Artificial Neural Networks - State-of-the-Art *

Hristo I. Toshev[1], Chavdar D. Korsemov[2], Stefan L. Koynov[3] and Vesselin P. Velichkov[4]

Abstract: - **The paper is a review of the state-of-the-art for the artificial neural networks. The background for their application begins with comparisons with the biological prototypes and the classical von-Neuman architecture. Artificial neural networks prototypes in biology and their technological equivalent are presented Advantages and disadvantages of neural networks are given.**

Keywords - **artificial neural networks; artificial neural network properties; artificial neural network applications.**

I. INTRODUCTION - MATHEMATICAL EQUATIONS, NEURAL NETWORKS AND FUZZY SYSTEMS

The description of (and knowledge of) a system, according to [1] can be done in three different ways, namely: with mathematical equations, distributed parameters (neural networks – NNs), and linguistic rules.

Despite their obvious simplicity mathematical equations are not applicable to complex systems because of some similarly complex reasons. In this approach, besides defining the exact relationship between the varying parameters, also the time variable has to be taken into account (speaking of time-dependent systems).

The latter inconvenience seems to be overcome in the case of linguistic rules, which can be easily modified. They include even badly defined languages, which allow controversial conclusions from one and the same fact.

Mathematical equations and “crisp” rules of the type “if-then” describe an algorithm of a process manifestly. NNs describe an algorithm of a process in a hidden way. They are atypical example of distributed data processing. Artificial NNs can be examined as iterative systems in two different typical example of distributed data processing. Artificial NNs can be examined as iterative systems in two different ways.On one hand they are structurally - iterative, as their structure is a simple iteration of their components; in this way they can be interpreted as a structured graph. On the other hand artificial NNs are algorithmically -iterative, the aim of the algorithm being to define the centroids of different classes if the system is self-organizing. If the system is not self-organizing, the centroids of its classes are “taught” by a teacher. The drawbacks of NNs come from the unpredictability of the system for every moment in time, as well as from the non-implicit convergence for every separate case (because of which, instead of global convergence**,** an asymptotic convergence is usually pursued); even small variations of input data cause new knowledge to be learned by the network. Finally, it must be said that NNs give in to fuzzy systems in terms of ease of development. However, in spite of its disadvantages, NNs along with fuzzy systems are much more suitable for solving badly defined problems, compared to the systems with mathematical equations.

The merging of fuzzy systems and NNs into so-called “adaptive fuzzy systems” is a comparatively old technique in the field ([2]-[11]). Contemporary tendencies are towards merging fuzzy logic with chaos and NNs. Example of a fuzzy system is the modeling of human brain and an example of chaos in terms of informatics and NNs are the nonlinear dynamic processes of the complex neural nets in the human brain.

II. ANN ARTIFICIAL NEURAL NETWORKS PROTOTYPES IN BIOLOGY AND THEIR TECHNOLOGICAL EQUIVALENT

Comparison between biological neural network system and Fon Neumann computer

The upper equivalence follows from [12]. Next are listed the main components of the Fon Neumann architecture (fNA), followed by the characteristics of the classical computer architecture and of the biological neural system (BNS).

Processor. While the fNA-processor is complicated, BNS processor is simple.

Memory. While fNA-memory is divided from fNA-processor and is localized and content addressable, BNS-memory is integrated into the BNS-processor and is distributed and content addressable.

Processing. While fNA-processing is centralized, sequential and is based on the programs in memory, BNS-processing is distributed, parallel and self-learnt.

Reliability. While fNA-reliability is very vulnerable, BNS-reliability is robust.

* This study is partly supported by the Ministry of Education and Science, National Science Fund, contract No. I-1302 / 2003, Sofia, Bulgaria.

[1] Hristo I. Toshev is with the Institute of Information Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl. 29A, 1113 Sofia, Bulgaria, E-mail: toshev@iinf.bas.bg

[2] Chavdar D. Korsemov is with the Institute of Information Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl. 29A, 1113 Sofia, Bulgaria, E-mail: korsemov@iinf.bas.bg

[3] Stefan L. Koynov is with the Institute of Information Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl. 29A, 1113 Sofia, Bulgaria, E-mail: baruch@iinf.bas.bg

Vesselin P. Velichkov is with the Institute of Information Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl. 29A, 1113 Sofia, Bulgaria, E-mail: vesko@iinf.bas.bg

Information representation and basic tasks. Information in fNA is in the form of numbers and symbols and this is why the main problems of fNA are in the data representation. In contrast to fNA, information in BNS has impulse-analog representation, and the main problems come from analysis/recognition of input data.

