

Embedded Patterns, Indirect Couplings with Randomness, and Memory Capacity in Neural Networks

Algis GARLIAUSKAS

*Institute of Mathematics and Informatics
Akademijos 4, LT-06883, Vilnius, Lithuania
e-mail: galgis_1@ktl.mii.lt*

Received: November 2008; accepted: May 2009

Abstract. In the present paper, the neural networks theory based on presumptions of the Ising model is considered. Indirect couplings, the Dirac distributions and the corrected Hebb rule are introduced and analyzed. The embedded patterns memorized in a neural network and the indirect couplings are considered as random. Apart from the complex theory based on Dirac distributions the simplified stationary mean field equations and their solutions taking into account an ergodicity of the average overlap and the indirect order parameter are presented. The modeling results are demonstrated to corroborate theoretical statements and applied aspects.

Keywords: neural network, free energy density, Dirac distribution, memory capacity, ergodicity.

1. Introduction

The behavior of particles in statistical physics for ferromagnetic materials is described as usual by the mean field theory. The theory corresponds to a large number of particles or spins in the limit of a dimension. The possible fluctuations are automatically leveled in the thermodynamic limit, given the good conditions to find a complete solution. The mean field theory can form rather rich structures with different consequences.

It is known, that the ferromagnetic alloys can achieve the following phases: paramagnetic, ferromagnetic (antiferromagnetic), spin glasses, and some time mixed states (Amit *et al.*, 1985b; Edwards and Anderson, 1975; Kirkpatrick and Sherrington, 1978; Sherrington and Kirkpatrick, 1975). The spinglass transition was studied on the basis of practical experiments and Monte Carlo simulation (Kasuya, 1956; Lee and Young, 2003; Nakamura and Endoh, 2002; Petrákovskii *et al.*, 2001). Among different phases there exists a phase transition which is revealed itself at a certain critical temperature. The phase transition line, which divides the spin glass domain from the paramagnetic phase or the ferromagnetic one, is nonlinear as a complicated phase transition limit (Katzgraber and Young, 2005; Krzakala, 2005). The line existence and form the mostly depends on the strength of the external field, for example, even in small fields there is no line in one- or three-dimensional spin glasses (Young and Katzgraber, 2004).

At a higher temperature than critical, there is a paramagnetic phase: the spins are not polarized. Below the critical temperature, the system of spins drops into a new phase called a spinglass phase. In this phase, ergodicity as well as symmetry of the original system are broken. The Boltzmann measure is arranged in several states of the spin configuration space called pure states. Spins tend to freeze in a certain preferred direction. This direction depends both on the spin state and on the spin number. Thus, the spinglass phase is in an ordered phase.

Besides temperature in the particle physics, a communication among the spins plays an important part, when the spins are situated in the neighboring places as well as found far from each other. If to take an affair with the alloy AuFe, the concentration of spins (Fe) in the metallic host (Au) is rather low and the distance among spins is fairly large, i.e., their interactions are only weakly correlated. Apart from this, the so-called the Rudermann–Kittel–Kasuya–Yosida (RKKY) (Rudermann and Kittel, 1954; Kasuya, 1956; Yosida, 1957) interaction is known when the spins interact mainly indirectly.

The direct ferromagnetic interactions of couplings and the indirect RKKY ones contain randomness, as a rule. In Hemmen's (1982) work, the ferromagnetic coupling has been taken on the average as random, expressed by two independent, identically distributed variables. A more general case has been studied in the Provost and Valee (1983) paper, where there are no limitations of random variables.

In the present article, the ferromagnetic and Ising-like models are considered in the general case, where ferromagnetic couplings are not averaged as in Amit *et al.* (1985), Hemmen (1982) but they are random patterns and the system is similar to a neural network (NN). A special attention will be paid to memorization of the system and ergodicity properties of the random coupling variables. In Section 1, the state of the considered problem is discussed. Section 2 is devoted to definition of the free energy density function. Its solutions are considered in Section 3 taking into account the Dirac distributions. Sections 4, 5 discuss the stationary statement of the free energy function and modeling results.

