Comparative Analysis of Genetic and Evolutionary Approaches in the Study of Multilayer Models of the Artificial Neural Networks

Hristo I. Toshev[1] and Chavdar D. Korsemov[2]

Abstract - **The paper makes comparative analysis of artificial neural networks and of evolutionary computation.Multilayer models are introduced with some of their main features from a genetic or evolutionary viewpoint. For the purpose of the study mathematic formalism is applied by penalty functions, based on an apriori set number of the next generations.**

Keywords - **genetic approaches, evolutionary computation, artificial neural networks, multilayer models.**

I. Introduction

Multilayer models are introduced with some of their basic properties from the genetic and evolutionary points of view. Examples of such models are the generalized evolutionary artificial neural network design model (GEANNDM) and the artificial neural network design model (ANNDM). The chosen ANN models (including the design-, the optimization- and the learning- phases) are animated starting from an evolutionary computation (EC) background, including genetic synthesis. A mathematical formalism is presented for the task of multilayer model investigation by means of penalty functions based on a predefined number of a series of generations.

II. The ANN Models Methodology

About The Physical Nature Of The ANN Application Tasks

Any specific task has its own concrete physical nature. Therefore the common aspects of the different *categories* or *clusters* from the ANN-application tasks loose from their distinctness. The categorization and the clusterization of the ANN-application tasks may serve as a true prerequisite for the *identification of the common properties of the tasks*. This will guarantee the enhancement and the acceleration of the ANN-design process.

About The Initial Setting of the Possible ANN Architecture

The *initial* setting of the ANN architecture is independent of the ANN parameter optimization and the following training. Therefore it may be done independently of the other sequences in the ANN-design process. The authors present a deductive tree per any of the two ANN architecture types (feedforward and recurrent) in Fig. 1 and Fig. 2. The affirmative branches follow [1] and the most outstanding feature of the affirmative architecture is chosen by the authors for the corresponding fork condition.

1°. If the condition for *linear separability* is valid then the initial architecture is a *single-layer perceptron* else go to 2°.

2°. If the condition for an *approximate optimum* is valid then the initial architecture is a *multilayer perceptron* (*stochastic approximation* is actual) else go to 3°.

3°. If the task is a *multivariable interpolation* then the initial architecture is with *radial basis functions* (*statistical approximation* is actual) else the algorithm ***stops***.

Fig. 1. Feedforward ANN Architecture Initial Setting Algorithm

1°. If the *speed is not critical* to the model performance then the initial architecture is a *Hopfield network* else go to 2°.

2°. If the condition for *pattern stability* is valid then the initital architectture is an *ART* else go to 3°.

3°. If the *dimensions of the input and of the output are equal* then the initial architecture is a *Willshaw - von der Malsburg map* else go to 4°.

4°. If the condition for *data compression* is valid then the initial architecture is a *Cohonen map* else the algorithm ***stops***.

Fig. 2. Recurrent ANN Architecture Initial Setting Algorithm

About The Retraining of the ANN And About Its Upgrade

Based on the genetic and evolutionary approaches the authors offer a *step-by-step optimization* instead of the *element-by-element optimization* (as it is the case with the usual ANN-design process). They consider this aspect as one of the most significant in the ANNDM model. The optimization versions may be defined in different and

[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

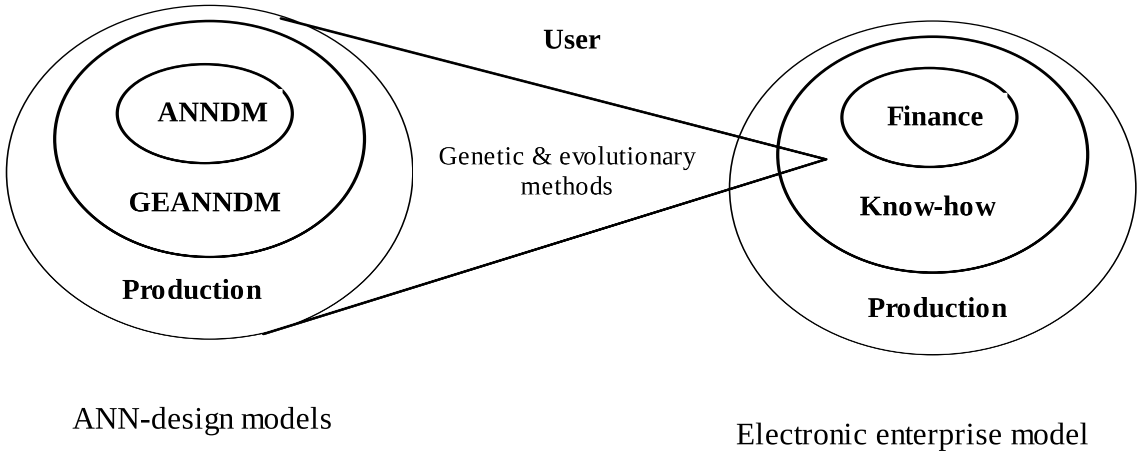

Fig. 3. Interrelations Between Different Multilayer Models

complementary ways which include the simple parameter variations as a private case; these versions are seen to be based on the already mentioned approaches which are not only more powerful but they also include the immediate participation of the designer in the man-machine dialog.