Work environment. While in fNA it is well defined, the work environment in BNS is badly defined and its limits have to be accepted by default.

Models of evolutionary computation in NNs

According to the approach of evolutionary computation, which has gathered much popularity lately, NNs are one of the main representatives of the FOE-model (filogenesis, ontogenesis, epigenesis) of simulated computation [13].

After they reach a certain level of complexity, living organisms develop highly specialized processes, which allow an individual to integrate the huge amount of interactions with environment. In relation to living organisms such processes are referred to as epigenesis.

Three systems in living organisms represent epigenesis: nervous system, immune system and endocrine system. From the three systems, the one receiving greatest attention from researchers is the nervous system. A typical example in that respect is the nervous system, which according to [14] includes 10^{10} neurons and 10^{14} synapses, if compared to the four symbol genome, which has length $3 \bullet 10^{9}$, according to [13]. The immune system is a prototype of systems for finding software bugs [15], of mobile robots controllers [16], and of immune systems of computers [17]. The endocrine system is composed of a large number of gland tissues, which release directly into the blood stream, hormones, controlling bodily functions as reproduction, etc. From functional point of view this system is similar to some extent to the nervous system in the sense that both of them help the individual's adaptation to changes in environment.

Artificial NNs are the technical equivalent of the three above mentioned biological systems, their synaptic weights and their topology changing in response to outside influence. The paradigm of artificial NNs is a technical realization of the epigenetic axis of the FOE-model and as such it acquires amazing results, which very often are equal to, if not exceeding, the results of the traditional methods. Artificial NNs are applied mainly in programming and in small number of cases in hardware equipment. Some of their applications in practice are: data analysis, function approximation, associations, inter- and intra class categorizations, recognition and classification of images, data compression, prediction and control [12].

In recent years the attention of researchers has been centered around evolutionary artificial NNs, which beside the epigenetic axis of the FOE-model also include its filogenetic axis. According to [18] and [19], filogenesis in living nature covers the evolution of the species itself. The technical solution from NNs's theory point of view is a population of NNs, in which the evolution is realized on global (population's) level, while learning is realized on individual level (i.e. single NN level). Examples in that respect are the works of Liu and Yao [20], Nolfi et al. [21] and Yao [22], although they are completely out-of-date. Another interesting example is the effect of Baldwin, who demonstrates the complex interaction between filogenesis and epigenesis.

The application of the latter process for the uses of simulated computation is examined in [23] and [24]. The FE-plane (filogenesis-epigenesis) in nature is also connected with the process of human linguistic learning i.e. to what extent the linguistic ability is inbred (filogenetic) or acquired (epigenetic).

A short historical review on the subject discussed can be found in [25], and information on this process applied in artificial systems is given in [26]-[28].

Concept of artificial NN

The artificial neural network (NN) is a high-level parallel structure of interconnected simple elements (neurons), which are hierarchically organized. The information in NN is stored in distributed way as weights and links between neurons, and its acquisition is realized through association. The algorithms defining the changes in weights are called training rules. The main models of NNs are examined in [29-31]. The trainable networks have possessed the flexibility, necessary for solving of complex dynamic problems. Such NNs have to be able to adapt on two clearly separable levels: by changing of the dynamics of the inter-neuronal interactions (which is normally achieved through change of the synaptic weights) and by changing the actual topology of the network. Topology modification proves to be a successful solution to the stability-flexibility dilemma, i.e. how a trainable system may remember the already acquired knowledge, while at the same time it continues to add new knowledge [32].

III. Advantages and Disadvantages of NN

Advantages

The strongest, and at the same time the most specific features of a NN are: adaptivity, robustness to fuzzy and noisy input, reliability:

Adaptivity. There is no inlaid, concrete algorithm in a NN; the training of a NN is done by example and it can continue during the use of the NN in real applications.

Robustness to fuzzy and noisy input. The NN finds the "true" path by association.

Reliability. Even the removing of some neurons is not fatal as the information in the NN is distributed.

Disadvantages

The main problems when working with artificial NNs are: the curse of dimensionality, number of layers and number of neurons in each layer (network design problems), selection of energy (step) function along with interpretation on physical level, convergence of the network in case of complex definition domains of input vectors (classification problems), time for training/self-organization.

