INFORMATICA, 2010, Vol. 21, No. 3, 339–348

An Expansion of the Neural Network Theory by Introducing Hebb Postulate

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

Received: June 2009; accepted: April 2010

Abstract. In the presented paper, some issues of the fundamental classical mechanics theory in the sense of Ising physics are introduced into the applied neural network area. The expansion of the neural networks theory is based primarily on introducing Hebb postulate into the mean field theory as an instrument of analysis of complex systems. Appropriate propositions and a theorem with proofs were proposed. In addition, some computational background is presented and discussed.

Keywords: Hamiltonian, Ising model, Hebb postulate, neural network, memory capacity.

1. Introduction

The statistical physics of particles can be elucidated from the point of view of the classical mechanics as well as a quantum-mechanics approach. The quantum-mechanics included itself the classical one in the sense of a limit case. In the quantum-mechanics, moving of an electron is described by a wave function, while in the classical mechanics, the electron is considered as a material particle moving along the trajectory defined by the equations of a motion (Landau and Lifschitz, 1989). The particle of the system is given by the moment and its projection and in the quantum-mechanics the moment acquires a special sense, i.e., a quantum number, that defines states of the system under their transformation properties relative to cycling of the system coordinates. Any particle possesses a spin moment or simply a spin and an orbital moment for characterizing a particle motion in the space.

In literature, most cases are demonstrated by diluted magnetic alloys, such as AuFe or CuMn, with a magnetic impurity concentration. The main condition seems to be a locally random competition between ferromagnetic and antiferromagnetic forces. These physical examples with surprising properties – sharp cusp in the susceptibility and others are studied regarding theoretical positions (Edwards and Anderson, 1975; Fischer, 1975; Kirkpatrick and Sherrington, 1978; Sherrington and Kirkpatrick, 1975; Sherrington and Southern, 1975) as well as experimentally (Canella and Mydosh, 1972). In order to show the existence of a spinglass phase Edwards and Anderson (EA) (Edwards and Anderson, 1975) have introduced a simple model in which a classical Heisenberg model with random Gaussian distribution of exchange interactions centered at zero is considered. Fisher (1975), Sherrington and Southern (1975) have investigated a quantum extension of the EA model. A class of infinite-ranged random models has been widely investigated by Sherrington and Kirkpatrick (1975, 1978) by the mean field theory, order parameters and phase diagrams. For this model, the extended mean field theory of Thouless *et al.* (1977) is discussed under physically sensible low temperature prediction. The issues of investigation of finite-ranged random models are characterized by Katzgraber *et al.* (2001, 2005), and the survey article of Garliauskas (2005).

The mean field spin-glass models with ferromagnetic couplings depending on two random variables are discussed by Hemmen (1982) and an extended case with random p-vectors and an ergodicity hypothesis is examined by Provost and Vallee (1983). Here the spins of ferromagnet are taken on a regular lattice and the interactions are as random (Fischer, 1975). Some other problems such as the influence of indirect couplings, random embedded patterns and other ordering parameters for neural networks are discussed by Garliauskas (2009).

In this article, the main attention is paid to the collective behavior of neural networks (NN) the theoretical background of which is like a statistical behavior of ferromagnet or, in general, like the statistical mechanics founded by infinite-ranged Ising Hamiltonians. Configurations of the neurons like the spin system are defined as random elements and memories with the help of couplings in the thermodynamic limit (Amari and Maginu, 1988; Amit *et al.*, 1985; Katzgraber and Young, 2005; Sompolinsky, 1986 and many others).

In this work, an attention is also paid to the mathematical foundation of an artificial neural network that is tightly connected with storing information, extends the ideas of the article (Garliauskas, 2006). The possible advantages and benefits of the theory are an including more realistic presentation of the Hebb postulate providing the accumulated effect of the learning on the synaptic connections ad simplifying the solution of free energy function as a conventional surface instead of saddle-like point one.

In Section 2, the Hamiltonian functional as the fundamental statistical mechanic theory is discussed. A specific Hamiltonian is considered here for neural network systems. In Section 3.1, different definitions, propositions, and notions, necessary for more accurate proofs of the methodology proposed , are considered. Section 3.2 is aimed at the determination theorem and its proof. The additional computational background is presented in Section 4.

