❐   510

Distributed memory of neural networks and the problem of the intelligence`s essence

Ibragim E. Suleimenov[1], Dinara K. Matrassulova[2], Inabat Moldakhan[2], Yelizaveta S. Vitulyova[2], Sherniyaz B. Kabdushev[3], Akhat S. Bakirov[2]

[1]V. I. Vernadsky Crimean Federal University, Simferopol, Republic of Crimea
[2]Almaty University of Power Engineering and Telecommunications, Almaty, Republic of Kazakhstan
[3]National engineering Academy of Republic of Kazakhstan, Almaty, Republic of Kazakhstan

Article Info

ABSTRACT

The question of the nature of the distributed memory of neural networks is considered. Since the memory capacity of a neural network depends on the presence of feedback in its structure this question requires further study. It is shown that the neural networks without feedbacks can be exhaustively described based on analogy with the algorithms of noiseproof coding. For such networks the use of the term "memory" is not justified at all. Moreover, functioning of such networks obeys the analog of Shannon formula, first obtained in this paper. This formula allows to specify in advance the number of images that a neural network can recognize for a given code distance between them. It is shown that in the case of artificial neural networks with negative feedback it is really justified to talk about a distributed memory network. It is also shown that in this case the boundary between distributed memory of a neural network and information storage mechanisms in such elements as RS-triggers is diffuse. For the given example a specific formula is obtained, which connects the number of possible states of outputs of the network (and, hence, the capacity of its memory) with the number of its elements.

Corresponding Author:

Akhat S Bakirov
Almaty University of Power Engineering and Telecommunications
127/1 Baitursynuly st, Almaty, Republic of Kazakhstan
Email: axatmr@mail.ru

1. INTRODUCTION

It is often emphasized that memory of artificial neural networks (ANN) is distributed in current literature [1], [2]. It is pertinent to emphasize that this conclusion was actually made on an empirical basis, as well as a significant part of other conclusions that in one way or another relate to the properties of ANNs [3], [4]. Until now, there are no algorithms that would allow calculating the weight coefficients of a specific ANNs, based on the requirements connected with the solving problem. Most of the results related to ANNs are in fact the results of various computer experiments [5], [6].

Accordingly, the conclusions that are made in the neuroscience, as a rule, are the result of empirical generalization [7], i.e. most of the conclusions suffer from the same drawback as the conclusions that are made on a purely empirical basis, without the use of certain methodological concepts. As a result, in the sources, especially popular [8], [9], very often neural networks operating on the basis of empirically established algorithms are opposed to software products that use explicitly prescribed algorithms.

Accordingly, the question of what can be considered the memory of neural networks, from the point of view of a consistent scientific methodology, today de facto remains open, especially if we consider it in

connection with the problem of the essence of intelligence [10], [11]. Indeed, if we consider a neural network with a concrete set of weight coefficients, then, strictly speaking, there is no need to talk about its memory at all. There is an object that performs specified operations in accordance with the structure of links and their characteristics. It is shown that the issue of distributed memory of neural networks is far from trivial in this report. Effects that can really be interpreted as the emergence of distributed memory occur only when the elements of the ANNs are covered by feedback. The results obtained are applied to the analysis of transpersonal information objects that arise in the global communication madia, i.e. the prerequisites for consistent interpretation of the essence of social consciousness and the collective unconscious are propesed.

We also emphasize that the proposed approach also allows one to obtain results that are of interest for the theory of neural networks. Namely, the question of how many images a neural network can recognize remains largely open [12], [13]. More precisely, the statistical properties of neural networks are still insufficiently studied, which, among other things, is since the overwhelming majority of the results obtained in this area are de facto an interpretation of computer experiments [5], [6]. Our proposed approach, which uses an analogy between neural networks and error-correcting coding algorithms, allows us to solve this problem. Specifically, we have obtained an analogue of Shannon's formula, which describes the limiting capabilities of the ANN in terms of pattern recognition.

This result, among other things, shows that, contrary to the opinion of many authors [14], [15], the use of the term "memory" in relation to neural networks in which there are no feedbacks is not completely correct. Memory as such appears only when feedbacks are formed in the neural network. Moreover, if the feedbacks are negative, then the line between the ANN memory and the information storage mechanisms that are currently implemented based on logical elements is erased.

2. RESEARCH METHOD
2.1. Comparison of neural networks with electronic circuits based on logical elements

The main research method used in this work is as follows. A comparison is made between ANNs that do not contain feedbacks and ANNs that contain negative feedbacks. The introduction of ANNs containing feedback into consideration allows us to show that there is no fundamental difference between ANNs built on the basis of formal neurons and classical electronic circuits assembled on the basis of logical elements. The essence of the research method used is illustrated by the following example.