2. Free Energy Density of NN as an Ising System

The mean field theory for ferromagnet and spinglasses allow us to provide an exact solution to a constant infinite-ranged exchange couplings in the thermodynamic limit. For the disordered system, an analogous infinite-ranged problem was defined and solved by Sherrington and Kirkpatrick (1975). Various thermodynamic quantities and competition properties were studied.

In this section, we try to include unaveraged couplings, that are variable, and random and RKKY interactions accepted as other than interpreted in Hemmen (1982), and to apply this theory to the neural network system. One may assume that the neurons of NN are on a regular lattice and interactions via random synapses are related.

The Hamiltonian expressing the direct and indirect interactions among neurons in NN is as follows:

$$H = -1/2 \sum_{(ij)}^N J_{ij} S_i S_j - 1/2 \sum_{(ij)}^N \sum_{(\rho\nu)}^r \eta_{\rho\nu} \zeta_i^\rho \zeta_j^\nu S_i S_j \quad (i \neq j), \quad (1)$$

where J_{ij} are direct couplings between i and j neurons, $\eta_{\rho\nu}$ are indirect couplings, $\|\eta_{\rho\nu}\|$ is a symmetric regular matrix, ζ_i^ρ and ζ_j^ν are components ($\rho, \nu = 1, 2, \dots, r$) of the $\vec{\xi}_i$ and $\vec{\zeta}_j$ that are N random independent equally distributed vectors with mean zero and variance one, S_i and S_j are the state or potential variables of the i th and j th neurons.

The free energy density according to the Boltzmann distribution

$$f = -1/\beta \lim_{N \rightarrow \infty} N^{-1} \ln(Z), \quad (2)$$

where $Z = \sum_{\{S\}} \exp(-\beta H)$ is the partition function as the normalization constant of the Boltzmann distribution, $\beta = 1/T$ (T is the temperature).

The other variables will be defined in the next two sections devoted to study of the mean field theory solutions and ergodicity.

3. Solutions of the Functional with Dirac Distributions

The free energy density in the thermodynamic limit can be defined as leading to the Dirac distributions of general quantities of an overlapping function and other ordering parameters.

Let us introduce them. First, the direct couplings among neurons are expressed over the embedded patterns $\{\xi_i\}$ or $\{\xi_j\}$ for the i th or j th neuron (Hebb (1949) postulate) as follows

$$J_{ij} = \frac{1}{p} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu, \quad (3)$$

where p is the number of embedded patterns.

Second, the vector of order parameters \vec{m} whose components are an external source conjugate to the product $\xi_i^\mu S_i$ is introduced. \vec{m} is the vector of the average overlap between the local magnetization S_i and ξ . The average overlap of one \vec{m} component is presented as follows

$$m^\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu S_i, \quad (4)$$

that is, $\vec{m} \equiv \{m^1, m^2, \dots, m^p\}$.

Third, the components of the indirect order parameter vector \vec{q} are presented

$$q^\nu = \frac{1}{N} \sum_{i=1}^N \zeta_i^\nu S_i, \quad (5)$$

that is, $\vec{q} \equiv \{q^1, q^2, \dots, q^r\}$.

The free energy is given by

$$-\beta f = \lim_{N \rightarrow \infty} N^{-1} \ln T_r \exp(-\beta H). \quad (6)$$

The trace, T_r , is a finite sum over all neuron configurations.

In order to calculate free energy density one includes the $p + r$ Dirac distributions $\delta(Nm^\mu - \sum_{i=1}^N \xi_i^\mu S_i)$ ($\mu = 1, 2, \dots, p$) and $\delta(Nq^\nu - \sum_{i=1}^N \zeta_i^\nu S_i)$ ($\nu = 1, 2, \dots, r$). The expansion of δ functions in Fourier integrals (over τ^ν and t^ν) leads to factorization of the trace operation.