2. Hamiltonian Functionals

The main characteristic of the statistical physics as well as the neural network theory is the total energy function of a spin system expressed by a spin Hamiltonian. In the quantum-mechanical case, the spin Hamiltonian for a regular system is presented as

$$H = -\frac{1}{2}\sum_{(i,j)\in N} J_{ij}\left[S_i^z S_j^z + \gamma\left(S_i^x S_j^x + S_i^y S_j^y\right)\right] - \mu h \sum_{i\in N} S_i^z \quad (i \neq j), \tag{1}$$

where J_{ij} is a coupling of the spin system, $\vec{S}_i$. $\vec{S}_j$ are quantum-mechanical spin operators with the components x, y, z on the ith and jth sites, μ is a magnetic moment, and h is the external magnetic field. The x component S_i^x and S_j^x of vectors $\vec{S}_i$ and $\vec{S}_j$ have eigenvalues $S, S-1, S-2, \ldots, -S$, and the y and z components are similar. γ is a constant; for $\gamma = 0$, it is coincident to the Ising model and, for $\gamma = 1$, to the Heisenberg model (Griffits and Lebowitz, 1968).

The surface of Hamiltonian (1) is very complicated under an the influence of non-linearities of the couplings J_{ij} and the products of vectors $\vec{S}_i$ and $\vec{S}_j$. There are such as valleys, barriers, multiple local minima, complex boundary conditions. Another case of Hamiltonian form is accepted as a more complicated one under the influence random anisotropy of the mixed-spin Ising model (Vieira *et al.*, 2001). According to Vieira *et al.* (2001) the mixed-spin Ising model is represented by a two-sublattice system with the variables $\sigma = \pm 1$ and $S = 0, \pm 1$ on sublattices A and B, respectively. The most general spin Hamiltonian is described as

$$H_{AB} = -J \sum_{i \in A, j \in B} \sigma_i S_j + D \sum_{j \in B} S_j^2, \tag{2}$$

where J is a parameter of ferromagnetic coupling, the second member of (2) determines the spin field with the parameter $D > 0$, A and B are the sets of sublattices.

The competition between ferromagnetic exchange and anisotropy leads to the appearance of critical lines and a tricritical point location (Vieira *et al.*, 2001).

To be specific, for the neural system we shall consider a system with the Hamiltonian

$$H_0 = -\frac{1}{2} \sum_{(i,j) \in N} J_{ij} S_i S_j - h \sum_{i \in N} S_i \quad (i \neq j), \tag{3}$$

where $S_i = 2S_i^z = \pm 1$. Here instead σ of (2) we return to S_i, S_j with new appropriate sense. The first sums are over the nearest – neighbor pairs of sites. The coupling constant J_{ij} between two neighboring neurons depends on the actual pair of the states S_i and S_j considered. The constant J_{ij} can be either positive (ferromagnetic interaction or neuron excitation) or negative (antiferromagnetic interaction or neuron inhibition).

Furthermore, Hamiltonians (1), (3) will be used in simplified form for the case where is no the external magnetic field.

The behavior of spins, when the coupling strength is random, is considered by the Boltzmann distribution

$$P_J(S) = \frac{1}{Z_J} e^{-\beta H(J,S)}, \tag{4}$$

where

$$Z_J = \sum_{\{S\}} e^{-\beta H(J,S)} \tag{5}$$

is a partition function. Then the free energy density

$$f = -\frac{1}{\beta N}\ln(Z_J) \tag{6}$$

is self-averaging in the thermodynamic limit. The $\beta = 1/T$, i.e., the inverse of temperature. In dynamics, starting from an arbitrary initial configuration, the system evolves monotonically decreasing the value of H and leads to the limit of a steady state, which is the local extremum.

