1990 International Symposium on Information Theory and Its Applications (ISITA'90)
Hawaii, U.S.A., November 27-30, 1990

# A STOCHASTIC ASSOCIATIVE MEMORY USING MUTUALLY CONNECTED NEURAL NETWORK

Kenji NAKAYAMA and Masanobu TONOMURA

Dept. of Electrical and Computer Eng., Faculty of Tech., Kanazawa Univ.
2-40-20, Kodatsuno, Kanazawa, 920 JAPAN

**ABSTRACT**

This paper presents a new stochastic associative memory. Contents of the memory are classified into the several sets. Relations among elements in two sets are memorized on equilibrium states of a mutually connected neural network. From partial information, the related elements are recollected with probability. Stochastic dynamics is accomplished by randomly selecting a unit for changing network state. Distribution of the equilibrium states is obtained by repeating this dynamics so many times.

## I. INTRODUCTION

Artificial neural networks, consisting of a great number of nonlinear units, and connections among them, has lately attracted considerable attention. They have been applied to pattern recognition, coding, classifiers, feature extraction, and associative memories (Ref.1).

Association is very important intelligence of a human being (Ref.2). There are two basic functions for an associative memory. One of them is to recollect the related matters from the specified information. The other is to determine the order of recollection for several candidates.

The former function requires a product of the several memorized patterns. One model has been reported (Ref.3,4). However, this model seems to be insufficient for recollection from the partial information.

The second function requires a measure to evaluate intensity of association. One useful measure may be probability. This leads a stochastic associative memory. The Boltzmann machine (Ref.5) could be one approach. It, however, takes a long learning time due to simulated annealing. Furthermore, a temperature scheduling problem is very complicated.

In this paper, we propose a new type of stochastic associative memory, taking the above two basic functions into account. It consists of mutually connected neural networks, with asymmetrical connections. Contents of the memory are classified into three kinds of information sets, including categories, articles and attributes. Relations between elements in two sets are stored in a single network. In order to relate elements in three sets, a plural number of the networks are combined. Stochastic dynamics is accomplished by randomly selecting a unit for changing the network state. Examples of computer simulation are demonstrated.

## II. INFORMATION CLASSIFICATION

In a human brain, many things are not independently memorized, but they are related. Therefore, information classification is essential for an artificial associative memory. We employ three kinds of information sets, including category, article and attribute for this purpose.

## III. STRUCTURE AND FUNCTIONS

### Elemental Networks

The elemental network is a mutually connected neural network, as shown in Fig.1. Connection weights are generally asymmetrical. One elemental network stores relations among elements in two information sets. One unit corresponds to one element. The unit takes two levels '1' and '0'. The level '1' means the corresponding element is memorized or is associated.

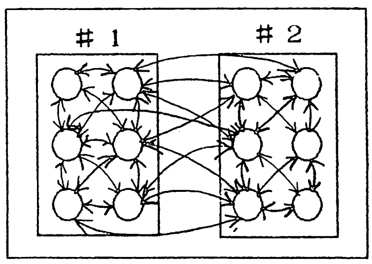


Fig.1 Elemental network, relating the information sets #1 and #2.

Relations among elements can be expressed by patterns of units, whose level is '1'. This pattern is called 'memory pattern' in this paper. A plural number of memory patterns, expressed by $P(m)$, $m=1 \sim M$, are stored on equilibrium states in a single elemental network.

## Network State Transition

Let N be the number of units in the elemental network. Input and output for the ith unit, in the nth transition step, are denoted by $u_i(n)$ and $v_i(n)$. They are related by

$$u_i(n) = \sum_{j=1}^{N} w_{ij}(n) v_j(n) \quad (1)$$

$$v_i(n+1) = f(u_i(n)) \quad (2)$$

where, $w_{ij}(n)$ is a connection weight from the ith unit to the jth unit. A nonlinear function f() is defined by

$$v_i(n+1)=1, \quad u_i(n)>T \quad (3a)$$

$$v_i(n+1)=v_i(n), \quad u_i(n)=T \quad (3b)$$

$$v_i(n+1)=0, \quad u_i(n)<T. \quad (3c)$$

T is a threshold level.

## Combined Network

An entire network, expressing all relations, is constructed by combining several elemental networks. One example is shown in Fig.2. The network Net-1 stores relations between the sets #1 and #2, and Net-2 realizes relations between the sets #2 and #3, respectively. The entire associative memory consists of Net-1 and Net-2.