Hamiltonian (1) with regard to (3–5), and the Fourier transformation of Dirac distributions assumes the following expression

$$\begin{aligned} H = & -\frac{1}{2p}(N^2 \vec{m}^2 - pN) - \frac{1}{2}N^2(\vec{q}\eta\vec{q}) \\ & + \frac{N}{2}\vec{m} \sum_{\mu=1}^p \sum_{i=1}^N \xi_i^\mu S_i + N(\vec{\zeta}\eta\vec{q}) \frac{1}{N} \sum_{i=1}^N S_i. \end{aligned} \quad (7)$$

First and fourth members in right-hand side are formed from the first member of (1) regarding with (3) and (4). Second and fifth members are followed from the second member of (1) and formula (6),

Then the Boltzmann partition function

$$\begin{aligned} Z = & T_r \exp(-\beta H) \\ = & T_r \int \prod_{(\mu,\nu)} m^\mu q^\nu \exp \left[\frac{\beta N^2}{2p} (\vec{m}^2 - pN) + \frac{\beta}{2} N^2 (\vec{q}\eta\vec{q}) \right] \\ & \times \exp \left[-\vec{m} \sum_{\mu=1}^p \sum_{i=1}^N \xi_i^\mu S_i - (\vec{\zeta}\eta\vec{q}) \frac{1}{N} \sum_{i=1}^N S_i \right] \\ & \times \delta \left(Nm^\mu - \sum_{i=1}^N \xi_i^\mu S_i \right) \delta \left(Nq^\nu - \sum_{i=1}^N \zeta_i^\nu S_i \right) \\ = & \left(\frac{1}{2} \right)^{p+r} \int \prod_{(\mu,\nu)} m^\mu \tau^\mu q^\nu t^\nu \exp \left[\frac{\beta}{2p} (N^2 \vec{m}^2 - pN) + \frac{\beta}{2} N^2 (\vec{q}\eta\vec{q}) \right] \\ & \times \exp [N(\vec{m}\vec{\tau} + \vec{q}\vec{t})] \\ & - \sum_{i=1}^N \ln 2 \cosh \left[\frac{N\beta}{p} (\vec{m}\vec{\xi}_i) + N\beta(\vec{\zeta}\eta\vec{q}) + (\vec{\xi}_i\vec{\tau}) + (\vec{\zeta}\vec{t}) \right]. \end{aligned} \quad (8)$$

There are used the multiple integrals, differentials of which are the order parameters m^μ , q^ν , and auxiliary variables r^μ and t^ν followed from an integration of δ -functions.

The free energy density

$$\begin{aligned} f &= -\frac{1}{\beta N} \ln(Z) \\ &= -\frac{(p+r)}{\beta N} \ln 2 + \frac{1}{2} - \frac{N}{2p} \bar{m}^2 (\bar{q}\eta\bar{q}) - \frac{N}{\beta} \bar{m}\bar{\tau} + \bar{q}\bar{t} \\ &\quad + \frac{1}{\beta N} \sum_{i=1}^N \ln 2 \cosh \left[\frac{N\beta}{p} (\bar{m}\vec{\xi}_i) + N\beta(\zeta\eta\bar{q}) + (\vec{\xi}_i\bar{\tau}) + (\vec{\zeta}\bar{t}) \right]. \end{aligned} \quad (9)$$

Note that the last member of (8) and (9) expresses the condition of the ergodicity or selfaveraging. According to the Birkhoff ergodic theorem, the average of all values of vectors $\vec{\xi}_i$ at the same time is equal to the average of one $\vec{\xi}$ over time. It means that

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \ln 2 \cosh \left[\frac{N\beta}{p} (\bar{m}\vec{\xi}_i) + N\beta(\zeta\eta\bar{q}) + (\vec{\xi}_i\bar{\tau}) + (\vec{\zeta}\bar{t}) \right] \\ = \ll \ln 2 \cosh \left[\frac{N\beta}{p} (\bar{m}\vec{\xi}) + N\beta(\zeta\eta\bar{q}) + (\vec{\xi}\bar{\tau}) + (\vec{\zeta}\bar{t}) \right] \gg, \end{aligned} \quad (10)$$

where $\ll \cdot \gg$ means the average with respect to the probability measure $d\mu(\vec{\xi})$ of one of the $\vec{\xi}_i$. The free energy density can be presented

$$\begin{aligned} f &= -\frac{(p+r)}{\beta N} \ln 2 + \frac{1}{2} - \frac{N}{2p} \bar{m}^2 (\bar{q}\eta\bar{q}) - \frac{N}{\beta} \bar{m}\bar{\tau} + \bar{q}\bar{t} \\ &\quad + \frac{1}{\beta} \ll \ln 2 \cosh \left[\frac{N\beta}{p} (\bar{m}\vec{\xi}) + N\beta(\zeta\eta\bar{q}) + (\vec{\xi}\bar{\tau}) + (\vec{\zeta}\bar{t}) \right] \gg. \end{aligned} \quad (11)$$

