Information storage and retrieval in spin-glass like neural networks

L. Personnaz, I. Guyon, G. Dreyfus

Classification
Physics Abstracts
05.90

Information storage and retrieval in spin-glass like neural networks

L. Personnaz, I. Guyon and G. Dreyfus

Ecole Supérieure de Physique et de Chimie Industrielles de la Ville de Paris,
Laboratoire d'Electronique, 10, rue Vauquelin, 75005 Paris, France

Résumé. — Dans la perspective de la réalisation de mémoires associatives à l'aide de réseaux de neurones, nous étudions la relation entre la structure d'un réseau et ses états attracteurs; nous montrons que, quel que soit l'ensemble des états que l'on désire mémoriser, il est généralement possible de calculer tous les paramètres du réseau de façon à assurer la stabilité de ces états. Le formalisme des verres de spins conduit à des résultats particulièrement simples qui permettent, dans certains cas, d'évaluer analytiquement leur attractivité.

Abstract. — The link between the structure of a neural network and its attractor states is investigated, with a view to designing associative memories based on such networks. It is shown that, for any preassigned set of states to be memorized, the parameters of the network can be completely calculated in most cases so as to guarantee the stability of these states. The spin glass formulation of the neural network problem leads to particularly simple results which, in some cases, allow an analytical evaluation of the attractivity of the memorized states.

1. Introduction.

It is more and more widely admitted that there is a deep connection between the behaviour of complex systems and that of spin glasses; the conceptual tools that were developed in statistical physics begin to be applied in various domains of engineering and optimization [1-4]. Large networks of non-linear elements have been shown to exhibit such intriguing properties as learning, information storage and retrieval, due mainly to the presence of attractors [5, 6]. Specifically, networks of highly stylized neurons have been shown to behave as distributed associative memories [7-9], which might be used in the context of pattern recognition and image processing, since such highly regular structures can be implemented with a high degree of integration in VLSI devices and may be fault tolerant.

We shall first recall briefly the definition of the neural network under investigation. We shall derive, in a compact from, the general stability condition. In a subsequent section, the synthesis of a network, that is, the determination of its parameters so that a given set of states be stable, shall be investigated; the spin-glass formulation is shown to be efficient and to lead to simple results about the stability and the attractivity of the memorized states. The relation between this approach and the Hopfield model is discussed.

2. Neural net model.

In the present paper, we use the McCulloch-Pitts model of neurons [10], whereby a neuron is a two-state, threshold element having several inputs (synaptic junctions) and one output (axon). The state of a neuron i at time t is represented by a variable λ_i which can have the numerical value H (high) if the neuron is active, and the value L (low) if the neuron is inactive. The values of H and L which will be considered in this paper are either $(H = 1, L = 0)$ or $(H = 1, L = -1)$. The value of λ_i is transmitted to other neurons by the axon. The state of a neuron at time $t + \tau$ depends on the states of its inputs at time t in the following way : each synaptic junction being characterized by its synaptic strength, the neuron performs the sum of the input signals weighted by the synaptic strengths (i.e., computes the « membrane potential »), and compares it to a threshold value θ_i. If the sum is larger than the threshold, the neuron goes to (or remains in) the active state; if the sum is smaller than the threshold, the neuron goes to (or remains in) the inactive state; if the sum is equal to the threshold, the neuron remains in its previous state :

$$\sum_{j=1}^{n} C_{ij} \lambda_j(t) > \theta_i \Rightarrow \lambda_i(t + \tau) = H$$

$$\sum_{j=1}^{n} C_{ij} \lambda_j(t) < \theta_i \Rightarrow \lambda_i(t + \tau) = L \qquad (1)$$

$$\sum_{j=1}^{n} C_{ij} \lambda_j(t) = \theta_i \Rightarrow \lambda_i(t + \tau) = \lambda_i(t) .$$

C_{ij} is the strength of the synaptic junction of neuron i receiving information from neuron j, and n is the number of synaptic junctions of neuron i.

We consider a network of n such neurons, without « sensory inputs » : we assume that the system has been set into an initial state, and we are interested in the spontaneous evolution of the network. The neurons reevaluate their membrane potentials synchronously at intervals of time τ. Since any neuron can be connected to any other neuron, the set of synaptic strengths can be put in the form of a square matrix C, the dimension of which is the number n of neurons.