3. The Mathematical Background of the Generalized NN Model

The physiological background of the neuro-spin behavior is briefly reviewed. In the Hopfield model (Hopfield, 1982) and others, each neuron is presented as an Ising spin with two possible states: "up" and "down" positions, i.e., the neuron fires an electrochemical signal in case its potential exceeds the threshold value. The neurons are interconnected by synaptic couplings of strength J_{ij}, which define the synaptic potential of a signal fired by the jth neuron to the postsynaptic potential of the ith neuron. The potential can be either positive (excitation) or negative (inhibition). Do not paying an attention that the neural network is presented by weekly realistic condition that every neuron is connected to every other one and the neurons will interacted indirectly, the theoretical background made under these suggestions do not diminish the practical importance.

3.1. *Notions, Propositions and Definitions*

The interactions between postsynaptic potentials and presynaptic ones are described by the Hebb rule or postulate (Hebb, 1949). The Hebb postulate is formulated as the following notation.

Notion. The Hebb postulate reads: "When an axon of cell A is near enough to excitate a cell B and repeatedly or persistently takes part in firing it, some growth process of metabolic change takes place in one or both cells such that A's efficiency as one of the cells firing B is enhanced" (Hebb, 1949). This definition is not strict. It is established by practical neurophysiological experiments and was further formalized as follows.

PROPOSITION 1. Based on *Notion* the next simple formula

$$\Delta J_{ij} = k S_i S_j \tag{7}$$

follows.

Here ΔJ_{ij} is a change of the synaptic weight J_{ij} between the ith and jth neurons, which depends on the conjunctive presence of the presynaptic potential S_j and the postsynaptic one S_i, and k is the learning rate.

PROPOSITION 2. This proposition expresses the Hebb postulate through an accumulated effect as follows

$$J_{ij} = \frac{1}{p}\sum_{\mu=1}^{p} \xi_i^{\mu}\xi_j^{\mu}, \tag{8}$$

where p is the number of patterns $\{\xi_i\}, \{\xi_j\}$ as the embedded images, besides, the patterns are random with equal probability for $\xi_i^{\mu},\ \xi_j^{\mu} = \pm 1$.

The authors of papers (Amit *et al.*, 1985; Sompolinsky, 1986), and others use formula (8) divided by N (number of neurons) instead of the number of patterns p, which, to my mind, leads to extremal reduction of coupling strengths J_{ij} when N is increased. This correction will be in the lead of ideas in this issue.

DEFINITION 1. The memory capacity α in the neural network as an Ising spin system is a relation

$$\alpha = p/N. \tag{9}$$

Supposedly, one is mind, that the memory capacity depends on an increase of the number of synapses rather than an increase the number of neurons with the same percentage (Little and Shaw, 1978).

DEFINITION 2. The function S: $R \to [-1, 1]$ is a squashing function if it is non-decreasing and satisfies

$$\lim S(u)_{u\to+\infty} = 1, \qquad \lim S(u)_{u\to-\infty} = -1. \tag{10}$$

The squashing functions include the signum function defined by

$$\operatorname{sgn} S(u) = \begin{cases} 1, & \text{if } u \geqslant 1, \\ -1, & \text{if } u < -1 \end{cases} \tag{11}$$

or the signum function with a threshold h

$$\operatorname{sgn} S(u-h) = \begin{cases} 1, & \text{if } u \geqslant h, \\ -1, & \text{if } u < h. \end{cases} \tag{12}$$

Other squashing functions are known: the sigmoidal function, the ramp or saturating function, the sigmoidal cosine squashing function.

Lemma 1. *If* S_i, ξ_i^{μ} *are squashing signum functions* (11) *there exists a relation*

$$\sum_{\mu=1}^{p}\sum_{ij}^{N} \xi_i^{\mu} S_i \xi_j^{\mu} S_j = \sum_{\mu=1}^{p}\left[\left(\sum_{i=1}^{N}\xi_i^{\mu} S_i\right)^2 - pN\right] \quad (i \neq j). \tag{13}$$

The proof of the lemma is obvious. We can easily verify that, given different values of p and N.

DEFINITION 3. The overlap function the m^μ of μ component of the vector $\vec{m}$ is an average of the product of the pattern ξ_i^μ and state variable S_i , i.e.,

$$m^\mu = \frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu S_i. \tag{14}$$

3.2. *The Free Energy Density Solution*

The free energy density function including its complex Hamiltonian can form a complicated surface with valleys and barriers. We give below some mathematical background with computational experiments in the next section. Let us formulate the following theorem.

