ELSEVIER Information Sciences 110 (1998) 179–193

INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL

The majority theorem of centralized multiple BAMs networks [1]

Chua-Chin Wang [a,*], Hon-Son Don [b]

[a] *Department of Electrical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan 80424, ROC, Taiwan*

[b] *Department of Electrical Engineering, National Chung-Hsing University, Taichung, Taiwan 40227, ROC, Taiwan*

Received 8 February 1995; received in revised form 17 June 1997; accepted 20 August 1997

Abstract

A method for modeling the learning of belief combination in evidential reasoning using a neural network is presented. A centralized network composed of multiple bi-directional associative memories (BAMs) sharing a single output array of neurons is proposed to process the uncertainty management of many pieces of evidence simultaneously. The convergence properties of the multi-BAM network are proved. The combination process of evidence is considered as a resonant process through the multi-BAM networks. Most important of all, a majority rule of decision making in presentation of multiple evidence is also found by the study of signal-noise-ratio (SNR) of the multi-BAM network. Some simulation examples are given. The result is coherent with the intuition of reasoning. © 1998 Elsevier Science Inc. All rights reserved.

Keywords: Bidirectional associative memory; Multi-BAM; Operation modes; Evidential reasoning; Belief function

1. Introduction

Neural networks have been drawing increasing interest as powerful tools to solve different tasks, [1,2]. An *associative memory* is one type of neural network

* Corresponding author. Tel.: 886 7525 2000 ext. 4144; fax: 886 7525 4199.

[1] This research was partially supported by National Science Council under grant NSC 84-0408-E-110-003.

PII: S0020-0255(97)10037-8

which essentially is a single functional layer or slab that associates one set of vectors with another set of vectors. Kosko [3,4] proposed a two-level nonlinear network, *bidirectional associative memory* (BAM), which extends a one-directional process to a two-directional process. One beneficial characteristic of the BAM is its ability to recall stored pattern pairs in the presence of noise. Wang et al. [5] discovered some flaws in Kosko's coding scheme and proposed two alternatives, *multiple training* and *dummy augmentation*, both of which enhance BAM's ability to find the global minimum.

Artificial intelligence researchers are exploring the utilization of neural networks in an expert system to handle uncertainty management [6–9]. The utilization of neural networks in evidential reasoning is the essence of the reasoning process, which is a forth-and-back reverberation process for uncertainty management. Assume people are given some evidence to "estimate" the credibility and the plausibility of a hypothesis, then the thinking behavior of people is to take some evidence into consideration and then gradually use the rest of given information to adjust the uncertainty so that a possibly optimal "estimate" for the credibility or the plausibility of the hypothesis can be reached. It is analogous to a system, which is excited by some input information, tending to reach a local minimum of the global system energy. The learning algorithm of the Boltzmann machine is a good instance of this kind of forth-and-back process. The feasibility of evidential reasoning by the neural networks has attracted much attention [6–8]. Since Shastri's method [8] has to use very complex nodes, it is more likely to utilize clusters of neurons instead of single neurons. Furthermore, since Shastri's weights are based on relative frequency of occurrence, they are only roughly compatible with the biologically plausible Hebbian learning rule [9]. Hsu et al., has implemented Dempster–Shafer's rule (D–S theory) using neural networks [3]. However, Hsu's method is not appropriate for the management of uncertainty reasoning because of the intrinsic shortcomings of the D–S theory [10,11]. In addition to the above works, the utilization of other neural networks for evidential reasoning can be found in [1,2,6,9,12].

Among the problems of evidential reasoning, conflicts caused by sequential programming and partial dependency are pretty hard to be fully resolved [10,13]. The basic reason is all of the traditional methods for evidential reasoning are developed for two pieces of evidence. Thus, when there are more than two pieces of evidence, conflicts will happen if the combination orders are different [11]. Wang et al., pointed out the importance of simultaneously processing many pieces of evidence [11], and we further proposed a method using multiple BAM structure to handle the demand of combining many evidence at the same time [10]. Because the relationship of evidence and the hypothesis is always referred to be an IF-and-THEN relationship. Hence, this IF-and-THEN format can be easily transformed into numbers which can be stored in memories, more specifically, associative memories. If people intend to evaluate the degree of a piece of evidence supporting a hypothesis, then they simply present the evidence to the memories to the recalled information. If there are more than a piece of evidence, then they can present all of the evidence and see the result of their common output. If the more evidence support one hypothesis, the result should be drawn closer to this hypothesis.

We also discuss the majority rule of decision making for handling many evidence at the same time. The majority rule means if more than half of the presented evidence support one hypothesis, though the rest of the presented evidence do not, the belief combination of all of these evidence must be dominated by the hypothesis. The bounds of a majority factor, k, is found and proved by the SNR approach [14]. The meaningful contribution of the majority rule is to predict the recall of certain common output patterns. This rule is intuitively in accordance with the human reasoning.

The rest of this paper is structured as follows: Section 2.1 briefly introduces Kosko's previous work on the BAM; Section 2.2 introduces the discrete multi-BAM network and discusses its convergence property to recall pattern pairs; Section 2.3 establishes the foundation of the majority rule of the multi-BAM network by using the SNR analysis approach; Section 3 shows some examples to illustrate the theories given in Section 2; and a conclusion is given in Section 4.