The state of the neural network is defined by the states of all its neurons : it can be expressed as a vector λ, the components of which are the states of the neurons. Similarly, a threshold vector θ can be defined. A state of the network is said to be stable if it is invariant in time.

The collective behaviour of such nonlinear networks has been shown to exhibit associative properties because their stable states can act as attractors [8]. The system thus embodies an associative memory [11, 12], which is able to retrieve a full information if given a distorted or incomplete version of it : assume that the system has been set into a state which is different from all stable states; if the latter act as attractors, the system will evolve (according to relations (1)) until it reaches one of them. The stable (memorized) states, if any, depend on the synaptic matrix and on the thresholds of the neurons. The computation of the synaptic matrix and, possibly, of the thresholds, can thus be considered as a learning process. However, the link between these parameters and the stable states is far from being clearly understood. The main problem can be formulated as follows : given a set of states to be memorized, hereinafter referred to as prototype states, how should the synaptic matrix and the thresholds be chosen so as to retrieve these states as efficiently as possible ? Obviously, the minimum requirement is that the prototype states be stable; very desirable features would be :

  i) the fact that prototype states are stable and act as attractors,
  ii) the absence of spurious stable states, or, at least, their predictability.

Up to now, the networks of the type discussed here have been studied essentially by numerical simulations. By making use of Cooper's storage prescription [16] for computing the synaptic

matrix from a given set of prototype states, and by taking all thresholds equal to zero, it was shown by J. Hopfield [8] that the prototype states are not always stable, and that many other stable states, different from the prototype ones, arise; these phenomena tend to decrease the efficiency of information retrieval. An upper limit for the storage capacity was determined. One of the merits of the Hopfield approach is that it establishes a striking similarity between neural nets and spin glass systems, so that the analytic tools developed for the study of spin glasses may be used in the investigation of such networks.

Nevertheless, powerful as statistical methods can be, deterministic predictions are necessary for such practical purposes as, for instance, implementing an associative memory in silicon for performing such a task as pattern recognition. The present paper is an attempt at finding analytical results which may give some insight into the behaviour of the network and lead to the design of associative memories satisfying the minimum requirement for such devices to be usable : the states to be memorized must be stable.

3. General stability condition.

The stability of a state of the network is expressed as :

$$\left(\sum_{j=1}^{n} C_{ij} \lambda_j - \theta_i \right) \sigma_i \geqslant 0 \qquad (2a)$$

for all $i = 1, 2, ..., n$, where σ_i is defined by :

$$\sigma_i = 1 \qquad \text{if} \quad \lambda_i = H$$
$$\sigma_i = -1 \qquad \text{if} \quad \lambda_i = L.$$

This condition can be expressed in an alternate form : a state described by a vector λ is stable if and only if there exists a diagonal matrix A, with all its elements positive or zero, satisfying the following relation :

$$C\lambda - \theta = A\sigma. \qquad (2b)$$

Two problems arise in practise, namely, the analysis and the synthesis of a neural network :

i) the analysis of a network can be expressed as follows : given a network, described by the numerical values of H and L, the matrix C and the vector θ, what is the behaviour of the network ? The answer to this question lies in relations (1) and (2a) as far as the evolution and the stability of the states are concerned. Therefore, relation (2b) is not likely to be a very efficient tool for the analysis of the network;

ii) the synthesis of a network can be expressed as follows : given a set of prototype states $\{ \lambda^k \}$, design a network that guarantees the stability of these states; therefore, the following relation :

$$C\lambda^k - \theta = A^k \sigma^k \qquad (3)$$

should be satisfied for all k. The designer can choose freely any suitable set of matrices A^k, provided they are diagonal with all elements positive or zero. Once these matrices are chosen, relation (3) can be used to compute C and θ; the convenience of this relation arises from the fact that it allows to design an efficient threshold network, with preassigned stable states, using simple linear algebra. Applications of this formulation will be presented in the next sections.