The simplest system capable of storing information is RS-trigger, which includes two elements in Figure 1(a). This trigger can be in two different states, provided that there is no signal at its input. Consequently, this system really stores information, while those ANNs that operate only on the basis of direct signal transformation (for example, a feedforward ANNs), strictly speaking, do not perform such an operation. It is easy to see that topologically of the RS-trigger coincides with the circuit of the hopfield neuroprocessor as shown in Figure 1(b), provided that the latter includes only two elements. This observation itself makes us assume that in the case when the elements of the ANNs are covered by negative feedbacks, then some structural elements appear in them that are capable of storing information in the same sense in which the information is stored by the RS-trigger. Consequently, for such ANNs, the question of the nature of memory ceases to be trivial; more precisely, there is no clear boundary between the case of a "separate memory cell" and what is called distributed memory of ANN in current literature on neural networks [16].
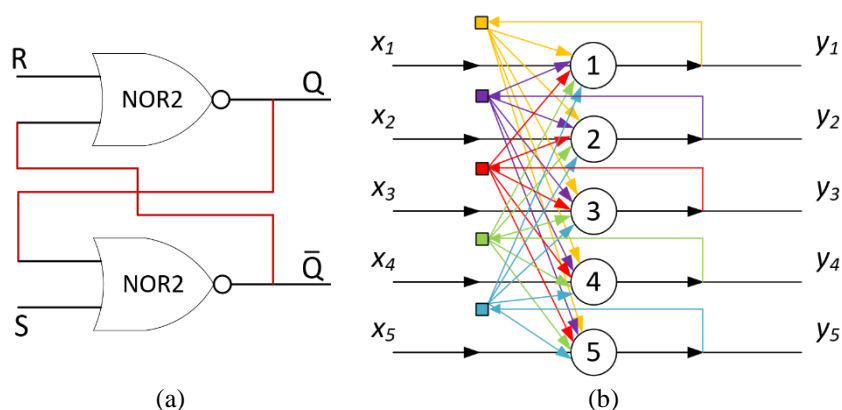

(a)   (b)

Figure 1. Diagram of the RS-trigger assembled on (a) the NOR logic elements and (b) hopfield's neuroprocessor-a case of five neurons

This issue is not only of highly specialized interest, but also has a direct bearing on the problem of the essence of intelligence. Recently we report [10], [11] that the intellect should first of all be considered as a kind of information processing system. It was also emphasized there that all discussions regarding which systems can be attributed to artificial intelligence, and which cannot, should be considered pointless, as long as the essence of intelligence as such remains undisclosed. Namely, all definitions of intelligence that are contained in the literature, including in the humanitarian, are essentially descriptive in nature, that is, they actually only list the features characteristic of a very specific kind of intelligence-human one. However, there is no reason to assume that intelligence must necessarily be close to human, respectively, any attempts to reveal its essence, starting from such tools as the Turing test, seem methodologically unfounded. Moreover, in [8] it was emphasized that there is every reason to consider nature of intelligence as a result of self-organization processes in the infocommunication space [17].

Namely, consider two people communicating with each other. It is customary to say that here two individuals enter into a dialogue, but this is very rough approximation. In fact, the exchange of signals between neurons that make up the brain of each of the interlocutors takes place. Continuing this logic, it is easy to conclude about the existence of a global communication network.

This example alone shows how important the issue of distributed memory of neural networks is, including from the point of view of interpreting the essence of intelligence. Indeed, if the memory of a neural network is distributed, in the strict sense of the word, then the collective neural network formed by individuals through interpersonal communication should be qualitatively different from a simple collection of separate fragments, each of which is localized within the brain of individuals.

Further, if the "information capabilities" of the aggregate neural network significantly exceed the information capabilities of the aggregate of its constituent fragments, but taken separately, then the global neural network can generate a new quality. In particular, it can, in one sense or another, store information that is only indirectly related to the memory of individuals. Similar conclusions can be traced in the humanities literature, in particular, the social consciousness, as stated in the sociological literature, is not reduced to the consciousness of individual individuals, but really responds to the emergence of a new quality [18]. Collective unconscious [19], mentality [20] and other collective effects associated with interpersonal communications are known well too.

All these questions are more than controversial; however, the above judgments are enough to demonstrate how important the question of the true nature of the distributed memory of neural networks is. As noted above, it is impossible to shed light on these issues through numerical experiments alone; a consistent methodology is needed here to reveal the essence of distributed memory. However, before proceeding to the construction of network models that demonstrate the true effects of distributed memory, we should consider, for comparison, networks where such effects are obviously absent. It is on this basis that the main research method of this work is built.

We first consider an ANN, in which there are no feedbacks, and prove that such a network can be exhaustively described on the basis of analogy with algorithms for error-correcting coding. This allows us to show that, strictly speaking, the term "memory" is not applicable to such networks at all. Moreover, as will be shown below, such networks can be exhaustively described based on analogy with error-correcting coding algorithms.

2.2. ANN research method by comparison with error-correcting coding algorithms

The essence of the used method is as: neural networks, which functioning is described by a direct functional dependence between the set of logical variables $\vec{U}_{i0}$ applied to the inputs and the set of logical variables describing the state of the outputs $\vec{U}_i$

$$\vec{U}_i = \vec{F}(\vec{U}_{i0}) \tag{1}$$

can be named networks with direct information processing. The obvious and most important example of such networks is feed-forward networks that do not have feedbacks.

The "logic" of neural networks of this type can be very simply disclosed [21], if a comparison is made between neural networks and error-correcting codes.

Let us consider a certain "image", which, possibly, contains a certain number of errors (here and below, an image is understood as a set of binary variables, each of which can take the values 0 or 1). Such "images" can encode images, voice signals, as well as many other things.