2. Framework of the multi-BAM network

2.1. Theory and structure of a BAM network

BAM was first introduced by Kosko [3,4]. The basic structure of a BAM is shown in Fig. 1. An array of n neurons, a_i's, in the bottom of the network, are input units; and another array of p neurons, b_j's, in the top of the network, are output units. An $n \times p$ matrix M is interpreted as a matrix of synapses between the input neurons, a_{ij}'s, and output neurons, b_{ij}'s. Matrix entry m_{ij} is a synaptic connection between a_i and b_j. The sign of m_{ij} determines the status of the connection: excitatory if $m_{ij} > 0$, inhibitory if $m_{ij} < 0$. The magnitude of m_{ij} determines the strength of the connection. The formation of the matrix M is based on the following operations. Given N training sample pairs, which are

$$\{(A_1, B_1), (A_2, B_2), \ldots, (A_N, B_N)\},$$

where

$$A_i = (a_{i1}, a_{i2}, \ldots, a_{in}), \qquad B_i = (b_{i1}, b_{i2}, \ldots, b_{ip}).$$

a_{ij} and b_{ij} are of either *on* or *off* status. The on and off status can be represented by $\{1, 0\}$ in the binary mode or $\{1, -1\}$ in the bipolar mode. The matrix M is constructed as

$$M = \sum_{i=1}^{N} X_i^{\mathrm{T}} Y_i, \tag{1}$$

where X_i and Y_i are the bipolar mode of A_i and B_i, respectively.

Suppose we wish to recall one of the nearest (A_i, B_i) pairs from the network if any input pattern A is presented to the network. Starting with this initial condition A which is closer to one stored pattern A_i than any other pattern, Kosko suggested the following forth-and-back operation to recall the B which is relatively close to B_i. By using a Lyapunov energy function, Kosco also proved that the forth-and-back reverberation process will converge to a final pair (A_i, B_i) with local minimum energy [3,4]. The BAM system energy is defined as

$$E(A, B) = -A M B^{\mathrm{T}}. \tag{2}$$

2.2. Discrete multi-BAM network

For evidential reasoning, many pieces of evidence could be presented to the processor simultaneously. This motivates us to explore the feasibility of organizing more than one BAM to handle the uncertainty of evidence. We propose a multi-BAM neural network, which is shown in Fig. 2. In the multi-

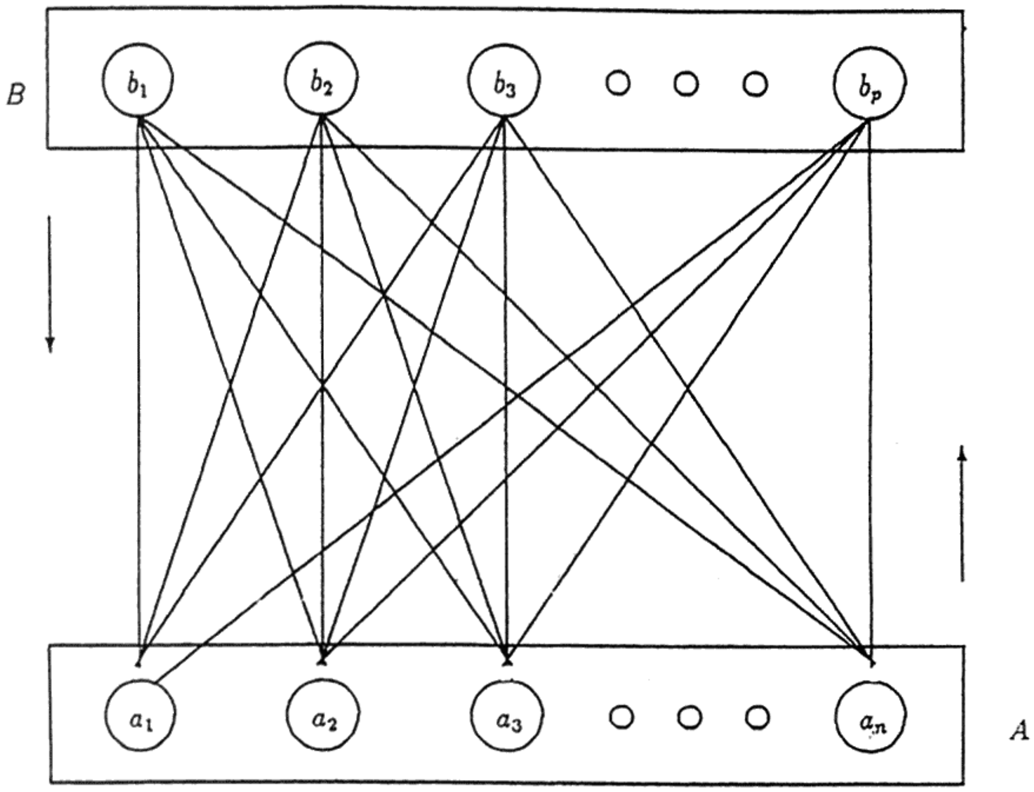

Fig. 1. The configuration of the BAM network.

BAM network, there are L BAMs with their individual array of input neurons, but these L BAMs share the only one array of output neurons.