4. Synthesis of a neural network.

In this section, we show how to synthesize a neural network, i.e. to determine the synaptic matrix and, if necessary, the thresholds, in order to memorize a given set of p prototype states. In a first part, a general design method is proposed; it can be applied whatever the chosen values

of H and L. It is shown that a synaptic matrix which guarantees the stability of the prototype vectors can be readily calculated. In a second section, we address the synthesis problem in the special case where $H = 1$ and $L = -1$ (spin-glass formulation); with a simple choice of the thresholds, the synaptic matrix can also be easily computed. It is shown that Cooper's storage prescription guarantees the stability of the prototype states only if the latter are orthogonal. Finally, the Hopfield model is discussed in view of our results.

4.1 GENERAL FORMULATION. — If the p prototype states are denoted by λ^k, the stability condition (3) for a prototype state is :

$$C\lambda^k = \mathbf{f}^k,$$

where $\mathbf{f}^k = A^k \sigma^k + \theta$.

In the most general case, the thresholds and the matrices A^k can be chosen freely (the latter are subjected to the condition of being positive diagonal). If the vectors λ^k are linearly independent, the solution of the above matrix equation is given by [13] :

$$C = F\Lambda^I, \tag{4}$$

where F is the matrix whose columns are the vectors $\mathbf{f}^k$
  Λ is the matrix whose columns are the vectors λ^k,
and Λ^I is the pseudo-inverse of matrix Λ :

$$\Lambda^I = (\Lambda^T \Lambda)^{-1} \Lambda^T.$$

The pseudo-inverse is easily computed recursively, by introducing the prototype vectors sequentially, without matrix inversion [14]. This allows to add a new prototype vector once a synaptic matrix has been computed without having to compute the whole matrix again : this is similar to most learning processes.

The meaning of the synaptic matrix C is the following : applying Λ^I to a vector λ yields the p coefficients of the orthogonal projection of λ in the subspace spanned by the λ^k's; therefore, $C\lambda$ is a linear combination of the $\mathbf{f}^k$'s with the above coefficients. In general, this matrix is not symmetric. This method completely determines a neural network which satisfies the minimum requirement that the prototype states be stable. It should be noticed that there is no condition on the $\{ \mathbf{f}^k \}$ family.

Two cases may arise in practise :

a) Networks in which the thresholds are predetermined. Technological considerations may lead to a particular choice of the thresholds, whereas the matrices A^k can still be chosen freely. C can therefore be computed by relation (4).

b) Networks in which the thresholds are not predetermined. The interpretation of matrix C is made simpler if one chooses A^k and θ so that :

$$\mathbf{f}^k = \lambda^k.$$

The stability condition (3) thus becomes :

$$C\lambda^k = \lambda^k. \tag{5}$$

In this case, C is the projection matrix of the n-dimensional state space into the subspace spanned by the prototype vectors :

$$C = \Lambda\Lambda^I. \tag{6}$$

The synaptic matrix is symmetric (which is highly questionable from a biological standpoint).

If the prototype vectors are orthonormal, one has :

$$\Lambda^T\Lambda = I,$$

where I is the identity matrix. Matrix C thus reduces to :

$$C = AA^{\mathrm{T}}.$$

A full discussion of the behaviour of neural networks in this general formulation lies beyond the scope of this paper. We shall now restrict our study to a case which is of special interest and is closely related to previous approaches.

4.2 SPIN-GLASS FORMULATION. — In most classical models of neural networks, the values of H and L are taken as 1 and 0 respectively. However, the striking analogy with spin glasses leads us to consider directly state vectors $\lambda = \sigma$, for which $H = 1$ and $L = -1$ [15]. In this section, we show that several simple analytical results can be obtained with such a formulation.

The stability condition for the p prototype state vectors σ^k may be rewritten in the form of relation (5) if one makes the simple choice :

$$A^k = I \quad \text{for all } k, \quad \text{and} \quad \theta = 0.$$

Therefore, the stability equation reduces to :

$$C\sigma^k = \sigma^k. \tag{7}$$

If the prototype vectors are linearly independent, the non-trivial solution is given by relation (6) :

$$C = \Sigma\Sigma^{\mathrm{I}}$$

where Σ is the matrix whose columns are the vectors σ^k.

