the memorized patterns are local minima of the energy function. In the next section we show that the energy of the network decreases monotonically with time in the case of random purely sequential dynamics and so the nonlinear 2^n-element number neural networks with $n < 3$ must end up in a state of equilibrium, which is stable against changes in the state of any single neuron.

2. The energy function of the model

The 2^n-element numbers [4] are introduced into the neural network so as to form the discrete 2^n-element number neural network models [5,6]. In the model, each neuron is assumed to be a multistate one, such as 2-state (± 1) for the real number, i.e., the Hopfield model, 4-state ($\pm 1 \pm i$) for the complex number, 16-state ($\pm 1 \pm i \pm j \pm k$) for the Hamilton number and 256-state ($\pm 1 \pm i \pm j \pm k \pm e \pm ie \pm je \pm ke$) for the Cayley number.

Suppose there are N neurons and M patterns S^μ ($\mu = 1, 2, \ldots, M$) stored in the network. The weight of the synaptic interconnection is given by the extended Hebbian learning rule:

$$J_{mn} = \sum_{\mu=1}^{M} S_m^\mu (S_n^\mu)^*, \quad m \neq n; \qquad J_{mm} = 0. \tag{1}$$

Here $(S_n^\mu)^*$ denotes the conjugate of S_n^μ. For the 2^n-element number, the equation $(\alpha\beta)^* = \beta^* \alpha^*$ holds. So we have $J_{mn}^* = J_{nm}$.

The interaction of a neuron with other neurons is defined by introducing a field at the location of another neuron through the synaptic interconnection. The total local field h_m at the location of a neuron m is given by

$$h_m = \sum_{n=1}^{N} J_{mn} S_n. \tag{2}$$

Obviously, the diagonal term $m = n$, i.e. the self-energy, is not included in the local field. The local field is actually a 2^n-element number field.

The dynamics of the neuron depends only on its local field and the evolution of the model will be analyzed under a range of discrete-time dynamics differing in their degree of parallelism, but all having the same rule for updating individual neurons

$$S_m(t+1) = \Theta(h_m(t)). \tag{3}$$

Here the function Θ is the generalized step function. The operating rule of the function $\Theta(\alpha)$ is as follows: whenever a real or imaginary component of α is nonnegative, a positive unit is drawn out for the corresponding component of $\Theta(\alpha)$; otherwise, a negative unit drawn out. The degree of parallelism of the updating is allowed to vary from random purely sequential to fully parallel. For purely sequential dynamics, in some randomly chosen sequence for m, the individual neurons m are computed only once. For fully parallel dynamics, all neurons assume new states in parallel and at the same moment.

The neural networks are strong nonlinear systems. So the stability of the dynamical evolution is an important question for the networks. Similarly to the Hopfield model, an energy function $E(s)$ is introduced to describe the properties of the 2^n-element number neural network:

$$E(s) = -\frac{1}{2}\sum_{m=1}^{N} s_m^* h_m = -\frac{1}{2}\sum_{m,n=1}^{N} s_m^*(J_{mn}s_n). \tag{4}$$

It is known that some properties of real number are lost gradually with the increase of n. It is not possible to compare two complex numbers in order to establish which is the larger of the two. Although the Hamilton numbers obey the combination law, they do not obey the exchange law of multiplication. For the Cayley numbers, they obey neither the combination law nor the exchange law of multiplication. The loss of calculation properties with the increase of n has an effect upon the 2^n-element number neural network models. For example, the sequence of positions above in Eqs. (1)-(4) is important for $n > 1$, i.e. for the Hamilton and Cayley number neural network; furthermore, for the Cayley number neural network, the order of calculation in a equation is important, and so the above expression of energy function is unsuitable to describe the Cayley number neural network model. Now we first prove that expression (4) ensures that the energy function is real for the 2^n-element number neural network with $n < 3$. Noting that the conjugate of the energy function is

$$E^* = -\frac{1}{2}\left(\sum_{m=1}^{N} s_m^* h_m\right)^* = -\frac{1}{2}\sum_{m,n=1}^{N}[s_m^*(J_{mn}s_n)]^* = -\frac{1}{2}\sum_{m,n=1}^{N}(s_m^* J_{mn})s_n, \tag{5}$$

with $n < 3$, the 2^n-element numbers obey the combination law, and hence

$$E^* = -\frac{1}{2}\sum_{m,n=1}^{N} s_m^*(J_{mn}s_n) = E. \tag{6}$$

So expression (4) is a suitable definition for the energy function of the 2^n-element number neural network with $n < 3$.