Suppose there are L BAMs in the network, and (A_{qi}, B_i) is a training pair of matrix $M_q, 1 \leqslant q \leqslant L$. Each matrix M_q is formulated similarly to Eq. (1) by A_{qi} and B_i. In the synchronous mode, L input patterns are presented to the L input arrays of the network, respectively, and fed through the individual M_q's to produce the individual B'_q's. Since every single entry of every B'_q expresses how strong it wants to determine the on or off status of the corresponding output neuron, the final status of every output neuron should be determined by the sum of every B'_q. Therefore, the entire operation is

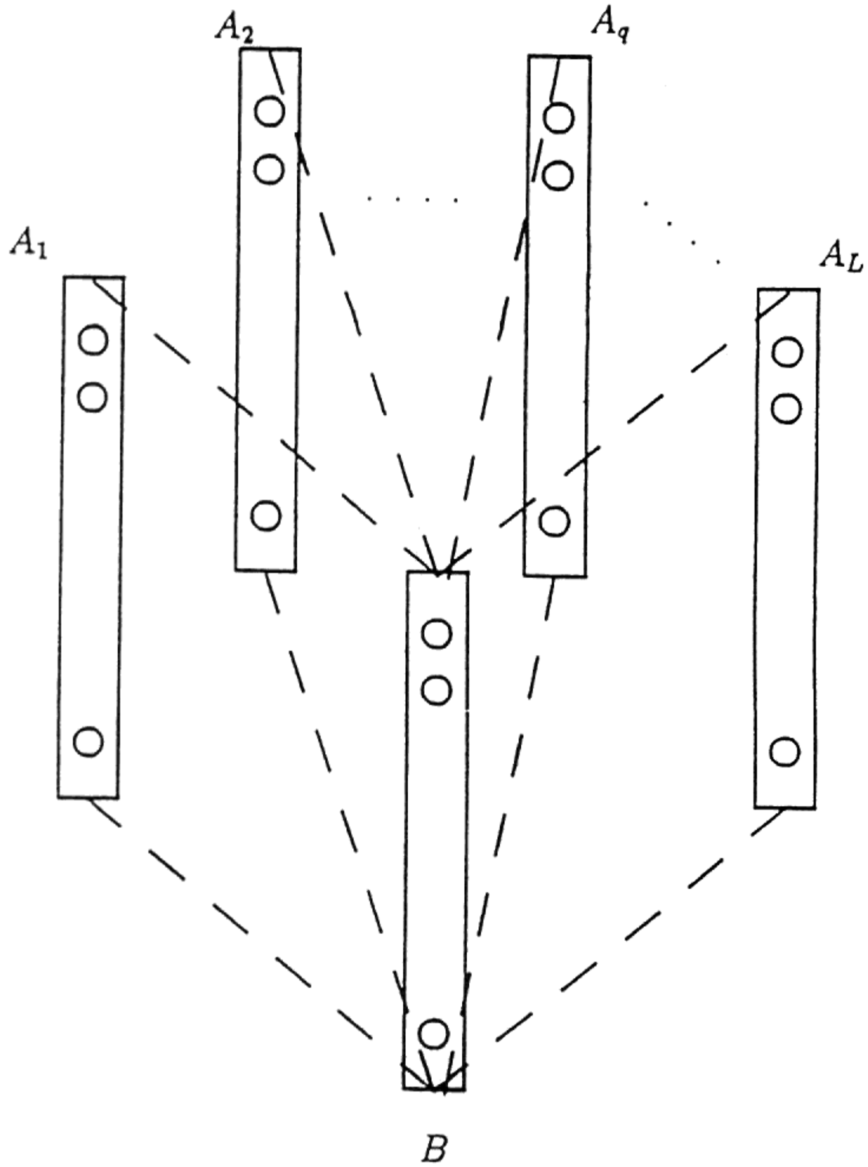

Fig. 2. The structure of a multiple-BAM network.

$$\left(A_q \rightarrow M_q \rightarrow B'_q,\ B' = \sum_{q=1}^{L} B'_q\right)$$
$$(A'_q \leftarrow M_q^{\mathrm{T}} \leftarrow B')$$
$$\vdots$$
$$\left(A_{qf} \rightarrow M_q \rightarrow B_{qf},\ B_f = \sum_{q=1}^{L} B_{qf}\right)$$
$$(A_{qf} \leftarrow M_q^{\mathrm{T}} \leftarrow B_f)$$
$$\vdots$$

where $A_{q\mathrm{f}}$ and B_{f} are the final state of the qth BAM. Ideally, the individual $(A_{q\mathrm{f}}, B_{b\mathrm{f}})$ pair will be relatively close to one of the training pairs of the qth BAM.

The following theorem guarantees the convergence and stability of a multi-BAM network.

Theorem 1. *Given a multi-BAM network with L BAMs as described above, its overall Lyapunov energy defined as*

$$E = \sum_{q=1}^{L} E_q(A_q, B).$$

The total energy will converge to a local minimal value.

Proof. Let the ith neuron of input array of the qth BAM be denoted as a_{qi}, and the jth neuron of the common output array B be denoted as b_j. Henceforth, we can derive the following equations by the matrix inner product operations.

$$\sum_q A_q M_q^j = \sum_q \sum_i a_{qi} m_{qij}, \tag{3}$$

$$BM_q^{\mathrm{T}i} = \sum_j b_j m_{qij}, \tag{4}$$

where M_q^j is the jth column of M_q, and $M_q^{\mathrm{T}i}$ is the ith row (column) of $M_q(M_q^{\mathrm{T}})$. Eq. (3) shows the input sum to b_j, and Eq. (4) is the input sum fed backward to a_{qi}. Zero is taken as the threshold for all neurons. Therefore, the threshold functions for the neurons of the network are

$$a_{qi} = \begin{cases} 1 & \text{if } BM_q^{\mathrm{T}i} > 0, \\ 0 & \text{if } BM_q^{\mathrm{T}i} < 0, \end{cases} \tag{5}$$

$$b_j = \begin{cases} 1 & \text{if } \sum_q A_q M_q^j > 0, \\ 0 & \text{if } \sum_q A_q M_q^j < 0. \end{cases} \tag{6}$$

