NEURAL NETWORK DESIGN FOR EFFICIENT INFORMATION RETRIEVAL

L. PERSONNAZ, I. GUYON and G. DREYFUS
E.S.P.C.I., Laboratoire d'Electronique
10, rue Vauquelin
75005 PARIS

The ability of neural networks to store and retrieve information has been investigated for many years. A renewed interest has been triggered by the analogy between neural networks and spin glasses which was pointed out by W.A. Little et al.[1] and J. Hopfield[2]. Such systems would be potentially useful autoassociative memories "if any prescribed set of states could be made the stable states of the system"[2] ; however, the storage prescription (derived from Hebb's law) which was used by both authors did not meet this requirement, so that the information retrieval properties of neural networks based on this law were not fully satisfactory. In the present paper, a generalization of Hebb's law is derived so as to guarantee, under fairly general conditions, the retrieval of the stored information (autoassociative memory). Illustrative examples are presented.

DESCRIPTION OF THE NETWORK : we consider a fully connected network of n McCulloch-Pitts formal neurons, with simultaneous, parallel operations. Each neuron is a binary state threshold device having n inputs (the states of all neurons) and one output (its own state). At time t, the state of neuron i is represented by a binary variable $\sigma_i(t)$ which can take the numerical values of +1 or -1. In order to determine its next state $\sigma_i(t+\tau)$, the neuron i performs a weighted sum of its inputs and compares it to a threshold value θ_i :

$$\sum_{j=1}^{n} C_{ij}\sigma_j(t) \begin{cases} > \theta_i \Rightarrow \sigma_i(t+\tau) = +1 \\ < \theta_i \Rightarrow \sigma_i(t+\tau) = -1 \\ = \theta_i \Rightarrow \sigma_i(t+\tau) = \sigma_i(t) \end{cases}$$

Therefore, the parameters of the network are the (n,n) matrix C of the weights $\{C_{ij}\}$ and the (n) vector $\underline{\theta}$ of the threshold values $\{\theta_i\}$. The state of the network is defined by an

NATO ASI Series, Vol. F20
Disordered Systems and Biological Organization
Edited by E. Bienenstock et al.

(n) vector $\underline{\sigma}$, the components of which are the states $\{\sigma_i\}$ of all the neurons.

NEURAL NETWORKS AS AUTOASSOCIATIVE MEMORIES : for the network to be useful as an autoassociative memory, it should have the following behaviour : if the network is set into a given state $\underline{\sigma} \in \{-1,+1\}^n$ (representing a distorted or incomplete information), it should evolve until it reaches a stable state (representing the full information to be retrieved). The problem of designing a neural network acting as an autoassociative memory is therefore : given a set of p prototype states to be memorized, how should the parameters C and $\underline{\theta}$ be chosen so as to retrieve these states as faithfully as possible ? Obviously, the minimum requirement is that the prototype states be stable ; very desirable features would be :

(i) the fact that prototype states be stable and act as attractors,

(ii) the absence of cycles,

(iii) the absence of spurious stable states or, at least, their predictability.

In the following, we show how to design networks embodying these three features.

THE GENERALIZED HEBB'S LAW : the general stability condition of a network can be expressed as follows[3] if all thresholds are taken equal to zero : a given state $\underline{\sigma}$ is stable if and only if there exists a diagonal matrix A, with all elements positive or zero, such that one has :

$$C\underline{\sigma} = A\underline{\sigma} \qquad (1)$$

In order to make the p prototype states $\{\underline{\sigma}^k\}$ stable, relation (1) must hold true for all of them :

$$C\underline{\sigma}^k = A^k\underline{\sigma}^k$$

A^k being a diagonal matrix with all its elements positive or zero. The general problem of finding C has been addressed in Ref. 3. In the present paper we take A^k=I for all k. If the prototype states are linearly independent, the solution is given by : $C=\Sigma\Sigma^I$ (2)

where Σ is the matrix whose columns are the prototype vectors $\{\underline{\sigma}^k\}$ and $\Sigma^I = (\Sigma^T\Sigma)^{-1}\Sigma^T$ is the pseudoinverse[4] of matrix Σ. The matrix C is symmetric. It is the orthogonal projection matrix into the subspace spanned by the prototype vectors. It can be computed either directly from relation (2), or recursively[5], without matrix inversion, by introducing each prototype vector once (and only once). This recursive computation is typical of a learning process.

One should notice that relation (2) is a generalized form of the classical Hebb's law $C=(1/n)\Sigma\Sigma^T$: in the special case where the prototype states are orthogonal, relation (2) reduces exactly to Hebb's law. Therefore, since our storage prescription guarantees the retrieval of the prototype states even if the latter are not orthogonal, it is a generalization of Hebb's law.

PROPERTIES OF THE GENERALIZED HEBB'S LAW : besides the fundamental feature of guaranteeing the stability of the prototype states, we show some additional properties of such networks : the absence of cycles, the characterization of the spurious stable states and the attractivity of the prototype states.

i- Absence of cycles. The "energy" of the network in the state $\underline{\sigma}$ can be defined by analogy with spin glasses in the absence of external field: $E=-\frac{1}{2}\underline{\sigma}^T C\underline{\sigma}$.

Assume that the system is in an unstable state $\underline{\sigma}$ and that the next parallel iteration drives it to a state $\underline{\sigma}'$. It can be shown[6] that : $\underline{\sigma}^T C\underline{\sigma} < \underline{\sigma}'^T C\underline{\sigma} < \underline{\sigma}'^T C\underline{\sigma}'$

Therefore, the energy is an ever decreasing function, thus preventing any cycle to occur even under parallel operation.