We can include the diagonal terms $m = n$ in (1) without prejudice, letting $J_{mm} = \sum_{\mu=1}^{M} S_m^\mu (S_m^\mu)^*$. Because of $s_m s_m^* = 2^n$, only a constant term, i.e. the self-energy, is added to the energy, which does not depend on the states of the neurons:

$$E(s)_{\text{self}} = -\frac{1}{2}\sum_{m=1}^{N}\sum_{m=1}^{N} s_m^* J_{mm} s_m = -\tfrac{1}{2}\cdot 2^{2n} NM. \tag{7}$$

This additional self-energy constant does not influence the dynamical evolution of the system and so can be dropped.

3. The energy function of the stored patterns

If the configuration of the network corresponds to a stored pattern s^ν, the energy function is expressed by

$$E(s^\nu) = -\frac{1}{2}\sum_{m\neq n}^{N}\sum_{\mu=1}^{M} s_m^{\nu *} s_m^\mu s_n^{\mu *} s_n^\nu = -\frac{1}{2}\left(2^{2n}N(N-1) + \sum_{m\neq n}\sum_{\mu\neq\nu} s_m^{\nu *} s_m^\mu s_n^{\mu *} s_n^\nu\right). \tag{8}$$

If we expand the second term of the right-hand side of Eq. (8) in terms of the constituent real and imaginary parts, it becomes a sum of $(M-1)(N-1)(2^n)^4$ terms. For the complex number model, one half of terms are real and the rest are the imaginary, and must cancel out each other. For the Hamilton number model, a quarter of terms are the real and the rest are imaginary (i, j, k), and must be vanish. So the second term is actually a sum of $(M-1)(N-1)(2^n)^3$ real terms. Assuming that the patterns are uncorrelated and the component values $s_{mi}^\mu = \pm 1$ $(0 \leq i < 2^n)$ are equally likely, according to the laws of statistics, one can find that the second term is of order $2^{3n}(M-1)(N-1)^{1/2}$, and therefore the energy becomes

$$E(s^\nu) \approx -2^{2n-1}N(N-1) + O(x|_{2^{3n-1}(M-1)(N-1)}), \tag{9}$$

where the function $O(x|_\sigma)$ is a Gaussian distribution with expectation value zero and standard deviation value σ

$$O(x|_\sigma) = (1/\sqrt{2\pi}\sigma)\exp(-x^2/2\sigma^2). \tag{10}$$

The influence of all other patterns $(\mu \neq \nu)$ causes a slight shift in the total energy. As long as $M \ll N$, i.e. when the number of stored patterns is much smaller than the total number of neurons in the network, the influence term will most likely not affect the total energy. If the state s has only one bit of wrong orientation compared with the stored pattern s^ν, without loss of generality, assume that one of the components in the first neuron is incorrectly oriented,

$$s_1 = s_1^\nu + \Delta s_1; \qquad s_m = s_m^\nu, \quad m > 1. \tag{11}$$

With only one bit wrong, one can obtain the relationship

$$s_1^{\nu *}\Delta s_1 + \Delta s_1^* s_1^\nu = -4. \tag{12}$$

For example with the Hamilton number, if $s_1^\nu = 1+i+j+k$ and $s_1 = 1-i+j+k$, we have $\Delta s_1 = -2i$ and so the above equation holds. Now if the configuration s is put into the network, the fluctuating term is replaced by another fluctuating term, which does not result in an essential modification of its distribution. However, the term with $\mu = \nu$ suffers a substantial change:

$$\begin{aligned}\sum_{m\neq n}^{N} s_m^* s_m^\nu s_n^{\nu *} s_n &= \sum_{m\neq n>1}^{N} s_m^* s_m^\nu s_n^{\nu *} s_n + \sum_{m>1}^{N} s_m^* s_m^\nu s_1^{\nu *}(s_1^\nu + \Delta s_1) \\ &\quad + \sum_{n>1}^{N}(s_1^{\nu *} + \Delta s_1^*) s_1^\nu s_n^{\nu *} s_n + (s_1^{\nu *} + \Delta s_1^*) s_1^\nu s_1^{\nu *}(s_1^\nu + \Delta s_1) \\ &= 2^{2n} N(N-1) + 2^n (N-1)(s_1^{\nu *}\Delta s_1 + \Delta s_1^* s_1^\nu) + 2^n \Delta s_1^* \Delta s_1 \\ &= 2^{2n} N(N-1) - 2^{n+2}(N-2), \end{aligned} \tag{13}$$

i.e. the energy of the configuration rises in proportion to the extent of its deviation from the stored pattern:

$$\begin{aligned} E(s) &\approx -2^{2n-1} N(N-1) + 2^{n+1}(N-2) + O(x|_{2^{3n-1}(M-1)(N-1)}) \\ &= E(s^\nu) + 2^{n+1}(N-2). \end{aligned} \tag{14}$$

The memorized patterns are therefore the local minima of the energy function $E(s)$.

4. The dynamical process

The energy function of a physical system is closely related to its dynamics. We can see that the energy function $E(s)$ of the 2^n-element number neural networks with $n < 3$ never increases with time in the case of random sequential updating of the neuron states. With a configuration $s(t)$, the energy of the network at time t is

$$E(s(t)) = -\frac{1}{2}\sum_{m\neq n}^{N}\sum_{\mu=1}^{M} s_m^* s_m^\mu s_n^{\mu *} s_n. \tag{15}$$

During the next period, i.e. time $t+1$, if the state of the first neuron is chosen and changed to another 2^n-element number neuron state according to the dynamics, i.e.

$$s_1(t+1) = \Theta(h_1(t)) = s_1(t) + \Delta s_1,$$

its contribution to the change of the network energy is

$$\begin{aligned}\Delta E = E(t+1) - E(t) &= -\frac{1}{2}\left(\sum_{m\neq 1}^{N}\sum_{\mu=1}^{M} s_m^* s_m^\mu s_1^{\mu *}\Delta s_1 + \sum_{m\neq 1}^{N}\sum_{\mu=1}^{M}\Delta s_1^* s_1^\mu s_m^{\mu *} s_m\right) \\ &= -\frac{1}{2}\left(\sum_{m\neq 1}^{N}\sum_{\mu=1}^{M}(s_1^\mu s_m^{\mu *} s_m)^*\right)\Delta s_1 - \tfrac{1}{2}\Delta s_1^*\left(\sum_{m\neq 1}^{N}\sum_{\mu=1}^{M}(s_1^\mu s_m^{\mu *} s_m)\right)\end{aligned}$$

$$= -\tfrac{1}{2}(h_1^* \Delta s_1 + \Delta s_1^* h_1). \tag{16}$$

For $n = 1$, the complex number neural network model, the states and the local fields of the neurons are all complex, so that

$$\Delta s_1 = \Delta s_1^R + i\Delta s_1^I h_1 = h_1^R + ih_1^I. \tag{17}$$

Thus

$$\Delta E = -(h_1^R \Delta s_1^R + h_1^I \Delta s_1^I). \tag{18}$$

For $n = 2$, the Hamilton number neural network model,

$$\Delta s_1 = \Delta s_{10} + i\Delta s_{11} + j\Delta s_{12} + k\Delta s_{13}, \qquad h_1 = h_{10} + ih_{11} + jh_{12} + kh_{13}, \tag{19}$$

so that

$$\Delta E = -\sum_{i=0}^{3} (h_{1i} \Delta s_{1i}). \tag{20}$$

Therefore from the rule of the dynamics, i.e. Eq. (3), one can see that $h_1^* \Delta s_1 + \Delta s_1^* h_1$ is always nonegative. For example with the Hamilton number, if $h_{11} > 0$, Δs_{11} is always nonegative. Hence for the 2^n-element number neural network with $n < 3$, we have

$$\Delta E \leq 0. \tag{21}$$

The energy contribution of a given neuron decreases monotonically with time. Since the energy function $E(s)$ is bounded from below, this implies that the network dynamics must reach a stationary state that corresponds to a local minimum of the energy function. The networks thus always end up in a state of equilibrium, which is stable against changes in the state of any single neuron.