In fact the alternative approach is designed for the effective application of genetic and evolutionary methods. This means that the *fixed* values are replaced by *possible* values. Besides it is possible to output the mathematical formulation of the actual constraints, of the possible crossovers, of the number and of the composition of the generations of the solution elements (i. e. of the admissible given number of done iterations) and also of any possible statistical information. In this sense the authors consider the ANNDM model as a synthesis of a software for mathematical modeling (e. g. MatLab) and of a software for ANN modeling and study (e. g. NeuralWorks Explorer, NeuralWorks Professional II, [2]) *plus additional properties*.

It is clear that this approach lowers the obstacles mainly in the stages of identification of the ANN-application task and also of identification of the training paradigm (supervised, unsupervised or hybrid). The psychological comfort is strongly increased in the form of a proximity and a friendliness for the designer and the apparent reduction of the possible design expenses is already a reality. Also it is quite possible to determine the data representation, the architecture and the training algorithm as a consequence of the identified application task and of the training paradigm respectively.

III. THE MULTILAYER MODELS AS A CONVENIENT BASIS FOR APPLICATIONS OF GENETIC AND EVOLUTIONARY METHODS

The Relations Between The Multilayer Models

Analogical to the presented GEANNDM model is the model of an enterprise for electronic ware shown in Fig. 3. The usual approach may deny the analogy between the two models. Anyway the authors state that both triple-layer models are isomorphic to a number of features the first one being the number of the layers of course (the most evident difference being the transform of the "user-designer" couple into the enterprise working staff). In fact the GEANNDM model is just one of the many possible links of the whole know-how enterprise layer.

Corollary 1. The multilayer model approach makes possible the investigation not only of separate models, but also of relationships between different multilayer models.

Genetic And Evolutionary Approaches To Multilayer Models

The multilayer model search space is denoted with S and $F \subseteq S$ denotes the *feasible* search subspaces. Every such model has its peculiarities or *constraints* which are set by the physical nature of the concrete task. Applied to the two presented models this means that the constraints are defined by the up-to-date development of the science and also by the actual *possibilities* like the investments, production base, qualifications and skill of the working staff. Another feature of the serial manufacture is the *continuity* in the new production models concerning different aggregates and systems of the older models (from the point of view of the genetic and the evolutionary approaches it means that all these possibilities define in a complementary way such properties like the penalties in the evolutionary methods and also the crossovers and the mutations in the new generations for the genetic methods).

Corollary 2 (**for serial manufactures**). Most important are the different generations of serial models (from the genetic point of view) and the *constraints of the design and of the serial production* (from the evolutionary point of view).

Mathematical Formulation Of The Multilayer Models And The Constraints

The authors present the following approach. Let us choose in a random manner two neighbouring layers, the one is relatively peripheral and the other is the relative core. From the point of view of the possible optimizations the default numbering is a basis to accept that the relatively peripheral layer is identified with the whole search space S and that the relative core represents the feasible subspaces F of the solutions. In turn this is a convenient basis for application of the developed evolutionary approaches, namely the penalty functions (the constraints). The constraints may be set in a bidirectional manner as inward or outward constraints concerning the final layer in this micro-bilayer submodel. Therefore it is an option to define the multilayer model in any of the two possible manners with respect to the penalties.

Corollary 3. The multilayer models are a natural basis for application of the evolutionary methods, namely the penalties (the constraints) defined in any of the two layers in the micro-bilayer submodels and oriented to the other layer in them.

It is possible that the default numbering might insert some inconvenience and especially in the case when the physical nature of the layer which is the relative core becomes very important for some aspect of the multilayer model. Then it is possible to make a *version* of the basic model with the *inverse numbering* of at least two neighbouring layers after which the investigation may proceed in the usual way.

Corollary 4. The constraints and the penalties may be introduced by default numberings or by inverse numberings of (versions of) the basic multilayer model. The reason for the inverse numbering usually lies in the physical nature of the relative core which induces *outward* constraints. *It is the physical nature of every layer which makes it unique in the whole multilayer model.*

The Penalty Functions In The Multilayer Models

The authors introduce the mathematical description of the multilayer model such that the penalty functions may be classified in three types: 1) *inside* any concrete layer, 2) *between* any two layers in a single multilayer model, 3) *between any two multilayer models*:

$$\mathrm{eval}(\overline{X}) = f(\overline{X}) + \sum_{l=1}^{L} a_l \left[\lambda(t) \sum_{j=1}^{m} f_j^2(\overline{X}) \right]^l \qquad (1)$$

where:

$\mathrm{eval}(\overline{X})$ — feasible and unfeasible solutions if $\overline{X} \in F$ is the optimal solution of the general non-linear programming model with continuous variables;

$f(\overline{X})$ — goal function for optimization;