The energy of a BAM is defined by Eq. (2), which will decrease along discrete trajectories in the space $\{0,1\}^n \times \{0,1\}^p$. This property can be shown by the fact that changes in state variables result in $\Delta E < 0$. Similarly, we can show that the multi-BAM network has the same property to guarantee the convergence of the network energy. The overall energy of the multi-BAM network is expressed as

$$E = \sum_q E(A_q, B) = -\sum_q \sum_i \sum_j a_{qi} b_j m_{qij}$$

$$= -\sum_{q \neq r} \sum_{i \neq k} \sum_j a_{qi} b_j m_{qij} - a_{rk} \sum_j b_j m_{rkj}, \tag{7}$$

$$= -\sum_q \sum_i \sum_{j \neq k} a_{qi} b_j m_{qij} - b_k \sum_q \sum_i a_{qi} m_{qik}. \tag{8}$$

We also note that $\Delta a_{qi}, \Delta b_j \in \{-1,0,1\}$ for the binary mode, and $\Delta a_{qi}, \Delta b_j \in \{-2,0,2\}$ for bipolar mode, thus implying that only the nonzero state changes have to be taken into consideration. The energy change $\Delta E = E_2 - E_1$ caused by the state change Δa_{rk} can be derived from Eq. (5), which is

$$\frac{\Delta E}{\Delta a_{rk}} = -\sum_j b_j m_{rkj} = -BM_r^{\mathrm{T}k}. \tag{9}$$

The right-hand side of Eq. (9) is the negative input sum to neuron a_{rk} shown in Eq. (3). Henceforth, if $\Delta a_{rk} > 0$, Eq. (5) shows $BM_r^{\mathrm{T}k} > 0$, and then it is concluded that $\Delta E = -\Delta a_{rk} BM_r^{\mathrm{T}k} < 0$. On the other hand, if $\Delta a_{rk} < 0$, Eq. (5) shows that $BM_r^{\mathrm{T}k} < 0$, and thus $\Delta E = -\Delta a_{rk} BM_r^{\mathrm{T}k} < 0$ still holds. Similarly, the energy change due to the other state change b_k is expressed as

$$\frac{\Delta E}{\Delta b_k} = -\sum_q \sum_i a_{qi} m_{qik} = -\sum_q A_q M_q^k. \tag{10}$$

Again we can recognize that the right portion of the above equation is the negative input sum to neuron b_k form Eq. (4). Thus $\Delta b_k > 0$ only if $\sum_q A_q M_q^k > 0$, and $\Delta b_k < 0$ only if $\sum_q A_q M_q^k < 0$. Hence $\Delta E = -\Delta b_k \sum_q A_q M_q^k < 0$ in either of the above two cases.

The previous discussion ensures $\Delta E < 0$, which indicates the overall energy of the forth-and-back reverberation process of the multi-BAM network will decrease in the space $\{0,1\}^n \times \{0,1\}^p$, as we expected.

2.3. A majority rule of decision making for the multi-BAM network

2.3.1. The majority rule of the strict case

According to the discussion in the previous sections, every single BAM tends to store their own pattern pairs in the local minimums of their network, respectively. Assume there are L single BAMs consisting of a multiple BAM network, and these BAMs share a single output array of processing units. If these individual BAMs are activated by respective input patterns and they don't "agree" to have the same conclusion, i.e., the same output pattern, what will be the final result of the whole network. This problem is like a reasoning mechanism which takes many evidence into consideration at the same time in order to reach an optimal estimation of the hypothesis.

Hence, we formulate the entire problem as follows: Given a multi-BAM network composed of L single BAMs, what is the minimal majority factor $\rho, \rho \in [0,1]$, to make ρL BAMs, which are vowing a common output pattern and the other BAMs are not, dominate the common output? In other words, we are interested in exploring the lower bound of the ρL which can force the output pattern to be their common output pattern. Note that in fact the ρL denotes an integer, $Ceiling(\rho L)$, which is the smallest integer larger than ρL. In the following text, we simply use the ρL without any loss of robustness.

Before we discuss the lower bound of ρL, we have to study an extreme case in which a upper bound of ρL will be derived. In the rest discussion of this section, Xs and Ys are the bipolar mode of As and Bs, respectively. Assume the pattern pairs, $(X_{11}, Y_r), (X_{21}, Y_r), \ldots, (X_{\rho L_1}, Y_r)$, are encoded in 1st to ρLth BAMs, respectively, and pattern pairs, $(X_{(\rho L+1)1}, Y_s), \ldots, (X_{L1}, Y_s)$, are stored in $(\rho L+1)$th to Lth BAMs, respectively. Thus, when input patterns, $X_{11}, X_{21}, \ldots, X_{L1}$, are presented at the input array of each individual BAM, what would be the result of output?