The main task that typical artificial neural networks perform is to reconstruct an image that contains errors based on a certain learning procedure. In other words, in accordance with Figure 2, there is a well-defined mapping of a certain set of binary variables onto its subsets. The error correction codes widely used in the telecommunications industry [22], [23] essentially solve a similar problem. The main idea of constructing such codes is based on the use of redundant information. So, to transmit one of 2^4 code

combinations, the binary (7,4) Hamming code uses 7 binary variables [24]. To simplify, one can say that 7 bits are used in order to convey the same amount of information that could be transmitted (with the deliberate absence of errors) using only 4 bits. From the point of view of set theory as shown in Figure 3, the use of error correction code can be interpreted as; the set A of all possible code combinations (in the example of the binary (7,4) Hamming code considered above-2^7) is divided into subsets A_i, the number of which is equal to the number of code combinations r, treated as a code with an error corrected (in this example, 2^4). Any a ∈ A_i is associated with a code combination with an absent error from the set B.
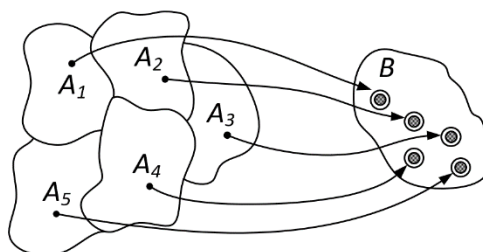
Figure 2. Morphism (surjection) of the set A into the set B, which defines the partitions of the set A into subsets A_i, each of which corresponds to a certain codeword with an absent error
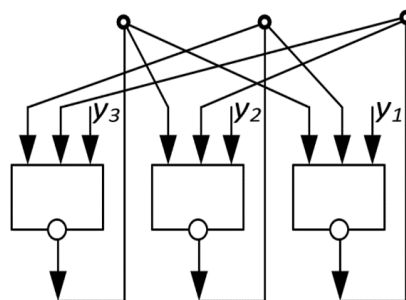

Figure 3. Transition from Hopfield neuroprocessor circuit to analog RS-trigger circuit on three elements

The neural network pattern recognition procedure can be viewed from exactly the same positions. As is known [25], [26], this procedure is as: at the inputs of the neurons that form the first layer of the network (for definiteness, we will consider feedforward networks), a set of binary variables is fed, interpreted as a recognizable pattern, possibly containing errors. At the outputs of the neurons of the last layer of the network (it is assumed that this network is trained on the corresponding set of images), a set of signals is formed that make up the initial image, which does not contain errors. It can be seen that the diagram in Figure 2 is also applicable for this situation.

The analogy with error-correcting coding makes a clear estimate of the number of patterns that can be recognized/reconstructed by the neural network. Of course, their architecture can be different, and the values of the weight coefficients can also be different. However, the indicated analogy makes it possible to make estimates based on the consideration of code distances; these estimates do not depend on the specifics of the ANNs of the type under consideration and, therefore, are very general in nature.

3. RESULTS
3.1. Analogue of Shannon's formula for ANN without feedback

To estimate the number of sequences capable of displaying all possible sequences of N characters with an admissible number of errors m, as shown in (2) is valid:

$$k_{N,m} = \frac{2^N}{1+\sum_1^m C_N^i} \qquad (2)$$

where m is the number of errors to be corrected, and C_N^i is the binomial coefficient.

The base 2 logarithm of $k_{N,m}$ allows us to estimate the number of symbols in binary sequences, the set of which provides coverage of the entire set of N-valued sequences with m admissible errors. Otherwise,

from the point of view of the goals of this work, the limiting case $N \gg m$ is of interest: the number of neurons in the considered network is significant, and the number of permissible deviations m is relatively small.

The following approximation was obtained for the asymptotic of the sum of binomial coefficients under the conditions $N \to \infty, \frac{m}{N} = o(1)$ in [27].

$$1 + \sum_{i=1}^{m} C_{N_0}^i \sim C_{N_0}^m \frac{N-m}{N-2m} \tag{3}$$

Table 1 shows the results of calculating the quantities $S = 1 + \sum_{1}^{m} C_{N_0}^i$ and $S_1 = 1 + \sum_{1}^{m-1} C_N^i$ from N, for $N = 2^k$ and $m = N/4$, as well as the values of the ratios S to the binomial coefficient $C_N^{N/4}$ and the ratios of these values to each other. It can be seen that as N increases, the results of direct calculation quickly approach the estimate given by (3).