Theorem. *Let Ω be a set of N, $N \in \Omega$, and Θ a set of p, $p \in \Theta$, then the free energy density* (6) *with the partition function of type* (5)

$$Z = \exp\left(\frac{\beta}{2p}\sum_{(i,j)\in N}\xi_i^\mu S_i \xi_j^\mu S_j\right), \quad (i \neq j) \tag{15}$$

and corresponding to the Proposition 2*, Definitions* 2 *and* 3*, and the lemma has a local optimal solution*

$$\vec{m} = \frac{1}{N}\sum_{i=1}^{N}\vec{\xi}_i \tanh\left(\frac{\beta N}{p}\vec{m}\vec{\xi}_i\right). \tag{16}$$

Proof. Let us represent Hamiltonian (3) without the external field

$$H_0 = -\frac{1}{2}\sum_{(i,j)\in N}J_{ij}S_iS_j, \quad (i \neq j).$$

Then in view of Proposition 2, Definitions 2 and 3, Lemma and the multiplicative sum $\sum_{\pm 1}\cdot\sum_{\pm 1}\cdots\sum_{\pm 1}$ passing to $\sum_{i=1}^{N}\cosh(\frac{\beta N}{p}\vec{m}\vec{\xi}_i)$, the partition function Z (15) becomes

$$Z = \exp\left(\frac{\beta N}{2}\right)\exp\left(-\frac{\beta N^2}{2p}\vec{m}^2 + \sum_{i=1}^{N}\ln 2\cosh\left(\frac{\beta N}{p}\vec{m}\vec{\xi}_i\right)\right).$$

According to (5) the free energy function

$$f = -\frac{1}{N\beta}\ln Z = -\frac{1}{2} + \frac{N}{2p}\vec{m}^2 - \frac{1}{N\beta}\sum_{i=1}^{N}\ln 2\cosh\left(\frac{\beta N}{p}\vec{m}\vec{\xi}_i\right). \tag{17}$$

Then the extremum (minimum) of the free energy function is obtained after differentiating the overlap parameter and equating it to zero:

$$\vec{m} = 1/N \sum_{i=1}^{N} \vec{\xi}_i \tanh\left(\frac{\beta N}{p} \vec{m}\vec{\xi}_i\right). \tag{18}$$

This is the necessary condition of the minimum free energy function. The computation results in next section show that the minimum is global.

To corroborate the sufficient condition, we take one component of the vector $\vec{m}$. Then the second derivative is found

$$\frac{\mathrm{d}^2 f}{\mathrm{d}(m^\mu)^2} = \frac{N}{p} - \frac{\beta N}{p^2} \sum_{i=1}^{N} (\xi_i^\mu)^2 \left(1 - \tanh^2\left(\frac{\beta N}{p} m^\mu \xi_i^\mu\right)\right)$$

and it is positive, because the derivative over $\tanh(\cdot)$ is positive and is equal to or is smaller than one. The β and p are always positive and $\beta < p$. The $(\xi_i^\mu)^2$ is always equal to one because the vector projections $\xi_i^\mu = \pm 1$. The proof of the theorem is complete.

4. Computational Background

A computational analysis is necessary because, if we consider different ranges of parameters, there are some doubts, always or not always, for example, whether the necessary or sufficient conditions of the theorem (Section 3) will be completely fulfilled. Besides, it is of interest how individual parameters influence the main output functions. To this end, let us introduce the parameter α (9), memory capacity in NN, into the free energy density equation (17) written for one component of vectors $\vec{m}$ and $\vec{\xi}_i$. Then we obtain

$$f = -\frac{1}{2} + \frac{1}{2\alpha}\left(m^\mu\right)^2 - \frac{1}{\beta N} \sum_{i=1}^{N} \ln 2\cosh\left(\frac{\beta}{\alpha} m^\mu \xi_i^\mu\right), \tag{19}$$

and write down the second derivative over m^μ

$$\frac{\mathrm{d}^2 f}{\mathrm{d}(m^\mu)^2} = \frac{1}{\alpha}\left[1 - \frac{\beta}{N\alpha} \sum_{i=1}^{N} (\xi_i^\mu)^2 \left(1 - \tanh^2\left(\frac{\beta}{\alpha} m^\mu \xi_i^\mu\right)\right)\right]. \tag{20}$$

