A COGNITIVE APPROACH TO NEURAL NETWORK MODEL BASED ON THE COMMUNICATION SYSTEM BY AN INFORMATION CRITERIA

A. Horiushkina, S. Seker, and U. Korkmaz

Abstract— In this paper presents the structural analogy between the artificial neural net (ANN) architecture and communication system principles. Thus, a feed forward auto-associative ANN is assumed like a communication system and the channel capacity is defined as an information measure related to the number of hidden nodes which are in one hidden layer of the neural net. Here, the channel capacity formula which is given for the communication channel is represented as information a criterion that is connected with Shannon's entropy. Hence it is defined as information capacity under the interpretation of cognitive capacity and set theory approach is also used to define the ANN in manner of cognitive system.

Keywords — *Neural Nets, Communication Channel, Learning System, Shannon's Entropy, Cognitive Systems.*

1. INTRODUCTION

THE rapid development of technology over the past decade in the neural network topology area, has determined that one of the most important problems here is the determination of number of hidden nodes and layers in the neural topology. It is shown that there are a lot of studies in the theoretical and application area [1-3]. As studies show, at present time this is still an open problem, but its importance is reduced due to the development of hardware technologies, and then the time for computing data is reduced. However, technology which are using today goes to different fields for example quantum computing and artificial cognitive system. Especially, the cognitive system engineering applications will be point of focus at next generations of the technology. The cognitive system consists of: Sensing, learning, communication, logic and decision making. In this sense, a neural network can be considered as learning system which is consistent with the data, and also it can be interpreted from this framework and it is presented as follows. This study defines the artificial neural networks (ANN) as a cognitive structure and its hidden layer is described as communication channel and hence, interpreted as an information processor unit. In this manner ANN's cognitive interpretation is given.

2. ARTIFICAL NEURAL NETWORK (ANN) AND A BASIC COMMUNICATION SYSTEM

An artificial Neural Network can be related with the Communication system as below.

2.1. The Topology of ANN

An ANN-topology can be defined as parallel massively distributed system [4,5] by the following figure 1:

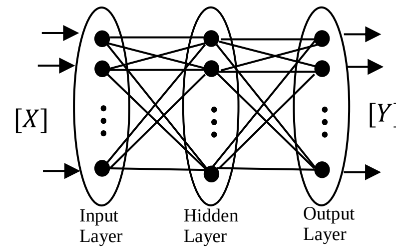

Fig.1. An auto associative neurol network

As indicated in figure 1 it is in the form of auto-associative, namely its input nodes number equals to output nodes number. Where, the input and output matrices are respectively defined with the same dimensions as below:

$$X \triangleq \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1M} \\ \vdots & \vdots & & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NM} \end{bmatrix}_{N\times M}$$

and

$$Y \triangleq \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1M} \\ \vdots & \vdots & & \vdots \\ y_{N1} & y_{N2} & \cdots & y_{NM} \end{bmatrix}_{N\times M}$$

Here the neural topology is accepted with only one hidden layer to get the simplicity and the number of the hidden nodes taken place in the hidden layer is considered as an unknown parameter (variable). This unknown parameter, which is the number of hidden nodes, is represented as channel capacity of the Communication system.

2.2. Communication Channel

Basic structure of a communication system is highly similar with the neural topology [6] as shown in Figure 1 and 2,

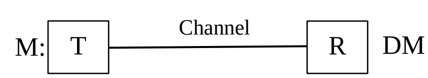

Fig. 2. A Communication System.

where

M: Modulator, T: Transmitter, R: Receiver and DM: Demodulator;

The communication channel can be represented in the form of wire, air, radio link etc.

The channel capacity is described by the Shannon's formula as follows

$$C_{max} = Blog_2[1 + SNR] \tag{1}$$

Here,

C_{max}: Maximum channel capacity and $[C_{max}]$: $bit/_s$

B: Band-width of the channel,

Alla Horiushkina, is Computer science, Universytet Morski w Gdyni, Gdynia, Poland, (e-mail: aegoriushkina@gmail.com).

Serhat Seker, is with Department of Electrical Engineering, Istanbul Technical University, Istanbul, Turkey, (e-mail: sekers@itu.edu.tr).

Ufuk Korkmaz, is with Department of Electrical Engineering, Istanbul Technical University, Istanbul, Turkey, (e-mail: ufukkorkmaz@itu.edu.tr).

Manuscript received Mar 1, 2018; accepted Apr 1, 2018.
Digital Object Identifier:

SNR: Signal to Noise Ratio.