Table 1. Results of numerical calculations to check the adequacy of the asymptotic estimate (4)

N	S	S_1	$C_N^{N/4}$	$C_N^{N/4}/S$	S_1/S
8	37	9	28	0.757	0.24
16	2517	697	1820	0.723	0.277
32	15033173	4514873	10518300	0.6997	0.3
64	7.13*1014	2.25*1014	4.89*1014	0.685	0.315
128	2.18*1030	7.07*1029	1.48*1030	0.676	0.324
256	2.83*1061	9.3*1060	1.9*1061	0.671	0.328
512	6.67*10123	2.2*10123	4.5*10123	0.669	0.33
1024	5.2*10248	1.7*10248	3.5*10248	0.6679	0.332

Thus, for the number of code sequences that can correct the number of errors up to 25% of the number of characters, can be written in (4):

$$k \sim \frac{2}{3} \frac{2^N}{C_N^m} \tag{4}$$

Further, with a large number of characters in the sequence, it is permissible to use the Stirling as shown in (5).

$$m! \approx \sqrt{2\pi m} \left(\frac{m}{e}\right)^m \tag{5}$$

where e is the base of natural logarithms. Accordingly, for the case $m = N/4$, we have

$$C_N^m \approx \frac{1}{\sqrt{\frac{3}{8}\pi N}} \frac{2^{2N}}{3^{\frac{3}{4}N}} \tag{6}$$

Using (2), we obtain

$$\log_2 k \sim \frac{1}{2} \log_2 \left(\frac{\pi N}{6}\right) + N \left(\frac{3}{4} \log_2 3 - 1\right) \tag{7}$$

Relationship (7) shows that for large N, the value $\log_2 k$, which determines the possibilities of information compression due to the factor of permissible errors, begins to depend on N almost linearly. In particular, this means that an increase in the degree of information compression due to the factor of permissible errors with an increase in the number of symbols in a code sequence is achieved only at relatively small N. Thus, with 25% of permissible errors, 16-digit sequences can actually be transmitted by a 5-digit code, but for large N, the degree of information compression practically ceases to depend on N. Indeed, the ratio $\frac{\log_2 k}{N}$, can be taken as the degree of information compression, which, as shown in (7), remains constant with high accuracy at large N

$$\frac{\log_2 k}{N} \sim \left(\frac{3}{4} \log_2 3 - 1\right) = 0,189 \tag{8}$$

In (12) admits a natural generalization to the case of any multiplier q specifying the percentage of acceptable errors; $m = qN$

Using again the Stirling formula for the asymptotic behavior of the binomial coefficient, we obtain

$$\log_2 k_{N,qN} \sim N[1 + q \log_2 q + (1-q) \log_2(1-q)] + \log_2 \frac{1-2q}{1-q} + \frac{1}{2}\left(\log_2 N + \log_2 2\pi q(1-q)\right) \qquad (9)$$

For large N, the constant terms, as well as the term that depends on N logarithmically, can be neglected in this formula. Consequently, the degree of information compression and, in the general case, does not depend on N

$$\frac{\log_2 k_{N,qN}}{N} \sim 1 + q \log_2 q + (1-q) \log_2(1-q) \qquad (10)$$

The last two terms on the right-hand side of (10) coincide with the formula for the Shannon (informational) entropy H of a sequence of binary signals, taken with the opposite sign;

$$H = -q \log_2 q + (1-q) \log_2(1-q) \qquad (11)$$

Interpreting q as the probability of an error in a sequence of binary symbols, we can write

$$\frac{\log_2 k_{N,qN}}{N} \sim 1 - H(q) \qquad (12)$$

In (12), which can be interpreted as an analogue of Shannon's formula, obtained to describe the operation of neural networks, has an extremely transparent meaning: information entropy is a measure of uncertainty introduced due to the appearance of errors. If such uncertainty is introduced artificially, i.e. the appearance of errors with a frequency q is acceptable (since they are interpreted as permissible deviations), then the measure of information compression provided due to this factor should be determined by the entropy factor.

The material of this section proves that for "direct action" ANNs the conclusion about their distributed memory is valid only with very large reservations. More precisely, as shown in (12) shows that with an increase in the number of elements of a network of this type, a new quality cannot arise in it. A new quality appears only when the elements of the ANNs are caught up in a feedback loop. The proof of this statement is given below on the example of ANNs, which are direct analogs of the RS-trigger.

3.2. The specificity of the ANN's distributed memory: the importance of negative feedbacks

The main feature of the RS-trigger, which allows it to store information, is the inverse output. Accordingly, a circuit similar to an RS-trigger, but containing three elements, will have the form shown in Figure 3. This circuit is simultaneously analogous to both the RS-trigger and the Hopfield neuroprocessor. The inputs of each of its elements are fed, firstly, external control signals, and secondly, inverted signals taken from the two remaining elements of the system.

It is convenient to consider the case when each logical element included in this system performs the following operation:

−   at the output of the system, a logical unit is formed when the total number of logical units at its input is equal to zero or one;
−   a logical zero is formed at the output of an element when the total number of logical ones at its input is two or three.

Provided that all the elements available in the circuit shown in Figure 3 form a signal at the output that corresponds to either a logical zero or a logical one, this description is exhaustive. One can see that in such a situation, the number of stable states of the system is exactly three. One can see that the nature of the operations performed by the elements of the considered scheme fully corresponds to the classical formula describing the functioning of an individual neuron too.

