Hierarchical multilayer model of memory

Vik. S. Dotsenko
L. D. Landau Institute of Theoretical Physics, Academy of Sciences of the USSR

(Submitted 15 May 1986)
Pis'ma Zh. Eksp. Teor. Fiz. **44**, No. 3, 151–153 (10 August 1986)

A multilayer spin model of memory with a hierarchical structure is constructed. Not only individual patterns but also entire classes of correlated patterns can be contained in the memory in this model. The model corresponds in structure to the visual cortex of the brain.

One of the rapidly developing fields of application of the theory of spin glasses is the construction of spin models of memory, which work as systems of an associative memory and are remote analogs of the neuron networks of the brain.[1]

The idea of storing patterns with the help of spin models can be summarized as follows. We assume a system of N Ising spins which is described by the Hamiltonian

$$H = -\sum_{ij} J_{ij}\sigma_i\sigma_j \qquad (1)$$

and we assume that there are p spin configurations (patterns) $\{\sigma_i^{(s)}\}$, $s = 1,2,..., p$, which are to be stored. These configurations are not correlated with each other. For this interaction between spins, we should choose J_{ij} to be

$$J_{ij} = \frac{1}{N} J_0 \sum_{s=1}^{p} \sigma_i^{(s)}\sigma_j^{(s)} \quad . \qquad (2)$$

For a sufficiently small value of the ratio $\alpha = p/N$ ($\lesssim 0.14$), the model in (1), which is governed by the equations of ordinary relaxation dynamics, will function in a certain temperature range as an associative-memory system.[2,3] In other words, the system will relax toward a particular pattern stored in its memory, depending on which of the configurations $\{\sigma_i^{(s)}\}$ is most similar to the initial configuration.

As the number of patterns to be stored increases, they "interfere" with each other: Many local minima appear in the free energy which do not correspond to any one of the stored patterns $\{\sigma_i^{(s)}\}$. As a result, the "working quality" of the associative memory suffers: The effective attraction region of the stored patterns becomes narrower. There exists a critical curve $\alpha(T)$ such that at $\alpha > \alpha(T)$ there is no memory at all: The system goes into a disordered spin-glass state.[2,3] In other words, if the number of patterns is large enough, the spin-spin interactions J_{ij} in (2) are effectively random, and the model in (1) becomes the Sherrington-Kirkpatrick model[4] of a spin glass.

There is the possibility, however, of altering the memorization algorithm in (2) to construct a model whose memory will contain an exponential number of *correlated* configurations. This deviation is based on the recently discovered hierarchical structure of spin states corresponding to energy minima of the Sherrington-Kirkpatrick model of a spin glass.[5,6] The idea is that the patterns to be stored are aligned in a

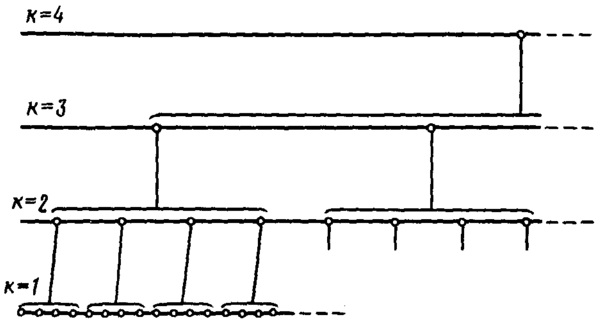

FIG. 1. Hierarchical organization of the interaction of spins between layers.

family tree in increasing order of "resemblance," and the spin system itself is broken up into a hierarchy of nested clusters. At each level of the hierarchy, the clusters independently "memorize" their own fraction of the patterns of the corresponding level of the family tree.[7]

In the present letter we offer a different model, based on the same principle. Its structure is considerably simpler and makes it possible to see explicitly the various classes (family trees) of the stored states. This new model is of further interest in that it is analogous in many ways to the hierarchical structure of the visual cortex of the brain.[8]

We assume a system of K layers of Ising spins. In each layer there are N_k spins ($k = 1,2,...,K$). We group the spins in the layers into clusters $\{\Omega_{i_k}\}$ in such a way that each cluster of level k contains Ω_k spins. The interaction between layers is assumed to be of such a nature that all the spins belonging to a given cluster in layer $k-1$ interact with one and only one spin in layer k (Fig. 1). The number of spins thus decreases in each successive layer: $N_k/\mathrm{N}_{k+1} = \Omega_k$. A spin of layer k interacts in a ferromagnetic fashion with the resultant moment of a cluster of layer $k-1$ in such a way that their spins are the same: $\mathrm{sign}(\Sigma_{i\in\Omega_{k-1}} \sigma_{i_{k-1}}) = \sigma_{i_k}$.