Suppose the Y_r is the output pattern that we are looking for, i.e., it is deemed as the signal. By the SNR approach, [14], the evolution equations, Eqs. (5) and (6), for BAM can be rewritten in a bitwise manner.

$$
\begin{aligned}
y_j &= \operatorname{sgn}\left(\sum_q X_q M_q^j\right) = \operatorname{sgn}\left(\sum_{q=1}^{L}\sum_{i=1}^{n}\sum_{k=1}^{N}(x_{qi}\,x_{qki})y_{kj}\right)\\
&= \operatorname{sgn}\left(\sum_{q=1}^{\rho L}\sum_{i=1}^{n}\sum_{k=1}^{N}(x_{qri}\,x_{qki})y_{kj} + \sum_{q=\rho L+1}^{L}\sum_{i=1}^{n}\sum_{k=1}^{N}(x_{qsi}\,x_{qki})y_{kj}\right)\\
&= \operatorname{sgn}\left(\sum_{q=1}^{\rho L}\sum_{i=1}^{n}(x_{qri}\,x_{qri})y_{rj} + \sum_{q=1}^{\rho L}\sum_{i=1}^{n}\sum_{k\neq r}^{N}(x_{qri}\,x_{qki})y_{kj}\right.\\
&\quad \left. + \sum_{q=\rho L+1}^{L}\sum_{i=1}^{n}(x_{qsi}\,x_{qsi})\,y_{sj} + \sum_{q=\rho L+1}^{L}\sum_{i=1}^{n}\sum_{k\neq s}^{N}(x_{qsi}\,x_{qki})y_{kj}\right)\\
&= \operatorname{sgn}\left(\rho L\,n\,y_{rj} + (1-\rho)Lny_{sj} + \sum_{q=1}^{\rho L}\sum_{i=1}^{n}\sum_{k\neq r}^{N}(x_{qri}\,x_{qki})y_{kj}\right.\\
&\quad \left. + \sum_{q=\rho L+1}^{L}\sum_{i=1}^{n}\sum_{k\neq s}^{N}(x_{qsi}\,x_{qki})y_{kj}\right), \qquad (11)
\end{aligned}
$$

where the first term is the desired pattern, called signal, and the rest terms are deemed as noise. On the other hand, the largest power of noise, which means all of the rest $(1-\rho)L$ BAMs support another output pattern Y_s, is

$$\begin{aligned}\text{Noise} &= n^2(1-\rho)L + \rho L(n-2)^2(N-1) + (1-\rho)L(n-2)^2(N-1)\\ &= n^2(1-\rho)L + (n-2)^2(N-1).\end{aligned} \tag{12}$$

And the power of the signal is

$$\text{Signal} = \rho L\, n^2.$$

In the above noise power equation, we assume not only all of the rest $(1-\rho)L$ BAMs support another output pattern, but also this pattern is the closest pattern to the desired one. If the desired output is intended to be recalled, then the sufficient condition is the Signal > Noise according to the SNR approach. Thus we can conclude the lower bonds for this definite recall condition of ρ, called *strict lower bound*, is

$$\begin{aligned}&\rho L n^2 > n^2(1-\rho)L + (n-2)^2(N-1),\\ &\rho > \frac{1}{2} + \frac{(n-2)^2(N-1)}{2n^2}.\end{aligned} \tag{13}$$

Note that this lower bound of ρ means any ρ bigger than this threshold can force the output pattern to be the common desired output pattern of the ρL BAMs in the network. However, if the bound of Eq. (13) is larger than 1, it means even all of the BAMs support one output pattern, there is no guarantee to recall this common output pattern.

2.3.2. *The majority rule of the general case*

In the above extreme case, we assume all of the rest $(1-\rho)L$ BAMs support another same output pattern which is only one bit Hamming distance away from the desired pattern. Generally speaking, however, most of the reasoning problems won't be this special. We will consider a general case in which the ρL BAMs still support a common output pattern, but the rest $(1-\rho)L$ BAMs do not support the same output pattern, i.e., they individually support their own output patterns, respectively. Based upon this assumption, then we can derive the following results.

First of all, we have to discuss what the capacity of a single BAM really is in terms of a signal-to-noise ratio approach. According to Kosko's formulation [3],

$$Y = X\,M = (X\,X_h^{\mathrm{T}})Y_h + \sum_{i\neq h}^{N}(X\,X_i^{\mathrm{T}})Y_i, \tag{14}$$

$$= nY_h + \sum_{i\neq h}(X\,X_i^{\mathrm{T}})Y_i. \tag{15}$$

It is natural to assume that the stored pattern pairs are drawn from $\{-1,1\}^n$ with uniform probability. Hence, the first term corresponds to the signal, Signal, which has the power n^2. The second term, the noise, has a zero mean and the variance as follows:

$$E[v_1^2] = 2\sum_{k=0}^{n-1}[n-2(k+1)]^2\left(\frac{1}{2}\right)^{n-1}C_k^{n-1} = 2\sum_{k=0}^{m}[m-2k-1]^2\left(\frac{1}{2}\right)^{m}C_k^m,$$

$$m = n-1.$$

Assume

$$f(x,y) = \sum_{k=0}^{m} x^k y^{m-k} C_k^m = (x+y)^m,$$

$$h(y) = \frac{1}{y} f(y^{-1},y) = \sum_k y^{m-2k-1} C_k^m.$$

Then, we can get

$$yh'(y) = \sum_k (m-2k-1) y^{m-2k-1} C_k^m,$$

$$(yh'(y))' = \sum_k (m-2k-1)^2 y^{m-2k-2} C_k^m.$$

If $y=1$, then

$$(h'(1))' = \sum_k (m-2k-1)^2 C_k^m.$$

It is trivial to derive $(yh'(y))'$ and substitute in $y=1$. The result is

$$(h'(1))' = 2(m+1)2^{m-1} = n2^{n-1}.$$

Hence the signal to noise ratio, i.e., the capacity, of Kosko's BAM is

$$\mathrm{SNR}_{\mathrm{BAM}} = \frac{n^2}{2(N-1)(1/2)^{n-1}2^{n-1}n} = \frac{n}{2(N-1)}. \tag{16}$$