Indeed, if we assume that logical zero and logical one corresponds to zero and one signal levels, then these operations are described by (13).

$$Y_j = \theta \left(\sum_{i \neq j}^m \bar{Y}_i + X_j - \frac{m}{2}\right) \qquad (13)$$

where $Y_j = 0,1$ is a variable describing the state of the output of the j−th element, $\bar{Y}_i$ is its inverse value, $X_j = 0,1$ is a variable describing the state of the input of the j−th element, $\theta(x)$ is the Heaviside function,

$$\theta(x) = \begin{cases} 0, x < 0 \\ 1, x \geq 0 \end{cases} \qquad (14)$$

We emphasize that, in accordance with (13), the shift for all network elements is the same and equal to $\frac{3}{2}$. All possible stable states of the system, the diagram of which is shown in Figure 3, correspond to the situation when a logical unit is formed at the output of one of the elements, which can be chosen arbitrarily, and logical zeros are formed at the output of the other two.

The generalization of (13) to the case of a network containing an arbitrary odd number of elements of the type under consideration has the form

$$Y_j = \theta \left(\sum_{i \neq j}^{m} \bar{Y}_i + X_j - \frac{m}{2} \right) \qquad (15)$$

where m is the number of elements in the system.

A system of the type under consideration, containing five elements, is stable when the number of elements, at the output of which a logical unit is formed, is equal to two. Two elements, at the output of which a logical unit is formed, can be chosen arbitrarily due to the complete symmetry of the circuit, as shown in Figure 4.
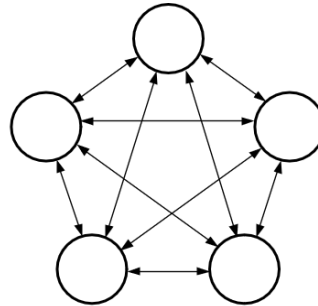
Figure 4. Circuit topology Figure 2, highlighting its completely symmetrical character

Thus, if in the first example (Figure 3) the number of possible states was equal to three, then in the second example (Figure 5) the number of possible logical stable states n already reaches ten $n = C_5^2$. This is what allows us to speak about the existence of a real distributed memory of neural networks, provided that negative feedbacks can take place in such networks (which corresponds to the use of inverted signals fed to the inputs of circuit elements in Figure 4.
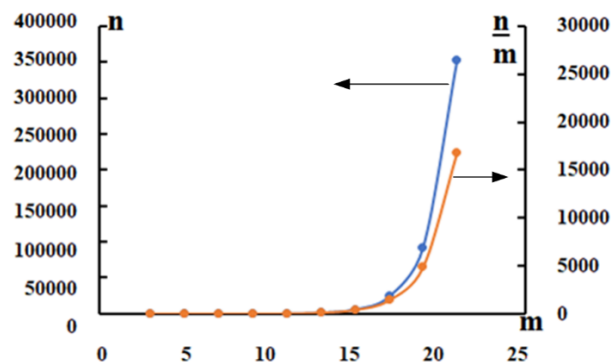

Figure 5. Dependence of the number of admissible states of an analogue of RS-trigger containing m elements (left axis) and the number of admissible states per one element (right axis) on the number m

We emphasize that the considered analogy between the RS-trigger and the Hopfield neuroprocessor (which is legitimate, at least if we talk about the topology of the circuits of the type considered above) allows us to assert that the informational capabilities of an individual element that is part of a neural network significantly increase if the number of neurons in the ANN increases.

More precisely, the number of admissible states per one element of the system grows nonlinearly, with an increase in the number of elements in the system. The general formula for the number of stable states n in a completely symmetric analogue of RS-trigger containing m elements, which can be derived by considering circuits similar to those shown in Figure 5, can be written as follows.

$$n = C_m^{\frac{m-1}{2}}$$ (16)

It is assumed that the number of elements m in the system is odd. Result (16) was obtained directly on the basis of (15). The graph of the dependence given by (23) is shown in Figure 5. The same figure shows the dependence of the number of admissible stable states $\frac{n}{m}$ per one element of the system on m. It can be seen that the "information capabilities" of an individual element really increase significantly as the number of elements in a network of the type under consideration increases.

Thus, it can be argued that if feedbacks are implemented in a neural network, then it can acquire very nontrivial properties. In particular, in such a system, memory can be implemented, interpreted in the same sense in which memory is inherent in the classic element of storage systems-a trigger.

From the point of view of methodology, an example of fully symmetric analogs of RS-trigger, first of all, shows that the border between artificial neural networks (the term is understood in the broadest sense of the word) and systems with lumped memory is at least diffuse. If we endow the elements that make up the neural network with inverters, or rather, if negative feedbacks can be realized in such a system, then a very nontrivial distributed memory appears in it. If there are no such elements, then in accordance with the conclusions made in the previous section, the neural network cannot form a new quality-it is nothing more than a system that solves the same problem that is solved in the theory of error-correcting coding.

4. DISCUSSION

Let us emphasize that the question of the dependence of the memory capacity of a neural network on the number of elements is fundamental. If the memory capacity of a neural network depends on the number of its elements nonlinearly (which corresponds to the model built in the previous section), then the global neural network, which is formed due to interpersonal communications, should acquire some additional qualities that are not reduced to the properties of individual fragments of the neural network that localized within the brain of each of the individuals.

We emphasize again that this thesis has already been formulated in the humanitarian literature. Particuliarly, it is proved that public consciousness is not reduced to the consciousness of individuals. This is something qualitatively different. It should be more correct to say that there is a transpersonal level of information processing, the properties of which are still poorly understood. At this level of information processing, some non-trivial information entities may well be formed. Indeed, if the exchange of signals between neurons localized within the brain of an individual leads to the emergence of intelligence and the human mind, then it is quite reasonable to assume that the exchange of signals in the global network generates information objects that also have very nontrivial properties.