It is considerably better to solve integral (8) by the Monte Carlo technique and to neglect the fluctuations in the thermodynamic limit; it will be equivalent to $n \rightarrow 0$ in the mean field theory (Sherrington and Kirkpatrick, 1975). The modeling of free energy density (8) is achieved by a steepest-descent integration at the beginning with (τ^μ, t^ν) then with (m^μ, g^ν) . Equations (9) or (11), the solution of which determines a saddle point relative (τ^μ, t^ν) , are a generalization of theories Hemmen (1982), and Provost and Vallee (1983).

4. The Stationary Statement of the Free Energy Function

Though the solutions, based on Dirac distributions, are more accurate, however for the qualitative evaluation suffices to apply more simplified technique represented below.

Omitting some arrangements of formulas (7) and (8), the Boltzmann partition function can be presented as

$$Z = \exp \left\{ \frac{1}{2} - \frac{N^2}{2p} \bar{m}^2 - \frac{N^2}{2} (\bar{q} \eta \bar{q}) + \sum_{i=1}^N \ln 2 \cosh \left[\frac{N\beta}{p} (\bar{m} \xi_i) + N\beta (\bar{\zeta} \eta \bar{q}) \right] \right\}. \quad (12)$$

The free energy density reduces to

$$f = -\frac{1}{N\beta} \ln Z_0 = -\frac{1}{N\beta} \left[\frac{1}{2} - \frac{N^2}{2p} \bar{m}^2 - \frac{N^2}{2} (\bar{q} \eta \bar{q}) + \sum_{i=1}^N \ln 2 \cosh \left[\frac{N\beta}{p} (\bar{m} \xi_i) + N\beta (\bar{\zeta} \eta \bar{q}) \right] \right]. \quad (13)$$

Then the local optimum equations are as follows

$$\bar{m} = \frac{1}{N} \sum_{i=1}^N \xi_i \tanh \left[\frac{N\beta}{p} (\bar{m} \xi_i) + N\beta (\bar{\zeta} \eta \bar{q}) \right]. \quad (14a)$$

$$\bar{q} = \frac{1}{N} \sum_{i=1}^N \bar{\zeta} \tanh \left[\frac{N\beta}{p} (\bar{m} \xi_i) + N\beta (\bar{\zeta} \eta \bar{q}) \right]. \quad (14b)$$

With respect to ergodicity

$$\bar{m} = \ll \xi \tanh \left[\frac{N\beta}{p} (\bar{m} \xi) + N\beta (\bar{\zeta} \eta \bar{q}) \right] \gg. \quad (15a)$$

$$\bar{q} = \ll \bar{\zeta} \tanh \left[\frac{N\beta}{p} (\bar{m} \xi) + N\beta (\bar{\zeta} \eta \bar{q}) \right] \gg. \quad (15b)$$

All solutions (15a) and (15b) will be more predictive at the critical temperature T_c ; below T_c their behavior can face a complex problem of choosing the states that are energetically favored.

5. Some Modeling Results

The main task of brief modeling is to answer a question whether the solutions (14a), (14b) or (15a), (15b) of free energy density (13) achieve the saddle point or the local (global) minimum point.

The second question raised is what behavior of the overlapping vector \bar{m} and vector \bar{q} will be dependent on the random patterns ξ_i and $\bar{\zeta}$.

In the first task, we have calculated the free energy function (13) in a simplified way; the random parameters were fixed and function (13) was analyzed for one component of the vectors \vec{m} and \vec{q} . It suffices to evaluate the character of the free energy function surface.

The modeling results of the free energy function are graphically shown in Fig. 1. We see that the local minimum also is the global one with approximately $m_{\min} = 0.6$, $q_{\min} = 0.1$, and there is no saddle point what has been mentioned in Amit *et al.* (1985a), Hemmen (1982), and Provost and Valee (1983).

The second modeling task was more complicated to solve. We have modeled the five-dimensional vectors \vec{m} and \vec{q} , afterwards found their projections and moduli dependent on the variable parameter β . Moduli of vectors were averaged for random parameters $\{\xi_i\}$ and $\{\zeta\}$ and expressed as that: $Mm = \ll \|\vec{m}\| \gg$ and $Mq = \ll \|\vec{q}\| \gg$.