3. SHANNON'S INFORMATION AND ANN-TOPOLOGY

Analysis of the literature [1-3] showed that C. Shannon was defined the channel capacity by a mathematical method as indicated in Eq. (1). The Shannon's approach can be generated from the Boltzmann's entropy [7] approach as shown in

$$H = k\ln(W). \tag{2}$$

Here W is the number of macroscopic states in the system while indicating the constant k is Boltzmann's constant. C. Shannon developed this concept in the probabilistic manner and it was given as written below:

$$H = -\sum_{i=1}^{N} p_i \, log_2\,(p_i). \tag{3}$$

Where, N is number of the possible states and p_i is (i=1, 2,…, N) also probability value which is connected with each case (state).

3.1. ANN Topology and Channel Capacity

An auto-associative neural network topology is considered as shown in the figure 1, it is input-output pairs has same dimensions as vector sizes, in terms of the number of training patterns, the following assumptions is taken into account:

Assumption:

The input matrix can be defined as a square matrix

$$\underset{\sim}{\mathrm{I}} \triangleq \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ \vdots & \vdots & & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NN} \end{bmatrix}_{N\times N}$$

The probability of each of elements is appeared in the matrix $\underset{\sim}{\mathrm{I}}$ is defined with equally likelihood cases:

$$p_i \triangleq p(x_i) = \frac{1}{N}$$

Hence, the Shannon's formula which is given by equation (1) can be written as

$$H = -\sum_{i=1}^{N} p_i \, log_2\,(p_i) = -N\left(\frac{1}{N} log_2\left(\frac{1}{N}\right)\right),$$

$$H = \; -N log_2\left(\frac{1}{N}\right) = log_2(N). \tag{4}$$

And it is approximated to number of hidden nodes M, using proportionality:

$$M \propto H = log_2\,(N) \tag{5}$$

Therefore, this proportionality, the factor J used in Equation (6) and it is an integer.

As a result of this calculation, the equation (1) that is Shannon's channel capacity can be extended to neural topology in the form of

$$M = J\,log_2\,(N). \tag{6}$$

In this formula, N is a number of states or number of elements of input vector. Factor J is also proportionality constant and M is also number of the hidden nodes in the one hidden layer [1,2]. This is also the equality of channel capacity which is defined in the Communication Systems. Hence, it is described as an information amount (Inf.) depending on the topology [1, 2,8].

4. DEFINITION OF A COGNITINE SYSTEM

A cognitive system is related with cognition and the cognition is also a combination of the Perception, Attention, Memory and Intelligent. In general manner, the cognitive system is defined with the other abilities between the perception and action. These are realized on the environment by means of the feedback mechanism as shown in the following figure 3 [9].

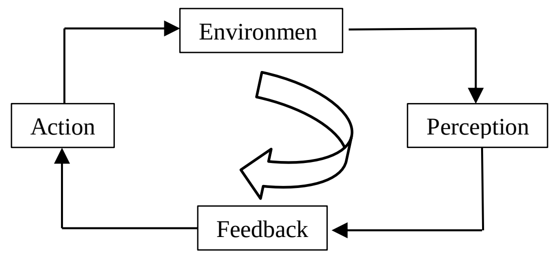

Fig. 3. Cognitive System: Dynamical Perception-action cycle.

This perceptron-action cycle gets to provide the information about the environment. Hence, the information is gained by going on one cycle to the next. This is a Cognitive Dynamical System.

In this cycle, the action is an executive subsystem and it plays role of the memory received from the environment. In this manner, it is presented as an executive memory that learns from actions on the environment. From another side, the perceptual memory is connected with the learning from observations of the environment. More general form of the memory system is "working memory" which provides the means between the perceptual memory and executive memory. The whole memory system provides the prediction of actions taken by the system while learning from the system [10, 11].

Attention is another function of the cognitive system and it is a structure which is built on the perception and memory system. In this manner it is a subsystem of the perception and memory. Intelligence is a basic function of cognitive system and it is based upon the perception, memory and attention as an upper system of 4-functions. These 4-functions of the cognition mechanism provide the decision making to do the intelligent choices under the uncertainties. In this sense a cognitive system is defined by the following Venn-diagram:

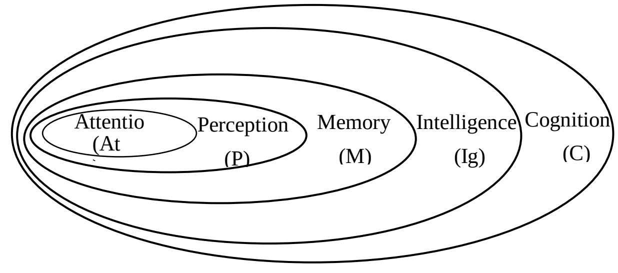

Fig. 4. Venn-diagram of cognitive system.