This conclusion, in turn, cannot but force us to reconsider the views on the formation of human intelligence and its evolution (even throughout written history). Indeed, the collective "component" of human intelligence cannot but depend on how society is organized. Accordingly, there is every reason to believe that the history of mankind can be viewed from the standpoint of the co-evolution of human intelligence and social structure. At a minimum, the organization of society in the most significant way affects the nature of interpersonal communications, and, consequently, the information objects that are formed at the transpersonal level of information processing.

In modern conditions, when many political scientists, not without reason, talk about cardinal transformations of the world order [28] and telecommunication technologies are developing more and more rapidly, it is permissible to conclude that, in fact, we are talking about changes in the structure of global infocommunication space. This cannot but affect the evolution of human intelligence as such, in accordance with the conclusions made in this work.

Simplifying, if the previous transformations of the world order (for example, those associated with the first and second industrial revolutions) practically did not affect human intelligence, today the situation is fundamentally different. The formation of human-machine systems (and this is how the rapid development of social online networks and the growth of their influence on society should be interpreted) cannot but have a serious impact on the fundamental base of civilization-human intelligence. De facto, we are talking about the fact that social online networks cannot but strengthen the collective component of intelligence.

Of course, all these phenomena still require a comprehensive study; their detailed consideration is obviously beyond the scope of a separate work. However, it should be emphasized that the very fact of

recognizing the existence of a transpersonal level of information processing is a significant step forward. At a minimum, sociology acquires an additional tool that allows one to interpret and quantitatively describe those phenomena that were previously discussed only at the phenomenological level of understanding.

Further, outside the scope of the work remains the question of how negative feedbacks correspond to real physiological processes in the human brain. However, for the purposes of this work, this question is not decisive. Indeed, the emergence of distributed memory in society because of interpersonal information exchange can also be interpreted based on simplified models, when an individual is considered as an analogue of a neuron.

A corresponding example was presented in [29], where it was shown that there are conditions under which any voting council is converted into a kind of analogue of a neural network. This example is informative because it allows one to prove that under certain conditions, a decision on a particular issue is made not by a set of Council members, but by a neural network formed from them, which itself has a certain distributed memory.

For this kind of models, the question of the possibility of forming negative feedbacks is obviously not worth it since the mutual influence of council members on each other can certainly have both a positive and a negative sign. As was emphasized, there is a visible visual proof of this. Thus, a member of the dissertation council can vote against a dissertation submitted by a student of his personal enemy, even when it deserves a positive attitude towards itself.

5. CONCLUSION

Thus, the question of the nature of the distributed memory of neural networks turns out to be closely related to the question of the essence of intelligence. Namely, from the conclusion that the memory capacity of a biological neural network can nonlinearly depend on the number of elements, i.e. if the encompassing neural network can perform more complex operations than its parts taken separately, the existence of a transpersonal level of information processing follows. From this, in turn, it follows that human intelligence should be viewed from the standpoint of dialectics-it is simultaneously formed both through the process of signal exchange between neurons localized within the individual's brain and through the exchange of signals between such relatively independent fragments of the global neural network, which can be identified with the noosphere. Otherwise, human intelligence is dialectical in nature, it is determined by both individual and collective principles. This allows us to conclude that the problem of the essence of intelligence is inextricably linked with the question of the nature of the distributed memory of artificial neural networks. Mathematical models proving that the capacity of the distributed memory of a neural network can nonlinearly depend on the number of elements in the system become a tool for studying transpersonal information structures, i.e. one of the most important "components" of human intelligence. Models of this type presented in this work show, however, that a nonlinear dependence of the memory capacity of a distributed neural network on the number of its elements can take place only when these elements are covered by negative feedbacks. If there are no such connections in the system, the above dependence quickly approaches linear as the number of elements in the network increases. Such networks allow direct analysis based on comparison with error-correcting codes; moreover, the formula expressing the number of corrected errors in terms of Shannon's entropy is valid. Consequently, a new quality that makes it possible to approach the modeling of real intelligence can appear only in neural networks that have negative feedbacks. Such networks, however, are also dialectical in nature: their memory can be both distributed and concentrated at the same time. The proof of this, in particular, is that the circuit of the classical RS-trigger is topologically equivalent to the circuit of the Hopfield neuroprocessor, which contains two elements.