The modeling results, shown in Fig. 2, corroborate the increase of averages of vector moduli, depending on an inverse temperature ($\beta = 1/T$). It is shown that both global minimum parameters possess a saturation zone. This is probably challenged by the presence of hyperbolic tangents in both solutions. Another version of this modeling case was aimed at termination of the same vector moduli dependency on a number of patterns p which characterize a memory capacity ($\alpha = p/N$) in neural networks. In separate about memory capacity and information retrieval from an associative memory in neural networks has been considered in Garliuskas (2005, 2007). Some results of modeling are reflected in Fig. 3.

The vector moduli are decreasing dependent on p at a fixed number of neurons N . With a view to memorize more patterns in the neural network it is necessary to increase N significantly.

Note that the behavior of these two vector moduli is identical with the only difference in the quantified values.

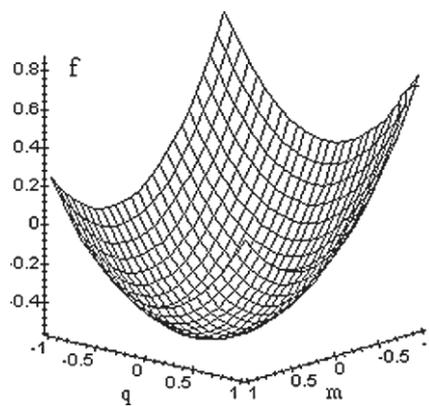


Fig. 1. The surface of the free energy density function.

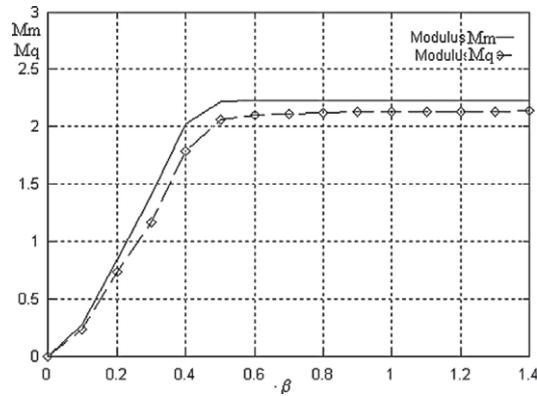


Fig. 2. The averaged moduli of vectors \vec{m} and \vec{q} dependency on the inverse temperature parameter β .

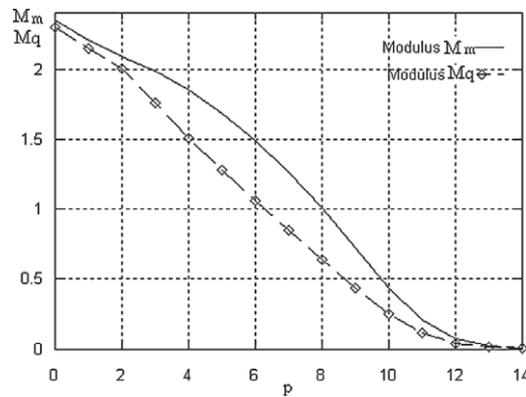


Fig. 3. The averaged moduli of vectors \vec{m} and \vec{q} dependency on the number of embedded patterns p .

6. Conclusions

In conclusions it can be noted that:

1. Introducing of the Hebb rule, embedded random patterns by an overlap parameter, and a vector of indirect order parameters allows to generalize the mean field theory applied to NN. Thus the development of applied statistical mechanics in an idealized Ising presentation for the neural network theory can enrich the theory itself and some applied areas.
2. The numerical modeling results show that the overlapping macroparameter decreases when number of embedded patterns are increasing at fixed number of neurons in the network.
3. After some numerical modeling, it was established that the free energy density optimal solution is global minimum and no saddle point as supposed.
4. Further findings of the best neural network technologies with randomness and fuzziness, based on statistical physics, could be useful and fruitful.