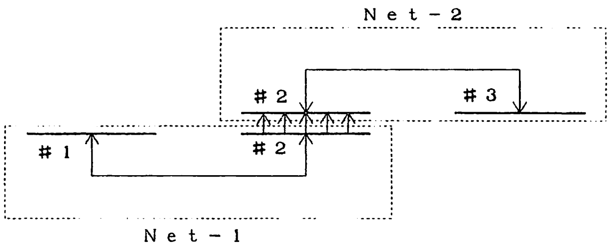


Fig.2 Combined form of two elemental networks Net-1 and Net-2.

In this model, association from the set #1 to #3 is carried out through Net-1 and Net-2. the details will be discussed in Sec.V.

Although it is possible to employ the third network for directly relating between #1 and #3. However, we must consider lack of information provided in advance. In this case, the above multi-stage association process is required.

## Functions

The following functions are taken into account.

(1)Association from partial information:

The memorized elements are recollected from the specified partial information. Probabilities are provided to express association intensity.

(2)Association from combined information:

After learning relations between two sets on a network, recollection from the elements in several information sets is performed.

(3)Self-organization:

When relations between the sets #1 and #2, and between the sets #2 and #3 are provided. They are stored in Net-1 and Net-2, respectively. Relations between the sets #1 and #3 are estimated through Net-1 and Net-2, as described previously. A new elemental network Net-3, realizing the estimation results, is self-organized.

# IV. LEARNING ALGORITHM

Connection weights are adjusted so as to store all memory patterns on equilibrium states. A learning algorithm for the elemental networks is descried here.

## Initial Guess

Initial connection weights are determined to be zero. Since connection weights are adapted by an additive rule, like the Hebbian rule (Ref.1), this initial guess does not affect a learning process. Furthermore, connection weights between unrelated units are desirable to remain zero.

## State of Network

During a learning process, the network state is fixed to memory patterns P(m), m=1～M.

## Threshold Level for Unit Activation

After a learning process is completed, undesirable equilibrium states can exist. Under this situation, recollection from partial information requires to stabilize the equilibrium state. In other words, a valley of an energy function, should be deepened. For this purpose, the threshold level for the jth unit is chosen as follows:

$$\text{If } v_j(m)=1 \text{ in } P(m), \text{ then } T=+\theta \quad (4a)$$

$$\text{If } v_j(m)=0 \text{ in } P(m), \text{ then } T=-\theta \quad (4b)$$

$$\theta > 0 \quad (4c)$$

$\theta$ is taken to be a large number.

## Updating Connection Weights

(1)The network state is fixed to a memory pattern P(m). Connection weights are updated by

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n) \quad (5)$$

Furthermore, $\Delta w_{ij}(n)$ is determined by

(a)$v_j(m)=1$ and $u_i(m)=1$

$$\text{If } u_j(m) \geqq +\theta, \text{ then } \Delta w_{ij}(n)=0 \quad (6a)$$

$$\text{If } u_j(m) < +\theta, \text{ then } \Delta w_{ij}(n)=+\Delta w \quad (6b)$$

where $\Delta w>0$

(b)$v_j(m)=0$ and $u_i(m)=1$

$$\text{If } u_j(m) \leqq -\theta, \text{ then } \Delta w_{ij}(n)=0 \quad (7a)$$

$$\text{If } u_j(m) > -\theta, \text{ then } \Delta w_{ij}(n)=-\Delta w \quad (7b)$$

(c)$v_j(m)=1$ or 0 and $u_i(m)=0$

$$\Delta w_{ij}(n)=0 \quad (8)$$

All connection weights are updated simultaneously for P(m).

(2)The above adaptation is repeated for all memory patterns P(m), m=1～M. This process is counted as one iteration in the learning process.

(3)The above processes (1) and (2) are repeated until $\Delta w_{ij}(n)$ becomes zero. During the learning process, $\Delta w$ is gradually decreased, in order to suppress vibration, which may occur in a mutually connected neural network.

## V. ASSOCIATION PROCESS

### Single Association Process

First, a single association process on the elemental networks is described. In the initial network state, the units, corresponding to the specified elements, take a high level '1', and the rests take a low level '0'. Network dynamics follows Eqs.(1)～(3).