It should be noticed that the prototype states and their linear combinations belonging to $\{-1,+1\}^n$ (if any) are the states of lowest possible energy (they are identical to the Mattis states referred to by D. Amit et al. in the present book).

ii- Spurious stable states. Consider a stable state $\underline{\sigma}$, which is possibly a non prototype state. Since C is the matrix of the orthogonal projection into the subspace spanned by the prototype vectors, $C\underline{\sigma}$ is a linear combination of the prototype vectors.

Since $\underline{\sigma}$ is stable, from relation (1), there exists a positive diagonal matrix A such that :

$$C\underline{\sigma} = A\underline{\sigma}$$

Vector $\underline{\sigma}$ is therefore a "normalized" combination of the prototype states. Thus, any stable state is a normalized combination of the prototype states.

iii- Attractivity of orthogonal prototype states. If the prototype states are orthogonal, it has been shown[6] that any state lying within a Hamming distance of n/2p of a given prototype state will converge to that state in one iteration. Similarly, the opposite of any prototype state has a radius of attraction of n/2p.

EXAMPLES : we present an example which illustrates the efficiency of the generalized Hebb's law : a neural network designed after this law is used for error correction purposes. The titles of scientific journals have been chosen as prototype patterns. Each alphabetic character has been coded on six bits. The prototype states are shown in the upper left block of the figure. Each example in the other two blocks has three lines : the first one is the initial state ; the second and third lines are the final states reached by neural networks designed with the generalized Hebb's law and with Hebb's law respectively. These two networks have exactly the same structure; the only difference between them is the analytic expression of the matrix C. As was mentioned above, the numerical computation of that matrix is performed in both cases by an algorithm which yields the exact result (within roundoff errors) after a finite number of steps (equal to the number of prototype vectors). In the upper right block, the retrieval of the prototype states is attempted ; as expected, all the prototype states are retrieved in the first case, whereas several are forgotten in the second case. The lower block shows the error correction properties ; obviously, the generalized Hebb's law is much more efficient than Hebb's law for error correction.

CONCLUSION : we have proposed a generalization of Hebb's law which guarantees a perfect retrieval of the information stored in a neural network. This approach has enabled us to demonstrate the absence of cycles under parallel iteration conditions, to evaluate the attractivity of the prototype states, and to clarify the nature of the spurious stable states. The examples that are presented show that such networks exhibit reliable properties of autoassociative memories.

(1) W.A. LITTLE , G.L. SHAW, Math. Biosciences **39**, 281 (1978).

(2) J. J. HOPFIELD, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).

(3) L. PERSONNAZ, I. GUYON, G. DREYFUS, J. Physique Lett. **46**, L-359 (1985).

(4) A. ALBERT, Regression and the Moore-Penrose pseudoinverse (Academic Press, 1972).

(5) T.N.E. GREVILLE, SIAM Rev. **2**, 5 (1960).

(6) To be published.

PROTOTYPE STATES ****************
PHYSICAL REVIEW LETTERS
JOURNAL OF NEUROBIOLOGY
BIOLOGICAL CYBERNETICS
BULL. OF MATH BIOPHYSICS
J. PHYSICAL OCEANOGRAPHY
PROG. THEOR. PHYS. KYOTO
SPECULAT. SCI. TECHNOL.
INDIAN J PURE APPL. MATH

PHYSICAL REVIEW LETTERS	J. PHYSICAL OCEANOGRAPHY
PHYSICAL REVIEW LETTERS PHYSICAL REVIEW LETTERS	J. PHYSICAL OCEANOGRAPHY J. PHYSICAL OCEANOGRAPHY
JOURNAL OF NEUROBIOLOGY	PROG. THEOR. PHYS. KYOTO
JOURNAL OF NEUROBIOLOGY JOURNAL OF NEUROBIOLOGY	PROG. THEOR. PHYS. KYOTO PROG. THEOR. PHYS. KYOTO
BIOLOGICAL CYBERNETICS	SPECULAT. SCI. TECHNOL.
BIOLOGICAL CYBERNETICS CPPKKEAD P RH QODPLORT	SPECULAT. SCI. TECHNOL. SPECULAT. SCI. TECHNOL.
BULL. OF MATH BIOPHYSICS	INDIAN J PURE APPL. MATH
BULL. OF MATH BIOPHYSICS COPKK TE O RE BPRIPLOCT	INDIAN J PURE APPL. MATH INDIAN J PURE APPL. MATH

PHISICAL REFIEX LETTER	BIOL. CYBERN.	MROG. CHEOR. PHYS. TOKYO
PHYSICAL REVIEW LETTERS PHYRICAI REVIEW LETTERS	BIOLOGICAL CYBERNETICS COPKK TE O RE BPRIPLOCT	PROG. THEOR. PHYS. KYOTO PROG. THEOR. PHYS. KYOTO
PHYSICAL REV. LET.	DIOLORICAL CIBERNETICS	SPECULATION ET TECHNOL.
PHYSICAL REVIEW LETTERS DOQKKFAI O RE BPRIQYPCT	BIOLOGICAL CYBERNETICS CPPKKEAD P RH QODPLORT	SPECULAT. SCI. TECHNOL. DOQKK TH O RE BPRLPYPCT
JØNRNEL AF NOURABIALAGI	J. PHICINAL OCEAN.	INDIEN I PERE EPPL. NATH
JOURNAL OF NEUROBIOLOGY COPKK TE O RE BPRIPLOCT	J. PHYSICAL OCEANOGRAPHY DOQKK TH O RE BPRLPYPCT	INDIAN J PURE APPL. MATH INDIAN J PURE APPL. MATH