References

- Amit, D.J., Gutfreund, H., Sompolinsky, H. (1985a). Spin-glass model of neural networks. *Phys. Rev. Lett.*, 32(2), 1007–1018.
- Amit, D.J., Gutfreund, H., Sompolinsky, H. (1985b). Storing infinite numbers of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.*, 55(14), 1530–1533.
- Edwards, S.F., Anderson, P.W. (1975). Theory of spin glasses. *J. Phys. F: Metal. Phys.*, 5, 965–974.
- Garliauskas, A. (2005). An influence of nonlinearities to storage capacity of neural networks. *Informatica*, 16(2), 159–174.
- Garliauskas, A. (2007). Information conveyed by the neural network systems and its applied significance. *Informatica*, 18(2), 159–174.
- Hebb, D.O. (1949). *The Organization of Behavior*. Wiley, New York.
- Hemmen, J.L. (1982). Classical spin-glass model. *Phys. Rev. Lett.*, 49(6), 409–412.
- Kasuya, T. (1956). A theory of metallic ferro- and antiferromagnetism on Zener's model. *Prog. Theor. Phys.*, 16(1), 45–57.
- Katzgraber, H.G., Young, A.P. (2005). Probing the Almeida–Thouless line from the mean-field model, *cond-mat.0507138 v3(2005)*, 1–6.
- Katzgraber, H.G., Palassini, M., Young, A.P. (2001). Monte-Carlo simulations of spin-glasses at low temperatures, *Phys. Rev. B*, 63, 184422.
- Kirkpatrick, S., Sherrington, D. (1978). Infinite ranged models of spin-glasses. *Phys. Rev. Lett. B*, 17(11), 4381–4403.
- Krzakala, F. (2005). Zero temperature phase diagram of infinite connectivity spin glasses, *cond-mat.0409449 v3(2005)*, 1–4.
- Lee, L.W., Young, A.P. (2003). Single spin- and chiral-glass transition in vector spin glasses in three dimensions. *Phys. Rev. Lett.*, 90, 227203, arXiv: cond-mat/0302371, 1–5.
- Nakamura, T., Endoh, S. (2002). Single spin- and chiral-glass transition in a $\pm 1J$ Heisenberg spin-glass model in tree dimensions. *J. Phys. Soc. Jpn.*, 71, 2113, arXiv: cond-mat/0110017, 1–4.
- Petrakovskii, G.A. et al. (2001). Spin-glass state in $CuGa_2$. *Phys. Rev. B*, 63, 184425.
- Provost, J.P., Vallee, G. (1983). Ergodicity of the coupling constants and the symmetric n -replicas trick for a class of mean-field spin-glass models. *Phys. Rev. Lett.*, 50(8), 598–600.
- Rudermann, M.A., Kittel, C. (1954). Indirect exchange coupling of nuclear magnetic moments by conduction electrons. *Phys. Rev.*, 96(1), 99–102.
- Sherrington, D., Kirkpatrick, S. (1975). Solvable model of a spin glasses. *Phys. Rev. Lett.*, 35, 1792–1795.
- Yosida, K. (1957). Magnetic properties of Cu-Mn alloys. *Phys. Rev.*, 106(5), 893–898.
- Young, A.P., Katzgraber, H.G. (2004). Absence of an Almeida–Thouless line in three-dimensional spin glasses. *cond-mat.0407031 v2(2004)*, 1–4.

A. Garliauskas received his habil. dr. degree of technical sciences from the Computer Center, the Department of the USSR Academy of Sciences, Novosibirsk, USSR, in 1977. He is a senior researcher of the Institute of Mathematics and Informatics, a chairman of the Informatics Department of the Lithuanian Science Society and a member of its council. His research interest includes neuroinformatics methodology, control problems and development of neural networks algorithms, recognition, chaos processes.

Įterptieji šablonai, netiesioginiai atsiktiniai ryšiai ir atminties inlumas neurotinkluose

Algis GARLIAUSKAS

Straipsnyje nagrinėjama neurotinklų teorija, grindžiama Izingo modeliu. Analizuojami netiesioginiai neuronų ryšiai, Dirako skirstinių taikymai ir korekcija įvedant Hebo taisyklę. Įsimenamieji įterptieji šablonai neurotinkle ir netiesioginiai ryšiai imami atsiktiniais. Teorinių teiginių patvirtinimui pateikiami skaitmeninio modeliavimo rezultatai.