The threshold level T is decreased from that used in the learning process, in order to stabilize the equilibrium state. Actually, T is set to be zero. Therefore, Eq.3 is modified as follows:

$$v_j(n+1)=1, \quad u_j(n)>0 \qquad (9a)$$
$$v_j(n+1)=v_j(n), \quad u_j(n)=0 \qquad (9b)$$
$$v_j(n+1)=0, \quad u_j(n)<0 \qquad (9c)$$

When all units no longer change their state, then the network settles down onto an equilibrium state. In this state, the units having a high level '1', besides the initially specified, express associated elements.

### Stochastic Association Process

If the initial network state locates on the way to one deep valley of an energy function, then the state will monotonously converge into that valley.

On the other hand, if it locates on a saddle point among several valleys as shown in Fig.3, the final destination is dependent on not only the specified partial information, but also the order of selecting units for changing the state. By randomizing the order of selections, we can introduce random shake effects on a convergence route.

Thus, statistical association is accomplished by randomly selecting units, and changing network state unit by unit.

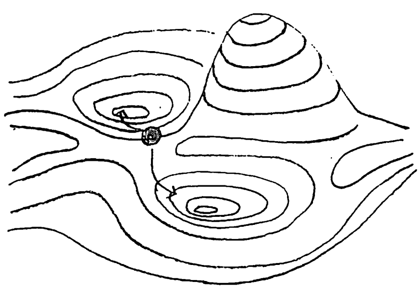

Fig.3 Energy function in network state space. A black ball indicates the initial state.

This random state transition is repeated so many times by changing the order of selections. Let the number of the resulting equilibrium states be K, and the number of the equilibrium states, where the jth unit takes a high level '1', be $K_j$. Then probability of associating the element corresponding to the jth unit is given by $K_j/K$ (x100 %). This probability is also called 'association rate' in this paper.

### Consistency of Specified Information

If the specified elements are consistent with each other, their initial state are fixed during the association process. In this case, the resulting equilibrium state includes all specified elements.

On the other hand, if the specified elements are not consistent with each other, the equilibrium states, including all such elements are not adjusted. Therefore, it is not desirable to fix state of the corresponding units. They are also selected for changing their state. Some part of the specified elements still remain in the resulting equilibrium state. The associated elements are related to the remaining elements. What part of the specified elements will remain is dependent on not only the initial state but also the order of selecting units.

### Unexpected Association

Two kinds of equilibrium states, expressing a memorized pattern and a non-memorized pattern can exist. The latter is not exactly related to the specified information, but it is partially related, or indirectly related. This kinds of equilibrium states produce vague association with low probability. It may be some time meaningful. Because a human being also do such association.

### Multi-Network Association

Relations, provided in advance, are not always complete. So, the following situation may occur. Relations between #1 and #2, and between #2 and #3 are provided, and are stored in the elemental networks Net-1 and Net-2, respectively. On the other hand, a relation between the sets #1 and #3 is not provided. This model was shown in Fig.2. Under this situation, estimating the related elements in #3, from the specified elements in #1, requires association stretch over several networks. The following two methods are proposed.

(1) Net-1 is initially set to be the specified elements in #1. After the statistical dynamics on Net-1, the elements in #2 are associated with probabilities. Next, these probabilities are transferred to the units of #2 in Net-2. They are used as the initial activation levels. After the statistical dynamics, the related elements in #3 are associated with probabilities.

(2) First, a single association is carried out on Net-1. The unit numbers of the associated elements are transferred to #2 in Net-2. State of the same units in Net-2 are initialized to be '1'. A single association is also carried out on Net-2, resulting one equilibrium. By repeating the above association, a set of equilibrium states results. From this set, probabilities for the associated elements in #3 are calculated.

## VI. SIMULATION

Information sets consist of categories, articles and attributes. As a whole, 14 categories, 20 articles and 166 attributes are included. In the learning process, $\Delta w$ is decreased as follows: $\Delta w=1$, $n=1\sim50$, $\Delta w=0.5$, $n=51\sim100$, $\Delta w=0.2$, $n=101\sim150$. Furthermore, $\theta$ is chosen to be 40. In the association process, the threshold level is zero. The number of equilibrium states is 1000.

Fig.4 shows examples for association of vegetables (articles) from their attributes, color, shape, size and taste. The relations between articles and attributes, corresponding to the Fig.4. are listed in Table 1. By increasing the specified attributes, the associated vegetables are gradually limited. The association rates in each step are reasonable compared with Table 1.