Two cases of particular interest may be considered :

a) The prototype vectors are orthogonal : $(\Sigma^{\mathrm{T}} \Sigma)^{-1}$ reduces to $(1/n)\, I$, so that matrix C is nothing but the storage prescription advocated by L. Cooper [16] :

$$C = (1/n)\, \Sigma\Sigma^{\mathrm{T}}. \tag{8}$$

b) The prototype states are random vectors the components of which are $+ 1$ or $- 1$ with probability 1/2. In the thermodynamic limit this case is similar to the previous one with an additional special property : the number of 1's appearing in the prototype vectors is equal to the number of $- 1$'s. It should be mentioned that this thermodynamic limit case can be simulated conveniently with a finite number of neurons, provided n is an integer power of 2, if the prototype states are a subset of the Walsh functions.

In case (a) or in the thermodynamic limit of case (b), the behaviour of the stable states as attractors can be studied analytically; the following results, the demonstration of which will be given in a more detailed paper, can be derived :

i) The key parameter for the behaviour of the network is the Hamming distance $H(\sigma, \sigma')$ between states σ and σ' (i.e. the number of neurons whose state should be changed in order to turn state σ into state σ') or, more specifically, the quantity $n - 2\, H(\sigma, \sigma')$: this is not surprising since it is equivalent of the overlap between spin glass states [17]. All the prototype states are equidistant, the distance between them being $n/2$.

ii) Any prototype state has a radius of attraction which is at least equal to $n/2\, p$: any state lying within a distance of $n/2\, p$ of any prototype state will be attracted by that state. Therefore, the « attractivity » of a prototype state can be evaluated by the ratio of the minimum number of states that it will attract to the total number of states $(1/2)^n \sum_{k=0}^{n/2\, p} C_n^k.$

Consequently, if p becomes of the order of $n/2$, the attractivity falls sharply (or, equivalently, the number of spurious stable states increases).

iii) Similarly, the negative of any prototype state has a radius of attraction $n/2\, p$.

4.3 RELATION TO THE HOPFIELD MODEL. — We now have the necessary ingredients to discuss the relation of the present approach to the Hopfield model as described in reference [8]. In this model, the values of H and L are taken equal to 1 and 0 respectively; therefore, the stability of the prototype vectors is expressed by relation (3). The state vector λ is related to σ by :

$$\lambda = (1/2)\,(\sigma + 1)\,,$$

where **1** is a vector the components of which are all equal to 1. Thus, relation (3) can be rewritten as :

$$C\sigma^k = 2\,A^k\,\sigma^k + 2\,\theta - C1\,.$$

The resolution of this equation is simplified if one notices that it is exactly identical to relation (7) if one takes :

$$A^k = (1/2)\,I \quad \text{for all } k \quad \text{and} \quad 2\,\theta = C1\,.$$

In this situation, the synthesis of the network that guarantees the stability of the linearly independent prototype states consists in :

i) determining the synaptic matrix from the given prototype states, after relation :

$$C = \Sigma\Sigma^{\mathrm{I}}\,,$$

ii) determining the thresholds from that matrix :

$$\theta = (1/2)\,C1\,. \tag{9}$$

If the prototype states are orthogonal, the relation defining the synaptic matrix is identical to Cooper's storage prescription (8). If, in addition, the prototype vectors have equal numbers of + 1's and − 1's, they are orthogonal to vector **1**; therefore, $C1 = \mathbf{0}$: after relation (9), all the thresholds are equal to zero. In the numerical simulations presented in reference [8], J. Hopfield uses prototype state vectors λ^k which have a large number of randomly chosen components; therefore, the corresponding σ^k vectors are nearly orthogonal and have almost equal numbers of − 1's and + 1's. Thus, Hopfield's choice of Cooper's storage prescription and of thresholds equal to zero is very appropriate [18] with the particular choice of prototype vectors that was made in reference [8]. However, for real applications in pattern recognition for instance, the prototype vectors represent an image to be memorized : in general, they will not be orthogonal nor quasi orthogonal, and they will not have equal or almost equal numbers of 0's and 1's; therefore, their stability is by no means guaranteed if a Hopfield network is used. Conversely, in our approach, the synaptic matrix and the thresholds can be computed very easily so as to memorize faithfully any set of prototype states { σ^k }, the only restriction being that they must be linearly independent.