Where *At*: Attention and $At \subset P \subset M \subset Ig \subset C$

$$M \supset (At \cup P) \tag{7}$$
$$Ig \supset (At \cup P \cup M) \tag{8}$$
$$C \supset (At \cup P \cup M \cup Ig) \tag{9}$$

5. CONCLUDING REMARKS AND COGNITIVE INTERPRETATION

The comparisons on the cognitive system (COGS) and artificial Neural Network (ANN) can be representing as shows bellow:

$$ANN{:}\, f(I, O, W, L)$$
$$COGS{:}\, f(At, Ac, P, M, Ig)$$

Fig. 5. The displaying between ANN and COGS.

Here ANN is a function of I: Input; O: Output; W: Weight and L: Learning as well as architecture for COGS it becomes At: Attention; Ac: Action; P: Perceptron; M: Memory and Ig: Intelligence.

Thus, ANN provides the one-to-one relationship with COGS. But, the connection between L and Ig is dashed line because of the connection in weak sense because Ig is not identity to the learning. However, using the communication channel capacity approach, it is connected with the number of the hidden nodes in a one-hidden layer and hence, this units play role of the Information processor and its capacity. Therefore, it is shown as below:

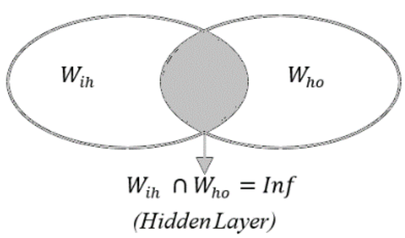

Fig. 6. The Venn diagram for getting the information (Inf) in the hidden layer.

As a result the functional structure of the trained ANN can be given as $ANN^*{:}\, f(I, O, W^*, L, Inf)$; *: after the training.

Then, ANN^* becomes a cognitive system under the consideration of following comparison:

$$ANN^*{:}\, f(I, O, W^*, L, Inf)$$
$$COGS{:}\, f(At, Ac, P, M, Ig)$$

Fig. 7. The displaying between ANN* and COGS.

Finally, the concept of the cognition can be attributed to the ANN-structure by means of the information capacity of the hidden nodes in ANN.

5.1. Displaying Compositions Between ANN and COGS Sets

In set theory, there are three types of the relationships. These are respectively injective, surjective and bijective. If the displaying considered as shown in figures (5) and (7), they can be represented in the form of the combinations.

$$Displaying : ANN \rightarrow ANN^* \rightarrow COGS$$

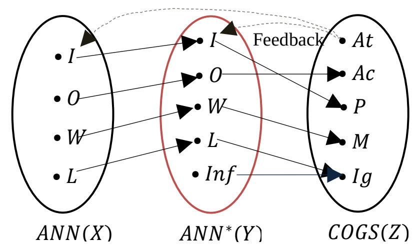

Fig. 8. The injective compositions of Displaying.

This is injective combination: ANN^* need not be injective, in this manner the relationship between ANN and $COGS$ is one-to-one relation with the reason of injective property. However, the element At, which shows the attention function, in the $COGS$ is not connected with any element in ANN or ANN^*. However, this function can be related with the input element of the ANN (or ANN^*). This provides a feedback connection of the cognitive system. To get this feedback, the order of the displaying is changed as follows:

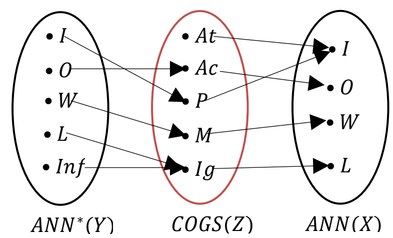

Fig. 9. The surjective compositions of Displaying.

Here the entire relationship is in the form of surjective composition. It means that if there is any relationship in a next way $f{:}\, Y \rightarrow Z$ it can be shown by the $f{:}\, Z \rightarrow Y$, then it can be interpreted as ANN^* also becomes Cognitive system while a Cognitive system is displaying to a ANN system. Hence, the cognitive systems can be interpreted by means of the ANN [12].

REFERENCES

[1] S. Seker, "An Analytical Approach Based on Information Theory for Neural Network Architecture", Proceeding of the 1993 International Joint Conference on Neural Networks, vol. 1, 1993, pp. 309-312.