## VII. CONCLUSIONS

A new stochastic associative memory has been proposed. Relations, provided in advance, are stored on the equilibrium states. By repeating network state transition so many times, statistical distribution of the equilibrium states is obtained. From this distribution, probabilities for the associated elements are calculated. Examples of computer simulation have demonstrated that reasonable association rates can be obtained.

## REFERENCES


[1]D.E.Rumelhart and J.L.McClelland,"Parallel Distributed Processing," MIT Press, 1986.

[2]T.Kohonen,"Self-Organization and Associative Memory," Springer-Verlag, 1984.

[3]Hirai,"A model of human associative processor (HASP)," IEEE Trans on SMC, vol.SMC-13, no.5, pp.851-857 1983.

[4]Y.Hirai,"Mutually linked HASP's: A solution for constraint-satisfaction problems by associative processing," IEEE Trans on SMC, vol.SMC-15, no.3, pp.432-442 1985.

[5]D.H.Ackley, G.E.Hinton and T.J.Sejinowski,"A learning algorithm for Boltzmann machine," Cognitive Sciences, no.9, pp.147-169 1985.


Table 1 Relations between articles (vegetables) and attributes. A part of the whole information is listed.

| Article | Attribute | | | | |
|---|---|---|---|---|---|
| | color | shape | size | taste | others |
| lettuce | white | with leaves | middle | watery | thin |
| cabbege | white | with leaves | large | lightness | heavy |
| Chinese cabbege | white | with leaves | large | watery | pickle, kimchi |
| Chinese nettle tree | white | parasol | small | lightness | bound together with egg |
| radish | white | long | large | pungent | pickle, foot |
| Welsh onion | white | long | middle | pungent | |
| bamboo shoot | white | | large | lightness | pimply |

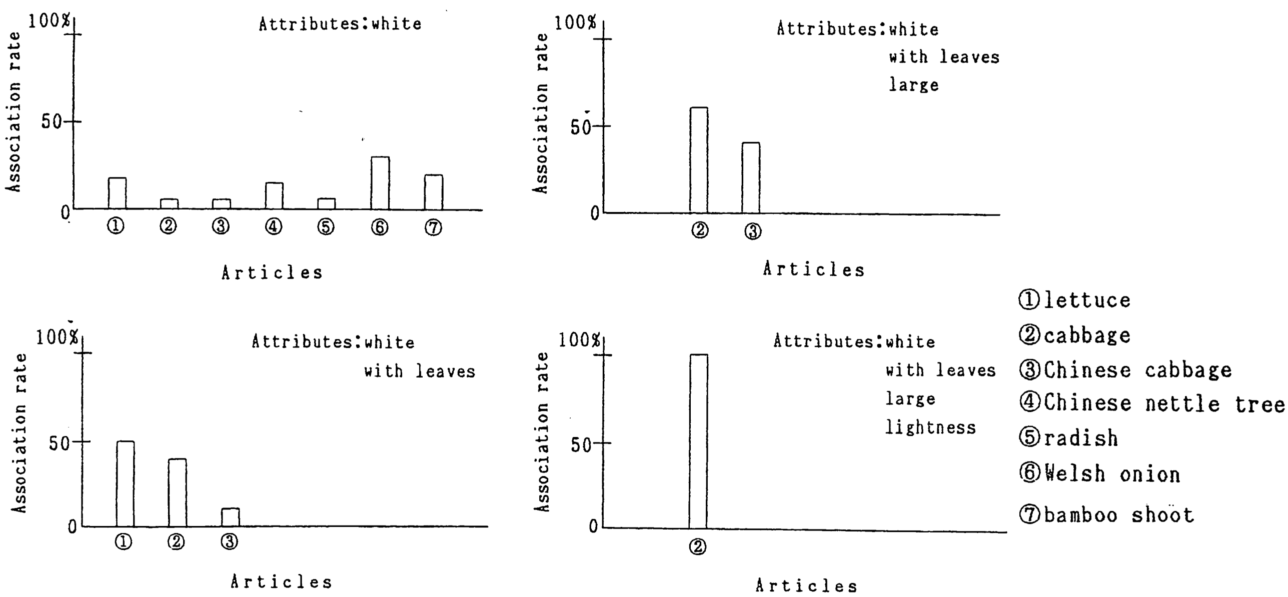


Fig.4 Examples for association of vegetables from their attributes.