The patterns to be stored, which will be constructed and exhibited in the lowest level ($k=1$), are classified in a hierarchical tree[7] (Fig. 2). Specifically, we say that some family of states belongs to a common "parent" state in the next (second) level of the hierarchy if the signs of the magnetizations of the spin clusters $\{\Omega_{i_1}\}$ are the same

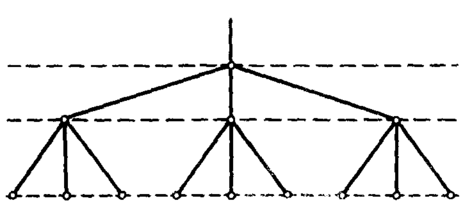

FIG. 2. Hierarchical classification of spin states in a family tree.

in all of these states. In other words, the parent state of such a family is given by some spin configuration in the second layer.

In going from layer to layer we can thus construct a tree or class of correlated states in this manner (Fig. 2). The number of levels of the hierarchy is equal to the number of layers. The number of classes (or trees) is the number of spin configurations in the last layer.

The memorization algorithm is as follows. We introduce the spin interactions in each layer only within clusters:

$$J_{ij} = \frac{1}{\Omega_k} J_0 \sum_{s=1}^{p_k} \sigma_{i_k}^{(s)} \sigma_{j_k}^{(s)} \; ; \qquad (i_k, j_k) \in \Omega_{i_k} . \tag{3}$$

Here $\{\sigma_{i_k}^{(s)}\}$ $(s = 1,2,...,p_k)$ are the spin configurations which are to be stored in the clusters of layer k. It should be noted that the spin states stored in this manner are constructed from the same set of units: a fixed set of states of the clusters. This set can of course be distinctive in each cluster and in each layer.

There are no clusters in the top layer, and the spin states (the bases of the trees) are stored as in the Hopfield model,[1] (2).

If this new model is to work well, the number of units p_k in each cluster must not be large: $p_k/\Omega_k < \alpha(T)$, where $\alpha(T)$ is the critical curve we mentioned earlier, below which there is a region of "good memory."[2,3] Correspondingly, there must be an analogous restriction on the number of trees (the number of states in the top layer): $p_K/N_K < \alpha(T)$.

The process of recognizing a pattern which is presented begins at the top layer, where the most general information—crude patterns—is stored. In other words, a particular tree is chosen. In each successive layer, a particular branch is then chosen, and the process is pursued in this fashion all the way to the last layer, where the exact state is established.

Curiously, a hierarchical organization of this type is characteristic of the visual cortex of the brain[8] and possibly of the entire cortex, since there is reason to believe that the cortex is uniform. The visual cortex contains six layers ($K = 6$), and a neuron of each successive layer of complexity collects information from a group of neurons of the preceding layer (as in Fig. 1). What happens here, however, is not simply a progressive coarsening of the pattern on the retina but also a sort of resolution into components. For example, a neuron of the second layer of complexity collects information from such a cluster of neurons of the simplest layer that it responds to only a line of a certain length and in a certain direction. An organization of this sort could of course be implemented in the model discussed here.

[1]J. J. Hopfield, Proc. Nat. Acad. Sci. USA **79**, 2554 (1982); **81**, 3088 (1984).
[2]D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. Lett. **55**, 1530 (1985).
[3]L. B. Ioffe and M. V. Feigelman, Europhys. Lett. **1** (1986).
[4]S. Kirkpatrick and D. Sherrington, Phys. Rev. B **17**, 4384 (1978).

[5]M. Mezard *et al.*, Phys. Rev. Lett. **52**, 1156 (1984).
[6]M. Virasoro and M. Mezard, Preprint, 1985.
[7]Vik. S. Dotsenko, J. Phys. **C18**, L1017 (1985).
[8]D. H. Hubel and T. N. Wiesel, Proc. Royal Soc. London **B198**, 1 (1977).

Translated by Dave Parsons