$\lambda(t)$ — updated every generation t in the following way [5]:

$$\lambda(t+1) = \begin{cases} (1/\beta_1) \cdot \lambda(t), \text{ if } \overline{B}(i) \in F \text{ for all } t-k+1 \le i \le t \\ \beta_2 \cdot \lambda(t), \text{ if } \overline{B}(i) \in S - F \text{ for all } t-k+1 \le i \le t \\ \lambda(t), \quad \text{else} \end{cases}$$

$f_j(\overline{X})$ — constraint violation measure for the *j*-th constraint such that [6]:

$$f_j(\overline{X}) = \begin{cases} \max\{0, g_j(\overline{X})\}, \text{ if } 1 \le j \le q \\ \left| h_j(\overline{X}) \right|, \quad \text{if } q+1 \le j \le m \end{cases}$$

Here $g_j(\overline{X}) \le 0, j = 1,\dots,q$ and $h_j(\overline{X}) = 0, j = q+1,\dots,m$ is a set of additional constraints $m \ge 0$ the intersection of which with the search space S defines the feasible set F.

l — indicator of the constraint type with upper bound $L = \{2|3\}$:

$$l = \begin{cases} 1: \text{inside a given layer (the lowest constraint level);} \\ 2: \text{between two layers inside a given multilayer model} \\ \quad \text{(the moddle constraint level);} \\ 3: \text{between two multilayer models (the highest const-} \\ \quad \text{raint level).} \end{cases}$$

a_l — coefficient array reflecting the weights of the different constraint levels in the formula. It is adjusted heuristically.

IV. Conclusions

Hierarchical structures of a pyramidal type interpreted with multilayer models are introduced with their advantages and limitations. Hierarchical systems are complemented in a natural way but not overlapped by structures without hierarchies. Both types of organizations arrange and represent *in different ways* objects from a *common* domain. The multilayer models provide the most relevant description of the ensemble of features of the ANN learning process.

References

[1] S. Haykin. *Neural networks*. Englewood Cliffs, Macmillan Publishing Co., 1994.

[2] D. Williams and J. B. Gomm. The introduction of neural network projects in a degree of electrical and electronic engineering. 12th Int. Conf. on Systems for Automation of Engineering and Research SAER'98, Proc. 19-20 September 1998, Varna - St. Konstantin resort, Bulgaria, pp. 102-106, 1998

[3] J. C. Bean and A. B. Hadj-Alouane. A dual genetic algorithm for bounded integer programs. Department of Industrial and Operations Engineering, The University of Michigan, TR 92-53, 1992.

[4] Zb. Michalewicz. The significance of the evaluation function in evolutionary algorithms. *Evolutionary algorithms*, L. D. Davis, K. De Jong, M. D. Vose, L. D. Whitley (eds.), Springer, 1999, pp. 151-166, 1999

[5] E. Falkenauer. Applying genetic algorithms to real-world problems. *Evolutionary algorithms*, L. D. Davis, K. De Jong, M. D. Vose, L. D. Whitley (eds.), Springer, 1999, pp. 65-88, 1999

[6] S. Koynov, Ch. Korsemov and H. Toshev. The design of the artificial neural networks as a basis for their generalized evolutionary model. 15th Int. Conf. on Syst. for Automation of Engineering and Research SAER'2001, Proc. 21-23 September 2001, Varna - St. Konstantin resort, Bulgaria, pp. 185-190, 2001

[7] S. Koynov, Ch. Korsemov and H. Toshev. The artificial-neural-networks-design generalized evolutionary model. 15th Int.

Conf. on Syst. for Automation of Engineering and Research SAER'2001, Proc. 21-23 September 2001, Varna - St. Konstantin resort, Bulgaria, pp. 191-195, 2001

[8] A. Jain. Artificial Neural Networks: A Tutorial. IEEE Computer, 1996, March, pp. 31-44, 1996

[9] D. B. Fogel, E. C. Wasson and E. M. Boughton. Evolving Neural Networks for Detecting Breast Cancer. Cancer Letters, v. 96, 1995, pp.49-53, 1995

[10] V. Ilchev "Transformation of Lineary Separable Sets in Hough Spase", *Comptes rendus de l' Academie Bulgare des Sciences*, Tome 56, 2003, № 1, pp. 31-36, 2003..

[11] L. Docheva, A. Bekiarski, "Analogue Neural Network Comparative Simulation by Means f MATLAB and PSPICE Software Produkts" , Proceedings of the XXXVIII International Scientific Conference on Information, Communication and Energy Systems and Technologies ICEST 2003, Sofia, pp.270-272, 2003.

[12] M. Bonilla, I. Baruch, J. M. Flores, M. M. Goire, B. Nenkova "Model Reference Adaptive Neural Control of a Variable Structure System", - In: *Cybernetics and Information Technologies*, V. 3, No. 2, pp. 3-14. 2003

[13] I. Baruch., J.M.Flores and B. Nenkova, "Adaptive Control Schemes Based on Recurrent Trainable Neural Networks", *Cybernetics and Information Technologies*, Vol. 4, No. 1, Sofia, pp. 89-102, 2004.