Curse of dimensionality. In most of the cases in practice, the input vectors of the NN are composed of hundreds or even thousands of components. This makes necessary a reduction of dimensions to be made. The latter is realized during preprocessing, when the highly informative components are separated from the less informative ones; data compression is a synonym of preprocessing. (There are situations in which it is necessary to artificially generate additional samples, as in the case of image recognition). Other methods for dimensionality decrease is the use of some symmetry in the training sequence or to use the principle of unification, i.e. to use the knowledge for a single object as valid for the rest of the objects of the respective class as well.

Network design problem. Another fundamental problem appears to be the problem of defining the number of hidden neurons, i.e. the number of the neurons positioned between the input and output layers. What makes this problem so crucial for the design of the NN is the fact that from the number of hidden neurons the properties of the network are defined in respect to the concrete application. If theoretically the addition extra neurons results in settling into a local minimum, then the removing of neurons corresponds to the case of convergent network.

Selection of energy function (physical level interpretation). Also very important proves to be the problem of how to escape the local minimums in our search of the global one. A number of physical in their essence methods are used. The more popular of them are the Hebb training and the Bolzman; beside them other methods are also used, among which – statistical, binary, neuro-dynamical, etc. On mathematical level the physical method can be explained as the selection of the type of the energy function.

Convergence of the network in complex definition domains of the input vectors (classification problems). From samples' classification/recognition point of view the problem of the complex definition domains of the input data is very essential. Very often instead of looking for absolute convergence, it is looked for asymptotic.

Training/self-organizing time. The training of a NN can be realized in two ways, each of them characterized by its own training time. Training can be done "per sample" (guarantees quicker result acquisition) or over the whole sample sequence – packet training (slower than the first case). Although in the case of packet training the weights' changes for each sample have to be kept in memory through the whole epoch, it is guaranteed that the error will slide down over the weight-plane. The time for packet training is called "epoch".

IV. Conclusions

The paper is a review of the state-of-the-art for the artificial neural networks. The background for their application begins with comparisons with the biological prototypes and the classical von-Neuman architecture. Artificial neural networks prototypes in biology and their technological equivalent are presented. Advantages and disadvantages of NN are given.

References

[1] T. Yamakawa, A fuzzy inference engine in nonlinear analog mode and its application to a fuzzy logic control, *IEEE Trans. on Neural Networks*, V. 4, N. 3, pp. 496-521, May 1993.

[2] T. Yamakawa, Japanese Patent Applications, Tokuganhei, No. 1-043600, Feb. 23, 1989, No. 1-133690, May 26, 1989

[3] T. Yamakawa A fuzzy neuron and its application to pattern recognition, Proc. Third IFSA Congress, Seattle, WA, Aug. 6-11, 1989, pp. 30-38, 1989

[4] T. Yamakawa Pattern recognition hardware system employing a fuzzy neuron, Proc. Int. Conf. Fuzzy Logic & Neural Networks, Iizuka, Japan, July 20-24, 1990, 943-948, 1990.

[5] T. Yamakawa A fuzzy neuron and its application to a hand-written character recognition system, Proc. IEEE Int. Symp. Circuits and Systems, Singapore, June 11-14, 1991, 1369-1372, 1991.

[6] Proc. Int. Conf. Fuzzy Logic and Neural Networks, Iizuka, Japan, July 20-24, 1990.

[7] T. Yamakawa A design algorithm of membership functions for a fuzzy neuron using example-based learning, Proc. IEEE Int. Conf. Fuzzy Systems, San Diego, CA, Mar. 8-12, 1992, pp.75-82, 1992.

[8] B. Kosko, Neural networks and fuzzy systems, Englewood Cliffs, NJ, Prentice Hall, 1992.

[9] B. Kosko, Neural networks for signal processing, Englewood Cliffs, NJ, Prentice Hall, 1991.

[10] Proc. A Second Int. Conf. Fuzzy Logic and Neural Networks, Iizuka, Japan, July 17-22, 1992.

[11] T. Yamakawa , *et al.*, A neo fuzzy neuron and its applications to system identification and prediction of the system behavior, Proc. Second Int. Conf. Fuzzy Logic & Neural Networks, Iizuka, Japan, July 17-22, 1992, pp. 477-484, 1992.