Now let us analyze the SNR of the multi-BAM network. The noise terms of Eq. (11) are the 2nd, 3rd, and 4th terms. The power of 2nd term is

$$\mathrm{Noise}_2 = (1-\rho)Ln^2.$$

As for the 3rd and 4th terms of Eq. (11), they are actually sums of $\rho L(N-1)$ and $(1-\rho)L(N-1)$ independent identically distributed random variables, respectively. Therefore, the variances of the 3rd and 4th terms are $\rho L(N-1)$ and $(1-\rho)L(N-1)$ times the variance of a single random variable. Hence, the power of the 3rd and 4th terms are, respectively,

$$\mathrm{Noise}_3 = E(v_1^2)\rho L, \quad \mathrm{Noise}_4 = E(v_1^2)(1-\rho)L.$$

In short, the total noise power can be represented as

$$\begin{aligned}\text{Noise} &= \text{Noise}_2 + \text{Noise}_3 + \text{Noise}_4 \\ &= (1-\rho)L\,n^2 + \rho L\,2(N-1)n + (1-\rho)L\,2(N-1)n \\ &= (1-\rho)L\,n^2 + L\,2(N-1)n.\end{aligned}$$

Therefore, if we want a stored pattern pair to be recalled in a multi-BAM network, its SNR must be greater than 1, i.e., the power of signal must be larger than that of noise.

$$\rho\,Ln^2 > (1-\rho)L\,n^2 + L\,2(N-1)n, \rho > \frac{1}{2} + \frac{N-1}{n}.$$

2.3.3. Theorem of the majority rule for a multi-BAM network

Given a multi-BAM network with L single BAMs, ρL BAMs support a same common output pattern, where $\rho \in [0,1]$. The condition for the output pattern of the network is the same as the one supported by the ρL BAMs is

$$\rho > \frac{1}{2} + \frac{1}{2\,\text{SNR}_{\text{BAM}}}. \tag{17}$$

If the lower bound in Eq. (17) is larger than 1, then it means even all of the BAMs in the network support one output pattern, there is no guarantee to recall this output pattern.

By the above theorem, please note because the SNR_{BAM} is larger than 0, the lower bound of k can be simplified to be $\frac{1}{2}$ which complies the human intuition. That is, if more than 50% of the evidence supports a hypothesis, then the reasoning result most likely would be the same as this hypothesis.

3. Simulation examples

In this section, we use some examples of belief combination in evidential reasoning to illustrate the theoretical results of the multi-BAM networks discussed in Section 2.

Example 1. This example is used to prove the bidirectional stability of multi-BAM networks. Given a 2-BAM network, the training pattern pairs for the first BAM of the network are as follows:

$$\begin{aligned}A_{11} &= (111010001100101111010100000000001), \\ A_{12} &= (111010001100101111010100000000010), \\ A_{13} &= (111010001100101111010100000000100), \\ A_{14} &= (111010001100101111010100000001000), \\ B_1 &= (11001100110010111100111100), \\ B_2 &= (11001101010010111100111100),\end{aligned}$$

$B_3 = (101011011100101101100011)$,

$B_4 = (100100011100101101100011)$,

the training pairs for the second BAM of the network are,

$A_{21} = (11110110101001101011101000000001)$,

$A_{22} = (11110110101001101011101000000010)$,

$A_{23} = (11110110101001101011101000000100)$,

$A_{24} = (11110110101001101011101000001000)$,

where the B_j's of the lower part of the network are the same as those of the upper part. Then, M_1 and M_2 can be computed by Eq. (1). Let A_{11} and A_{21} be employed as inputs of the two parts of the network, respectively. The computer simulation results of the 2-BAM network are the following:

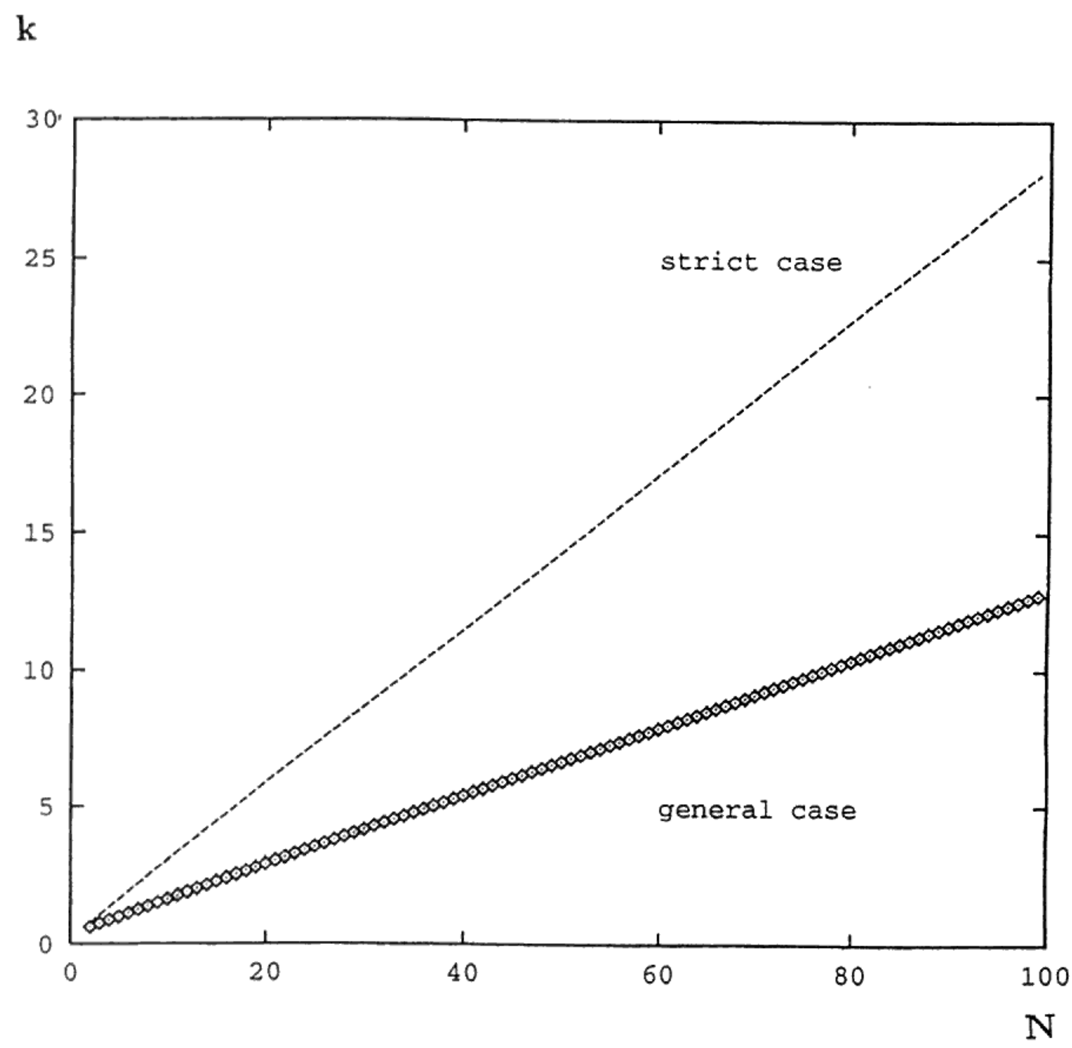

Fig. 3. The minimal ρ in the strict case and the general case.

$$\text{Iteration 1}, \quad E = -592,$$
$$\text{Iteration 2}, \quad E = -2640,$$
$$B_{\text{final}} = (1\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 1\ 1).$$

The network converges very quickly to a stable state, which verifies the prediction of Theorem 1. Note that the B_{final} is not equal to B_1. The reason is the strict lower bound ρ according to Eq. (13) is 1.76, which is larger than 1. This implies that even all of the BAMs support one output pattern, there is no guarantee to recall this common output pattern.

Example 2. This example illustrates the prediction of the majority rule of the multi-BAM network. First, let us use Fig. 3 to demonstrate the lower bound of ρ in the majority rule. In this example, $n = p = 8$. The numerical data is tabulated in Table 1.

According to the definition of lower bound ρ, ρ is a number between 0 and 1. That is, when the lower bound of ρ is larger than 1, there is always a chance that the network might fail to be converged to a desired output pattern even if all of the input patterns support that output pattern. In this example, $n = p = 8$. Hence, if each single BAM only stores 2 pattern pairs, it is guar-

Table 1
The minimal ρ in the strict case and the general case

N (number of pairs)	Strict case	General case
2	0.781250	0.625000
3	1.062500	0.750000
4	1.343750	0.875000
5	1.625000	1.000000
10	3.031250	1.625000
15	4.437500	2.250000
20	5.843750	2.875000
25	7.250000	3.500000
30	8.656250	4.125000
35	10.062500	4.750000
40	11.468750	5.375000
45	12.875000	6.000000
50	14.281250	6.625000
55	15.687500	7.250000
60	17.093750	7.875000
65	18.500000	8.500000
70	19.906250	9.125000
75	21.312500	9.750000
80	22.718750	10.375000
85	24.125000	11.000000
90	25.531250	11.625000
95	26.937500	12.250000