[2] E. Eryurek, and B.R. Upadhyaya, "Sensor validation for power plants using adaptive backpropagation neural network", IEEE Transactions on Nuclear Science, Vol. 37 , 1990, no. 2.

[3] G. Mirchandani, and W. Cao, "On Hidden Nodes for Neural Nets, *IEEE Transactions on Circuits ans Systems*, vol. 36, no. 5, 1989, pp.661-664.

[4] R.P. Lippmann, "An Introduction to Computing with Neural Nets", IEEE ASSP Mag., vol. 2, 1988, pp. 75-89.

[5] D.E. Rumelhart, J.L. McClelland and PDP Research Group, "Parallel Distributed Processing, Volume 1, Explorations in the Microstructure of Cognition: Foundations", MIT Press, 1988, pp.45-76.

[6] C.E. Shannon, A Mathematical Theory of Communication, Bell Systems Technical Journal, vol. 27, 1948, pp. 623-656.

[7] L. Boltzmann, (1886). "The Second Law of Thermodynamics" (pgs. 14-32; struggle quote, pg. 24, affinity and energy set free, pgs. 26-27). In B. McGinness, ed., Ludwig Boltzmann: Theoretical physics and Philosophical Problems: Selected Writings. Dordrecht, Netherlands: D. Reidel, 1974.

[8] M. van Gerven, and S. Bohte, "Artificial Neural Networks as Models of Neural Information Processing", Frantiers in Computational Neuroscience, vol. 11, 2017, pp. 1-2.

[9] S. Haykin, "Cognitive Dynamic Systems", Cambridge Univiversity Press, 2012, pp.14-42.

[10] E. Hollnagel, "Cognitive System Performance Measurement", In E. Hollnagel G. Mancini, and D.D. Woods (Eds.), Intelligent Decision Aids. New York: Springer-Verlag, 1986.

[11] E. Hollnagel, and D.D. Woods, "Cognitive System Engineering: New wine in the new bottles", International Journal of Man-Machine Studies, vol. 18, 1983, pp. 583-600.

[12] D.S. Chaplot, C. MacLellan, R. Salakhutdinov, amd K. Koedinger, "Learning Cognitive Models Using Neural Networks", arXiv: 1806.080665v1, (arxiv.org/pdf/1806.08065.pdf), June 2018.

BIOGRAPHIES

Alla Horiushkina was born in Ukraine. She received the PhD degree in Computer science from National Technical univerity "Kharkov Politechnical Institute" (NTU "KhPI") in 2016. She became associate professor in 2017 at the same university. Her current research interests are Signal Processing, Computational Intelligence, and Computational Cognition.

Serhat Seker started his education at Mathematics Engineering Department of Istanbul Technical University (ITU) and graduated from Electrical and Electronic Engineering Faculty where he started his career as a research assistant at Electrical Power Engineering Department in 1985. He got his master degree at ITU's Nuclear Energy Institute and his PhD at Electrical Engineering Division of the same university's Science and Technology Institute with his research titled as "Stochastic Signal Processing with Neural Network in Power Plant Monitoring." Dr. Seker studied during his PhD thesis at Energy Research Centre of the Netherlands (ECN) with his scholarship provided from Netherlands Organization for International Cooperation in Higher Education (NUFFIC) and worked on signal analysis techniques there. He was titled as Assistant Professor and Associate Professor at ITU in 1995 and 1996 respectively. Also, he worked on industrial signal processing at Nuclear Engineering Department and Maintenance and Reliability Centre of the University of Tennessee, Knoxville-USA, by getting the scholarship from the Scientific and Technological Research Council of Turkey (NATOTUBITAK B1) in 1997. He had many administrative duties at ITU in the previous years, including his vice dean position at Electrical and Electronic Engineering Faculty during 2001 - 2004. He was a department head in Electrical engineering department between 2004 and 2007. Also, he was dean of Technology Faculty and vice rector as well as the founding dean of Engineering Faculty at Kırklareli University between 2011 and 2013. Dr. Seker is still the Dean of the Electrical and Electronics Engineering Faculty at ITU since 2013.

Ufuk Korkmaz was born in Turkey. He received the BSc, MSc and PhD degrees from the Ondokuz Mayıs University (OMU), Physics Department, in 2006, 2010 and 2014 respectively. He starting researches as Post-Doc in Istanbul Technical University (ITU) in 2018. His research interests are IR and UV spectroscopy, X-ray single crystal diffraction, Understanding the nature of H bonds in supramolucular stracture, Quantum Mechanics and Quatum information theory.