[12] A. Jain, A. K., Artificial neural networks: a tutorial, *IEEE Computer*, Mar. 1996, pp. 31-44, 1996.

[13] M. Sipper, *et al.*, A phylogenetic, ontogenetic, and epigenetic view of bioinspired hardware systems, *IEEE Trans. on Evol. Comput.*, Apr. 1997, V. 1, N. 1, pp. 83-97, 1997.

[14] G. M. Shepherd, C. Koch, Introduction to synaptic circuits, The Synaptic Organization of the Brain (G. M. Shepherd, ed.), N.Y., Oxford University Press, 1990, 3-31, 1990.

[15] S. Xanthakis, ,*et al.*, Immune system and fault-tolerant computing, Evolution artificielle 94, 1995, Cepadues, cop., 1994.

[16] A. Ishiguro, *et al.*, Immunoid: an immunological approach to decentralized behavior arbitration of autonomous mobile robots, Parallel Problem Solving from Nature - PPSN IV (*Lecture notes in Computer Science*, V. 1141), H.-M. Voigt, W. Ebeling, I. Rechenberg, H.-P. Schwefel, Eds., Heidelberg, Springer-Verlag, 1996, 666-675, 1996.

[17] J. O. Kephart, A biologically inspired immune system for computers, Artificial Life IV, R. A. Brooks, P. Maes, Eds., Cambridge, MA, MIT Press, 1994, pp. 130-139, 1994.

[18] A. Danchin, A selective theory for the epigenetic specification of the monospecific antibody production in single cell lines, *Ann. Immunol. (Institut Pasteur)*, V. 127C, 1976, 787804, 1976.

[19] A. Danchin, Stabilization fonctionelle et epigenese: une approche biologique de la genese de l'identite individuelle, L'identite, J.-M. Benoist, Ed., Paris, France, Grasset, 1977, 185-221, 1977.

[20] Y. Liu, X. Yao, Evolutionary design of artificial neural networks with different nodes, Proc. IEEE Int. Conf. Evol. Comput. (ICEC'96), 1996, 670-675, 1996.

[21] S. Nolfi, *et al.*, Learning and evolution in neural networks, Adaptive Behavior, V. 3, N. 1, 1994, 5-28, 1994.

[22] X. Yao, Evolutionary artificial neural networks, *Int. J. Neural Syst.*, V. 4, N. 3, 1993, 203-222, 1993.

[23] G. E. Hinton, S. J. Nowlan, How learning can guide evolution, *Complex Syst.*, V. 1, 1987, 495-502, 1997.

[24] D. Ackley, M. Littman, Interactions between learning and evolution, *Artificial Life II (SFI Studies in the Sciences of Complexity*, V. X), C. G. Langton, C. Taylor, J. D. Farmer, S. Rasmussen, Eds., Redwood City, CA, Addison-Wesley, 1992, 487-509, 1992.

[25] D. Dennet, C., Darwin's dangerous idea: evolution and the meanings of life, N.Y., Simon & Schuster, 1995.

[26] G. M. Werner, M. G. Dyer, Evolution of communication in

artificial organisms, *Artificial Life II (SFI Studies in the Sciences of Complexity*, V. X), C. G. Langton, C. Taylor, J. D. Farmer, S. Rasmussen, Eds., Redwood City, CA, Addison-Wesley, 1992, 659-687, 1992.
[27] B. MacLennon, Synthetic ethology: an approach to the study of communication, *Artificial Life II (SFI Studies in the Sciences of Complexity*, V. X), C. G. Langton, C. Taylor, J. D. Farmer, S. Rasmussen, Eds., Redwood City, CA, Addison-Wesley, 1992, 631-658, 1992.
[28] L. Steels, A self-organizing spatial vocabulary, *Artificial Life*, V. 2, N. 3, 1995, 319-332, 1995.
[29] S. Grossberg, The adaptive brain I: cognition, learning, reinforcement, and rythm; The adaptive brain II: vision, speech, language and motor control, Elsevier/NorthHolland, Amsterdam, 1986.
[30] R. P. Lippmann, An introduction to computing with neural nets, *IEEE ASSP Magazine*, April 4-22, 1987.
[31] P. Wasserman, Neural computing - theory & practice, N.Y., 1989.
[32] G. Carpenter, S. Grossberg, The ART of adaptive pattern recognition by a self-organizing neural network, *IEEE Computer*, Mar. 1988, 77-88, 1988.