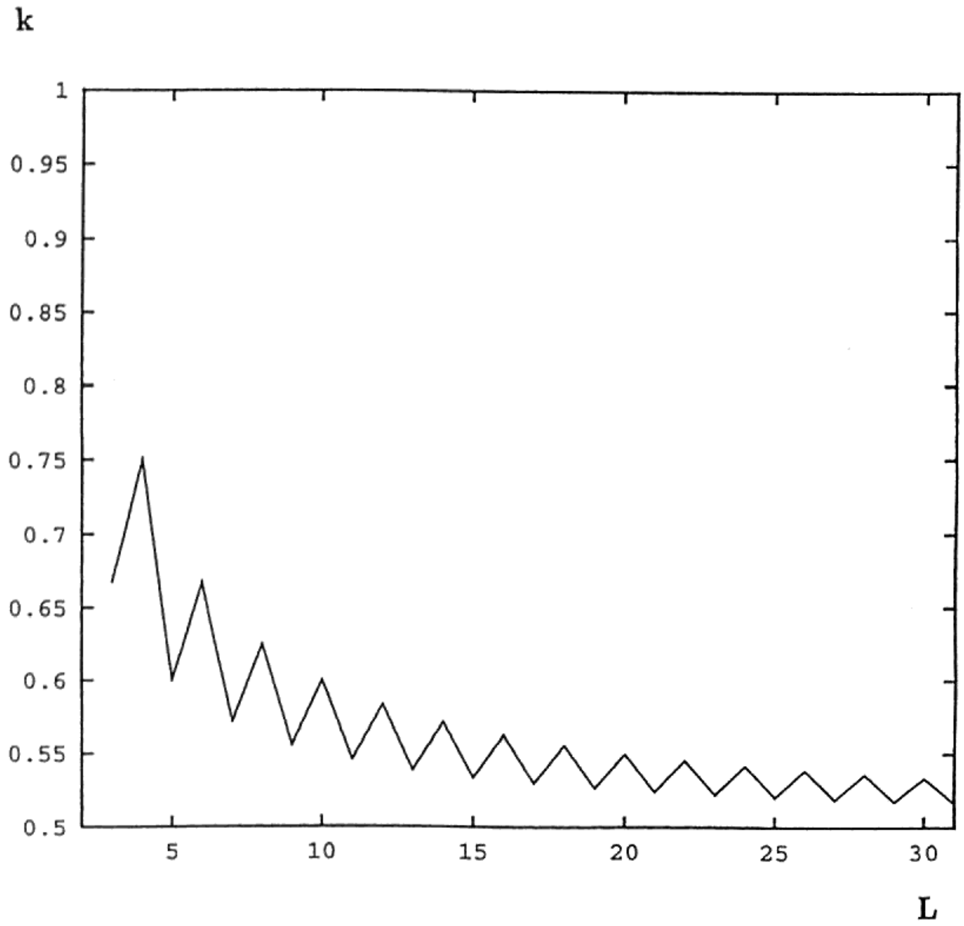

Fig. 4. The minimal ρ in multi-BAM networks with different L.

anteed that a common desired output pattern will be recalled if more than half of the BAMs intend to recall this pattern. If each individual BAM stores less than 5 pattern pairs, it is generally ensured the desired common patterns will be recalled. In the following, we use different multi-BAM network having different number of BAMs to verify the prediction of the majority rule. In the following

Table 2
The strict case and the general case using $\rho = \text{Ceiling}\left(\frac{L-1}{2}\right) + 1$

N (number of pairs)	$L=2$	$L=3$	$L=4$	$L=5$	$L=6$	$L=7$	$23 \geqslant L \geqslant 8$
2	S	S	S	S	S	S	S
3	S	S	S	S	F	F	F
4	S	S	S	S	F	F	F
5	S	S	S	F	F	F	F
6	F	F	F	F	F	F	F
7	F	F	F	F	F	F	F
$\geqslant 8$	F	F	F	F	F	F	F

S = Success, F = Fail.

simulation, Ceiling$(\frac{L-1}{2}) + 1$ of L BAMs support one common output pattern, and all of the pattern pairs are randomly generated. The ρ value of the simulation is shown in Fig. 4. The simulation results are tabulated in Table 2.

4. Conclusion

A multi-BAM neural network has been introduced for the belief combination in evidential reasoning. It is proved to be bidirectionally stable, which ensures the model's ability to reach a local energy minimum. The multi-BAM network facilitates the learning of the belief combination without the loss of robustness, and it also reduces the computation complexity. Two majority rules and their respective bounds for the majority factor, ρ, are presented. These rules will help researchers to use and predict the result of evidential reasoning. The majority rules comply with the intuition of evidential reasoning. This network provides the ability to process many evidence at the same time reaching a consented hypothesis.

References

[1] M. Eggers, T. Khuon, Neural network data fusion concepts and application, Inter. Joint Conf. Neural Network 2 (1990) 7–16.
[2] J. Rajapakse, R. Acharya, Multi sensor data fusion within hierarchical neural network, Inter. Joint Conf. Neural Network 2 (1990) 17–22.
[3] B. Kosko, Adaptive bidirectional associative memory, Appl. Opt. 26 (23) (1987) 4947–4960.
[4] B. Kosko, Bidirectional associative memory, IEEE Trans. Systems Man Cybernet 18 (1) (1988) 49–60.
[5] Y.-F. Wang, J.B. Cruz, Jr., J.H. Mulligan, Jr., Two coding strategies for bidirectional associative memory, IEEE Trans. Neural Network 1 (1) (1990).
[6] S.I. Gallant, Connectionist expert systems Communications of the ACM 31 (2) (1988).
[7] L.S. Hsu, H.H. Teh, S.C. Chan, K.F. Loe, Uncertainty Computation in Neural Networks, Applications of Artificial Networks, SPIE vol. 1294, 1990.
[8] L. Shastri, A connectionist approach to knowledge representation and limited inference, Cognitive Science, vol. 12, (1998) 331–392.
[9] M. Zeidenberg, Neural networks in artificial intelligence, Ellis Horwood, Chichester, UK, 1990.
[10] C.-C. Wang, H.-S. Don, A new approach to evidential reasoning by a potential model, Proceedings of the ISMM International Conference Computer Applications in Design, Simulation and Analysis, Las Vegas, 1991, pp. 192–195.
[11] C.-C. Wang, H.-S. Don, Evidential reasoning using neural networks, International Joint Conference on Neural Networks, Singapore, 1991, pp. 497–502.
[12] H. Tahani, J.M. Keller, Information fusion in computer vision using the fuzzy integral, IEEE Trans. Systems Man Cybernet 20 (3) (1990).
[13] G. Shafer, R. Logan, Implementing Dempster's rule for hierarchieal evidence, Artificial Intell. 33 (1987).
[14] C.-C. Wang, H.-S. Don, An analysis of high-capacity discrete exponential BAM, IEEE Trans. Neutral Networks 6 (2) (1995) 492–496.