REFERENCES

[1]   L. K. Hansen and P. Salamon, "Neural network ensembles," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 10, pp. 993–1001, 1990, doi: 10.1109/34.58871.
[2]   W. Rawat and Z. Wang, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review," *Neural Computation*, vol. 29, no. 9, pp. 2352–2449, Sep. 2017, doi: 10.1162/neco_a_00990.
[3]   R. I. Bendjillali, M. Beladgham, K. Merit, and A. Taleb-Ahmed, "Illumination-robust face recognition based on deep convolutional neural networks architectures," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 2, p. 1015, May 2020, doi: 10.11591/ijeecs.v18.i2.pp1015-1027.
[4]   A. Herlambang, P. W. Buana, and I. N. Piarsa, "Cloud-based architecture for face identification with deep learning using convolutional neural network," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 2, p. 811, Aug. 2021, doi: 10.11591/ijeecs.v23.i2.pp811-820.
[5]   A. Kurniawan and A. Harumwidiah, "An evaluation of the artificial neural network based on the estimation of daily average global solar radiation in the city of Surabaya," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 3, p. 1245, Jun. 2021, doi: 10.11591/ijeecs.v22.i3.pp1245-1250.
[6]   P. R. Iyer, S. R. Iyer, R. Ramesh, A. M.R., and K. N. Subramanya, "Adaptive real time traffic prediction using deep neural

networks," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 8, no. 2, p. 107, Jun. 2019, doi: 10.11591/ijai.v8.i2.pp107-119.

[7] S. Ibragim, B. Akhat, M. Dinara, G. Anastasiya, K. Mariya, and M. Grigoriy, "Example of the Use of Artificial Neural Network in the Educational Process," 2020, pp. 420–430.

[8] G. Anthes, "Lifelong learning in artificial neural networks," *Communications of the ACM*, vol. 62, no. 6, pp. 13–15, May 2019, doi: 10.1145/3323685.

[9] B. Yegnanarayana, *Artificial neural networks*. PHI Learning Pvt. Ltd., 2009.

[10] I. Suleimenov, A. Massalimova, A. Bakirov, and O. Gabrielyan, "Neural Networks and the Philosophy of Dialectical Positivism," *MATEC Web of Conferences*, vol. 214, p. 02002, Oct. 2018, doi: 10.1051/matecconf/201821402002.

[11] I. E. Suleimenov, O. A. Gabrielyan, A. S. Bakirov, and Y. S. Vitulyova, "Dialectical Understanding of Information in the Context of the Artificial Intelligence Problems," *IOP Conference Series: Materials Science and Engineering*, vol. 630, no. 1, p. 012007, Oct. 2019, doi: 10.1088/1757-899X/630/1/012007.

[12] S. España-Boquera, M. J. Castro-Bleda, J. Gorbe-Moya, and F. Zamora-Martinez, "Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 767–779, Apr. 2011, doi: 10.1109/TPAMI.2010.141.

[13] Q. Dou *et al.*, "Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1182–1195, May 2016, doi: 10.1109/TMI.2016.2528129.

[14] A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, and T. Lilicrap, "Meta-learning with memory-augmented neural networks," in *ICML'16: Proceedings of the 33rd International Conference on International Conference on Machine Learning*, 2016, pp. 1842–1850, doi: 10.5555/3045390.3045585.

[15] G. R. Fink, H. J. Markowitsch, M. Reinkemeier, T. Bruckbauer, J. Kessler, and W.-D. Heiss, "Cerebral Representation of One's Own Past: Neural Networks Involved in Autobiographical Memory," *The Journal of Neuroscience*, vol. 16, no. 13, pp. 4275–4282, Jul. 1996, doi: 10.1523/JNEUROSCI.16-13-04275.1996.

[16] H. Yoon, J. H. Nang, and S. R. Maeng, "Parallel simulation of multilayered neural networks on distributed-memory multiprocessors," *Microprocessing and Microprogramming*, vol. 29, no. 3, pp. 185–195, Oct. 1990, doi: 10.1016/0165-6074(90)90005-T.

[17] I. E. Suleimenov, Y. S. Vitulyova, A. S. Bakirov, and O. A. Gabrielyan, "Artificial Intelligence," in *Proceedings of the 2020 6th International Conference on Computer and Technology Applications*, Apr. 2020, pp. 22–25, doi: 10.1145/3397125.3397141.

[18] S. Baronett, *Journey into philosophy: An introduction with classic and contemporary readings*, 1st Editio. Routledge,Taylor & Francis, 2017.

[19] C. G. H. Jung, *Collected Works of C.G. Jung, Volume 4: Freud & Psychoanalysis*. Princeton University Press, 2014.

[20] J. Go and G. Lawson, Eds., *Global Historical Sociology*. Cambridge: Cambridge University Press, 2017.

[21] A. Bakirov and I. Suleimenov, "On the possibility of implementing artificial intelligence systems based on error-correcting code algorithms," *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 1, p. 83, 2021.

[22] J. Boiko, I. Kovtun, and S. Petrashchuk, "Productivity of telecommunication systems with modified signal-code constructions," in *2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, Oct. 2017, pp. 173–178, doi: 10.1109/INFOCOMMST.2017.8246374.

[23] B. S. G. Pillai, N. P. Anthapadmanabhan, D. Suvakovic, and H. Chow, "Multi-user encoding for forward error correction in passive optical networks," in *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2016, pp. 2467–2470, doi: 10.1109/ISCAS.2016.7539092.

[24] A. K. Singh, "Error detection and correction by hamming code," in *2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)*, Dec. 2016, pp. 35–37, doi: 10.1109/ICGTSPICC.2016.7955265.

[25] M. Koziarski and B. Cyganek, "Image recognition with deep neural networks in presence of noise – Dealing with and taking advantage of distortions," *Integrated Computer-Aided Engineering*, vol. 24, no. 4, pp. 337–349, Sep. 2017, doi: 10.3233/ICA-170551.