Note that $(\xi_i^\mu)^2$ always equals one.The change of the sign of its components does not influence $1 - \tanh^2(\cdot)$, as the derivative of $\tanh(\cdot)$ is always positive and smaller or equal to one.

We have performed some calculations using Maple 4 of the free energy function (19) and of the second derivative to check their behavior depending on the parameters α and m in their work ranges.

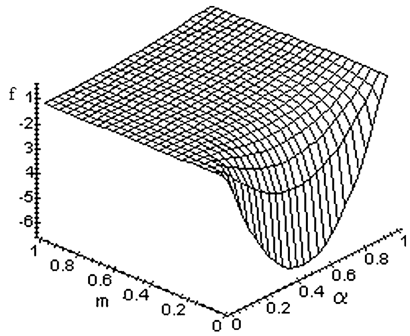

Fig. 1. The free energy density versus overlapping and memory capacity parameters.

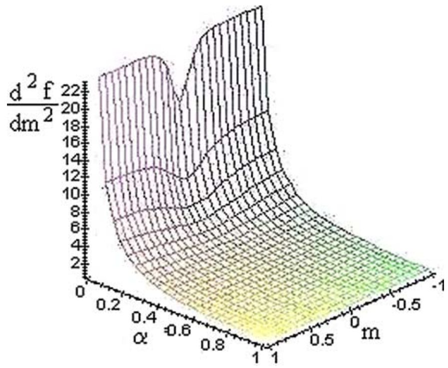

Fig. 2. The second derivative dependent on overlapping and memory capacity parameters.

The free energy surface is shown in Fig. 1 and its second derivative in Fig. 2. The free energy function in these ranges of parameters possesses a clearly expressed minimum which is a global one. The minimum values are limited to the boundary of parameters. Fig. 2 confirms the positiveness of the second derivative and corroborates the sufficient proof of the theorem (Section 3).

Further, let us additionally analyze the influence of temperature on the memory capacity in neural networks. The influence of nonlinearities and other factors has been studied in the articles (Amit *et al.*, 1985; Garliauskas, 2006). Taking into account that $\tanh(x)$ can be simply approximated as a linear one, i.e., $\tanh(x) = x + 0(x^3)$, the expression of overlapping function (18) can be written as follows

$$m^\mu = m^\mu \frac{\beta}{\alpha} \frac{1}{N} \sum_{i=1}^{N} |\xi_i^\mu|$$

in the narrow area of zero for one component of the vector $\vec{m}$. After m^μ reducing and expressing α we obtain

$$\alpha = \frac{1}{T/T_c} |\xi^\mu|,$$

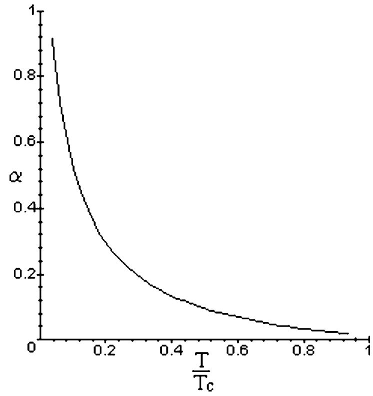

Fig. 3. The memory capacity versus temperature.

where $|\xi^\mu| = \frac{1}{N}|\sum_{i=1}^{N}\xi_i^\mu|$ is an absolute average of patterns, $\beta = T_c/T$, where T_c is the critical temperature.

The implies that the memory capacity is decreased nearing to the critical temperature and is increased when the temperature is lower than critical (Fig. 3).