Alternately, one should notice that if, for technical reasons, the thresholds must be taken equal to zero, the stability of any set of linearly independent prototype states can be guaranteed by computing C from relation (4) :

$$C = \Sigma A^{\mathrm{I}}\,.$$

As a final remark, a short discussion of the storage capacity may be useful. Two factors determine this capacity : the stability of the prototype states and their attractivity; obviously, if a prototype state is not stable, it cannot be recalled exactly; conversely, a prototype state can be stable, but with little attractivity; therefore, the problem is very complicated. In Hopfield's numerical experiments, it was shown that for $n = 100$ and $p = 5$ the 5 randomly chosen prototype states were stable; for $p = 15$, the probability of having no error in recall is approximately 20 %. About half of the prototype states are unstable. Therefore, the estimated storage capacity of 0.15 n seems

to stem both from the instability of the stable states and from the existence of attractive spurious states. In our approach, the stability of the prototype states is always guaranteed, so that the only limitation to the storage capacity is the decrease of the attractivity when the number of prototype states is increased.

5. Conclusion.

In the present paper, it has been shown that it is possible to express the stability condition of a state of a neural network by a matrix relation, which allows to address the problem of the synthesis of a network with simple tools of linear algebra. The results obtained by this method prove that it is possible to design a neural network when any family of linearly independent prototype states is given. By a proper choice of the synaptic matrix and, in some cases, of the activity thresholds, the stability of the prototype states can be insured. Moreover, the attractivity of these states can be evaluated analytically if the latter are orthogonal; this has been achieved by using the spin-glass formalism, which greatly simplifies the analysis of the network. Therefore, this approach has proved useful both for the understanding of the phenomena arising in networks of threshold elements and for the evaluation of potential engineering applications.

References

 [1] TOULOUSE, G., *Lecture Notes in Physics* **192** (1983) 2.
 [2] KIRKPATRICK, S., GELATT, C. D., VECCHI, M. P., *Science* **220** (1983) 671.
 [3] KIRKPATRICK, S., *J. Stat. Phys.* **34** (1983) 975.
 [4] SIARRY, P., DREYFUS, G., *J. Physique Lett.* **45** (1984) 39.
 [5] D'HUMIERES, D., HUBERMAN, B. A., *J. Stat. Phys.* **34** (1984) 361.
 [6] HUBERMAN, B. A., HOGG, T., *Phys. Rev. Lett.* **52** (1984) 1048.
 [7] AMARI, S., ARBIB, M. A., *Competition and Cooperation in Neural Nets* (Springer Verlag) 1982.
 [8] HOPFIELD, J. J., *Proc. Natl. Acad. Sci. USA* **79** (1982) 2554.
 [9] HOPFIELD, J. J., *Proc. Natl. Acad. Sci. USA* **81** (1984) 3088.
[10] McCULLOCH, W. S., PITTS, W., *Bull. Math. Biophys.* **5** (1943) 115.
[11] HINTON, G. E., ANDERSON, J. A., *Parallel Models of Associative memories* (L. Erlbaum) 1981.
[12] KOHONEN, T., *Self-organization and Associative Memory* (Springer Verlag) 1984.
[13] ALBERT, A., *Regression and the Moore-Penrose Pseudoinverse* (Academic Press) 1972.
[14] GREVILLE, T. N. E., *SIAM Rev.* **2** (1960) 5.
[15] PERETTO, P., *Biol. Cybern.* **50** (1984) 51.
[16] COOPER, L. N., LIBERMAN, F., OJA, E., *Biol. Cybern.* **33** (1979) 9.
[17] See for instance PARISI, G., *Phys. Rev. Lett.* **50** (1983) 1946.
[18] Note added in the revised version : since the paper was first submitted, it has been brought to our attention that relation (9) has been derived independently from statistical considerations by F. Fogelman and G. Weisbuch (private communication). These authors have shown by numerical simulations that the storage capacity of the Hopfield network is much improved if the thresholds are calculated from relation (9) instead of being taken equal to zero. This is an interesting confirmation of the validity of our approach.