[26] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 8697–8710, doi: 10.1109/CVPR.2018.00907.

[27] V. A. Sorokin, "Asymptotics for combinatorial sums related to the problem of the number of Latin rectangles," *Bulletin of the Bashkir State Pedagogical University*, no. 3, pp. 32–42, 2016, (In russian, online: https://cyberleninka.ru/article/n/metodika-postroeniya-asimptotik-dlya-kombinatornyh-summ-soderzhaschih-binomialnye-koeffitsienty).

[28] Jacques Sapir, "Are We on the Verge of a Major Transformation of the Global Economy?," *Studies on Russian Economic Development*, vol. 31, no. 6, pp. 606–620, Nov. 2020, doi: 10.1134/S1075700720060118.

[29] Y. S. Vitulyova, A. S. Bakirov, S. T. Baipakbayeva, and I. E. Suleimenov, "Interpretation of the category of 'complex' in terms of dialectical positivism," *IOP Conference Series: Materials Science and Engineering*, vol. 946, no. 1, p. 012004, Oct. 2020, doi: 10.1088/1757-899X/946/1/012004.

BIOGRAPHIES OF AUTHORS

Ibragim E. Suleimenov 🆔 📊 SC Ⓟ Professor of the Crimean Federal University named after V.I. Vernadsky (until 2020-Professor of the Almaty University of Energy and Communications). Graduated from the Physics Department of the Leningrad University named after A.A. Zhdanov in 1986; defended his thesis for the degree of candidate of physical and mathematical sciences in 1989 at the same university. In 2000, he defended his thesis for the degree of Doctor of Chemical Sciences at the Al-Farabi Kazakh National University. Academician of the National Engineering Academy of the Republic of Kazakhstan (since 2016), full professor (since 2018) according to the official certificate of the Ministry of Education and Science of the Republic of Kazakhstan. Actively develops interdisciplinary cooperation, including between natural science and humanities. He pays considerable attention to the interdisciplinary study of intelligence, both using mathematical models and at the level of philosophical interpretation. He can be contacted at email: esenych@yandex.kz.

Dinara K. Matrassulova (ID) (g) (SC) (P) is a PhD student of the Almaty University of Power Engineering and Telecommunications. In 2015 she received a bachelor's degree and in 2017 a master's degree in "Radio engineering, electronics and telecommunications" at the Almaty University of Power Engineering and Telecommunications. She works in the telecommunications company Kcell JCS, as a senior specialist in the field of fixed internet. Actively studies telecommunications, networks, artificial intelligence, neural networks, signal processing. She can be contacted at email: dinara.kutlimuratovna@gmail.com.

Inabat Moldakhan (ID) (g) (SC) (P) is a PhD student at the Almaty University of Power Engineering and Telecommunications. She received her master's degree in 2011 with a degree in radio engineering and communications at the Almaty University of Power Engineering and Telecommunications. From 2010 to 2019, she worked at KaR-Tel LLP (a mobile operator operating under the Beeline brand) as a leading engineer of the radio network monitoring and control department. At the moment she is engaged in research in the field of radio engineering, electronics and telecommunications in accordance with the topic of her PhD thesis "Development of new approaches to digital signal processing based on ternary logic". She can be contacted at email: imoldakhan@gmail.com.

Yelizaveta S. Vitulyova (ID) (g) (SC) (P) is a PhD student at the Almaty University of Power Engineering and Telecommunications after Gumarbek Daukeyev (AUPET). She received her master's degree in 2016 with a degree in radio engineering and communications at the AUPET. From 2016 to present she worked at AUPET as a senior lecturer of the department of Radio engineering, electronics and telecommunications. At the moment she is engaged in research in the field of radio engineering, electronics and telecommunications in accordance with the topic of her PhD thesis "Post-industrial paradigm of development of infocommunication segment in the military-industrial complex of the Republic of Kazakhstan". She can be contacted at email: lizavita@list.ru.

Sherniyaz B. Kabdushev (ID) (g) (SC) (P) is a PhD student at the Almaty University of Power Engineering and Telecommunications and Researcher at the National Engineering Academy of the Republic of Kazakhstan. He received his master's degree in 2015 with a degree in radio engineering and communications at the Almaty University of Power Engineering and Telecommunications. From 2013 to 2017, he worked at Kaztranscom JSC (a mobile operator) as a senior manager of the technical support of sales department and still have worked as a researcher at the National Engineering Academy of the Republic of Kazakhstan until now. He can be contacted at email: sherniyaz.kabdushev.hw@gmail.com.

Akhat S. Bakirov (ID) (g) (SC) (P) Senior Lecturer at the Almaty University of Power Engineering and Telecommunications. In 2015 he received a bachelor's degree and in 2017 a master's degree in "Radio engineering, electronics and telecommunications" at the Almaty University of Power Engineering and Telecommunications. In 2017-2020 was PhD student at the Almaty University of Power Engineering and Telecommunications. At the moment he is engaged in research in the field of radio engineering, electronics and telecommunications in accordance with the topic of his PhD thesis "Development of theoretical bases of methods of counteraction to modern forms of information warfare". He can be contacted at email: axatmr@mail.ru.