5. Conclusions

1. The inclusion of more realistic theoretical presentation of the Hebb postulate and an indirect couplings among neurons allows to reduce neural network theory and to do it closer to practical requirements of learning processes in the recognition theory.
2. In spite of literature sources, where mountain that the saddle point of the free energy function exists, we prove that after introducing proper Hebb postulate the energy function surface is simplified, and it is providing an existence of the minimum solution invoking the appropriate theorem.
3. Some modeling results confirm the theoretical findings into work ranges of parameters and an existence of global solution as well as a dependence of applied memory capacity of neural networks versus temperature.

REMARK. Certainly we understand, that only a small range of problems of theoretical neural network system has been studied here. The main investigations in future may be designated to search for global minimum of the free energy function with stochastic initial data, stability in dynamics, and chaos phenomenon.

References

Amari, S.I., Maginu, K. (1988). Statistical neurodynamics of associative memory. *Neural Netw.*, 1(1), 63–73.

Amit, D.J., Gutfreund, H., Sompolinsky, H. (1985). Spin-glass model of neural networks. *Phys. Rev. Lett.*, 32(2), 1007–1018.

Canella, V., Mydosh, J.A., (1972). Magnetic ordering in gold–iron alloys. *Phys. Rev. B*, 6, 4220–4237.

Edwards, S.F., Anderson, P.W. (1975). Theory of spin glasses. *J. Phys. F: Metal. Phys.*, 5, 965–974.

Fischer, K.H. (1975). Static properties of spin glasses. *Phys. Rev. Lett.*, 34(23), 1438–1441.

Garliauskas, A. (2005). An influence of nonlinearities to storage capacity of neural networks. *Informatica*, 16(2), 159–174.

Garliauskas, A. (2006). Nonlinearities in artificial neural systems interpreted as an application of Ising physics. *Nonlinear Anal.: Model. Control*, 11(4), 367–383.

Garliauskas, A. (2009). Embedded patterns, indirect couplings with randomness, and memory capacity in neural networks. *Informatica*, 20(4), 1–10.

Griffits, R.B., Lebowitz, J.L. (1968). Random spin systems: some rigorous results. *J. Math. Phys.*, 9(8), 1284–1292.

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.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA*, 79, 2445–2558.

Katzgraber, H.G., Young, A.P. (2005). Probing the Almeida–Thouless line from the mean-field model. *Cond-mat.*, 0507138 v3, 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.

Landau, L.D., Lifschitz, E.M. (1989). *Theoretical Physics: Quantum Mechanics (Non-Relative Theory)*, III, Nauka Press, Moscow (in Russian).

Little, W.A., Shaw, G.L. (1978). Analytic study of the memory storage capacity of a neural network. *Math. Biosci.*, 39, 281–290.

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.

Sherrington, D., Kirkpatrick, S. (1975). Solvable model of a spin glasses. *Phys. Rev. Lett.*, 35, 1792–1795.

Sherrington, D., Southern, B.W. (1975). Spin glass versus ferromagnet. *J. Phys. F: Metal. Phys.*, 5, L49–L53.

Sompolinsky, H. (1986). Neural networks with nonlinear synapses and static noise. *Phys. Rev. A*, 4(3), 2571–2574.

Thouless, S.J., Anderson, P.W., Palmer, R.G. (1977). Solution of solable model of a spin glass. *Philos. Mag.*, 35(3), 593–601.

Vieira, A.P., de Carvallio, J.K., Salinas, S.R. (2001). Phase-diagram of a random-anisotropy mixed-spin Ising model. *Phys. Rev. D*, 63, 184415.

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 researher of the Institute of Mathematics and Informatics, a chairman of the Informatics Department of the Lithuanian Sience Society and a member of its council. His research interest includes neuroinformatics methodology, control problems and development of neural networks algorithms, recognition, chaos processes.

Neurotinklų teorijos išplėtimas įvedant Hebo postulatą

Algis GARLIAUSKAS

Straipsnyje fundamentinės klasikinės mechanikos teorijos, Izingo fizikos prasme, pagrindai perkeliami į taikomąją neurotinklų sritį. Atliktas neurotinklų teorijos išplėtimas įvedant Hebo postulatą ir formuluojant bei įrodant atitinkamą teoremą. Duotas skaitmeninis teorijos pagrindimas.