The argument given above does not apply to the synchronous mode of operation of the neural network. Since all neurons assume new states in parallel and at the same moment, the contribution of an individual neuron to the energy function cannot be considered in isolation.

In the present section, it is proved that the memorized patterns are the local minima of the energy of the network in the case of $M \ll N$. So from the viewpoint of the energy function, one can conclude that in the case of random sequential dynamics the memorized patterns are the stable attractors of the 2^n-element number neural network models upon condition that the number of stored patterns is much smaller than the total number of neurons in the network. By using the signal-to-noise theory in Refs. [5,6], we also draw the above conclusion for the 2^n-element number neural network for the case of fully parallel dynamics.

5. Conclusion

In this paper, an energy function is defined for discribling the 2^n-element number neural networks with $n < 3$. The memorized patterns are local minima of the energy function when the number of stored patterns is much smaller than the total number of neurons in the network. In the case of random purely sequential dynamics, the energy of the network never increases with time and so the nonlinear 2^n-element number neural networks with $n < 3$ must end up in a state of equilibrium, which is stable against changes in the state of any single neuron. And so the memorized patterns are the stable attractors of the 2^n-element number neural network models. For the Cayley number neural network models ($n = 3$), can we or how we define a suitable energy function remains a problem. The difficulty is that many good properties of the calculation are lost.

The 2^n-element number neural network model can be used to recognize 2^{2^n}-level gray or color patterns, i.e. 4, 16, 256 levels gray or color patterns that are widely used in computers. There are some optical architectures which have been reported for the complex-valued vector-matrix multiplication [11,12]. For example, an optical 128-element 8-bit complex-valued vector-matrix multiplication architecture has been implemented by Mosca [12]. According to his idea, the 2^n-element number vector-matrix multiplication can be implemented in the optical domain [5]. Thus the 2^n-element number neural network model can be implemented by the optical-electronic system.

Acknowledgements

This work is supported by the National Natural Scientific Research Foundation, grant number: 19334032.

References

[1] I. Kanter, Potts-glass models of neural networks, Phys. Rev. A 37 (1988) 2739.
[2] A.J. Noest, Discrete-state phasor neural networks, Phys. Rev. A 38 (1988) 2196.
[3] H. Rieger, Storing an extensive number of gray-toned patterns in a neural network using multistate neurons, J. Phys. A 23 (1990) L1273.
[4] E.U. Conden, Handbook of Physics, 2nd Ed., Vol. 1 (New York, 1967) p. 22.
[5] J.W. Shuai, Z. X. Chen, R.T. Liu and B.X. Wu, The 2^n-element number neural network model: to recognize the multistate patterns, J. Modern Optics (accepted).
[6] J.W. Shuai, Z.X. Chen, R.T. Liu, and B.X. Wu, The Hamilton neural network model, Physica A (accepted).
[7] J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. USA 79 (1982) 2554.
[8] J.J. Hopfield, Neurons with graded response have collective computational properties like those of two-state neurons, Proc. Natl. Acad. Sci. USA 81 (1984) 3088.
[9] J. Bruck and J.W. Goodman, A generalized convergence theorem for neural networks and its applications in combinatorial optimization, IEEE Trans. Inform. Theory 34 (1988) 1089.
[10] C.M. Marcus and R.M. Westervelt, Dynamics of iterated-map neural networks, Phys. Rev. A 40 (1989) 501.
[11] E.P. Mosca, R.D. Griffin, F.P. Pursel and J.N. Lee, Acoustico-optical matrix-vector product processor: implementation issues, Appl. Opt. 28 (1989) 3843.
[12] H.X. Huang, L.R. Liu, Y.Z. Yin and L.Y. Zhao, Fast parallel complex discrete Fourier transforms using a multichannel optical correlator, Opt. Commun. 68 (1988) 408.